



Dissertação realizada sob a supervisão de:

**Professor Doutor Gabriel David**

**Professor Auxiliar do Departamento de Engenharia Electrotécnica e de Computadores**

**Faculdade de Engenharia da Universidade do Porto**

**Doutora Lígia Maria Ribeiro**

**Investigadora Auxiliar**

**Faculdade de Engenharia da Universidade do Porto**

## Resumo

O objectivo principal da presente dissertação é avaliar a aplicabilidade da Extração de Conhecimento em Bases de Dados a uma base de dados académica, tomando como estudo de caso a Faculdade de Engenharia da Universidade do Porto.

A metodologia seguida consistiu na aplicação de algumas técnicas de prospecção de dados (classificação e análise de ligações) aos dados relativos aos últimos cinco anos lectivos. A análise incidiu sobre as interdependências entre disciplinas e o perfil dos alunos baseado no desempenho no primeiro ano.

A partir dos modelos gerados (árvores e regras de decisão, e regras de associação) conseguem-se verificar interdependências entre disciplinas e a importância relativa que estas têm no sucesso escolar de um ou mais anos, fornecendo informação útil para apoio aos órgãos de gestão em medidas correctivas, como aconselhamento aos alunos, detecção precoce de problemas de insucesso e previsões para os anos lectivos seguintes.

**Palavras chave:** Extração de Conhecimento de Bases de Dados, Prospecção de Dados, Bases de Dados, Classificação, Árvores de Decisão, Regras de Decisão, Análise de Ligações, Regras de Associação.

## Abstract

The main goal of the current dissertation is to evaluate the applicability of Knowledge Discovery in Databases to an academic database, using as case study the *Faculdade de Engenharia da Universidade do Porto* (Faculty of Engineering of the University of Porto).

The followed methodology consisted in the application of some data mining techniques (classification and link analysis) to the last five years data. The analysis has been conducted on discipline interdependence and students profile based on first year performance.

From the generated models (decision trees and rules, and association rules) it is possible to verify interdependence between disciplines and the relative importance that they have on the students' success of one or more years, what constitutes useful information to support management in corrective tasks such as student counseling, earlier detection of failure problems and predictions for the following years.

**Keywords:** Knowledge Discovery in Databases, Data Mining, Databases, Classification, Decision Trees, Decision Rules, Link Analysis, Association Rules.



*à minha filha Tatiana  
ao meu marido Jorge  
aos meus pais Olívia e Marinho*

## Agradecimentos

Desejo agradecer a colaboração e apoio prestado por todos aqueles que, em diferentes momentos e por diversas razões, contribuíram para o desenvolvimento da presente tese.

Em particular, quero agradecer aos meus orientadores, Doutora Lígia Ribeiro e Professor Gabriel David, todo o trabalho de orientação e conselhos, sem os quais o resultado final nunca teria sido possível.

Gostaria de agradecer a todos os professores do Instituto Superior de Estudos Empresariais (actualmente Escola de Gestão do Porto) pela sua contribuição para a minha formação, e aos meus colegas, pela amizade e espírito de grupo.

Gostaria de agradecer ao Departamento de Matemática do Instituto Superior de Engenharia do Porto, em especial à Dra. Helena Vieira da Silva, pela oportunidade e apoio que me deram na realização deste mestrado, e também a todos os colegas de trabalho, em particular à Dra. Olga França.

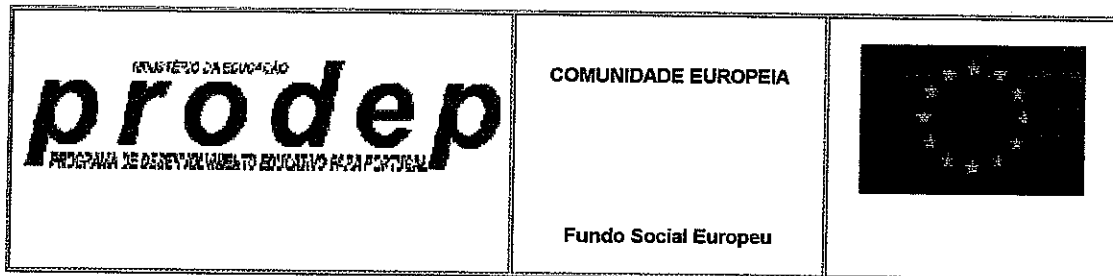
Um agradecimento particular à Doutora Lígia Ribeiro, visto que, sem a sua confiança e ajuda, não me seria possível frequentar este mestrado.

Quero ainda agradecer a toda a minha família pelo apoio e carinho que sempre me dedicaram, muito em especial aos meus queridos pais, Olívia e Marinho.

Finalmente, gostaria de agradecer ao meu marido Jorge, que é o último por ser um agradecimento demasiado grande, e como tal não existiriam páginas suficientes para o descrever. De qualquer forma, não gostaria de deixar de dizer obrigado pelo apoio e compreensão, em especial pela tua constante presença desde os momentos mais fáceis até aos piores, passando mesmo por aqueles em que queremos estar sozinhos. Obrigado ainda por teres abdicado de alguns sonhos apenas para estares presente.

## Apoio financeiro

A presente dissertação foi co-financiada pelo Fundo Social Europeu, ao abrigo da Medida 5/acção 5 - 2 do Programa de Desenvolvimento Educativo para Portugal - PRODEP, designada por Formação Avançada no Ensino Superior.



# Índice

Índice de Tabelas .....	11
Índice de Figuras .....	14
1. Introdução .....	19
1.1. Considerações gerais .....	19
1.2. Objectivos e enquadramento .....	21
1.3. Resultados atingidos .....	22
1.4. Plano da dissertação .....	22
2. Extracção de Conhecimento em Bases de Dados .....	24
2.1. Considerações gerais .....	24
2.2. Aplicações de ECBD e prospecção de dados .....	28
2.2.1. Automação da análise e catalogação de pesquisa do espaço .....	28
2.2.2. Selecção e relatórios sobre o que é interessante .....	29
2.2.3. Modelação de incerteza subjectiva em anotação de imagem .....	29
2.2.4. Predição de retorno de investimento em títulos .....	30
2.2.5. Encontrar associações em colecções de texto .....	30
2.2.6. Aprendizagem de padrões em imagens .....	31
2.2.7. Investigação empírica do fenómeno da expressão musical .....	32
2.2.8. <i>WebWatcher</i> : aprendizagem na <i>World Wide Web</i> .....	32
2.3. Algoritmos e técnicas de prospecção de dados .....	33
2.3.1. Classificação .....	34
2.3.1.1. Árvores e Regras de Decisão .....	35
2.3.1.2. Algoritmos C4.5 e C5.0 .....	37
2.3.1.3. Avaliação dos resultados .....	40
2.3.1.4. Evolução do algoritmo .....	41
2.3.2. Análise de ligações .....	42
2.3.2.1. Regras de Associação .....	42
2.3.2.2. Algoritmo Apriori .....	43
2.4. Ferramentas de prospecção de dados .....	47
2.4.1. Ferramentas de análise de ligações .....	47
2.4.2. Ferramentas de visualização de panorama .....	48
2.4.3. Ferramentas de prospecção de dados quantitativa .....	48

2.4.4. Clementine.....	49
3. Aplicação a uma base de dados académica.....	52
3.1. Preparação de dados.....	52
3.2. Objectivos da prospecção de dados .....	58
3.3. Metodologia.....	59
3.3.1. Árvores e regras de decisão .....	59
3.3.2. Regras de associação.....	60
3.3.3. Filtragem das regras de associação .....	62
3.4. Aplicação aos dados.....	67
3.4.1. Considerações gerais.....	67
3.4.2. Geração dos modelos .....	68
3.4.3. Teste dos modelos.....	70
4. Análise das árvores e regras de decisão.....	72
4.1. Primeiro ano .....	72
4.2. Segundo ano .....	76
4.3. Terceiro ano.....	83
4.4. Quarto ano .....	89
4.5. Quinto ano.....	96
4.6. Perfil dos alunos .....	99
4.7. Discussão dos resultados obtidos .....	103
5. Análise das regras de associação.....	106
5.1. Primeiro ano .....	106
5.2. Segundo ano .....	113
5.3. Terceiro ano.....	124
5.4. Quarto ano .....	132
5.5. Quinto ano.....	140
5.6. Perfil de alunos .....	142
5.7. Discussão de resultados .....	144
6 Conclusões .....	146
6.1. Discussão do trabalho realizado .....	146
6.2. Comparação entre os modelos gerados.....	147
6.3. Discussão de alternativas .....	149
6.4. Sugestões para trabalho futuro .....	150
Referências .....	151

ANEXOS .....	153
I. Árvores e Regras de Decisão .....	154
II. Análise de Árvores e Regras de Decisão .....	166
III. Regras de Associação .....	178

## Índice de Tabelas

Tabela 2.1 - Conjunto de treino para o PlayTennis .....	39
Tabela 2.2 – Conjunto de transações .....	43
Tabela 3.1 – Tabelas usadas na prospecção de dados.....	53
Tabela 3.2 – Áreas científicas .....	54
Tabela 3.3 - Disciplinas de Engenharia Informática e Computação .....	55
Tabela 4.1 -Modelos do C5.0 para algumas disciplinas do primeiro ano.....	74
Tabela 4.2 - Posição das disciplinas nas árvores de decisão.....	75
Tabela 4.3 - Percentagem de acertos para o primeiro ano .....	75
Tabela 4.4 - Modelos do C5.0 para algumas disciplinas do segundo ano .....	76
Tabela 4.5 - Posição das disciplinas nas árvores de decisão.....	77
Tabela 4.6 - Percentagem de acertos para o segundo ano.....	78
Tabela 4.7 - Modelos do C5.0 para algumas disciplinas do segundo ano .....	79
Tabela 4.8 - Modelos do C5.0 para algumas disciplinas do segundo ano .....	80
Tabela 4.9 - Posição das disciplinas nas árvores de decisão.....	82
Tabela 4.10 - Percentagem de acertos para o segundo ano.....	82
Tabela 4.11 - Modelos do C5.0 para algumas disciplinas do terceiro ano .....	84
Tabela 4.12 - Posição das disciplinas nas árvores de decisão.....	84
Tabela 4.13 - Percentagem de acertos para o terceiro ano.....	85
Tabela 4.14 - Modelos do C5.0 para algumas disciplinas do terceiro ano .....	87
Tabela 4.15 - Posição das disciplinas nas árvores de decisão.....	88
Tabela 4.16 - Percentagem de acertos para o terceiro ano.....	89
Tabela 4.17 - Modelos do C5.0 para algumas disciplinas do quarto ano .....	90
Tabela 4.18 - Posição das disciplinas nas árvores de decisão.....	91
Tabela 4.19 - Percentagem de acertos para o quarto ano.....	91
Tabela 4.20 - Modelos do C5.0 para as disciplinas do quarto ano.....	93
Tabela 4.21 - Posição das disciplinas nas árvores de decisão.....	94
Tabela 4.22 - Percentagem de acertos para o quarto ano.....	95
Tabela 4.23 - Modelos do C5.0 para as disciplinas do quinto ano.....	97
Tabela 4.24 - Posição das disciplinas nas árvores de decisão.....	98
Tabela 4.25 - Percentagens de acertos para o quinto ano .....	99
Tabela 4.26 – Resultados ao fim de cinco anos .....	99

Tabela 4.27 - Resultados ao fim de três anos.....	100
Tabela 4.28 - Resultados ao fim de três anos.....	100
Tabela 4.29 - Árvores e regras de decisão relativos ao perfil dos alunos.....	101
Tabela 4.30 - Árvores e regras de decisão relativos ao perfil dos alunos.....	102
Tabela 5.1 - Número de regras geradas para o primeiro ano.....	107
Tabela 5.2 - Relação entre as disciplinas do primeiro ano.....	113
Tabela 5.3 - Número de regras geradas para o segundo ano.....	114
Tabela 5.4 - Relação entre as disciplinas do segundo ano.....	119
Tabela 5.5 - Número de regras de associação para o segundo ano.....	120
Tabela 5.6 - Regras de associação para as disciplinas do segundo ano.....	121
Tabela 5.7 - Regras de associação para as disciplinas do segundo ano.....	122
Tabela 5.8 - Relação entre as disciplinas dos primeiro e segundo anos.....	122
Tabela 5.9 - Comparação entre as regras de associação.....	123
Tabela 5.10 - Número de regras geradas para o terceiro ano.....	124
Tabela 5.11 - Relação entre as disciplinas do terceiro ano.....	129
Tabela 5.12 - Número de regras de associação para o segundo ano.....	129
Tabela 5.13 - Regras de associação para o terceiro ano.....	130
Tabela 5.14 - Regras de associação para o terceiro ano.....	131
Tabela 5.15 - Relação entre as disciplinas dos segundo e terceiro anos.....	131
Tabela 5.16 - Comparação entre as regras de associação.....	132
Tabela 5.17 - Número de regras geradas para o terceiro ano.....	133
Tabela 5.18 - Relação entre as disciplinas do quarto ano.....	138
Tabela 5.19 - Número de regras de associação para o quarto ano.....	139
Tabela 5.20 - Relação entre as disciplinas dos terceiro e quarto anos.....	139
Tabela 5.21 - Comparação entre as regras de associação.....	140
Tabela 5.22 - Número de regras de associação para o quinto ano.....	141
Tabela 5.23 - Relação entre as disciplinas dos quarto e quinto anos.....	141
Tabela 5.24 - Melhores suporte e confiança das regras de associação.....	142
Tabela 5.25 - Regras de associação relativa às regras de associação.....	143
Tabela I.1 - Árvores e regras de decisão para o primeiro ano.....	155
Tabela I.2 - Árvores e regras de decisão para o segundo ano (semestral).....	156
Tabela I.3- Árvores e regras de decisão para o segundo ano (anual).....	157
Tabela I.4- Árvores e regras de decisão para o segundo ano (anual).....	158
Tabela I.5- Árvores e regras de decisão para o segundo ano (anual).....	159



Tabela I.6– Árvores e regras de decisão para o terceiro ano (semestral) .....	160
Tabela I.7– Árvores e regras de decisão para o terceiro ano (anual).....	161
Tabela I.8– Árvores e regras de decisão para o terceiro ano (anual).....	162
Tabela I.9– Árvores e regras de decisão para o quarto ano (semestral) .....	163
Tabela I.10– Árvores e regras de decisão para o quarto ano (anual).....	164
Tabela I.11– Árvores e regras de decisão para o quinto ano (anual).....	165
Tabela II.1– Análise das árvores e regras de decisão para o primeiro ano (sem.) ...	167
Tabela II.2– Análise das árvores e regras de decisão para o segundo ano (sem.)....	168
Tabela II.3– Análise das árvores e regras de decisão para o segundo ano (anual)...	169
Tabela II.4– Análise das árvores e regras de decisão para o segundo ano (anual)...	170
Tabela II.5 – Análise das árvores e regras de decisão para o terceiro ano (sem.)....	171
Tabela II.6–Análise das árvores e regras de decisão para o terceiro ano (anual).....	172
Tabela II.7– Análise das árvores e regras de decisão para o terceiro ano (anual)....	173
Tabela II.8– Análise das árvores e regras de decisão para o quarto ano (sem.).....	174
Tabela II.9– Análise das árvores e regras de decisão para o quarto ano (anual).....	175
Tabela II.10– Análise das árvores e regras de decisão para o quarto ano (anual)....	176
Tabela II.11– Análise das árvores e regras de decisão para o quinto ano (anual)....	177

## Índice de Figuras

Figura 2.1 - Passos no processo de ECBD .....	25
Figura 2.2 – Prospecção de dados e inteligência de negócio .....	26
Figura 2.3 – Árvore de decisão do <i>PlayTennis</i> .....	36
Figura 2.4 - Entropia em funções booleanas .....	38
Figura 2.5 - Algoritmo apriori .....	45
Figura 2.6 - Função apriori_gen .....	45
Figura 2.7 - Passos do algoritmo apriori .....	46
Figura 3.1 - Relações entre as tabelas .....	53
Figura 3.2 - Primeiro passo do algoritmo .....	63
Figura 3.3 - Segundo passo do algoritmo .....	63
Figura 3.4 – Regras de associação para a disciplina EIC1200 .....	64
Figura 3.5 - Primeiro passo do filtro .....	65
Figura 3.6 - Segundo passo do filtro .....	65
Figura 3.7 - Resultado final após aplicação do filtro .....	66
Figura 3.8 - Saída produzida pelo filtro .....	66
Figura 3.9 - Árvore de decisão para a disciplina EIC1200 .....	68
Figura 3.10 - Árvore de decisão para a disciplina EIC1200 .....	69
Figura 3.11 - Regras de decisão para a disciplina EIC1200 .....	70
Figura 3.12 - Análise de testes relativos à disciplina EIC1200 .....	71
Figura 5.1 - Regras de associação para a disciplina EIC1200 .....	107
Figura 5.2 - Regras de associação para a disciplina EIC1202 .....	109
Figura 5.3 - Regras de associação para a disciplina EIC1203 .....	110
Figura 5.4 - Regras de associação para a disciplina EIC1204 .....	110
Figura 5.5 - Regras de associação para a disciplina EIC1205 .....	111
Figura 5.6 - Regras de associação para a disciplina EIC1206 .....	112
Figura 5.7 - Regras de associação para a disciplina EIC2200 .....	114
Figura 5.8 - Regras de associação para a disciplina EIC2201 .....	115
Figura 5.9 - Regras de associação para a disciplina EIC2202 .....	116
Figura 5.10 - Regras de associação para a disciplina EIC2203 .....	117
Figura 5.11 - Regras de associação para a disciplina EIC2204 .....	118
Figura 5.12 - Regras de associação para a disciplina EIC2205 .....	118

Figura 5.13 - Regras de associação para a disciplina EIC3200.....	124
Figura 5.14 - Regras de associação para a disciplina EIC3201.....	125
Figura 5.15 - Regras de associação para a disciplina EIC3202.....	126
Figura 5.16 - Regras de associação para a disciplina EIC3203.....	127
Figura 5.17 - Regras de associação para a disciplina EIC3204.....	127
Figura 5.18 - Regras de associação para a disciplina EIC3205.....	128
Figura 5.19 - Regras de associação para a disciplina EIC4200.....	134
Figura 5.20 - Regras de associação para a disciplina EIC4201.....	134
Figura 5.21 - Regras de associação para a disciplina EIC4202.....	135
Figura 5.22 - Regras de associação para a disciplina EIC4203.....	136
Figura 5.23 - Regras de associação para a disciplina EIC4205.....	137
Figura III.1 - Regras de associação: disciplina EIC1200.....	179
Figura III.2 - Regras de associação para a disciplina EIC1200 após o filtro.....	179
Figura III.3 - Regras de associação: disciplina EIC1202.....	180
Figura III.4 - Regras de associação: disciplina EIC1202 após o filtro.....	180
Figura III.5 - Regras de associação: disciplina EIC1203.....	181
Figura III.6 - Regras de associação: disciplina EIC1203 após o filtro.....	181
Figura III.7 - Regras de associação: disciplina EIC1204.....	182
Figura III.8 - Regras de associação: disciplina EIC1204 após o filtro.....	182
Figura III.9 - Regras de associação: disciplina EIC1205.....	183
Figura III.10 - Regras de associação: disciplina EIC1205 após o filtro.....	183
Figura III.11 - Regras de associação: disciplina EIC1206.....	184
Figura III.12 - Regras de associação: disciplina EIC1206 após o filtro.....	184
Figura III.13 - Regras de associação: disciplina EIC2200 (sem.).....	185
Figura III.14 - Regras de associação: disciplina EIC2200 (sem.) após o filtro.....	185
Figura III.15 - Regras de associação: disciplina EIC2201 (sem.).....	186
Figura III.16 - Regras de associação: disciplina EIC2201 (sem.) após o filtro.....	186
Figura III.17 - Regras de associação: disciplina EIC2202 (sem.).....	187
Figura III.18 - Regras de associação: disciplina EIC2202 (sem.) após o filtro.....	187
Figura III.19 - Regras de associação: disciplina EIC2203 (sem.).....	188
Figura III.20 - Regras de associação: disciplina EIC2203 (sem.) após o filtro.....	188
Figura III.21 - Regras de associação: disciplina EIC2204 (sem.).....	189
Figura III.22 - Regras de associação: disciplina EIC2204 (sem.) após o filtro.....	189
Figura III.23 - Regras de associação: disciplina EIC2205 (sem.).....	190

Figura III.24 - Regras de associação: disciplina EIC2205 (sem.) após o filtro.....	190
Figura III.25 - Regras de associação: disciplina EIC2100 (anual) após o filtro .....	191
Figura III.26 - Regras de associação: disciplina EIC2101 (anual) após o filtro .....	192
Figura III.27 - Regras de associação: disciplina EIC2102 (anual) após o filtro .....	193
Figura III.28 - Regras de associação: disciplina EIC2103 (anual) após o filtro .....	194
Figura III.29 - Regras de associação: disciplina EIC2104 (anual) após o filtro .....	194
Figura III.30 - Regras de associação: disciplina EIC2105 (anual) após o filtro .....	195
Figura III.31 - Regras de associação: disciplina EIC2200 (anual) após o filtro .....	195
Figura III.32 - Regras de associação: disciplina EIC2201 (anual) após o filtro .....	196
Figura III.33 - Regras de associação: disciplina EIC2202 (anual) após o filtro .....	197
Figura III.34 - Regras de associação: disciplina EIC2202 (anual) após o filtro (cont.) .....	198
Figura III.35 - Regras de associação: disciplina EIC2203 (anual) após o filtro .....	198
Figura III.36 - Regras de associação: disciplina EIC2204 (anual) após o filtro .....	199
Figura III.37 - Regras de associação: disciplina EIC2205 (anual) após o filtro .....	200
Figura III.38 - Regras de associação: disciplina EIC3200 (sem.) .....	201
Figura III.39 - Regras de associação: disciplina EIC3200(sem.) após o filtro.....	201
Figura III.40 - Regras de associação: disciplina EIC3201 (sem.) .....	202
Figura III.41 - Regras de associação: disciplina EIC3201 (sem.) após o filtro.....	202
Figura III.42 - Regras de associação: disciplina EIC3202 (sem.) .....	203
Figura III.43 - Regras de associação: disciplina EIC3202 (sem.) após o filtro.....	203
Figura III.44 - Regras de associação: disciplina EIC3203 (sem.) .....	204
Figura III.45 - Regras de associação: disciplina EIC3203 (sem.) após o filtro.....	204
Figura III.46 - Regras de associação: disciplina EIC3204 (sem.) .....	205
Figura III.47 - Regras de associação: disciplina EIC3204 (sem.) após o filtro.....	205
Figura III.48 - Regras de associação: disciplina EIC3205 (sem.) .....	206
Figura III.49 - Regras de associação: disciplina EIC3205 (sem.) após o filtro.....	206
Figura III.50 - Regras de associação: disciplina EIC3100 (anual) após o filtro .....	207
Figura III.51 - Regras de associação: disciplina EIC3101 (anual) após o filtro .....	207
Figura III.52 - Regras de associação: disciplina EIC3102 (anual) após o filtro .....	208
Figura III.53 - Regras de associação: disciplina EIC3103 (anual) após o filtro .....	209
Figura III.54 - Regras de associação: disciplina EIC3104 (anual) após o filtro .....	209
Figura III.55 - Regras de associação: disciplina EIC3105 (anual) após o filtro .....	210
Figura III.56 - Regras de associação: disciplina EIC3200 (anual) após o filtro .....	210

Figura III.57 - Regras de associação: disciplina EIC3201 (anual) após o filtro .....	211
Figura III.58 - Regras de associação: disciplina EIC3202 (anual) após o filtro .....	212
Figura III.59 - Regras de associação: disciplina EIC3203 (anual) após o filtro .....	212
Figura III.60 - Regras de associação: disciplina EIC3203 (anual) após o filtro (cont.) .....	213
Figura III.61 - Regras de associação: disciplina EIC3204 (anual) após o filtro .....	214
Figura III.62 - Regras de associação: disciplina EIC3205 (anual) após o filtro .....	214
Figura III.63 - Regras de associação: disciplina EIC4200 (sem.) .....	215
Figura III.64 - Regras de associação: disciplina EIC4200 (sem.) após o filtro .....	215
Figura III.65 - Regras de associação: disciplina EIC4201 (sem.) .....	216
Figura III.66 - Regras de associação: disciplina EIC4201 (sem.) após o filtro .....	216
Figura III.67 - Regras de associação: disciplina EIC4202 (sem.) .....	217
Figura III.68 - Regras de associação: disciplina EIC4202 (sem.) após o filtro .....	217
Figura III.69 - Regras de associação: disciplina EIC4203 (sem.) .....	218
Figura III.70 - Regras de associação: disciplina EIC4203 (sem.) após o filtro .....	218
Figura III.71 - Regras de associação: disciplina EIC4205 (sem.) .....	219
Figura III.72 - Regras de associação: disciplina EIC4205 (sem.) após o filtro .....	219
Figura III.73 - Regras de associação: disciplina EIC4100 (anual) após o filtro .....	220
Figura III.74 - Regras de associação: disciplina EIC4100 (anual) após o filtro (cont.) .....	221
Figura III.75 - Regras de associação: disciplina EIC4101 (anual) após o filtro .....	222
Figura III.76 - Regras de associação: disciplina EIC4101 (anual) após o filtro (cont.) .....	223
Figura III.77 - Regras de associação: disciplina EIC4101 (anual) após o filtro (cont.) .....	224
Figura III.78 - Regras de associação: disciplina EIC4102 (anual) após o filtro .....	225
Figura III.79 - Regras de associação: disciplina EIC4103 (anual) após o filtro .....	226
Figura III.80 - Regras de associação: disciplina EIC4104 (anual) após o filtro .....	227
Figura III.81 - Regras de associação: disciplina EIC4104 (anual) após o filtro .....	228
Figura III.82 - Regras de associação: disciplina EIC4105 (anual) após o filtro .....	228
Figura III.83 - Regras de associação: disciplina EIC4200 (anual) após o filtro .....	229
Figura III.84 - Regras de associação: disciplina EIC4200 (anual) após o filtro (cont.) .....	230
Figura III.85 - Regras de associação: disciplina EIC4201 (anual) após o filtro .....	230

Figura III.86 - Regras de associação: disciplina EIC4202 (anual) após o filtro .....	231
Figura III.87 - Regras de associação: disciplina EIC4202 (anual) após o filtro (cont.) .....	232
Figura III.88 - Regras de associação: disciplina EIC4202 (anual) após o filtro (cont.) .....	233
Figura III.89 - Regras de associação: disciplina EIC4203 (anual) após o filtro .....	234
Figura III.90 - Regras de associação: disciplina EIC4205 (anual) após o filtro .....	235
Figura III.91 - Regras de associação: disciplina EIC5100 (anual) após o filtro .....	236
Figura III.92 - Regras de associação: disciplina EIC5101 (anual) após o filtro .....	237
Figura III.93 - Regras de associação: disciplina EIC5101 (anual) após o filtro (cont.) .....	238
Figura III.94 - Regras de associação: disciplina EIC5102 (anual) após o filtro .....	239
Figura III.95 - Regras de associação: disciplina EIC5103 (anual) após o filtro .....	240
Figura III.96 - Regras de associação: disciplina EIC5104 (anual) após o filtro .....	241
Figura III.97 - Regras de associação: disciplina EIC5106 (anual) após o filtro .....	242
Figura III.98 - Regras de associação: disciplina EIC5201 (anual) após o filtro .....	243
Figura III.99 - Regras de associação: disciplina EIC5201 (anual) após o filtro (cont.) .....	244

# Capítulo 1

## Introdução

Este capítulo contém uma motivação do tema de dissertação, a definição do seu objectivo, metodologia seguida e estrutura do documento.

### **1.1. Considerações gerais**

Com o aparecimento dos primeiros computadores chegou também uma nova forma de guardar informação inerente ao trabalho das organizações. Grandes quantidades de dados organizados e guardados fisicamente em grandes arquivos repletos de folhas de papel começaram a ser substituídos por ficheiros em formato digital, o que tornava a procura mais fácil e o armazenamento mais simples e compacto.

Nos anos sessenta começou a ser feita a utilização comercial das Bases de Dados (BD). Mas foi nos anos setenta, com o aparecimento dos Sistemas de Gestão de Bases de Dados (SGBD) relacionais que foi dado o grande impulso. Estes sistemas gerem e manipulam conjuntos de dados estruturados, segundo um modelo de dados assente na noção de relação e na teoria matemática de conjuntos. Estes sistemas disponibilizam actualmente ainda outras funcionalidades como acesso simultâneo, vistas, manutenção de integridade, etc... Para o conseguir, um SGBD suporta linguagens de:

- definição de dados, para criação e alteração da estrutura da BD;
- consulta de dados, para consulta e processamento dos dados armazenados;
- manipulação de dados, para acrescentar dados novos e modificar os existentes.

O grande sucesso dos SGBD relacionais assenta principalmente na simplicidade do modelo relacional e na generalização duma linguagem de manipulação simples e eficiente, o SQL (*Structured Query Language*), que é considerada a linguagem *standard* em BD.

Com o passar do tempo, as organizações foram acumulando grandes quantidades de dados. Estes dados, após terem sido gerados e utilizados para determinados fins específicos, eram vistos como contendo apenas informação histórica, aparentemente de pouca utilidade. Mas a informação escondida no meio dos grandes repositórios de dados pode ser de grande importância para a competitividade das organizações.

O conceito que hoje em dia as organizações têm das BD alterou-se significativamente. Assim, o objectivo é não só que estas representem de uma forma tão exacta quanto possível o meio em que estão inseridas como também se pretende obter delas mais e melhor informação. As perguntas que se pretendem ver respondidas tornaram-se mais complexas. Em vez de se querer saber “quem foram os clientes que compraram o produto X no último mês”, pretende-se saber “que outros produtos compram habitualmente os clientes quando compram o produto X”.

Este tipo de questões deu origem à área de Extração de Conhecimento de Bases de Dados (ECBD), vulgarmente conhecida por uma das partes que a compõem, a prospecção de dados (*Data Mining*).

A ECBD emergiu como uma nova disciplina, resultante de outras mais antigas: bases de dados, análise de dados e inteligência artificial. Pode-se dizer, num sentido mais vasto, que a ECBD é a aplicação de técnicas de inteligência artificial e de análise de dados a grandes bases de dados.

Assim, partindo das experiências com alguns casos de pequena dimensão, foram-se construindo alguns dos algoritmos mais conhecidos de inteligência artificial: redes neuronais, algoritmos genéticos, algoritmos de geração de regras de associação, árvores e regras de decisão, etc... Para testar estes algoritmos com um mais vasto conjunto de dados, os grande repositórios de dados guardados nas organizações tornaram-se a matéria-prima ideal.



## **1.2. Objectivos e enquadramento**

O objectivo principal da presente dissertação é avaliar a adequação do emprego da ECBD a uma base de dados académica, tomando como estudo de caso concreto a Faculdade de Engenharia da Universidade do Porto (FEUP). Apesar do âmbito local em que se insere o estudo, as conclusões tiradas poderão ser aplicáveis a outras escolas.

As bases de dados académicas contêm, entre outra informação, a relativa às classificações obtidas nas diversas disciplinas dos cursos. Esta informação está organizada em entidades (professores, alunos, cursos, disciplinas, etc...) e as ligações entre estas (inscrições dos alunos em disciplinas, leccionação das disciplinas por professores, etc...).

Com estes dados é possível confirmar alguma informação importante. Por exemplo, podem existir disciplinas cujo sucesso está directamente relacionado com o de outras. Algumas cadeiras, que são divididas em mais do que um semestre (por exemplo, Matemática I, II, III e IV), estarão, em princípio, bastante interligadas. No entanto pode-se questionar até que ponto elas poderão ser, ou não, dependentes entre si. Esta informação poderá ser importante para a implementação dum sistema de precedências entre disciplinas interligadas.

Uma outra perspectiva será a de descobrir relações que não são, à partida, previsíveis. Por exemplo, descobrir que uma disciplina de Inglês terá uma forte interligação com uma de Matemática, dificilmente estaria nas previsões de qualquer pessoa que se debruçasse sobre o problema.

Assim, esta informação pode ser crucial para o desenvolvimento dos planos curriculares e para uma rápida intervenção quando necessário. Tal como no mundo empresarial é importante reagir imediatamente para evitar descidas dos índices de produção ou dos lucros, também no mundo académico uma pronta intervenção pode evitar descidas no rendimento e sucesso escolar dos alunos.

### **1.3. Resultados atingidos**

Nesta dissertação, conseguiu-se obter alguns resultados que comprovam a adequação da aplicação da ECBD a uma base de dados académica. As regras de associação e as árvores e regras de decisão geradas conseguem descrever padrões nos dados, de forma a que estes possam ser usados e interpretados pelo utilizador final.

Os modelos gerados podem ser aplicados nas mais diversas tarefas de organização e gestão de qualquer instituição académica. No seu todo, e devido ao facto de serem de fácil percepção para o utilizador final, podem ser considerados como um sistema de apoio à decisão.

Existe, no entanto, um ponto a ressaltar. Quando temos uma hipótese, e utilizamos os algoritmos à nossa disposição para confirmá-la, conseguimos saber se é ou não válida. Quando aplicamos os algoritmos de modo a que sejam gerados resultados que, à partida, não se sabe o que se pode esperar deles, poderemos estar a descurar informação importante.

Assim, só se pode afirmar que se obtiveram resultados com informação importante, com consciência de que alguma informação pode ainda estar por descobrir no meio dos dados. Para além disso, alguns novos resultados poderão ser obtidos se, por exemplo, for introduzida nova informação que se venha a mostrar necessária.

### **1.4. Plano da dissertação**

A presente dissertação encontra-se dividida em seis capítulos, onde se descreve o trabalho desenvolvido, sendo dados alguns exemplos demonstrativos.

Após o actual capítulo de introdução, teremos, no próximo capítulo, a descrição do estado da arte no momento actual em ECBD. Para além dos conceitos fundamentais, descrevem-se algumas aplicações ao mundo real, os principais algoritmos de prospecção de dados existentes e os tipos de ferramentas disponíveis para auxiliar o processo de ECBD.

No terceiro capítulo, descreve-se a aplicação da ECBD à BD académica da FEUP, incluindo as diversas fases como a preparação dos dados, definição dos objectivos a alcançar, a metodologia seguida e a aplicação aos dados

No quarto capítulo, a análise baseia-se em árvores e regras de decisão. No quinto capítulo a análise recai sobre as regras de associação.

Finalmente, no sexto e último capítulo, apresentam-se as conclusões finais, faz-se uma discussão do trabalho realizado, referindo-se possíveis alternativas, bem como sugestões de trabalho futuro.

Em anexo são apresentados os modelos gerados nas análises do primeiro ao quinto ano. No anexo I apresentam-se as árvores e regras de derivação geradas, enquanto no anexo II temos a análise relativa a estes modelos. No anexo III apresentam-se as regras de associação antes e após a aplicação do filtro desenvolvido, com excepção na análise do completo, onde só se apresentam os resultados após o filtro.

# Capítulo 2

## Extracção de Conhecimento em Bases de Dados

Neste capítulo pretende-se desenvolver os principais tópicos relacionados com Extracção de Conhecimento em Bases de Dados (ECBD) e Prospecção de Dados (*Data Mining*). A próxima secção refere as principais diferenças entre estes dois conceitos e a sua relação com outros (*Data Warehouse*, OLAP, estatística, etc...). Na secção 2.2 faz-se a apresentação de algumas das aplicações ao mundo real. Na secção 2.3 descrevem-se alguns tipos de algoritmos e técnicas mais conhecidos nesta área. Na secção 2.4 descrevem-se alguns tipos de ferramentas de prospecção de dados disponíveis, descrevendo-se mais pormenorizadamente o Clementine, que vai ser usado no problema estudado.

### 2.1. Considerações gerais

Os conceitos ECBD e prospecção de dados estão intimamente ligados e são muitas vezes confundidos. Podem-se distinguir estes dois conceitos da seguinte forma: ECBD é definido como todo o processo de descoberta útil de conhecimento a partir de dados, enquanto prospecção de dados é uma das partes desse processo, que consiste na aplicação de algoritmos específicos para a extracção de padrões (modelos) a partir de dados [Fayyad et al., 1996b].

Pode-se ver a ECBD como o processo de usar algoritmos de prospecção de dados para extrair conhecimento a partir dos dados, em conjunto com qualquer pré-processamento, ensaio ou transformação dos mesmos. Prospecção de dados, por sua vez, é um passo no processo de ECBD, consistindo em alguns algoritmos que, sob limitações de eficiência computacional aceitáveis, produzem uma enumeração

particular de padrões sobre o conjunto de dados [Fayyad et al., 1996]. Podem ver-se os vários passos na figura 2.1.

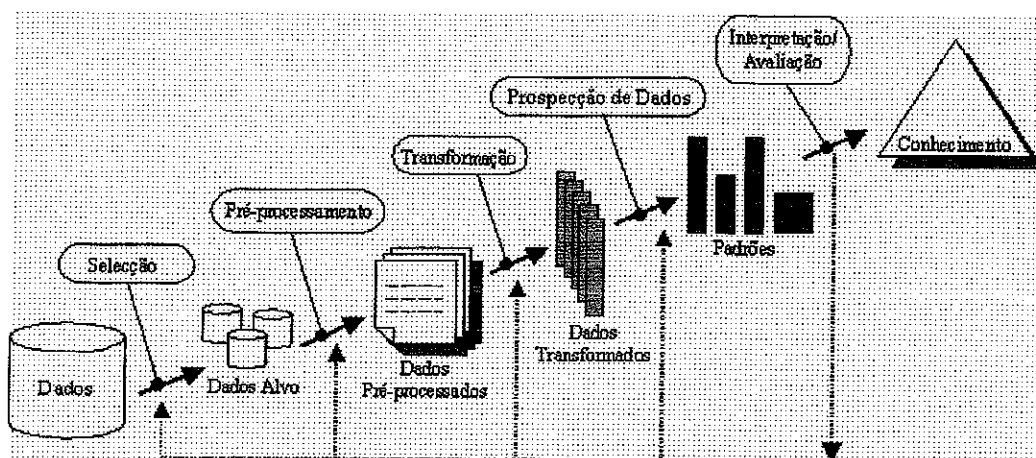


Figura 2.1 – Passos no processo de ECBD

Assim, pode-se enumerar um conjunto de passos básicos no processo de ECBD. Após o desenvolvimento da compreensão do domínio aplicacional, conhecimento anterior relevante e objectivos, é criado um conjunto de dados alvo, a partir da selecção de um conjunto de dados, ou de um subconjunto de variáveis ou amostras de dados, onde se deve fazer a extracção.

O passo seguinte consiste na limpeza de dados e pré-processamento, com operações básicas como remoção de ruído e *outliers*. A transformação dos dados consiste na redução e projecção dos mesmos, usando redução de dimensão, ou métodos de transformação para reduzir o número efectivo de variáveis em consideração ou encontrar representações invariantes para os dados.

A partir daqui pode-se decidir qual o objectivo do processo de ECBD (classificação, regressão, agrupamentos, etc...) e escolher os algoritmos de prospecção de dados a serem usados para procura de padrões nos dados. Após a sua aplicação e interpretação dos padrões encontrados, pode-se retornar a um dos passos anteriores para mais iterações.

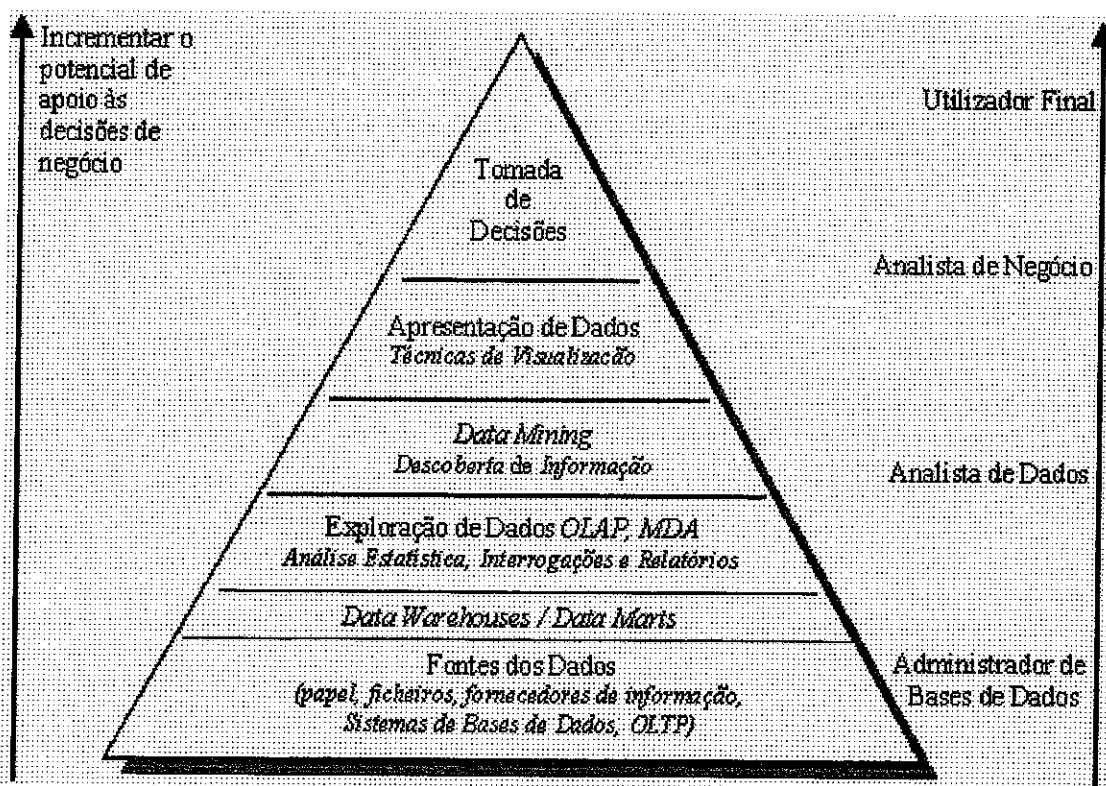
Finalmente, pode-se consolidar o conhecimento descoberto, incorporando-o no desempenho do sistema, ou simplesmente documentando-o e apresentando-o a partes interessadas, e ainda verificando e resolvendo conflitos potenciais em relação a conhecimento anterior do utilizador ou extraído a partir dos dados.

A maior parte da atenção é centralizada no passo da prospecção de dados. No entanto, os outros passos assumem um papel preponderante para que este seja realizado de forma eficiente. Por exemplo, em [Pyle, 1999], dá-se particular ênfase à preparação de dados, referindo que esta parte é muitas vezes esquecida, o que vai influenciar negativamente o desempenho da ECBD.

Em [Cabena et al., 1998], refere-se que as técnicas mais populares suportadas das ferramentas de prospecção de dados comerciais são, por ordem descendente: visualização, classificação, modelação com predição, segmentação, descoberta de associações e descoberta de padrões sequenciais.

Por outro lado, também é referido que o esforço a considerar na preparação de dados deve ser de aproximadamente 50%, estando a determinação dos objectivos de negócio a seguir, e finalmente, aproximadamente ao mesmo nível, as técnicas de prospecção de dados e a análise de resultados e assimilação de conhecimento.

Na figura 2.2, pode-se ver o posicionamento lógico, segundo os autores, das várias camadas tecnológicas.



**Figura 2.2 - Prospecção de Dados e inteligência de negócio**

Dois conceitos que aparecem muitas vezes ligados a ECBD e prospecção de dados são os de armazenamento de dados (*Data Warehousing*) e Processamento Analítico por Computador (*On-Line Analytical Processing*), geralmente conhecido pela sigla OLAP. Em [Fayyad et al., 1996b], *Data Warehousing* é definido como a tarefa das organizações de colecionar e limpar dados de transações, para torná-los disponíveis para análise por computador e apoio à decisão. Neste contexto, OLAP procura fornecer análise de dados multidimensional, que é superior ao SQL (*Structured Query Language*, linguagem *standard* de manipulação de dados mais usada em vários Sistemas de Gestão de Bases de Dados) no cálculo de sumários e quebras ao longo de várias dimensões. Pode-se ainda acrescentar a constatação de o processo de ECBD e OLAP serem vistos como duas facetas relacionadas de uma nova geração de ferramentas de extração e gestão de informação inteligentes [Fayyad et al., 1996].

Outra das questões que muitas vezes se coloca, principalmente a nível empresarial, é qual a diferença entre ECBD e estatística. Por outras palavras, o que as organizações querem saber é se vale a pena investir em ECBD e prospecção de dados, uma vez que já têm nas suas organizações pessoas a trabalhar na parte de análise de dados e estatística. Segundo [Fayyad et al., 1996b], a estatística tem muito em comum com ECBD: a estatística fornece uma linguagem e uma estrutura (*framework*) para quantificar a incerteza quando se tenta inferir padrões gerais a partir duma amostra particular numa população.

Em [Glymour et al., 1996], discute-se em pormenor a relação existente e a relevância de alguns conceitos da estatística para prospecção de dados. Prospecção de dados envolve quase sempre uma arquitectura de pesquisa que requer avaliação de hipóteses nas diversas fases da pesquisa, avaliação da saída resultante e o uso apropriado dos resultados. A estatística, apesar de não ter muito a oferecer na parte de compreensão das arquitecturas de pesquisa, pode ajudar na parte de avaliação de hipóteses ao longo da pesquisa, na avaliação dos resultados e compreensão do uso apropriado dos resultados, sendo importante para a prospecção de dados, tendo em conta a sua riqueza de procedimentos técnicos e resultados.

## 2.2. Aplicações de ECBD e prospecção de dados

Existem algumas aplicações que já usam ECBD e prospecção de dados. Nas subsecções a seguir serão referidas algumas delas, que se encontram descritos em detalhe em [Fayyad et al., 1996], nas subsecções de 2.2.1 a 2.2.4, e em [Michalski et al., 1998], nas subsecções de 2.2.5 a 2.2.8.

### 2.2.1. Automação da análise e catalogação de pesquisa do espaço

A primeira aplicação usa técnicas de aprendizagem de classificação para automação da redução e análise dum grande conjunto de dados astronómicos: *The Digital Palomar Observatory Sky Survey* (POSS-II). Espera-se que o conjunto de 3 *terabytes* em imagens contenham 2 mil milhões de objectos. Para a análise científica primária destes dados, é necessário detectar, medir e classificar todos os objectos, pelo que será necessária uma aproximação automatizada.

O sistema SKICAT (*Sky Image Cataloging and Analysis Tool*) integra técnicas para processamento de imagem, classificação de dados e gestão de bases de dados. Os algoritmos de aprendizagem são treinados para classificar os objectos detectados e podem classificar objectos, de classificação visual difícil, com grande exactidão. Uma grande quantidade de ferramentas de análise científica é fornecida para filtragem, manipulação, impressão e listagem dos dados.

No sistema SKICAT a aprendizagem desempenhou um papel importante, resolvendo um problema difícil e significativo do ponto de vista científico. Os principais benefícios desta aproximação são:

- aumento de rapidez na redução de dados;
- consistência da classificação;
- facilidade no acesso, análise e criação de nova informação dum conjunto de dados, de outro modo inatingível.



### 2.2.2. Selecção e relatórios sobre o que é interessante

Uma das áreas mais promissoras de ECBD é a análise automática de desvios. O sucesso nesta área depende da capacidade de identificação de alguns eventos relevantes entre os vários desvios potencialmente interessantes. A aproximação é determinar o interesse de um desvio pelo benefício potencial de uma acção relevante.

Esta aproximação foi implementada no sistema KEFIR (*Key Findings Reporter*), um sistema para descoberta e explicação de descobertas chave em bases de dados grandes e em constante mudança, estando correntemente a ser utilizado na análise de dados de cuidados de saúde.

O sistema faz uma escavação automática através dos dados, ao longo de dimensões múltiplas, para determinar os desvios mais interessantes de medidas quantitativas específicas, relativas aos seus valores prévios e esperados. Os desvios-chave são explicados a partir da sua relação com outros desvios nos dados e, quando apropriado, gera recomendações para acções em resposta a estes desvios.

O sistema KEFIR usa um *browser* WWW, para apresentar as suas descobertas num relatório em hipertexto, usando linguagem natural e gráficos de gestão.

### 2.2.3. Modelação de incerteza subjectiva em anotação de imagem

Examina-se aqui o problema de lidar com o conceito subjectivo de verdade no significado (*ground truth*) no contexto de descoberta de conhecimento de bases de dados de imagens e, em particular, o problema de avaliar o desempenho relativo da capacidade de detecção dos peritos humanos face à dos algoritmos.

Na prática, para aplicações de detecção remota (*remote-sensing*), a aquisição da verdade no significado é muitas vezes demasiado cara ou fisicamente impossível. Em vez disso, tem muitas vezes que se confiar nas opiniões subjectivas de peritos. Estes examinam visualmente as imagens e fornecem descrições subjectivas que, na sua essência, são estimativas com ruído da verdade no significado.

Neste contexto, foi apresentado um modelo probabilístico para explicar a segurança e tendência das descrições, e um esquema de máxima probabilidade para

estimativa de parâmetros. O modelo estimado combina várias opiniões para alcançar uma única estimativa consensual.

Estas estimativas formam a base para comparação do desempenho humano e do algoritmo de detecção na forma de características de operação do receptor (*Receiver Operation Characteristics, ROCs*) modificado. Um dos problemas onde esta aplicação foi testada foi a identificação de pequenos vulcões a partir de imagens de radar de Vénus.

#### **2.2.4. Predição de retorno de investimento em títulos**

A experiência com dados dos mercados de capitais sugere que o domínio pode ser efectivamente modelado por regras de classificação induzidas a partir de dados históricos disponíveis, para conseguir predições lucrativas para investimento em títulos.

Algumas técnicas de classificação, como a geração de regra minimal (R-MINI) e análise de característica contextual, parecem suficientemente robustas para extrair de forma consistente informação útil a partir de domínios com ruído, como os mercados financeiros.

Foram testados dados do conjunto S&P 500, sendo verificados os resultados das correlações e gestão de investimento simulado, baseado nas regras de classificação geradas por R-MINI. Foi estudada a forma como as regras para classificação podem ser efectivamente usadas para predição numérica e, eventualmente, para uma política de investimento.

#### **2.2.5. Encontrar associações em colecções de texto**

O sistema FACT (*Finding Associations in Collections of Text*) foi concebido para encontrar associações – padrões de co-ocorrência – entre palavras chave que rotulam itens numa colecção de documentos de texto. FACT usa uma perspectiva de descoberta de conhecimento centrada em interrogações, onde um pedido de descoberta de associações é visto como uma interrogação sobre o conjunto implícito de associações possíveis suportado por uma colecção de documentos.

Para além disso, quando o conhecimento profundo sobre as palavras chave que rotulam os documentos está disponível, FACT é capaz de usar esta informação no seu processo de descoberta, permitindo ao utilizador a especificação de restrições nos resultados desejados da interrogação em termos de conhecimento profundo. Além disso, a execução dum a interrogação de descoberta de associações é estruturada para que as restrições do conhecimento profundo possam ser exploradas na pesquisa dos resultados possíveis.

Em vez de pedir ao utilizador que especifique uma expressão de interrogação explícita na linguagem de interrogação de descoberta de associações, FACT apresenta ao utilizador um interface gráfico para a linguagem, tendo esta uma semântica bem definida para as acções de descoberta executadas pelo utilizador através do interface.

#### **2.2.6. Aprendizagem de padrões em imagens**

É possível usar a aprendizagem de padrões em imagens e sequências de imagens, e usar os padrões obtidos para interpretar novas imagens. Pode-se distinguir três pontos:

- interpretação semântica de imagens a cor de cenas de exterior;
- detecção de explosivos (*blasting caps*) em imagens raio-x de bagagens;
- reconhecimento de acções em sequências de imagens de vídeo.

Os processos de formação de imagens nesta área de problema, e a escolha de espaços de representação usado na aproximação apresentada, são os grandes desafios a ser actualmente estudados. Os resultados indicam a vantagem de aplicar aprendizagem à visão.

### **2.2.7. Investigação empírica do fenómeno da expressão musical**

As técnicas de aprendizagem automática são também aplicadas ao estudo dum fenómeno fundamental na música tonal: o desempenho musical expressivo. Pode-se definir o desempenho expressivo como a arte de adaptar uma peça musical tocando-a de forma ligeiramente diferente da pauta, variando continuamente certos parâmetros. Por exemplo, pode-se aumentar ou diminuir a velocidade, crescimentos altos ou suaves, colocação de micro-pausas entre eventos, etc... Existem algoritmos que induzem regras gerais de desempenho musical expressivo, a partir de exemplos de desempenhos reais realizados por músicos.

Motivado pela compreensão que o conhecimento geral sobre música assume um papel importante na forma como os seres humanos aprendem esta tarefa, podem-se apresentar duas alternativas na aproximação à aprendizagem baseada em conhecimento: aprendizagem ao nível da nota ou ao nível da estrutura. Em ambos os casos, o conhecimento do domínio fornecido ao aprendiz é baseado nas teorias estabelecidas da música tonal.

Os resultados experimentais mostram que ambas as aproximações levam a um melhoramento significativo dos resultados de aprendizagem, comparando com a aprendizagem puramente indutiva.

### **2.2.8. *WebWatcher*: aprendizagem na *World Wide Web***

*WebWatcher* é um assistente de procura de informação na *World Wide Web*. Este agente ajuda, de forma interactiva, os utilizadores a localizar a informação desejada usando conhecimento aprendido sobre quais as hiperligações que são mais prováveis de conduzir à informação pretendida.

O estudo tem incidido sobre:

- organizar o *WebWatcher* para fornecer conselhos interactivos aos utilizadores, enquanto faz as pesquisas, com ou sem sucesso, aos dados de treino do algoritmo;

- incorporar algoritmos de aprendizagem para automaticamente adquirir conhecimento para seleccionar a hiperligação apropriada, dados a actual página *Web* vista pelo utilizador e o seu objectivo de informação.

O algoritmo constrói, para cada página, ligação e objectivo, um vector com aproximadamente quinhentos e trinta valores binários relativamente a palavras, incluindo:

- duzentos sobre palavras sublinhadas na hiperligação;
- duzentos sobre palavras nas frases que contêm a hiperligação;
- cem sobre palavras nos cabeçalhos associados com a hiperligação;
- aproximadamente trinta sobre palavras usadas para definir o objectivo do utilizador.

Este vector, de tamanho fixo, pode ser usado nos principais algoritmos de aprendizagem automática (árvores e regras de decisão, regras de associação, etc...).

### **2.3. Algoritmos e técnicas de prospecção de dados**

Podem distinguir-se dois objectivos primários de prospecção de dados: predição e descrição. A predição consiste no uso de variáveis ou campos da base de dados para prever valores desconhecidos ou futuros de outras variáveis de interesse. A descrição destina-se a encontrar padrões, de fácil interpretação pelo utilizador final, que caracterizem os dados [Fayyad et al., 1996].

Os objectivos de predição e descrição são alcançados usando as principais técnicas de prospecção de dados: classificação, regressão, análise de ligações, agrupamento (*clustering*), sumário, modelação de dependências e análise sequencial.

Os dois primeiros, classificação e regressão, consistem em determinar uma função que associe a um item um determinado valor. No caso da classificação, esse valor será um elemento de entre várias classes pré-definidas, ao passo que na regressão se associa um valor de previsão real.

A análise de ligações determina relações entre campos numa tabela da base de dados. Um exemplo são as regras de associação [Agrawal, et al., 1993], que serão

descritas mais à frente. O objectivo é derivar correlações entre vários campos satisfazendo valores de suporte e confiança.

O agrupamento é uma tarefa descritiva comum, onde se tenta identificar um conjunto finito de categorias que descrevem os dados, podendo ser mutuamente exclusivos, hierárquicos ou com sobreposições.

O sumário envolve métodos que permitem encontrar uma descrição compacta para um subconjunto de dados.

A modelação de dependências consiste em encontrar um modelo que descreve dependências significativas entre variáveis. Estas existem a dois níveis: estrutural e quantitativo. No primeiro, especifica-se, muitas vezes na forma gráfica, quais as variáveis que são localmente dependentes entre si, enquanto no segundo se especifica a força das dependências usando uma escala numérica.

Finalmente, a análise sequencial permite modelar padrões sequenciais (como análise de séries temporais). O objectivo é modelar os estados do processo, gerando a sequência, ou extrair e mostrar desvios e tendências ao longo do tempo.

No caso em estudo, foram usadas duas técnicas de prospecção de dados: a classificação e a análise de ligações. Nas duas subsecções seguintes será feita uma abordagem destas duas técnicas, e dos algoritmos usados.

### **2.3.1. Classificação**

Como já foi definido, a classificação consiste em determinar uma função que associe a um item um determinado valor, que será um elemento de entre várias classes pré-definidas. As classes podem ser, por exemplo, resultados de aprovação em exames (“Aprovado” ou “Reprovado”), ou ainda mais específicos, dando uma apreciação qualitativa “Reprovado”, “Suficiente”, “Bom”, “Muito Bom”).

Considerando uma tabela de base de dados, tem-se um conjunto de atributos (também designados por campos ou colunas) a descrever cada caso, sendo estes elementos designados por tuplos (ou ainda por registos ou linhas). A classe de um tuplo é dada pelo valor de um atributo a prever, especificado pelo utilizador.

Considere-se o exemplo de uma tabela onde são guardados os resultados dos trabalhos práticos realizados a uma dada disciplina. Estes trabalhos contam para a nota final e permitem ter uma ideia do desempenho do aluno. Pretende-se saber,

dadas as notas dos trabalhos (que podem ser qualitativas ou quantitativas), se o aluno vai ter, ou não, aprovação à disciplina. Assim, a classe será o atributo resposta, que poderá tomar dois valores possíveis: “sim” ou “não”.

Na construção de um modelo de classificação, teremos a considerar dois conjuntos: de treino e de teste. O conjunto de treino é um conjunto de tuplos já classificados que dá origem ao modelo de classificação. O conjunto de teste é um conjunto também com os tuplos já classificados, onde é aplicado o modelo de classificação, podendo-se comparar as classes de cada tuplo com as classes previstas. Em alguns casos, pode-se dividir um grande conjunto de dados em dois, utilizando um como conjunto de treino e outro como conjunto de teste.

Existem vários modelos de classificação possíveis, como árvores e regras de decisão, ou redes neuronais. Para este estudo escolheram-se as árvores e as regras de decisão, uma vez que têm uma forma de representação bastante simples, o que torna o modelo inferido de fácil compreensão para o utilizador final.

### 2.3.1.1. Árvores e Regras de Decisão

Existem vários algoritmos que geram árvores e regras de decisão. Os algoritmos C4.5 [Quinlan, 1992] e o mais recente C5.0 (do mesmo autor, ainda não documentado) são os mais conhecidos, servindo de referência na área de aprendizagem. Outros métodos mais antigos, como o antecessor do C4.5, o ID3 [Quinlan, 1986] e o sistema CART (*Classification And Regression Trees*), descrito em [Breinman et al., 1984], são também várias vezes referidos.

As árvores de decisão são representadas na forma de uma árvore invertida, em que a raiz está no topo e as folhas no fundo. A raiz e os nós intermédios (isto é, os nós que não são folhas) representam atributos, dos quais saem ramos que representam as diferentes possibilidades de instanciação. As folhas representam a classe prevista.

A classificação de uma instância é feita de forma descendente, partindo da raiz até chegar a uma das folhas. Em cada passo verifica-se o valor do atributo, seguindo-se pelo ramo correspondente até ao próximo nó. Na figura 2.3 está um exemplo clássico: *PlayTennis*.

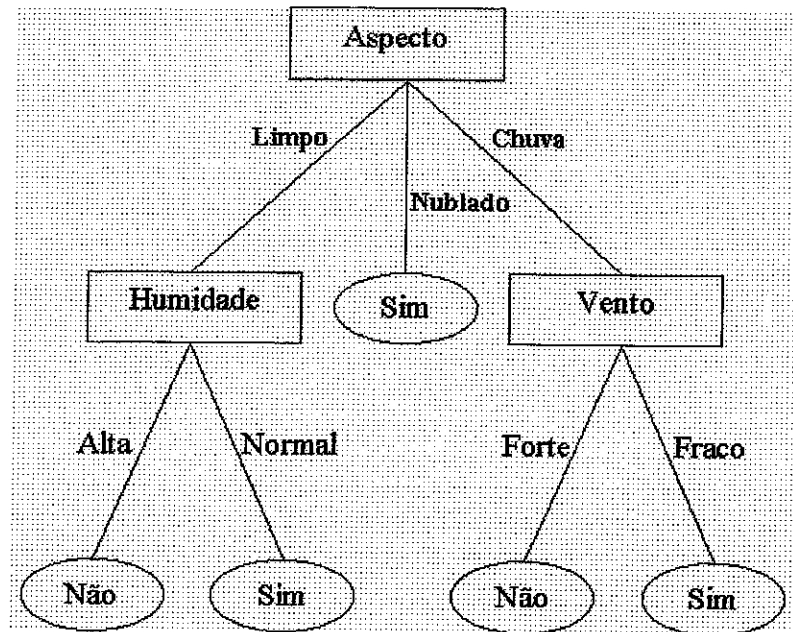


Figura 2.3 – Árvore de Decisão do PlayTennis

No exemplo *PlayTennis*, pretende-se saber se o tempo é ou não propício para praticar ténis. Considera-se um tuplo de quatro atributos: aspecto, humidade, temperatura e vento. Considere-se, por exemplo a seguinte instância:

<aspecto=limpo, humidade=normal, temperatura=quente, vento=forte>

Neste caso, a conclusão seria “sim”, uma vez que o aspecto do céu é limpo e a humidade normal.

Note-se que, na árvore decisão, o atributo temperatura não é nunca testado. Isto é uma indicação que este atributo não vai influir no processo de decisão.

A partir da árvore de decisão é fácil construir as regras de decisão. Estas regras serão da forma “se-então”, sendo o seu número igual ao número de folhas. Consideramos uma sequência  $(X_1, Y_1, X_2, Y_2, \dots, X_n, Y_n, C)$ , onde  $X_1$  é a raiz,  $C$  é uma folha, e cada  $(X_i, Y_i)$ ,  $i=1,2,\dots,n$ , é um par ordenado onde  $X_i$  é um nó e  $Y_i$  um ramo que parte desse nó. Esta sequência, que representa um caminho da raiz até uma folha, pode ser transformada na seguinte regra:

**Se  $X_1=Y_1$  e  $X_2=Y_2$  e ... e  $X_n=Y_n$  então  $C$**



No exemplo *PlayTennis*, a árvore de decisão dada originaria cinco regras:

Regra 1: Se aspecto=limpo e humidade=alta então classe=Não

Regra 2: Se aspecto=limpo e humidade=normal então classe=Sim

Regra 3: Se aspecto=nublado então classe=Sim

Regra 4: Se aspecto=chuva e vento=forte então classe=Não

Regra 5: Se aspecto=chuva e vento=fraco então classe=Sim

### 2.3.1.2. Algoritmos C4.5 e C5.0

Para a construção da árvore de decisão é preciso, em cada passo, escolher o atributo que vai ser testado em cada nó da árvore. Para tal é necessário uma boa medida quantitativa para definir o melhor dos atributos. Em [Mitchell, 1997], descreve-se uma propriedade estatística, chamada ganho de informação, que mede a eficiência de um dado atributo em separar os exemplos de treino, de acordo com a classificação pretendida.

Uma medida usada para estabelecer o ganho de informação é a entropia, que caracteriza a pureza dum conjunto arbitrário de exemplos. Considerando um conjunto  $S$ , contendo exemplos de duas classes (positivos e negativos), a entropia de  $S$  relativa a esta classificação booleana é dada por:

$$\text{Entropia}(S) = -p_+ \log_2 p_+ - p_- \log_2 p_-$$

sendo  $p_+$  a percentagem de casos positivos e  $p_-$  a percentagem de casos negativos.

Note-se que a entropia será zero se todos os elementos forem da mesma classe:

$$-\log_2 1 - 0 = 0$$

e será um se o número de elementos positivos e negativos forem iguais:

$$-\frac{1}{2} \log_2 \frac{1}{2} - \frac{1}{2} \log_2 \frac{1}{2} = -\frac{1}{2}(-1) - \frac{1}{2}(-1) = \frac{1}{2} + \frac{1}{2} = 1$$

De uma forma geral, quanto mais próximas forem as percentagens maior será a entropia (ver figura 2.4).

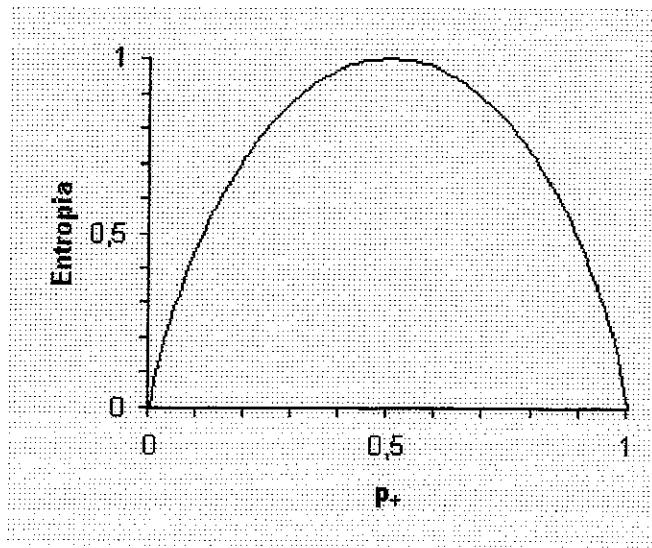


Figura 2.4 - Entropia em funções booleanas

Uma interpretação da entropia, a partir da teoria da informação, é que esta especifica o número de *bits* de informação necessário para codificar a classificação de um membro de S. Assim, quando a entropia é zero, todos os elementos são da mesma classe, logo não é preciso nenhum *bit*. Quando a entropia é um, é necessário um *bit* para codificar a informação (0 como negativo, 1 como positivo).

Generalizando para um conjunto de n classes,  $n \geq 2$ , temos:

$$\text{Entropia}(S) = \sum_{i=1}^n -p_i \log_2 p_i$$

sendo  $p_i$  a percentagem de elementos da classe de ordem  $i$ ,  $i=1, \dots, n$ .

Note-se que a base do logaritmo continua a ser dois, uma vez que estamos a falar da codificação esperada em *bits*. Por exemplo, se considerarmos quatro classes, isto é,  $n = 4$ , e consideramos que a percentagem de elementos de cada classe é igual

para todas, isto é,  $p_i = \frac{1}{4}$ ,  $i = 1, 2, 3, 4$ , teremos:

$$\begin{aligned} \text{Entropia}(S) &= \sum_{i=1}^4 -p_i \log_2 p_i = \sum_{i=1}^4 -\frac{1}{4} \log_2 \frac{1}{4} = 4 \left( -\frac{1}{4} \log_2 2^{-2} \right) = \\ &= 4 \left( -\frac{1}{4} \cdot (-2) \right) = 4 \cdot \frac{1}{2} = 2 \end{aligned}$$

Assim, necessitamos dois bits para codificar a informação relativa às quatro classes: 00, 01, 10 e 11.

A medida usada, o ganho de informação, é a redução esperada da entropia causada pela partição do conjunto de exemplos de acordo com um dado atributo. Assim, o ganho de informação  $\text{Ganho}(S, A)$ , de um determinado atributo  $A$ , relativo a um conjunto de exemplos  $S$ , é dado por:

$$\text{Ganho}(S, A) = \text{Entropia}(S) - \sum_{v \in \text{valores}(A)} \frac{|S_v|}{|S|} \text{Entropia}(S_v)$$

sendo  $\text{valores}(A)$  o conjunto de todos os valores possíveis do atributo  $A$ , e  $S_v$  o subconjunto de elementos de  $S$  para os quais o atributo  $A$  tem o valor  $v$  ( $S_v = \{s \in S : A(s) = v\}$ ).

Vejamos a forma como o algoritmo vai funcionar para o exemplo concreto *PlayTennis*. Consideremos os dados da tabela 2.1.

Tabela 2.1 - Conjunto de treino para o *PlayTennis*

Aspecto	Temperatura	Humidade	Vento	Classe
Limpo	Quente	Alta	Fraco	Não
Limpo	Quente	Alta	Forte	Não
Nublado	Quente	Alta	Fraco	Sim
Chuva	Média	Alta	Fraco	Sim
Chuva	Fria	Normal	Fraco	Sim
Chuva	Fria	Normal	Forte	Não
Nublado	Fria	Normal	Forte	Sim
Limpo	Média	Alta	Fraco	Não
Limpo	Fria	Normal	Fraco	Sim
Chuva	Média	Normal	Fraco	Sim
Limpo	Média	Normal	Forte	Sim
Nublado	Média	Alta	Forte	Sim
Nublado	Quente	Normal	Fraco	Sim
Chuva	Média	Alta	Forte	Não

Começamos por calcular a entropia:

$$\text{Entropia}(S) = -\frac{9}{14} \log_2 \frac{9}{14} - \frac{5}{14} \log_2 \frac{5}{14} = 0,940$$

Podemos agora calcular os ganhos de informação relativos a cada um dos atributos. Vejamos em pormenor o cálculo relativo ao atributo Aspecto:

$$\begin{aligned} \text{Ganho}(S, \text{Aspecto}) &= \text{Entropia}(S) - \sum_{v \in \{\text{Limpo}, \text{Nublado}, \text{Chuva}\}} \frac{|S_v|}{|S|} \text{Entropia}(S_v) = \\ &0,940 - \frac{5}{14} \text{Entropia}(S_{\text{Limpo}}) - \frac{4}{14} \text{Entropia}(S_{\text{Nublado}}) - \frac{5}{14} \text{Entropia}(S_{\text{Chuva}}) = \\ &0,940 - \frac{5}{14} \left( -\frac{2}{5} \log_2 \frac{2}{5} - \frac{3}{5} \log_2 \frac{3}{5} \right) - \frac{4}{14} \cdot 0 - \frac{5}{14} \left( -\frac{3}{5} \log_2 \frac{3}{5} - \frac{2}{5} \log_2 \frac{2}{5} \right) = 0,246 \end{aligned}$$

De forma análoga, tem-se:

$$\text{Ganho}(S, \text{Temperatura}) = 0,029$$

$$\text{Ganho}(S, \text{Humidade}) = 0,151$$

$$\text{Ganho}(S, \text{Vento}) = 0,048$$

Desta forma, o atributo escolhido será Aspecto. Cria-se assim o primeiro nó da árvore de decisão com o atributo Aspecto, e três ramos correspondentes a cada um dos valores possíveis: limpo, nublado e chuva.

No caso Aspecto = nublado, todos os exemplos são da classe sim. Assim, este ramo vai dar a uma folha, com a classe sim. Nos restantes casos vai-se continuar a aplicar o algoritmo aos subconjuntos  $S_{\text{Limpo}}$  e  $S_{\text{Chuva}}$ . O aspecto final da árvore é o da figura 2.3.

### 2.3.1.3. Avaliação dos resultados

A avaliação de árvores e regras de decisão pode ser feita de duas formas: qualitativa e quantitativa.

Na forma qualitativa, mais subjectiva, podemos considerar a legibilidade da árvore ou regras de decisão. A árvore ou as regras devem ser de fácil compreensão para um utilizador final no seu processo de decisão.

Na forma quantitativa pretende-se fazer uma estimativa do erro na classificação feita pelo modelo gerado. Esta estimativa mede-se pela percentagem de erros da classificação. O erro será sempre estimado pela razão entre o número de casos errados e o tamanho do conjunto de teste.

Os métodos mais comuns usados para estimar o erro destes algoritmos são:

- **separação dos dados em dois subconjuntos.** Uma parte dos dados é usada como conjunto de treino, gerando o modelo, sendo os restantes usados para teste;
- **validação cruzada.** O conjunto de dados  $C$  é dividido em  $n$  subconjuntos. Para cada subconjunto  $S$ , é usado o subconjunto  $C \setminus S$  como conjunto de treino, sendo  $S$  usado como conjunto de teste. Quanto maior for o valor de  $n$ , maior será a fiabilidade.
- **excluir um elemento.** Igual à validação cruzada quando  $n$  é igual ao número de tuplos. Para cada elemento, são usados todos os elementos, excepto esse, para treino, usando-se o excluído para teste.

#### 2.3.1.4. Evolução do algoritmo

O algoritmo ID3 original usa este algoritmo na sua forma mais simples, sendo depois estendido de modo a incorporar alguns melhoramentos (sendo renomeado para C4.5): evitar sobreajustamento dos dados (corte de nós em que o acréscimo de erro é reduzido e simplificação de regras), incorporação de valores contínuos, medidas alternativas para selecção de atributos, manipular exemplos com dados por preencher e atributos com custos diferentes.

A necessidade das medidas alternativas teve a ver com o facto de o ganho de informação usado no ID3 favorecer os atributos com mais valores possíveis. Como caso extremo, um atributo “data de nascimento” numa base de dados académica iria ter um ganho muito elevado mas iria, muito provavelmente, gerar uma árvore de decisão com vários ramos e apenas um nó: a raiz.

Assim, com o C4.5 surgiu uma medida alternativa que é o rácio de ganho (*gain ratio*), que penaliza este tipo de atributos, usando o termo dividir informação (*split information*), que é sensível à forma como o atributo divide uniformemente os dados:

$$\text{SplitInformation}(S, A) = - \sum_{i=1}^c \frac{|S_i|}{|S|} \log_2 \frac{|S_i|}{|S|}$$

Onde  $S_1$  até  $S_c$  são os  $c$  subconjuntos de exemplos resultantes da partição de  $S$  pelo atributo  $A$  com  $c$  valores possíveis. O rácio de ganho será:

$$\text{GainRatio}(S, A) = \frac{\text{Ganho}(S, A)}{\text{SplitInformation}(S, A)}$$

Relativamente às regras de decisão, procedeu-se, a partir do C4.5, à simplificação das mesmas, como uma forma de evitar o sobreajustamento dos dados. Assim, após a geração das regras a partir dos caminhos da raiz até às folhas das árvores, são retiradas todas as condições das premissas que resultem num aumento da precisão estimada. Após a simplificação, as regras são ordenadas pela precisão estimada, sendo a classificação realizada de acordo com esta ordem.

Do C4.5 para o C5.0 existem algumas alterações. A mais significativa tem a ver com as regras. Enquanto no C4.5 as regras são ordenadas (a primeira a satisfazer as premissas é a que é disparada), no C5.0 existe um mecanismo de votação, que selecciona a classe a atribuir ao novo caso.

### 2.3.2. Análise de ligações

A análise de ligações determina relações entre campos numa tabela da BD. O objectivo é derivar correlações entre vários campos satisfazendo valores de suporte e confiança. A forma mais conhecida de análise de ligações é a de regras de associação, ligada à análise de dados de cesto de compras (*basket analysis*).

#### 2.3.2.1. Regras de Associação

As regras de associação aparecem descritas, pela primeira vez, em [Agrawal, et al., 1993]. A ideia é encontrar associações entre diferentes itens de forma a que a presença de alguns itens num registo impliquem a presença de outros. Esta técnica é definida para um tipo especial de dados (vulgarmente conhecidos por *basket data*) onde um tuplo consiste num conjunto de atributos binários chamados itens. A cada tuplo chama-se uma transacção.

Tomando como exemplo um cesto de compras dum supermercado, cada produto existente no supermercado será um item: fruta, vegetais, leite, ovos, etc... Em cada transacção (compra) do cliente, teremos um tuplo com a informação dos produtos que são comprados (cada item com o valor verdadeiro ou falso). A partir

daqui podemos inferir algumas regras de associação como, por exemplo: “90% dos clientes que compram fruta e vegetais também compram leite”.

Formalmente, uma regra de associação é uma expressão da forma  $X \Rightarrow Y$ , onde  $X$  e  $Y$  são conjuntos de itens mutuamente exclusivos ( $X \cap Y = \emptyset$ ). O significado intuitivo deste tipo de regras é que as transações de bases de dados que contêm o conjunto  $X$  têm tendência a conter o conjunto  $Y$ .

### 2.3.2.2. Algoritmo Apriori

Para ajudar a descrever o algoritmo *apriori* [Agrawal & Srikant, 1994] considere-se o exemplo da tabela 2.2:

Tabela 2.2 – Conjunto de transações

Transações
134
235
1235
25

Uma regra de associação para este caso poderia ser  $\{23\} \Rightarrow \{5\}$ , isto é, quando o cliente compra os produtos 2 e 3, tende a comprar também o produto 5.

Cada regra de associação  $X \Rightarrow Y$  é caracterizada por dois factores: suporte e confiança. O suporte é a razão entre o número de tuplos que satisfazem  $X$  e  $Y$  simultaneamente e o número total  $n$  de tuplos:

$$\text{Suporte} = \frac{|X \cup Y|}{n}$$

A confiança é a razão entre o número de tuplos que satisfazem  $X$  e  $Y$  simultaneamente e o número total de tuplos que satisfazem  $X$ :

$$\text{Confiança} = \frac{|X \cup Y|}{|X|}$$

No exemplo anterior, teremos:

$$\text{Suporte}(\{23\} \Rightarrow \{5\}) = \frac{|\{235\}|}{n} = \frac{2}{4} = 0,5$$

$$\text{Confiança}(\{23\} \Rightarrow \{5\}) = \frac{|\{235\}|}{|\{23\}|} = \frac{2}{2} = 1$$

A descoberta de regras de associação realizada pelo algoritmo *Apriori*, consiste em extrair todas as regras em que o suporte e a confiança sejam superiores ou iguais a valores definidos pelo utilizador.

Assume-se, por simplicidade, que os itens numa transacção e os conjuntos de itens se encontram ordenados por ordem lexicográfica. Define-se tamanho como o número de itens de um conjunto. Um conjunto de itens de tamanho  $k$  é chamado  $k$ -conjunto de itens. Um  $k$ -conjunto  $c$  de itens é representado por  $c[1].c[2] \dots c[k]$ , tendo associado o suporte para este conjunto.

No algoritmo vamos usar as seguintes notações:

- $L_k$  – conjunto de grandes  $k$ -conjuntos de itens, isto é, com suporte acima do mínimo;
- $C_k$  – conjunto de  $k$ -conjuntos de itens candidatos (potencialmente grandes);
- $D$  – conjunto das transacções.

O primeiro passo do algoritmo consiste em contar o número de ocorrências de cada item para determinar  $L_1$ . No passo de ordem  $k$ , existem duas fases. Em primeiro lugar,  $L_{k-1}$  é usado para gerar  $k$ -conjuntos candidatos  $C_k$ . Em seguida, é contado o suporte dos candidatos em  $C_k$ . Na figura 2.5 está descrito o algoritmo *apriori*. Na figura 2.6 temos a função *apriori-gen*, que gera  $C_k$ .

Consideremos o exemplo da tabela 2.2. Vamos supor que o suporte mínimo é 50%, isto é 2 transacções. Teremos, então, os passos da figura 2.7.

A partir do conjunto de itens, serão geradas as regras de associação possíveis (a partir de  $L_2$  e  $L_3$ ):

$$\begin{aligned} &\{1\} \Rightarrow \{3\}; \quad \{3\} \Rightarrow \{1\}; \quad \{2\} \Rightarrow \{3\}; \quad \{3\} \Rightarrow \{2\}; \\ &\{2\} \Rightarrow \{5\}; \quad \{5\} \Rightarrow \{2\}; \quad \{3\} \Rightarrow \{5\}; \quad \{5\} \Rightarrow \{3\}; \\ &\{2\ 3\} \Rightarrow \{5\}; \quad \{2\ 5\} \Rightarrow \{3\}; \quad \{3\ 5\} \Rightarrow \{2\}. \end{aligned}$$



```

 $L_1 = \{\text{grandes 1-conjuntos de itens}\}$ 
para ( $k = 2; L_{k-1} \neq \emptyset; k++$ ) fazer
     $C_k = \text{apriori\_gen}(L_{k-1});$  // novos candidatos
    para todos transações  $t \in D$  fazer
         $C_t = \text{subconjunto}(C_k, t);$  // candidatos contidos em t
        para todos candidatos  $c \in C_t$  fazer
             $c.\text{count}++;$ 
        fim
         $L_k = \{c \in C_k : c.\text{count} \geq \text{suporte\_mínimo}\}$ 
    fim
 $\text{Resposta} = \bigcup_k L_k$ 

```

Figura 2.5 - Algoritmo apriori

Passo 1 (junção):

```

insert into  $C_k$ 
select  $p[1], p[2], \dots, p[k-1], q[k-1]$ 
from  $L_{k-1} p, L_{k-1} q$ 
where  $p[1] = q[1], \dots, p[k-2] = q[k-2], p[k-1] < q[k-1];$ 

```

Passo 2 (corte):

```

para todos conjuntos de itens  $c \in C_k$  fazer
    para todos  $(k-1)$ -subconjuntos de itens  $s$  de  $c$  fazer
        se  $(s \notin L_{k-1})$  então
            delete  $c$  from  $C_k;$ 

```

Figura 2.6 - Função apriori\_gen

Se colocarmos a confiança mínima em, por exemplo, 80%, teremos as seguintes regras de associação (todas com confiança 100%):

$\{1\} \Rightarrow \{3\}; \{2\} \Rightarrow \{5\}; \{5\} \Rightarrow \{2\}; \{2\ 3\} \Rightarrow \{5\}; \{3\ 5\} \Rightarrow \{2\}.$

Todas as regras têm a mesma confiança, mas duas têm maior suporte (75%):

$\{2\} \Rightarrow \{5\}; \{5\} \Rightarrow \{2\}.$

$L_1$		$C_2$	
Conj. itens	N.º Trans.	Conj. itens	N.º Trans.
{1}	2	{1 2}	1
{2}	3	{1 3}	2
{3}	3	{1 5}	1
{5}	3	{2 3}	2
		{2 5}	3
		{3 5}	2

$L_2$		$C_3$	
Conj. itens	N.º Trans.	Conj. itens	N.º Trans.
{1 3}	2	{2 3 5}	2
{2 3}	2		
{2 5}	3		
{3 5}	2		

$L_3$	
Conj. itens	N.º Trans.
{2 3 5}	2

Figura 2.7 - Passos do algoritmo apriori

Assim, estas regras são mais fortes que as regras  $\{2\ 3\} \Rightarrow \{5\}$  e  $\{3\ 5\} \Rightarrow \{2\}$ , respectivamente. Em ambos os casos, o uso de mais um atributo na premissa não aumenta a confiança, e tem menos suporte que as duas primeiras. Podemos, desta forma, desprezar estas últimas.

Como se confirmará mais à frente, se a uma regra acrescentarmos um atributo na sua premissa, o suporte da regras será menor ou igual, ao passo que a confiança aumentar. Assim, a partir do momento em que, ao acrescentarmos um atributo à premissa de uma regra, a confiança se mantiver e o suporte diminuir, desprezamos esta nova regras. Se, na mesma situação, o suporte se mantiver mas a confiança aumentar, desprezamos a regra inicial. Nos casos em que o suporte diminui e a confiança aumenta, terá que ser feita uma análise mais detalhada. Pode-se dar mais valor à confiança, ao suporte ou, por exemplo, a uma função que tenha por parâmetro estes dois factores.

## **2.4. Ferramentas de prospecção de dados**

Existem hoje em dia várias ferramentas implementando alguns dos algoritmos e métodos de prospecção de dados mais conhecidos. Em [Westphal & Blaxton, 1998], faz-se uma descrição muito completa sobre algumas das ferramentas que estão disponíveis, dividindo-as em três categorias:

- Ferramentas de análise de ligações;
- Ferramentas de visualização de panorama;
- Ferramentas de prospecção de dados quantitativa.

Nas três subsecções seguintes descreve-se sucintamente cada um destes tipos de ferramentas. Finalmente, na última subsecção, refere-se uma das ferramentas mais conhecidas ao nível de ECBD e prospecção de dados, que foi usada no estudo actual: o Clementine.

### **2.4.1. Ferramentas de análise de ligações**

A análise de ligações é o processo de construção de redes de objectos interconectados através de relacionamentos, de modo a expor padrões e tendências. Esta análise usa associações item a item para gerar redes de interacções e conexões a partir de conjuntos de dados definidos. Os diagramas de análise de ligações têm uma variedade de nomes: diagramas entidades-relações, redes conectadas, nós e ligações, grafos dirigidos, etc...

Os métodos de análise de ligações permitem a adição de dimensões a uma análise que as outras formas de visualização não suportam. Pela representação explícita das relações entre objectos, ganha-se uma perspectiva totalmente diferente de como os dados podem ser analisados e quais os tipos de padrões que podem ser descobertos.

Uma das desvantagens da análise de ligações é que o número agregado de registos que podem ser apresentados na maior parte dos diagramas é limitado, comparando com outros paradigmas de visualização. Assim, as análises tendem a centrar-se na verificação de subconjuntos de grandes conjuntos de dados.

Apesar disso, a análise de ligações fornece meios poderosos de fazer prospecção de dados visual, principalmente se se souber tirar partido das opções de saída de impressão, filtros e formatos de apresentação. Usados de forma adequada, estes sistemas permitem rapidamente identificar padrões, grupos emergentes e conexões entre gerações.

Algumas ferramentas de análise de ligações existentes são: NETMAP, Analyst's Notebook, Imagix 4D e Daisy.

### **2.4.2. Ferramentas de visualização de panorama**

As ferramentas de visualização de panorama descrevem ambientes nos quais os dados são colocados na forma de terrenos limitados geometricamente. Uma característica crítica deste tipo de visualização é que o posicionamento relativo dos elementos de dados dentro do terreno geométrico é usado para representar informação que é importante para a análise.

Os sistemas de visualização de panorama usam habitualmente representações abstractas em ambientes interactivos, com imagens virtuais a três dimensões para mostrar grandes quantidades de dados. Os panoramas podem ser efectivamente usados no modo de análise pró-activa, para encontrar tendências de alto nível, dentro de conjuntos de dados complexos. Isto é geralmente acompanhado por navegação nos panoramas e orientação visual dos dados para revelar tendências escondidas.

Algumas das ferramentas desenvolvidas usando este paradigma são direccionadas para aplicações de tempo real. As demonstrações criadas para alimentação de dados em tempo real são muitas vezes bastante poderosas na sua representação de valores paramétricos usando técnicas de animação e simulação.

Algumas das ferramentas de visualização de panorama existentes no mercado são: MineSet 2.0, Metaphor Mixer e In3D.

### **2.4.3. Ferramentas de prospecção de dados quantitativa**

As aproximações quantitativas são preferíveis e muitas vezes necessárias, quando se pretende obter estimativas do nível de significância ou de confiança. Além

disso, a análise quantitativa pode fornecer informação de sumário sobre diferenças de grupos globais e tendências, que pode ser usada para fornecer análises de primeira apanha (*first cut analyses*) de grandes conjuntos de dados. Nem sempre é necessário ter dados numéricos para conduzir análises quantitativas. Pode-se criar descritores quantitativos de dados qualitativos usando técnicas de abstracção de dados para criar variáveis temporárias ou contar dados a partir de rotinas de agregação.

Os diagramas quantitativos podem manipular volumes de dados extremamente grandes. Os diagramas da maioria das ferramentas de prospecção de dados quantitativas tendem a ter um aspecto tradicional e a poderem ser usados em agrupamentos (*clustering*), sumário e comparações de domínio. As implementações recentes aumentaram o alcance destes diagramas para fazer análise hiperdimensional. Estas características, adicionadas, fornecem ambientes ideais para a identificação de tendências lineares ou exponenciais ocorrendo nos conjuntos de dados.

Entre as ferramentas de prospecção de dados quantitativa encontram-se Enterprise Miner, Diamond, CrossGraphs e Clementine. Esta última foi usada na prospecção de dados da presente dissertação e será descrita em pormenor na próxima secção.

#### 2.4.4. Clementine

O sistema Clementine é uma das mais conhecidas ferramentas de prospecção de dados quantitativa. Este sistema contém uma variedade de funções de manipulação de dados e capacidades analíticas poderosas, num interface simples e bem desenhado. Este sistema corre nas principais plataformas, incluindo UNIX, Windows NT e VMS.

O sistema foi criado para analistas não técnicos que têm um conhecimento razoável sobre o seu conjunto de dados. Existem várias funções, desde a geração de gráficos simples e visualizações de dados, até análise com redes neuronais e indução de regras. O interface baseia-se na construção de modelos através do processo de arrastar e largar um conjunto de ícones, conectá-los e invocar procedimentos.

Uma das grandes vantagens do Clementine é a sua versatilidade, uma vez que inclui suporte para muitas das fases do processo de ECBD: acesso a dados, integração de vários tipos de dados, modelação, análise e apresentação de resultados.

Todas estas fases são transparentes ao nível do interface. Segundo [Westphal & Blaxton, 1998], esta é uma das poucas ferramentas existentes que possui este nível de integração.

Existem variadas opções no Clementine [Clementine, 1997]. Vamos aqui falar apenas das mais importantes e realçar os dois algoritmos de prospecção de dados que vamos utilizar: C5.0 e *apriori*.

Os ícones (ou nós) encontram-se divididos em seis conjuntos: fontes de dados, operações sobre registos, operações sobre campos, gráficos, modelação e saídas (*outputs*). Estes nós, na parte de baixo da janela da aplicação, podem ser arrastados para o painel principal. Os nós no painel principal podem ser ligados entre si de forma a definir um fluxo de dados. Os nós de fontes de dados são nós de entrada, os nós de gráficos e de saídas são nós de saída, e os nós de operações sobre registos e campos são nós intermédios. Os nós de modelação geram modelos (que aparecem na parte direita da janela da aplicação), os quais podem ser usados no processo de ECBD.

Os nós de fontes de dados permitem aceder directamente aos dados. Estes dados podem estar em ficheiros ASCII (de tamanho fixo ou variável), ficheiros exportados pelo Clementine ou SPSS, ou ainda através duma conexão ODBC, que permite ligar a vários tipos de Bases de Dados (no nosso caso foi usada esta conexão para ligar a uma base de dados em ORACLE).

Os nós de operações sobre registos fazem as operações habituais sobre registos em bases de dados, como selecção, união, ordenação, adição, selecção de elementos distintos e funções agregadas, e ainda um nó para obter uma amostra dos dados e outro para equilíbrio dos mesmos.

Os nós de operações sobre campos permitem filtrar campos, derivar novos campos e adicionar campos baseados em registos anteriores (útil para séries temporais). Existe ainda um nó para preenchimento de valores nulos, outro para transformar um campo em valor booleano e um último que permite alterar os tipos dos campos.

Os gráficos podem ser de vários tipos desde gráficos bidimensionais com uma ou mais funções, rede de ligações (dirigida ou não), histogramas, distribuições ou colecções.

Os modelos podem ser redes neuronais, redes de Kohonen, construção de regras, regressão, GRI (*Generalised Rule Induction*), e ainda os dois que já foram referidos (C5.0 e *apriori*).

Finalmente, temos os nós de saídas. Existem nós que permitem obter informação sobre os dados, como análise, relatório de qualidade dos dados e estatística. Outros nós permitem visualizar os dados, como tabelas e matrizes. Existem ainda outros nós que permitem exportar dados para ficheiros ASCII, folha de cálculo Excel, SPSS, e ainda através de ligação ODBC, que permite exportar para vários tipos de bases de dados (por exemplo, para ORACLE).

De uma forma geral a ECBD no Clementine processa-se em duas fases. Numa primeira fase, mais experimental, os dados são importados, realizam-se algumas operações sobre registos e campos, e os resultados podem ser visualizados graficamente ou, por exemplo, através duma tabela. Daqui podem observar-se algumas relações importantes nos dados. Numa segunda fase podem aplicar-se os algoritmos de prospecção de dados, para gerar os modelos finais. Os modelos gerados podem ser usados, posteriormente, para teste, aplicando-os aos mesmos dados ou a um conjunto de dados de teste.

O nó do algoritmo C5.0 tem várias opções. O tipo de saída pode ser regras ou árvores de decisão. O modo pode ser a construção do modelo ou validação cruzada. Podem agrupar-se valores simbólicos num mesmo ramo. Pode ainda usar-se o reforço (*boosting*) que consiste em produzir várias regras, construídas a partir dos erros das anteriores, sendo a resposta, na aplicação do modelo, dada por votação. O método pode favorecer a produção de regras com confiança máxima (com acerto em quase todo o conjunto de teste) ou deixá-las o mais geral possível. Pode ainda indicar-se o ruído percentual esperado dos dados.

O nó do algoritmo *apriori* também tem algumas opções. Existem as duas percentagens necessárias: o suporte e a confiança. Pode ainda determinar-se o número máximo de premissas admitidas numa regra e tratar os dados como sendo do tipo booleano.

# Capítulo 3

## Aplicação a uma base de dados académica

Na presente dissertação pretende-se estudar a aplicabilidade das técnicas de ECBD e prospecção de dados a uma base de dados académica. Usou-se para esse efeito a base de dados académica da Faculdade de Engenharia da Universidade do Porto (FEUP), onde são registados os dados relevantes relativos às entidades alunos, cursos e disciplinas, e às associações entre elas.

Na próxima secção descreve-se a forma como se procedeu à preparação dos dados. Na secção 3.2 são definidos os objectivos da prospecção de dados, na secção 3.3 apresenta-se a metodologia seguida e na última descreve-se a forma como se processou a prospecção de dados.

### **3.1. Preparação de dados**

A base de dados da FEUP, desenvolvida em ORACLE, guarda as informações académicas consideradas relevantes em qualquer instituição de ensino. Estes dados estão na base do sistema de informação universitário da FEUP (SIFEUP), que disponibiliza essa informação via *Internet*.

Na presente dissertação apenas foi utilizada uma parte dos dados, que se considerou mais relevante. Assim, após uma análise prévia, consideraram-se cinco tabelas, descritas na tabela 3.1 (os campos chave encontram-se sublinhados), estando as relações entre estas na figura 3.1. Todos os dados se referem ao período que decorreu entre os anos lectivos de 1994/95 e 1998/99.



Tabela 3.1 – Tabelas usadas na prospecção de dados

TABELA	ATRIBUTOS UTILIZADOS
ALUNOS	<ul style="list-style-type: none"> <li>• <u>NUMERO</u> (número mecanográfico);</li> <li>• CURSO (curso em que o aluno está inscrito).</li> </ul>
CURSOS	<ul style="list-style-type: none"> <li>• <u>CODIGO</u> (código do curso);</li> <li>• SIGLA (sigla do curso);</li> <li>• NOME (nome do curso).</li> </ul>
DISCIPLINAS	<ul style="list-style-type: none"> <li>• <u>CODIGO</u> (código da disciplina);</li> <li>• NOME (nome da disciplina);</li> <li>• CURSO (curso a que a disciplina pertence).</li> </ul>
DIS_CURRICULOS (Inscrição dos alunos em disciplinas)	<ul style="list-style-type: none"> <li>• <u>ALUNO</u> (número mecanográfico do aluno);</li> <li>• <u>DIS</u> (código da disciplina);</li> <li>• <u>A_LECTIVO</u> (ano lectivo de inscrição);</li> <li>• ANO (ano curricular da inscrição).</li> </ul>
ITENS_RES_DIS (Resultados das provas de avaliação efectuadas)	<ul style="list-style-type: none"> <li>• <u>ALUNO</u> (número mecanográfico do aluno);</li> <li>• <u>DIS</u> (código da disciplina);</li> <li>• <u>A_LECTIVO</u> (ano lectivo do resultado);</li> <li>• <u>EPOCA_RES</u> (época de exames do resultado);</li> <li>• RESULTADO (resultado da prova).</li> </ul>

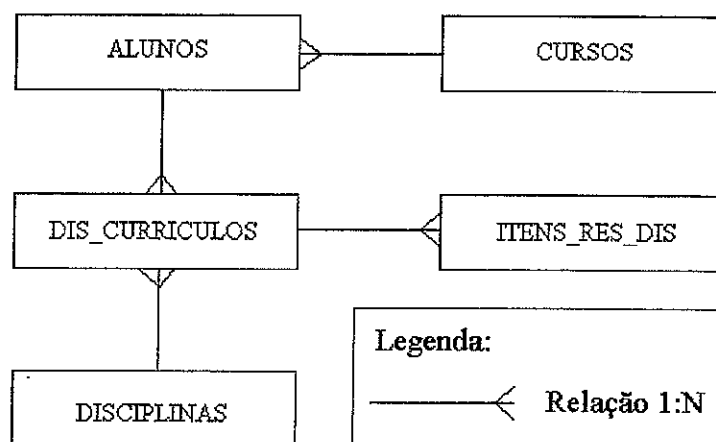


Figura 3.1 - Relações entre as tabelas

Dos cursos existentes, procedeu-se à escolha de um ou mais que pudessem ser representativos. Após alguma análise, acabou por ser escolhido o curso de Licenciatura em Engenharia Informática e Computação.

A razão para a escolha deste curso reside no facto de, durante os cinco anos a que se reporta esta análise, ter existido apenas uma mudança no Plano de Estudo. Essa mudança consistiu na troca de semestre entre duas disciplinas do primeiro ano: EIC1106 e EIC1206. Não existindo na base de dados informação completa sobre as equivalências e regimes de transição de planos de estudo, seria difícil conseguir fazer uma análise correcta nos outros cursos.

As disciplinas são todas semestrais e estão divididas em dezasseis áreas científicas. Na tabela 3.2 encontram-se as áreas científicas com as respectivas siglas e, na tabela 3.3, encontram-se descritas as disciplinas ordenadas por ano e semestre, contendo a sigla da área científica.

*Tabela 3.2 – Áreas científicas*

<b>SIGLA</b>	<b>ÁREA CIENTÍFICA</b>
AED	Algoritmos e Estruturas de Dados
ARQ	Arquitectura de Computadores
ASP	Aspectos Sociais e Profissionais
BDS	Bases de Dados e Sistemas de Tratamento de Informação
CNS	Computação Numérica e Simbólica
COM	Técnicas de Comunicação
FIS	Modelação Física
GMQ	Gestão e Métodos Quantitativos
IAR	Inteligência Artificial e Robótica
IHM	Interacção Homem-Máquina
IND	Informática Industrial
LPR	Linguagens de Programação
MAT	Matemática
MES	Metodologia e Engenharia de Software
SED	Sistemas Electrónicos Digitais
SOP	Sistemas Operativos

Tabela 3.3 - Disciplinas de Engenharia Informática e Computação

ANO	SEM.	CÓDIGO	NOME	ÁREA CIENT.	
1	1	EIC1100	Análise Matemática I	MAT	
		EIC1101	Algebra	MAT	
		EIC1102	Introdução à Programação I	LPR	
		EIC1103	Circuitos e Sistemas I	FIS	
		EIC1104	Electrónica Digital	SED	
		EIC1106	Técnicas de Comunicação e Inglês Funcional	COM	
	2	2	EIC1200	Análise Matemática II	MAT
			EIC1202	Introdução à Programação II	LPR
			EIC1203	Circuitos e Sistemas II	FIS
			EIC1204	Arquitectura de Computadores	ARQ
			EIC1205	Sistemas Digitais	SED
			EIC1206	Física Experimental	FIS
2	1	EIC2100	Física	FIS	
		EIC2101	Análise Numérica	CNS	
		EIC2102	Algoritmos e Estruturas de Dados I	AED	
		EIC2103	Sinais e Telecomunicações	FIS	
		EIC2104	Microprocessadores e Microcomputadores	ARQ	
		EIC2105	Laboratório de Circuitos e Sistemas	FIS	
	2	2	EIC2200	Sistemas de Informação e Bases de Dados	BDS
			EIC2201	Computação Gráfica e Interfaces	IHM
			EIC2202	Algoritmos e Estruturas de Dados II	AED
			EIC2203	Instrumentação Electrónica	SED
EIC2204	Estatística Aplicada	GMQ			
EIC2205	Laboratório de Arquitectura de Computadores	ARQ			
3	1	EIC3100	Novas Metodologias de Programação	MES	
		EIC3101	Sistemas Operativos Clássicos	SOP	
		EIC3102	Teoria da Computação I	MAT	
		EIC3103	Programação Paralela	AED	
		EIC3104	Investigação Operacional	GMQ	
		EIC3105	Laboratório de Computação Gráfica e Interfaces	IHM	
	2	2	EIC3200	Tecnologia de Sistemas de Gestão de Bases de Dados	BDS
			EIC3201	Compiladores	LPR
			EIC3202	Teoria da Computação II	MAT
			EIC3203	Introdução à Inteligência Artificial	IAR
EIC3204	Gestão de Empresas	GMQ			
EIC3205	Laboratório de Sistemas Operativos	SOP			
4	1	EIC4100	Engenharia de Software	MES	
		EIC4101	Redes de Comunicação de Dados	ARQ	
		EIC4102	Informática Industrial	IND	
		EIC4103	Sistemas Periciais e Robótica	IAR	
		EIC4104	Contabilidade	CMQ	
		EIC4105	Laboratório de Sistemas de Gestão de Bases de Dados	BDS	
	2	2	EIC4200	Computação Numérica e Simbólica	CNS
			EIC4201	Aspectos Sociais da Informática	ASP
			EIC4202	Análise de Investimentos	GMQ
			EIC4203	Laboratório de Informática Avançada	IAR
			EIC4204	Concepção e Fabrico Assistido por Computador	IND
			EIC4205	Sistemas de Comunicação Integrada	BDS
EIC4206	Sistemas Operativos Distribuídos	SOP			
5	1	EIC5100	Gestão de Projectos de Software	MES	
		EIC5101	Sistemas de Apoio a Decisão	GMQ	
		EIC5102	Seminários	ASP	
		EIC5103	Logística	IND	
		EIC5104	Planeamento e Gestão da Produção	IND	
		EIC5105	Sentidos Artificiais	IHM	
		EIC5106	Sistemas Multimédia	IHM	
	2	2	EIC5200	Projecto	IND
			EIC5201	Estágio	IND

Antes da aplicação das técnicas de prospecção de dados é necessário preparar convenientemente os dados. A primeira operação feita, tendo em conta que um aluno só pode estar inscrito num curso, foi dividir os dados por curso. Desta forma, o espaço de pesquisa diminui consideravelmente, tornando a prospecção mais eficiente. Poder-se-ia também pensar em fazer outras divisões lógicas, como por ano lectivo ou curricular. No entanto, esta divisão não pode ser feita, uma vez que o comportamento dos alunos em anos anteriores pode ser importante no processo de ECBD.

Assim, para cada ano curricular, do primeiro ao quinto, foram construídas novas tabelas, da seguinte forma:

- para os anos lectivos compreendidos entre 1994/95 e 1998/99, seleccionaram-se, a partir da tabela DIS\_CURRICULOS, os alunos que estavam inscritos nesse ano curricular;
- para estes alunos, foram seleccionados todos os resultados de exames obtidos em anos lectivos e curriculares iguais ou anteriores, a partir das tabelas DISC\_CURRICULOS e ITENS\_RES\_DIS, considerando um novo campo relativo à classificação na disciplina (0 ou 1, no caso em que se pretende apenas ter em conta a aprovação ou não à disciplina, ou “Muito Bom”, “Bom”, “Suficiente” e “Reprovado”, nos restantes casos);
- é construída uma tabela cruzada (*crosstab query*), sendo as linhas referentes aos alunos, as colunas referentes às disciplinas e o valor referente à classificação obtida na disciplina (este valor é obtido calculando o máximo de todos os resultados). Note-se que, em alguns casos, vão aparecer valores em branco, quando o aluno nunca foi a nenhum exame de determinada disciplina, sendo esses casos considerados como 0 ou “Reprovado”.

Desta forma, presume-se que a aprovação a uma disciplina implica a assimilação dos conhecimentos presentes no seu programa. Assim, considera-se que uma disciplina poderá ser influenciada por disciplinas a que o aluno já tenha obtido aprovação anteriormente.

Como já foi referido, houve uma alteração no plano de estudo, que resultou na troca de semestre entre de duas disciplinas do primeiro ano. Na mudança de

semestre ocorreu também uma mudança do código da disciplina. Assim, a disciplina EIC1105 passou do primeiro para o segundo semestre, passando a ter o código EIC1206, ao passo que a EIC1201 passou do segundo para o primeiro semestre, passando a ter o código EIC1106. Na fase de pré-processamento converteram-se os códigos anteriores para os novos.

Como os dados são gerados para vários anos lectivos, os alunos que foram avaliados em mais do que um ano aparecem apenas uma vez. Por exemplo, um aluno que tenha sido aprovado a todas as disciplinas do primeiro ano, pode ter conseguido esse objectivo em um, dois, três ou mais anos.

Por um lado, pode-se pensar que os resultados poderiam ser melhores se este aluno aparecesse na tabela tantas vezes quantas tivesse frequentado o primeiro ano. No entanto, esta abordagem mostra-se mais tendenciosa, se pensarmos que, por exemplo, ao longo de cinco anos, temos alunos que estão a frequentar o curso pelas primeira, segunda, terceira, quarta e quinta vezes, o que faria com que o desempenho dos últimos fosse contado cinco vezes mais que os dos primeiros. Assim, o peso do desempenho dos últimos sobrepor-se-ia ao dos primeiros.

Os testes preliminares, com os dados separados num conjunto de treino e noutro de teste, confirmaram que a análise seguida consegue melhor percentagem de acertos que esta última hipótese.

Construíram-se também tabelas para identificação do perfil dos alunos, considerando as classes “Muito Bom” (3) , “Bom” (2), “Suficiente” (1) e “Reprovado” (0), a indicarem notas superiores a quinze, entre catorze e quinze, entre dez e treze, e inferiores a dez, respectivamente.

Considerou-se, para cada aluno inscrito no primeiro ano em 1994/95, as médias das notas obtidas, incluindo as notas inferiores a dez e considerando as faltas como zero, dividindo em quatro perfis:

- F (Fracó), para média inferior a dez;
- S (Suficiente), para média entre dez e treze;
- B (Bom), para média entre catorze e quinze;
- MB (Muito Bom), para média superior a quinze.

Consideraram-se três classes, indicando a situação ao fim do quinto ano:

- C – Curso concluído;
- F – Finalista, isto é, com hipótese de concluir o curso no próximo ano;
- N – Não finalista.

Como estes dados se referem apenas aos alunos que entraram na licenciatura no ano lectivo de 1994/95 (primeiro ano da licenciatura), resolveu-se estender este estudo aos anos de 1995/96 e 1996/97, considerando o resultado ao fim de três anos. Consideraram-se, para esta abordagem quatro classes, indicando a situação ao fim do terceiro ano, fazendo correspondência ao quinto ano:

- C3 – Terceiro ano concluído;
- F32 – Inscrito no terceiro ano, com entre trinta e trinta e cinco disciplinas feitas;
- F31 – Inscrito no terceiro ano, com entre vinte e quatro, e vinte e nove disciplinas feitas;
- N3 – Não está inscrito no terceiro ano.

### **3.2. Objectivos da prospecção de dados**

Os dados, para além da sua função óbvia de gestão da informação relativa à actividade lectiva da FEUP, podem conter informação implícita relevante. Em muitos casos, com uma simples interrogação à base de dados, pode saber-se quais as disciplinas a que os alunos têm maiores dificuldades, obtendo as percentagens de aprovações. Assim, para prever quantos alunos serão aprovados a uma determinada disciplina, podem usar-se as percentagens dos anos anteriores, supondo que a tendência se mantém.

No entanto, as percentagens não se mantêm estáveis ao longo dos anos. Por outro lado estas percentagens, por si só, não fornecem nenhuma pista sobre as razões do desempenho de uma dada disciplina.

A primeira hipótese que é estudada é a existência de disciplinas que, pelo seu conteúdo, tenham um papel preponderante para o sucesso doutras. Por exemplo, parece natural que um aluno, que ainda não tenha obtido aprovação a Análise

Matemática I e esteja inscrito a Análise Matemática II, tenha muitas dificuldades em obter aprovação nesta disciplina.

As interdependências entre as disciplinas podem acontecer a vários níveis. São estudadas as dependências entre disciplinas do mesmo ano, de anos consecutivos ou entre disciplinas da mesma área científica. Levando em consideração estas hipóteses, a análise tradicional torna-se mais difícil. O uso de algoritmos de prospecção de dados pode tornar esta análise mais fácil e de melhor compreensão para o utilizador final.

A segunda análise efectuada visa a detecção precoce de perfis de alunos. Uma das hipóteses é ver quais as disciplinas que funcionam como indicadores para um bom desempenho escolar dos alunos. Pode-se verificar se é possível, no primeiro ano de um aluno, prever qual a situação em que este estará ao fim de três anos ou dos cinco de duração normal do curso.

### **3.3. Metodologia**

Vamos, de seguida, focar a metodologia utilizada, referindo as técnicas de prospecção de dados usadas e uma aplicação desenvolvida durante a presente dissertação, que permite reduzir o domínio de análise de regras de associação.

#### **3.3.1. Árvores e regras de decisão**

As árvores e regras de decisão são usados de um ponto de vista de classificação. Pretende-se, a partir de um conjunto de dados de teste, prever a classe de um determinado atributo. Para este estudo usou-se o algoritmo C5.0, um dos mais conhecidos que geram árvores e regras de decisão.

A classificação usando árvores e regras de decisão permite fazer um previsão para casos futuros. Desta forma é possível, por exemplo, prever a classe de atributos relativos a um ano lectivo, e conseguir prever quantos alunos se irão inscrever a cada disciplina no seguinte. Este tipo de previsão pode ser útil para conseguir ter uma ideia antecipada das necessidades de serviço docente para o ano lectivo seguinte, e

também para se poder analisar se as instalações estão a chegar a um ponto de saturação.

Apesar de se poder obter relações de interdependência entre disciplinas, como acontece com as regras de associação, o objectivo das árvores e regras de decisão é classificar efectivamente os casos de entre uma das classes possíveis.

Consideram-se dois tipos de classificação. A primeira, usa duas classes, 1 e 0 correspondentes à aprovação ou não numa disciplina, respectivamente. Foi ainda considerada a subdivisão em mais classes: “Muito Bom”, “Bom”, “Suficiente” e “Reprovado”.

### 3.3.2. Regras de associação

A técnica de selecção de regras de associação, usando o algoritmo *apriori*, permitem responder à questão de relacionar o sucesso em algumas disciplinas com o de uma outra. Assim é possível encontrar regras da forma: “os alunos que já obtiveram aprovação a Análise Matemática I e Álgebra também obtêm aprovação a Análise Matemática II”. Esta regras podem ser expressas de forma mais simplificada:  $AM\ I \ \& \ Alg \Rightarrow AM\ II$ .

Como se viu no capítulo anterior, a estas regras vão estar ligados dois parâmetros: o suporte, isto é, qual a percentagem de casos em que a premissa e a conclusão se verificam, simultaneamente, relativamente ao total de alunos, e a confiança, isto é, relativamente aos casos em que a premissa se verifica qual a percentagem em que a conclusão também se verifica. Considera-se um caso um aluno inscrito num determinado ano curricular. Estes dois parâmetros são importantes na interpretação de resultados. Por exemplo, considerem-se as três regras seguintes:

$$AM\ I \ \& \ Alg \Rightarrow AM\ II$$

$$AM\ I \Rightarrow AM\ II$$

$$Alg \Rightarrow AM\ II$$

É fácil de perceber que as duas últimas regras terão um suporte maior ou igual ao da primeira. Por outro lado, a confiança poderá variar em qualquer sentido. Se aumentar, estamos perante uma regra nitidamente mais forte. No entanto, a



confiança poderá ser menor ou igual. Qual será então a regra a considerar como mais representativa? De uma forma geral se tivermos as duas regras seguintes:

$$A \& B \Rightarrow C$$

$$A \Rightarrow C$$

teremos sempre que jogar com os dois parâmetros. Se na segunda regra o suporte se mantiver e a confiança diminuir, então a primeira regra é nitidamente mais representativa. Por outro lado, se o suporte aumenta e a confiança se mantém inalterada ou aumenta, então será a segunda a mais representativa. Se o suporte aumenta e a confiança diminui, será necessária uma função de avaliação da representatividade.

Outra questão que se pode colocar é relativamente a uma disciplina que tenha uma percentagem de aprovações próxima dos 100%. Neste caso, poderemos ter várias regras em que esta disciplina aparece na conclusão, com um elevado grau de confiança.

Todas estas questões terão que ser tratadas com bastante cuidado durante a fase da análise. A forma como são interpretados o suporte e a confiança é a chave para obter informação útil e correcta a partir das regras de associação.

Apesar de não ser esse o seu objectivo principal, as árvores e regras de decisão podem indicar também associações importantes. Por exemplo, se tivermos a seguinte regra de decisão:

$$\text{Se } A=1 \text{ então } B=1;$$

isto corresponderá à regra de associação:

$$A \Rightarrow B$$

Esta relação poderá ser importante na análise das regras de associação, uma vez que poderá ajudar a decidir qual a regra que poderá ser mais importante.

### 3.3.3. Filtragem das regras de associação

Como já vimos anteriormente, para cada regra  $X \Rightarrow Y$ , temos:

$$\text{Suporte} = \frac{|X \cup Y|}{n}$$

$$\text{Confiança} = \frac{|X \cup Y|}{|X|}$$

Uma regra que tenha apenas o melhor suporte ou a melhor confiança não será necessariamente a melhor. Poderão existir, por exemplo, outras regras muito próximas, relativamente à melhor confiança, sendo o suporte bastante melhor.

Assim sendo, após a análise de alguns casos, definiram-se alguns critérios para escolher, de entre as regras a considerar, aquela que será a mais importante em termos de informação. Note-se que, logo à partida, o algoritmo *apriori* obriga a que sejam definidos o suporte e a confiança mínimos, construindo um conjunto de regras que satisfaça estas restrições.

Como o número de regras de associação pode ser elevado, foi construído um programa auxiliar para filtrar as regras, ficando apenas aquelas que resultam do algoritmo descrito em seguida. Este programa conseguiu reduzir de forma acentuada o número de regras a considerar.

Após a geração destas regras, existem algumas que podem ser imediatamente eliminadas. Para isso vamos escolher as regras  $RS$  e  $RC$ , com maior suporte e maior confiança, respectivamente. No caso de haver mais do que uma regra com maior suporte ou com maior confiança, escolhe-se aquela em que o outro valor seja maior.

No caso de existir uma ou mais regras com maiores suporte e confiança, simultaneamente, estas serão as mais fortes. Note-se que as outras regras também poderão ser importantes, mas não tanto como estas.

Vamos debruçar-nos sobre os casos em que não existem regras com os melhores suporte e confiança, simultaneamente. Consideremos  $(S_{RS}, C_{RS})$  e  $(S_{RC}, C_{RC})$  como sendo os suporte e confiança relativos às regras  $RS$  e  $RC$ , respectivamente, e sejam  $S_{min}$  e  $C_{min}$  os suporte e confiança mínimos. Se  $RS$  e  $RC$  não tiverem valores de suporte e confiança iguais, teremos a situação representada no gráfico da figura 3.2.

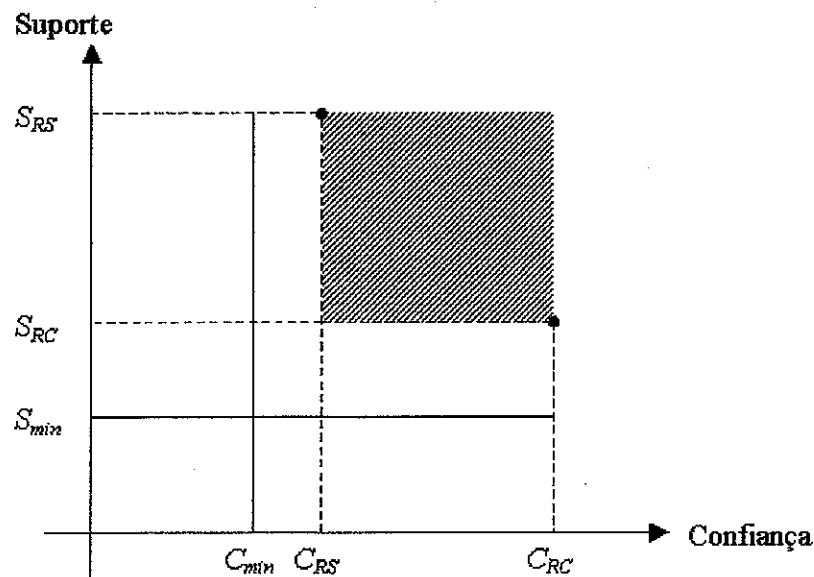


Figura 3.2 - Primeiro passo do algoritmo

Assim, apenas as regras que se encontrem dentro do domínio a sombreado na figura 3.2, incluindo os pontos referentes aos melhores suporte e confiança, serão consideradas.

Colocando de parte os pontos referentes aos melhores suporte e confiança, continua-se a aplicar o conceito para os restantes, considerando as regras com os segundos melhores suporte e confiança, *RS2* e *RC2*, respectivamente, como se pode ver na figura 3.3.

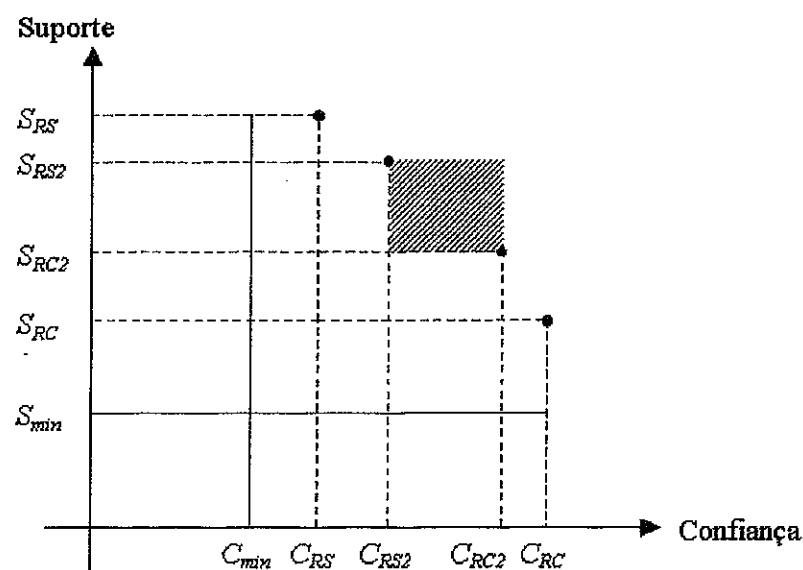


Figura 3.3 - Segundo passo do algoritmo

Este processo vai continuar até não existirem mais pontos a considerar. Os pontos que se obtêm referem-se às regras que têm melhor suporte e confiança em cada passo. Uma característica importante destes pontos é que as ordenações por suporte e confiança são inversas: a regra com melhor suporte é a que tem menor confiança, e vice versa.

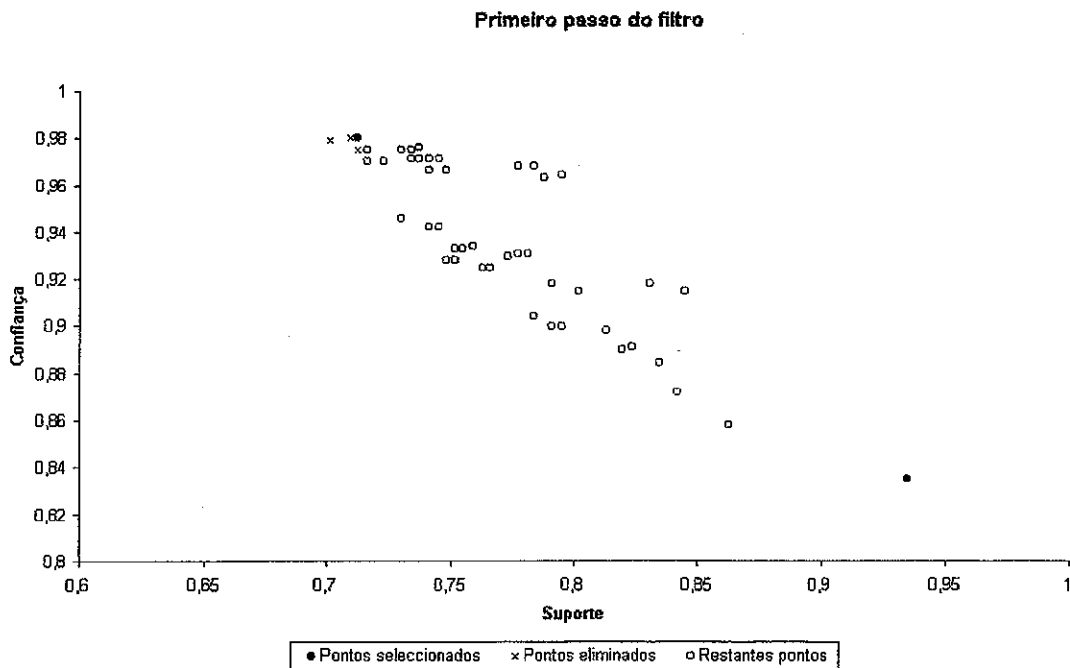
A partir daqui, ou se consideram critérios de desempate para escolha da regra mais importante (por exemplo, o produto do suporte pela confiança, que é uma das opções de ordenação do Clementine), ou se faz uma análise manual. Neste caso, o bom senso do analista será de grande importância.

Na figura 3.4 pode ver-se um conjunto de regras de associação gerado pelo *a priori* a partir do Clementine antes da aplicação do filtro (entre parêntesis estão o número de casos que suportam a regra e respectiva percentagem, e a confiança).

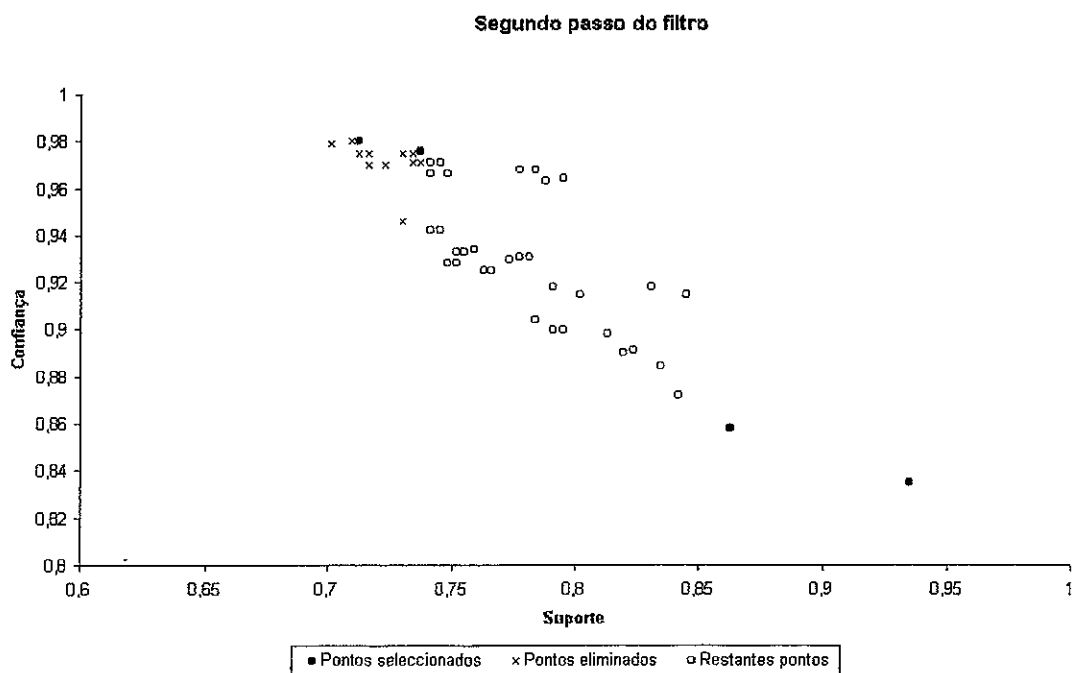
EIC1200	<=	EIC1100	(221:79.5%, 0.964)
EIC1200	<=	EIC1101 & EIC1100	(218:78.4%, 0.968)
EIC1200	<=	EIC1101 & EIC1102 & EIC1100	(206:74.1%, 0.971)
EIC1200	<=	EIC1101 & EIC1102 & EIC1103 & EIC1100	(198:71.2%, 0.98)
EIC1200	<=	EIC1101 & EIC1102 & EIC1103	(207:74.5%, 0.942)
EIC1200	<=	EIC1101 & EIC1102	(217:78.1%, 0.931)
EIC1200	<=	EIC1101 & EIC1103 & EIC1100	(204:73.4%, 0.975)
EIC1200	<=	EIC1101 & EIC1103	(216:77.7%, 0.931)
EIC1200	<=	EIC1101	(235:84.5%, 0.915)
EIC1200	<=	EIC1102 & EIC1100	(207:74.5%, 0.971)
EIC1200	<=	EIC1102 & EIC1103 & EIC1100	(198:71.2%, 0.98)
EIC1200	<=	EIC1102 & EIC1103	(213:76.6%, 0.925)
EIC1200	<=	EIC1102	(232:83.5%, 0.884)
EIC1200	<=	EIC1103 & EIC1100	(205:73.7%, 0.976)
EIC1200	<=	EIC1103	(229:82.4%, 0.891)
EIC1200	<=	EIC1104 & EIC1100	(208:74.8%, 0.966)
EIC1200	<=	EIC1104 & EIC1101 & EIC1100	(208:74.8%, 0.966)
EIC1200	<=	EIC1104 & EIC1101 & EIC1102 & EIC1100	(201:72.3%, 0.97)
EIC1200	<=	EIC1104 & EIC1101 & EIC1102 & EIC1103	(203:73.0%, 0.946)
EIC1200	<=	EIC1104 & EIC1101 & EIC1102	(211:75.9%, 0.934)
EIC1200	<=	EIC1104 & EIC1101 & EIC1103 & EIC1100	(199:71.6%, 0.975)
EIC1200	<=	EIC1104 & EIC1101 & EIC1103	(210:75.5%, 0.933)
EIC1200	<=	EIC1104 & EIC1101	(223:80.2%, 0.915)
EIC1200	<=	EIC1104 & EIC1102 & EIC1100	(201:72.3%, 0.97)
EIC1200	<=	EIC1104 & EIC1102 & EIC1103 & EIC1100	(195:70.1%, 0.979)
EIC1200	<=	EIC1104 & EIC1102 & EIC1103	(209:75.2%, 0.928)
EIC1200	<=	EIC1104 & EIC1102	(221:79.5%, 0.9)
EIC1200	<=	EIC1104 & EIC1103 & EIC1100	(199:71.6%, 0.975)
EIC1200	<=	EIC1104 & EIC1103	(220:79.1%, 0.9)
EIC1200	<=	EIC1104	(240:86.3%, 0.858)
EIC1200	<=	EIC1106 & EIC1100	(219:78.8%, 0.963)
EIC1200	<=	EIC1106 & EIC1101 & EIC1100	(216:77.7%, 0.968)
EIC1200	<=	EIC1106 & EIC1101 & EIC1102 & EIC1100	(204:73.4%, 0.971)
EIC1200	<=	EIC1106 & EIC1101 & EIC1102 & EIC1103	(206:74.1%, 0.942)
EIC1200	<=	EIC1106 & EIC1101 & EIC1102	(215:77.3%, 0.93)
EIC1200	<=	EIC1106 & EIC1101 & EIC1103 & EIC1100	(203:73.0%, 0.975)
EIC1200	<=	EIC1106 & EIC1101 & EIC1103	(215:77.3%, 0.93)
EIC1200	<=	EIC1106 & EIC1101	(231:83.1%, 0.918)
EIC1200	<=	EIC1106 & EIC1102 & EIC1100	(205:73.7%, 0.971)
EIC1200	<=	EIC1106 & EIC1102 & EIC1103 & EIC1100	(197:70.9%, 0.98)
EIC1200	<=	EIC1106 & EIC1102 & EIC1103	(212:76.3%, 0.925)
EIC1200	<=	EIC1106 & EIC1102	(228:82.0%, 0.89)
EIC1200	<=	EIC1106 & EIC1103 & EIC1100	(204:73.4%, 0.975)
EIC1200	<=	EIC1106 & EIC1103	(226:81.3%, 0.898)
EIC1200	<=	EIC1106 & EIC1104 & EIC1100	(206:74.1%, 0.966)
EIC1200	<=	EIC1106 & EIC1104 & EIC1101 & EIC1100	(206:74.1%, 0.966)
EIC1200	<=	EIC1106 & EIC1104 & EIC1101 & EIC1102	(209:75.2%, 0.933)
EIC1200	<=	EIC1106 & EIC1104 & EIC1101 & EIC1103	(209:75.2%, 0.933)
EIC1200	<=	EIC1106 & EIC1104 & EIC1101	(220:79.1%, 0.918)
EIC1200	<=	EIC1106 & EIC1104 & EIC1102 & EIC1100	(199:71.6%, 0.97)
EIC1200	<=	EIC1106 & EIC1104 & EIC1102 & EIC1103	(208:74.8%, 0.928)
EIC1200	<=	EIC1106 & EIC1104 & EIC1102	(218:78.4%, 0.904)
EIC1200	<=	EIC1106 & EIC1104 & EIC1103 & EIC1100	(198:71.2%, 0.975)
EIC1200	<=	EIC1106 & EIC1104 & EIC1103	(218:78.4%, 0.904)
EIC1200	<=	EIC1106 & EIC1104	(234:84.2%, 0.872)
EIC1200	<=	EIC1106	(260:93.5%, 0.835)

Figura 3.4 – Regras de associação para a disciplina EIC1200

Nas figuras 3.5 e 3.6 podem-se ver os dois primeiros passos da aplicação do filtro e, na figura 3.7, os pontos que se obtiveram da sua aplicação.



*Figura 3.5 - Primeiro passo do filtro*



*Figura 3.6 - Segundo passo do filtro*

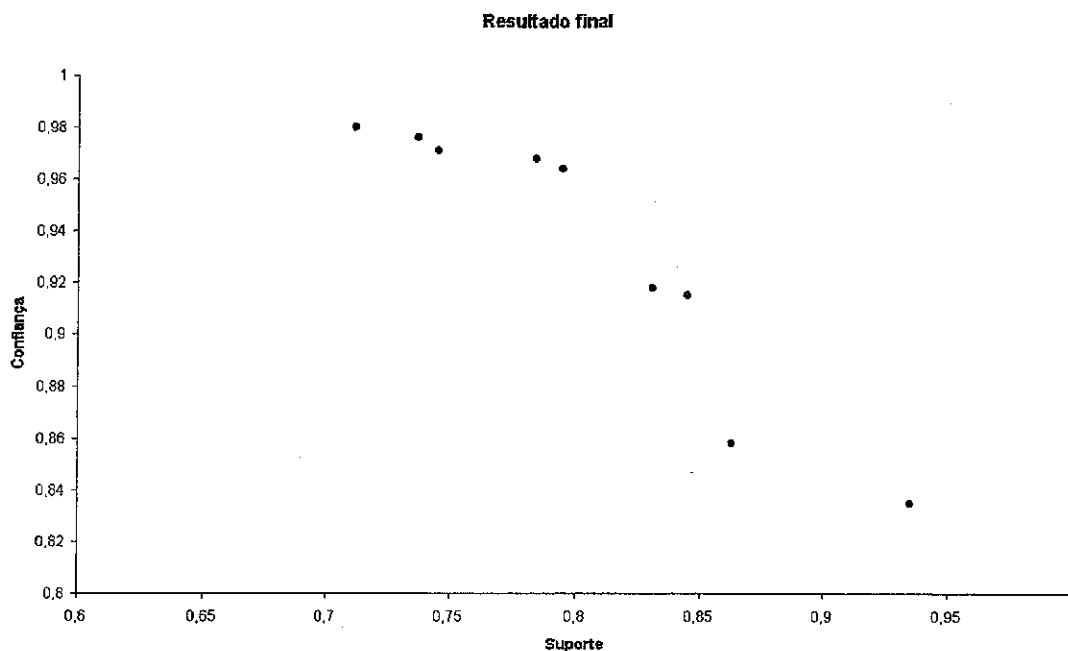


Figura 3.7 - Resultado final após aplicação do filtro

Assim, das cinquenta e seis regras inicialmente geradas, ficamos com apenas dez, que se encontram na figura 3.8, após a aplicação do filtro.

```

Melhor suporte: 260:93.5% (conf=83.5%)
Melhor confiança: 98.0% (sup=198:71.2%)

EIC1200 <= EIC1101 & EIC1102 & EIC1103 & EIC1100 (198:71.2%, 0.98)
EIC1200 <= EIC1102 & EIC1103 & EIC1100 (198:71.2%, 0.98)
EIC1200 <= EIC1106 (260:93.5%, 0.835)
EIC1200 <= EIC1103 & EIC1100 (205:73.7%, 0.976)
EIC1200 <= EIC1104 (240:86.3%, 0.858)
EIC1200 <= EIC1102 & EIC1100 (207:74.5%, 0.971)
EIC1200 <= EIC1101 (235:84.5%, 0.915)
EIC1200 <= EIC1101 & EIC1100 (218:78.4%, 0.968)
EIC1200 <= EIC1106 & EIC1101 (231:83.1%, 0.918)
EIC1200 <= EIC1100 (221:79.5%, 0.964)
    
```

Figura 3.8 - Saída produzida pelo filtro

De uma forma geral, a aplicação do filtro resulta na diminuição de aproximadamente 70% do número total de regras, nos casos estudados. Note-se que as regras foram habitualmente geradas com um suporte mínimo de 20% e uma confiança mínima de 80%. A alteração destes valores ou a sua aplicação a outros tipos de dados diferentes pode alterar a percentagem de diminuição do número total de regras.

### **3.4. Aplicação aos dados**

A prospecção de dados foi realizada usando a ferramenta Clementine. Desta forma é possível gerar os modelos pretendidos e, em alguns casos, aplicá-los directamente aos dados.

Para a geração de modelos, é necessário definir os campos em relação ao sentido dos dados: entrada, saída, ambos ou nenhum. No caso das regras de associação, existem várias disciplinas de saída, que irão aparecer na conclusão das regras, e várias de entrada, que poderão aparecer nas premissas. No caso das árvores e regras de decisão, temos apenas uma disciplina de saída, aquela que queremos classificar, e várias de entrada, que poderão aparecer na respectiva árvore ou regra de decisão.

#### **3.4.1. Considerações gerais**

Inicialmente foi necessário decidir quais as disciplinas que seriam usadas como entrada e saída. Analisando as várias hipóteses possíveis, chegou-se à conclusão que se poderiam gerar modelos das seguintes formas:

- geração de modelos para as disciplinas do ano curricular completo, usando as disciplinas do ano anterior;
- geração de modelos para as disciplinas do segundo semestre usando as do primeiro semestre do mesmo ano;
- geração dos modelos para a situação dos alunos ao fim de três ou cinco anos.

Esta última forma é independente das duas primeiras. Em relação a estas, ambas têm as suas limitações. No primeiro caso, será impossível classificar as disciplinas do primeiro ano, uma vez que não têm precedentes no plano de estudo. Uma hipótese a considerar seria a introdução na base de dados, para cada aluno, das suas notas de entrada e da prova específica, o que não existe actualmente.

O segundo caso depende da forma como está estruturado o calendário lectivo de cada escola. No caso da FEUP, os exames de recurso realizam-se após os da

época normal, antes do início do segundo semestre, o que possibilita este tipo de análise. Nas escolas onde os exames de recurso se realizam após o segundo semestre (habitualmente em Setembro) este tipo de análise não terá tanta fiabilidade, uma vez que as classificações terão que ser realizadas sem contemplar estes dados, para fornecerem informação em tempo útil.

### 3.4.2. Geração dos modelos

Os modelos são gerados usando o Clementine. Como já foi referido, antes da aplicação de cada algoritmo, são definidos quais os atributos de entrada (atributos que vão ser usados para geração do modelo) e atributos de saída (aqueles para os quais se vai tentar gerar o modelo).

Usando como nó inicial a leitura da tabela a ser tratada, referente a um determinado ano curricular, são gerados todos os modelos correspondentes, podendo também ser exportados para modo texto. O modelo gerado pelo algoritmo *apriori* usa todos os atributos de saída ao mesmo tempo, ao passo que, no caso do C5.0, tem-se sempre apenas um atributo de saída, de cada vez, aquele que se pretende classificar.

Os modelos gerados pelo C5.0 em Clementine podem ser guardados num formato interno do Clementine, para uso posterior com as outras tabelas, enquanto o modelo gerado pelo *apriori* não pode ser usado para interagir directamente com os dados. No caso do C5.0, o Clementine adiciona dois campos à tabela à qual o modelo é aplicado: um com o resultado da classificação e outro com a confiança dessa classificação. Assim, nos casos em que estamos a usar o conjunto de teste, podemos facilmente comparar a classe real com a classe prevista pelo C5.0.

Vejamos exemplos de modelos gerados no Clementine para a disciplina EIC1200 (Análise Matemática II). Já vimos, na figura 3.4, o exemplo das regras de associação. Na figura 3.9 temos um exemplo de uma árvore de decisão.

```
EIC1100 1 (110.0, 0.982) -> 1
EIC1100 0
  EIC1101 0 (20.0, 0.95) -> 0
  EIC1101 1
    EIC1103 1 (7.0, 0.714) -> 0
    EIC1103 0 (2.0, 1.0) -> 1
```

Figura 3.9 - Árvore de decisão para a disciplina EIC1200



Em cada linha tem-se a disciplina do nó e a classe a considerar. Nos casos em que o ramo leve a uma classificação, aparecem entre parêntesis o número de casos que satisfazem todo o percurso da raiz até à folha e a respectiva confiança, isto é, a razão entre os casos que correspondem à classe prevista e o número de casos dado, seguido da seta e da classe prevista. Por exemplo, na primeira linha temos cento e dez casos, tendo em 98,2% dos casos (108/110), a classificação 1. Esta percentagem não é de 100%, uma vez que, a partir da versão C4.5 do algoritmo, foi introduzido o corte de nós em que a redução do erro é reduzida, para evitar sobreajustamento dos dados. Nos outros casos, os filhos de cada nó aparecem em linhas mais abaixo e com uma tabulação à direita.

Esta árvore pode ser representada na forma gráfica mais habitual que se encontra na figura 3.10 (os dados relativos ao número de casos e à confiança são omitidos).

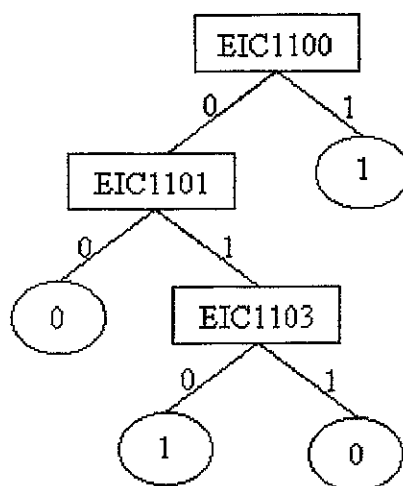


Figura 3.10 - Árvore de decisão para a disciplina EIC1200

Na figura 3.11 temos um exemplo de regras de decisão, obtidas a partir da árvore anterior, com a simplificação posterior.

Neste exemplo, temos também, entre parêntesis, o número de casos e, em vez da confiança, a estimativa da precisão dada pelo rácio de Laplace:

$$\frac{C+1}{T+2}$$

onde  $C$  é o número de classificações correctas da regra e  $T$  o número total de casos em que todas as condições da premissa são satisfeitas.

```
Rules for 1:
  Rule #1 for 1:
    if EIC1100 == 1
      then -> 1 (110, 0.973)

Rules for 0:
  Rule #1 for 0:
    if EIC1100 == 0
      then -> 0 (29, 0.806)

Default : -> 1
```

**Figura 3.11 - Regras de decisão para a disciplina EIC1200**

Por exemplo, se olharmos para a primeira regra relativa à classe 0, vemos que é equivalente a um caminho da árvore que já tínhamos verificado que continha cento e oito classificações correctas em cento e dez possíveis. Assim, a estimativa da precisão é dada por:

$$\frac{108+1}{110+2} = \frac{109}{112} \approx 0,973$$

As regras são agrupadas por classe, existindo ainda uma classe por defeito para os casos em que nenhuma das regras é aplicável.

### 3.4.3. Teste dos modelos

Após a geração dos modelos do C5.0, queremos ver qual o grau de confiança que podemos ter nos resultados alcançados. Assim, podemos fazer testes com tabelas em que os atributos a classificar estejam devidamente preenchidos, e comparar com os dados resultantes da classificação feita pelos modelos gerados pelo C5.0.

A parte dos testes também é realizada no Clementine. Para tal, a tabela com os dados de teste vai ser lida e, sequencialmente, vão ser aplicados todos os modelos gerados. Finalmente, a análise resultante das comparações é exportada para um ficheiro de texto (ver figura 3.12).

Este ficheiro contém vários dados de análise realizados automaticamente. Em primeiro lugar, é apresentada a comparação entre os campos \$C-EIC1200 (valor previsto para a disciplina EIC1200) e o EIC1200, indicando o número total de casos correctos e incorrectos, com a respectiva percentagem, seguido do total.

```
Results for output field EIC1200
Comparing $C-EIC1200 with EIC1200
Correct   :      128   ( 92.09%)
Wrong    :       11   (  7.91%)
Total     :      139
Coincidence Matrix
           $C-EIC1200
           0   1
0         23  9
1         2  105
```

**Figura 3.12 - Análise de testes relativos à disciplina EIC1200**

Em seguida, aparece a matriz de coincidência, onde as linhas representam a disciplina EIC1200 e as colunas o valor previsto em \$C-EIC1200. No exemplo da figura 3.12, a matriz de coincidência mostra que existem vinte e três casos em que o algoritmo C5.0 classificou os tuplos correctamente como 0 e cento e cinco como 1, tendo ainda classificado incorrectamente nove casos como 1 e dois como 0.

# Capítulo 4

## Análise das árvores e regras de decisão

No presente capítulo pretende-se fazer uma análise das árvores e regras de decisão geradas nos moldes anteriormente explicados. Relativamente ao primeiro ano é feita a classificação das disciplinas do segundo semestre tendo em conta as do primeiro. Do segundo ao quarto, faz-se não só a classificação do segundo semestre mas também a do ano completo, usando as disciplinas do ano anterior. No quinto ano faz-se apenas a classificação do ano completo, uma vez que o segundo semestre apenas tem uma disciplina (Projecto ou Estágio). Para a geração dos modelos, os dados foram divididos em duas partes aproximadamente iguais: uma para treino e outra para testes.

Assim as secções 4.1 até 4.5 descrevem os resultados do primeiro ao quinto ano, respectivamente. Na secção 4.6 é feita uma análise tendo em conta o perfil dos alunos baseados nas notas do primeiro ano. Finalmente, na última secção é feita uma discussão dos resultados obtidos.

### **4.1. Primeiro ano**

O primeiro ano é composto por doze disciplinas, seis em cada semestre. Reconhecem-se imediatamente algumas possíveis relações entre disciplinas: EIC1100 e EIC1200 (Análise Matemática I e II), EIC1102 e EIC1202 (Introdução à Programação I e II), e EIC1103 e EIC1203 (Circuitos e Sistemas I e II).

Tendo em conta as áreas científicas, pode-se tentar verificar se existem relações entre disciplinas dentro da mesma área científica, directamente relacionadas. As anteriores são nitidamente da mesma área. Para além dessas, temos ainda a

considerar possíveis relações entre as disciplinas EIC1103 e a EIC1206 (Física Experimental), da área científica de Modelação Física, entre as disciplinas EIC1104 (Electrónica Digital) e a EIC1205 (Sistemas Digitais), da área científica de Sistemas Electrónicos Digitais e entre as disciplinas EIC1101 (Álgebra) e a EIC1200, da área científica de Matemática.

Analisando as árvores e regras de decisão geradas pelo Clementine, na tabela 4.1, podemos constatar que as disciplinas EIC1100, EIC1102 e EIC1103, aparecem na raiz da árvore de decisão e isoladamente nas premissas das regras de decisão das disciplinas EIC1200, EIC1202 e EIC1203, respectivamente.

Por outro lado, não existe qualquer relação entre as disciplinas EIC1205 e EIC1104. Relativamente à disciplina EIC1206, só depende da EIC1103 no terceiro nível, aparecendo nas regras sempre acompanhada por outra disciplina. A disciplina EIC1101 aparece no segundo nível da árvore de decisão da EIC1200 mas, neste caso, a disciplina na raiz é da mesma área.

Para termos uma ideia mais geral da importância das disciplinas do primeiro semestre na previsão da classificação das do segundo semestre temos, na tabela 4.2, uma matriz em que as linhas correspondem às disciplinas do segundo semestre, as colunas às disciplinas do primeiro semestre e os valores correspondem à localização, na árvore de decisão relativa a cada uma das disciplinas do segundo semestre, das disciplinas do primeiro semestre (um para a raiz, dois para os filhos da raiz, etc...).

Constata-se desta forma que a disciplina EIC1102 (Introdução à Programação I) é a mais determinante na classificação das disciplinas do segundo semestre, aparecendo nas árvores de decisão em cinco das seis disciplinas do segundo semestre, três das quais na raiz. Por outro lado, a disciplina EIC1106 (Técnicas de Comunicação e Inglês Funcional) não tem qualquer influência nas disciplinas do segundo semestre.

Esta importância da disciplina não se justifica necessariamente pela importância real que o seu programa poderá ter nas restantes disciplinas. Uma hipótese mais viável é que a motivação principal dos alunos que entram no curso seja a da programação, uma vez que se trata de um curso de Engenharia Informática e Computação.

Tabela 4.1 - Modelos do C5.0 para algumas disciplinas do primeiro ano

Disciplina	Árvore de decisão	Regras de decisão
EIC1200	EIC1100 1 (110.0, 0.982) -> 1 EIC1100 0 EIC1101 0 (20.0, 0.95) -> 0 EIC1101 1 EIC1103 1 (7.0, 0.714) -> 0 EIC1103 0 (2.0, 1.0) -> 1	if EIC1100 == 1 then -> 1 (110, 0.973)  if EIC1100 == 0 then -> 0 (29, 0.806)  Default : -> 1
EIC1202	EIC1102 1 (115.0, 0.974) -> 1 EIC1102 0 EIC1101 0 (15.0, 0.867) -> 0 EIC1101 1 EIC1104 1 (6.0, 0.833) -> 1 EIC1104 0 (3.0, 1.0) -> 0	if EIC1101 == 1 and EIC1104 == 1 then -> 1 (112, 0.982)  if EIC1102 == 1 then -> 1 (115, 0.966)  if EIC1102 == 0 and EIC1104 == 0 then -> 0 (15, 0.824)  if EIC1101 == 0 and EIC1102 == 0 then -> 0 (15, 0.824)  Default : -> 1
EIC1203	EIC1103 1 EIC1101 1 (110.0, 0.936) -> 1 EIC1101 0 EIC1102 1 (3.0, 1.0) -> 0 EIC1102 0 (3.0, 0.667) -> 1 EIC1103 0 EIC1104 0 (15.0, 1.0) -> 0 EIC1104 1 EIC1102 1 (4.0, 1.0) -> 0 EIC1102 0 (4.0, 0.75) -> 1	if EIC1103 == 1 then -> 1 (116, 0.898)  if EIC1104 == 1 then -> 1 (118, 0.875)  if EIC1103 == 0 and EIC1104 == 0 then -> 0 (15, 0.941)  if EIC1102 == 1 and EIC1103 == 0 then -> 0 (7, 0.889)  Default : -> 1
EIC1205	EIC1102 1 (115.0, 0.965) -> 1 EIC1102 0 EIC1100 1 (7.0, 0.857) -> 1 EIC1100 0 (17.0, 0.706) -> 0	if EIC1100 == 1 then -> 1 (110, 0.964)  if EIC1102 == 1 then -> 1 (115, 0.957)  if EIC1100 == 0 and EIC1102 == 0 then -> 0 (17, 0.684)  Default : -> 1
EIC1206	EIC1101 1 (118.0, 0.915) -> 1 EIC1101 0 EIC1102 0 (15.0, 0.8) -> 0 EIC1102 1 EIC1103 1 (3.0, 1.0) -> 1 EIC1103 0 (3.0, 0.667) -> 0	if EIC1102 == 1 and EIC1103 == 1 then -> 1 (108, 0.927)  if EIC1101 == 1 then -> 1 (118, 0.908)  if EIC1101 == 0 and EIC1103 == 0 then -> 0 (15, 0.765)  if EIC1101 == 0 and EIC1102 == 0 then -> 0 (15, 0.765)  Default : -> 1

Tabela 4.2 - Posição das disciplinas nas árvores de decisão

	EIC1100	EIC1101	EIC1102	EIC1103	EIC1104	EIC1106
EIC1200	1	2	-	3	-	-
EIC1202	-	2	1	-	3	-
EIC1203	-	2	3	1	2	-
EIC1204	2	-	1	-	2	-
EIC1205	2	-	1	-	-	-
EIC1206	-	1	2	3	-	-

Assim, os alunos com aproveitamento a esta disciplina, poderão sentir uma motivação maior para o resto do ano, o que não sucederia se tivessem tido aprovação, por exemplo, apenas a disciplinas de Matemática e Física. Este último caso poderia ser visto como uma falta de vocação para o curso em que se inscreveram, o que se poderia traduzir numa desmotivação.

Uma das questões que se colocam é qual a qualidade das árvores e regras de decisão geradas pelo C5.0. Na tabela 4.3 apresentam-se as percentagens de acertos para o primeiro ano.

Tabela 4.3 - Percentagem de acertos para o primeiro ano

Disciplina	Árvore de decisão	Regras de decisão
EIC1200	92,09%	94,24%
EIC1202	91,37%	91,37%
EIC1203	91,37%	92,09%
EIC1204	87,77%	90,65%
EIC1205	91,37%	91,37%
EIC1206	89,93%	89,93%

Note-se que, em três dos casos, as classificações usando as regras são ligeiramente melhores que as feitas pelas árvores. As diferenças não são, no entanto, muito significativas, chegando os resultados a ser exactamente iguais nas restantes três.

Os valores obtidos revelam uma elevada percentagem de acertos. Nas árvores de decisão o pior resultado é da disciplina EIC1204, com 87,77% de classificações correctas, ao passo que nas regras de decisão o pior resultado é da disciplina EIC1206, com 89,93%.

## 4.2. Segundo ano

O segundo ano é composto por doze disciplinas, seis em cada semestre. Vamos começar por estudar a relação existente das disciplinas do segundo semestre relativamente ao primeiro.

Também aqui se pode pensar que as disciplinas EIC2102 e EIC2202 (Algoritmos e Estruturas de Dados I e II) vão estar directamente relacionadas. Tendo em conta as áreas científicas, podemos também verificar se existe alguma relação entre as disciplinas EIC2104 (Microprocessadores e Microcomputadores) e EIC2205 (Laboratório de Arquitectura de Computadores), da área científica de Arquitectura de Computadores. Na tabela 4.4 podemos ver os modelos gerados pelo C5.0.

*Tabela 4.4 - Modelos do C5.0 para algumas disciplinas do segundo ano*

Disciplina	Árvore de decisão	Regras de decisão
EIC2202	EIC2102 1 (76.0, 0.934) -> 1 EIC2102 0 (31.0, 0.935) -> 0	if EIC2102 == 1 then -> 1 (76, 0.923) if EIC2102 == 0 then -> 0 (31, 0.909) Default : -> 1
EIC2205	EIC2103 1 (79.0, 1.0) -> 1 EIC2103 0 EIC2104 1 (10.0, 0.8) -> 1 EIC2104 0 (18.0, 0.833) -> 0	if EIC2103 == 1 then -> 1 (79, 0.988)  if EIC2104 == 1 then -> 1 (85, 0.966)  if EIC2103 == 0 and EIC2104 == 0 then -> 0 (18, 0.8)  Default : -> 1

Como se pode ver na tabela 4.4 tanto a árvore como as regras da disciplina EIC2202 dependem apenas da disciplina EIC2102. Relativamente à disciplina EIC2205, temos uma dependência da EIC2104, mas que se encontra no segundo nível da árvore de decisão, sendo preterida pela EIC2103. Se olharmos para as regras de decisão é possível verificar que existem duas regras, com estas duas disciplinas a aparecerem isoladamente na premissa. O que sucede é que a regra utilizando a



disciplina EIC2104 cobre um maior número de casos mas a sua confiança em termos de classificação é menor.

Para termos uma ideia de qual a importância de cada disciplina do primeiro semestre na classificação das do segundo, construímos uma tabela semelhante à que havíamos feito para o primeiro ano, em que as disciplinas do primeiro semestre aparecem nas colunas e as do segundo nas linhas, tendo como valor a profundidade em que aparecem no nó. A tabela 4.5 mostra-nos estas relações.

Neste caso existem duas disciplinas que, em conjunto, são determinantes para o segundo semestre: EIC2102 (Algoritmos e Estruturas de Dados I) e EIC2103 (Sinais e Telecomunicações). Cada uma delas aparece duas vezes na raiz de árvores de decisão geradas, aparecendo ainda a última no nível de profundidade dois doutra árvore. Menos importante parece ser a disciplina EIC2100 (Física) que só aparece uma vez no nível dois duma das árvores

*Tabela 4.5 - Posição das disciplinas nas árvores de decisão*

	EIC2100	EIC2101	EIC2102	EIC2103	EIC2104	EIC2105
EIC2200	-	-	-	2	-	1
EIC2201	-	1	-	-	2	-
EIC2202	-	-	1	-	-	-
EIC2203	2	-	-	1	-	-
EIC2204	-	-	1	-	-	2
EIC2205	-	-	-	1	2	-

Após a geração dos modelos de classificação, como já aconteceu no primeiro ano, foram realizados testes com os restantes dados. A análise encontra-se na tabela 4.6.

Contrariamente ao que aconteceu no primeiro ano, os resultados dos testes foram iguais para as árvores e as regras. O pior resultado é da disciplina EIC2204, com 84,11% de classificações correctas.

Tabela 4.6 - Percentagem de acertos para o segundo ano

Disciplina	Árvore de decisão	Regras de decisão
EIC2200	87,85%	87,85%
EIC2201	91,59%	91,59%
EIC2202	93,46%	93,46%
EIC2203	91,59%	91,59%
EIC2204	84,11%	84,11%
EIC2205	93,46%	93,46%

No segundo ano já é possível gerar modelos tendo em conta os resultados do ano anterior. Existem várias disciplinas pertencentes à mesma área científica nos dois primeiros anos.

Assim, na área científica de Modelação Física temos um total de seis disciplinas, três em cada ano: EIC1103 (Circuitos e Sistemas I), EIC1203 (Circuitos e Sistemas II) e EIC1206 (Física Experimental), no primeiro ano, EIC2100 (Física), EIC2103 (Sinais e Telecomunicações) e EIC2105 (Laboratório de Circuitos e Sistemas), no segundo ano.

Outra área científica é a de Arquitectura de Computadores, com EIC1204 (Arquitectura de Computadores), no primeiro ano, EIC2104 (Microprocessadores e Microcomputadores) e EIC2205 (Laboratório de Arquitectura de Computadores), no segundo ano. Temos ainda a área científica de Sistemas Electrónicos Digitais, com EIC1104 (Electrónica Digital) e EIC1205 (Sistemas Digitais), no primeiro ano, e EIC2203 (Instrumentação Electrónica), no segundo ano.

Na tabela 4.7 e 4.8, temos os modelos gerados para estas disciplinas do segundo ano.

Vamos começar pelas disciplinas de Modelação Física. A disciplina EIC2100 depende das disciplinas EIC1204 e EIC1203. A primeira é a que parece na raiz da árvore de decisão e não pertence à mesma área científica, mas à de Arquitectura de Computadores. A EIC1203 é da mesma área científica e aparece no segundo nível, aparecendo ainda em duas regras, uma delas isoladamente.

Tabela 4.7 - Modelos do C5.0 para algumas disciplinas do segundo ano

Disciplina	Árvore de decisão	Regras de decisão
EIC2100	EIC1204 0 (21.0, 0.952) -> 0 EIC1204 1 EIC1203 1 (78.0, 0.872) -> 1 EIC1203 0 (8.0, 0.875) -> 0	if EIC1203 == 1 and EIC1204 == 1 then -> 1 (78, 0.862)  if EIC1204 == 0 then -> 0 (21, 0.913)  if EIC1203 == 0 then -> 0 (20, 0.909)  Default : -> 1
EIC2103	EIC1203 0 (20.0, 0.95) -> 0 EIC1203 1 EIC1104 0 (3.0, 1.0) -> 0 EIC1104 1 EIC1102 1 (81.0, 0.951) -> 1 EIC1102 0 (3.0, 0.667) -> 0	if EIC1102 == 1 and EIC1104 == 1 and EIC1203 == 1 then -> 1 (81, 0.94)  if EIC1104 == 0 then -> 0 (11, 0.923)  if EIC1203 == 0 then -> 0 (20, 0.909)  if EIC1102 == 0 then -> 0 (13, 0.867)  Default : -> 1
EIC2104	EIC1204 0 (21.0, 0.714) -> 0 EIC1204 1 EIC1104 1 (84.0, 0.94) -> 1 EIC1104 0 (2.0, 1.0) -> 0	if EIC1104 == 1 and EIC1204 == 1 then -> 1 (84, 0.93)  if EIC1104 == 0 then -> 0 (11, 0.769)  if EIC1204 == 0 then -> 0 (21, 0.696)  Default : -> 1
EIC2105	EIC1104 0 EIC1100 0 (7.0, 1.0) -> 0 EIC1100 1 EIC1103 1 (2.0, 1.0) -> 0 EIC1103 0 (2.0, 1.0) -> 1 EIC1104 1 EIC1103 1 (88.0, 0.932) -> 1 EIC1103 0 EIC1101 1 (3.0, 1.0) -> 0 EIC1101 0 EIC1204 1 (2.0, 1.0) -> 1 EIC1204 0 (3.0, 1.0) -> 0	if EIC1103 == 1 and EIC1104 == 1 then -> 1 (88, 0.922)  if EIC1104 == 1 and EIC1204 == 1 then -> 1 (84, 0.919)  if EIC1100 == 1 and EIC1103 == 0 and EIC1104 == 0 then -> 1 (2, 0.75)  if EIC1103 == 1 and EIC1104 == 0 then -> 0 (4, 0.833)  if EIC1101 == 1 and EIC1103 == 0 and EIC1104 == 1 then -> 0 (3, 0.8)  if EIC1103 == 0 then -> 0 (15, 0.706)  Default : -> 1

Tabela 4.8 - Modelos do C5.0 para algumas disciplinas do segundo ano

Disciplina	Árvore de decisão	Regras de decisão
EIC2203	<pre> EIC1206 0 (17.0, 0.941) -&gt; 0 EIC1206 1   EIC1104 0 (3.0, 1.0) -&gt; 0   EIC1104 1     EIC1203 1 (80.0, 0.975) -&gt; 1     EIC1203 0       EIC1100 1 (3.0, 0.667) -&gt; 1       EIC1100 0 (4.0, 1.0) -&gt; 0 </pre>	<pre> if EIC1104 == 1 and EIC1203 == 1 and EIC1206 == 1 then -&gt; 1 (80, 0.963)  if EIC1100 == 1 and EIC1104 == 1 and EIC1206 == 1 then -&gt; 1 (79, 0.951)  if EIC1100 == 0 and EIC1203 == 0 then -&gt; 0 (14, 0.938)  if EIC1104 == 0 then -&gt; 0 (11, 0.923)  if EIC1206 == 0 then -&gt; 0 (17, 0.895)  Default : -&gt; 1 </pre>
EIC2205	<pre> EIC1204 1 (86.0, 0.965) -&gt; 1 EIC1204 0   EIC1203 0 (12.0, 0.833) -&gt; 0   EIC1203 1     EIC1104 1 (7.0, 0.714) -&gt; 1     EIC1104 0 (2.0, 1.0) -&gt; 0 </pre>	<pre> if EIC1104 == 1 and EIC1203 == 1 then -&gt; 1 (84, 0.965)  if EIC1204 == 1 then -&gt; 1 (86, 0.955)  if EIC1104 == 0 and EIC1203 == 1 then -&gt; 0 (3, 0.8)  if EIC1203 == 0 and EIC1204 == 0 then -&gt; 0 (12, 0.786)  Default : -&gt; 1 </pre>

Relativamente à disciplina EIC2103, existe uma dependência principal da disciplina EIC1203, aparecendo esta na raiz da árvore de decisão e em duas das regras. A disciplina EIC2105 depende da EIC1103, mas a um nível mais baixo na árvore de derivação (uma vez no nível dois e outra no nível três). Aparece ainda em cinco das regras de decisão, uma delas isoladamente. A disciplina que aparece na raiz da árvore de decisão é a EIC1104, pertencente à área científica de Sistemas Electrónicos Digitais.

Relativamente às disciplinas de Arquitectura de Computadores, as disciplinas EIC2104 e EIC2205 dependem principalmente da disciplina EIC1204, da mesma área. Esta aparece na raiz das árvores de derivação de ambas, aparecendo ainda nas regras de decisão.

Finalmente, olhando para as disciplinas da área científica de Sistemas Electrónicos Digitais, nota-se que esta depende principalmente da disciplina EIC1206, da área científica de Modelação Física. Este facto é curioso, uma vez que esta disciplina não tem qualquer influência nas outras disciplinas da mesma área científica. No nível dois da árvore de decisão aparece a disciplina EIC1104, do grupo

de Sistemas Electrónicos Digitais, que aparece também em três regras de decisão, uma das quais isoladamente.

Nota-se, nestes resultados, que apenas as disciplinas da área científica de Arquitectura de Computadores dependem principalmente das disciplinas da mesma. No entanto, verifica-se entre estas três áreas alguma interdependência, uma vez que as disciplinas da área científica de Modelação Física dependem das outras duas áreas, enquanto a disciplina da área de Sistemas Electrónicos Digitais depende duma de Modelação Física.

No entanto, pode não se tratar de uma interdependência forte, mas antes dever-se ao facto de estas três áreas científicas, em conjunto cobrirem metade do número de disciplinas dos dois primeiros anos. Neste caso, não havendo uma dependência forte entre as disciplinas da mesma área, o C5.0 gerará modelos em que as dependências principais são de outras disciplinas, havendo uma maior probabilidade de estas se reflectirem numa destas três.

Em seguida vamos construir a matriz com as disciplinas do primeiro ano nas colunas e as do segundo nas linhas, sendo os valores a profundidade com que as primeiras aparecem nas árvores de decisão das segundas. Na tabela 4.9 podemos ver os resultados desta matriz.

Relativamente às disciplinas do primeiro ano que vão aparecer nas árvores de decisão, constata-se que a disciplina EIC1204 (Arquitectura de Computadores) é a mais influente, aparecendo em oito das doze árvores de decisão, quatro das quais na raiz. Relativamente perto encontram-se as disciplinas EIC1203 (Circuitos e Sistemas II), aparecendo sete vezes, três das quais na raiz, e EIC1104 (Electrónica Digital), que também aparece sete vezes, mas apenas uma vez na raiz. Em conjunto estas três disciplinas, representantes destas três áreas científicas, aparecem em onze das doze árvores. Duas disciplinas não aparecem em nenhuma árvore: EIC1106 (Técnicas de Comunicação e Inglês Funcional) e EIC1205 (Sistemas Digitais).

Vejamos de seguida as percentagens de acertos nos testes realizados, juntamente com as relativas à previsão da classificação para o segundo semestre. Estes dados encontram-se na tabela 4.10.

Tabela 4.9 - Posição das disciplinas nas árvores de decisão

	EIC1100	EIC1101	EIC1102	EIC1103	EIC1104	EIC1106	EIC1200	EIC1202	EIC1203	EIC1204	EIC1205	EIC1206
EIC2100	-	-	-	-	-	-	-	-	2	1	-	-
EIC2101	-	-	-	-	2	-	1	-	-	-	-	-
EIC2102	-	-	2	-	-	-	-	-	1	3	-	-
EIC2103	-	-	3	-	2	-	-	-	1	-	-	-
EIC2104	-	-	-	-	2	-	-	-	-	1	-	-
EIC2105	2	3	-	2	1	-	-	-	-	4	-	-
EIC2200	1	-	-	2	-	-	3	3	-	2	-	-
EIC2201	-	-	2	-	-	-	-	-	-	-	-	1
EIC2202	-	-	-	-	2	-	4	-	1	3	-	5
EIC2203	4	-	-	-	2	-	-	-	3	-	-	1
EIC2204	-	3	-	-	-	-	-	-	2	1	-	-
EIC2205	-	-	-	-	3	-	-	-	2	1	-	-

Tabela 4.10 - Percentagem de acertos para o segundo ano

Disciplina	Árvore de decisão	Regras de decisão	Prev. 2.º semestre
EIC2100	88,79%	88,79%	-
EIC2101	87,85%	87,85%	-
EIC2102	84,11%	84,11%	-
EIC2103	92,52%	92,52%	-
EIC2104	90,65%	90,65%	-
EIC2105	94,39%	94,39%	-
EIC2200	90,65%	87,85%	87,85%
EIC2201	89,72%	89,72%	91,59%
EIC2202	82,24%	82,24%	93,46%
EIC2203	94,39%	94,39%	91,59%
EIC2204	86,92%	86,92%	84,11%
EIC2205	89,72%	89,72%	93,46%

Neste caso, apenas na disciplina EIC2200 as árvores obtêm melhor percentagem de acertos que as regras, com 90,65% contra 87,85%. Nos restantes casos, as percentagens são iguais. A pior percentagem é a relativa à disciplina EIC2202, com 82,24%. Ainda relativamente à disciplina EIC2202, também se constata que a percentagem de acertos é bastante inferior na classificação do ano completo (82,24% contra 93,46%).

Verifica-se que cada um das formas de previsão da classificação é melhor em 50% das disciplinas. Note-se, no entanto, que em relação à disciplina EIC2200, haveria um empate se considerássemos as regras de decisão da classificação usando o ano anterior (exactamente 87,85%), pendendo o melhor resultado para a classificação usando o primeiro semestre.

Parece, no entanto, que não se pode falar de uma forma ser melhor que a outra. O que se depreende é que algumas disciplinas são mais dependentes das do primeiro semestre como, por exemplo, quando se tem duas disciplinas complementares, que poderiam ser dadas como uma única anual. Outras disciplinas serão mais dependentes de disciplinas de anos anteriores quando, por exemplo, englobam um conjunto de conhecimentos de base muito vasto.

### **4.3. Terceiro ano**

O terceiro ano é composto por doze disciplinas, seis em cada semestre. Considerando a previsão das disciplinas do segundo semestre usando as do primeiro, também neste ano se pode prever que exista alguma relação entre as disciplinas EIC3102 e EIC3202 (Teoria da Computação I e II), ambas da área científica de Matemática.

Analisando as disciplinas da mesma área científica, podemos considerar, em Gestão e Métodos Quantitativos, a relação entre EIC3104 (Investigação Operacional) e EIC3204 (Gestão de Empresas) e, em Sistemas Operativos, a relação entre EIC3101 (Sistemas Operativos Clássicos) e EIC3205 (Laboratório de Sistemas Operativos).

Na tabela 4.11, temos as árvores e regras de decisão relativas a estas disciplinas.

**Tabela 4.11 - Modelos do C5.0 para algumas disciplinas do terceiro ano**

Disciplina	Árvore de decisão	Regras de decisão
EIC3202	EIC3103 1 (59.0, 0.915) -> 1 EIC3103 0 (12.0, 1.0) -> 0	if EIC3103 == 1 then -> 1 (59, 0.902)  if EIC3103 == 0 then -> 0 (12, 0.929)  Default : -> 1
EIC3204	EIC3100 1 (57.0, 0.947) -> 1 EIC3100 0 EIC3103 1 (4.0, 1.0) -> 0 EIC3103 0 (10.0, 0.6) -> 1	if EIC3100 == 1 then -> 1 (57, 0.932)  if EIC3100 == 0 and EIC3103 == 0 then -> 1 (10, 0.583)  if EIC3100 == 0 and EIC3103 == 1 then -> 0 (4, 0.833)  Default : -> 1
EIC3205	EIC3105 1 (58.0, 0.966) -> 1 EIC3105 0 (13.0, 0.846) -> 0	if EIC3105 == 1 then -> 1 (58, 0.95)  if EIC3105 == 0 then -> 0 (13, 0.8)  Default : -> 1

Contrariamente ao esperado, a disciplina EIC3202 não depende da EIC3102, mas sim da EIC3103 (Programação Paralela), que pertence à área científica de Algoritmos e Estruturas de Dados.

Também em relação às outras disciplinas não existe qualquer relação. A disciplina EIC3204 depende das EIC3100 (Novas Metodologias de Programação) e EIC3103 (Programação Paralela), ao passo que a EIC3205 depende da EIC3105 (Laboratório de Computação Gráfica e Interfaces).

Estes dados mostram relações que não seriam esperadas. Vejamos, na tabela 4.12, a relação entre disciplinas do primeiro semestre e as árvores de decisão.

**Tabela 4.12 - Posição das disciplinas nas árvores de decisão**

	EIC3100	EIC3101	EIC3102	EIC3103	EIC3104	EIC3105
EIC3200	-	-	-	-	-	1
EIC3201	1	-	-	-	-	-
EIC3202	-	-	-	1	-	-
EIC3203	1	-	-	2	-	-
EIC3204	1	-	-	2	-	-
EIC3205	-	-	-	-	-	1



Verifica-se que existem três disciplinas que em conjunto cobrem todas as árvores, e três que não aparecem em nenhuma. Assim, a disciplina EIC3100 (Novas Metodologias de Programação) será a mais influente aparecendo em três árvores de decisão, sempre na raiz. As disciplinas EIC3105 (Laboratório de Computação Gráfica e Interfaces), que aparece duas vezes, ambas na raiz, e EIC3103 (Programação Paralela), que aparece três vezes, uma delas na raiz, são as restantes.

As disciplinas que não aparecem em nenhuma árvore são as EIC3101 (Sistemas Operativos Clássicos), EIC3102 (Teoria da Computação I) e EIC3104 (Investigação Operacional). Estas são precisamente as disciplinas que nós queríamos comparar.

A análise dos modelos mostra uma diminuição na percentagem de acertos, como se pode ver na tabela 4.13. Os valores percentuais, iguais para as árvores e regras de decisão, situam-se entre os 83,10% e os 88,73%. As razões desta diminuição poderão estar no menor número de casos usados para teste e para treino (setenta e um contra cento e sete no segundo ano).

*Tabela 4.13 - Percentagem de acertos para o terceiro ano*

Disciplina	Árvore de decisão	Regras de decisão
EIC3200	83,10%	83,10%
EIC3201	88,73%	88,73%
EIC3202	88,73%	88,73%
EIC3203	88,73%	88,73%
EIC3204	84,51%	84,51%
EIC3205	88,73%	88,73%

Por muito que estes dados possam parecer estranhos, é preciso não esquecer que o C5.0 gera árvores e regras de decisão que permitem rapidamente classificar um determinado exemplo, tendo em conta os outros atributos, mas não revela todas as relações existentes. Por outro lado, as regras de associação mostram todas as associações que tenham determinados suporte e confiança mínimos. Assim, quando falarmos das regras de associação, poderemos discutir com mais alguns dados esta questão.

Passando à classificação do ano completo, vamos ter várias disciplinas do segundo que poderão estar relacionadas com disciplinas do terceiro ano da mesma área científica.

Começando pela área científica de Algoritmos e Estruturas de Dados, temos EIC2102 e EIC2202 (Algoritmos e Estruturas de Dados I e II), do segundo ano, e a EIC3103 (Programação Paralela), do terceiro. Na área científica de Bases de Dados e Sistemas de Tratamento de Informação, consideramos as disciplinas EIC2200 (Sistemas de Informação e Bases de Dados), do segundo ano, e EIC3200 (Tecnologias de Sistemas de Gestão de Bases de Dados), do terceiro. Na área científica de Interação Homem-Máquina, temos as disciplinas EIC2201 (Computação Gráfica e Interfaces), do segundo ano, e EIC3105 (Laboratório de Computação Gráfica e Interfaces), do terceiro. Finalmente, na área científica de Gestão e Métodos e Quantitativos, temos as disciplinas EIC2204 (Estatística Aplicada), do segundo ano, e EIC3104 (Investigação Operacional) e EIC3204 (Gestão de Empresas), do terceiro

Na tabela 4.14 podemos ver as árvores e regras de decisão relativa a estas disciplinas.

Começando pela área científica de Algoritmos e Estruturas de Dados, constata-se que a disciplina EIC3103 não depende de disciplinas da mesma área, mas apenas da EIC2100 (Física) da área de Modelação Física. Também na área científica de Bases de Dados e Sistemas de Tratamento de Informação não existe qualquer relação entre as disciplinas da mesma área. Neste último caso, a disciplina EIC3200 depende, pela ordem seguinte, das disciplinas EIC2202, EIC2100 e EIC2101 (Análise Numérica). Outra área científica onde não existe qualquer relação é a de Interação Homem-Máquina, uma vez que a disciplina EIC3105 depende apenas da disciplina EIC2202.

Na área científica de Gestão e Métodos e Quantitativos já existe uma maior consistência, relativamente à relação entre as disciplinas da mesma área. No caso da disciplina EIC3104, a disciplina EIC2204 aparece apenas no nível três da árvore de decisão, mas também aparece em duas das regras, uma delas isoladamente. Relativamente à disciplina EIC3204 já existe uma relação mais forte, uma vez que esta disciplina depende exclusivamente da disciplina EIC2204.

Tabela 4.14 - Modelos do C5.0 para algumas disciplinas do terceiro ano

Disciplina	Árvore de decisão	Regras de decisão
EIC3103	EIC2100 1 (59.0, 0.966) -> 1 EIC2100 0 (12.0, 0.833) -> 0	if EIC2100 == 1 then -> 1 (59, 0.951)  if EIC2100 == 0 then -> 0 (12, 0.786)  Default : -> 1
EIC3104	EIC2102 0 (9.0, 1.0) -> 0 EIC2102 1 EIC2103 0 (3.0, 1.0) -> 0 EIC2103 1 EIC2204 1 (54.0, 0.963) -> 1 EIC2204 0 (5.0, 0.6) -> 0	if EIC2102 == 1 and EIC2103 == 1 and EIC2204 == 1 then -> 1 (54, 0.946)  if EIC2102 == 0 then -> 0 (9, 0.909)  if EIC2103 == 0 then -> 0 (8, 0.9)  if EIC2204 == 0 then -> 0 (11, 0.769)  Default : -> 1
EIC3105	EIC2202 1 (61.0, 0.934) -> 1 EIC2202 0 (10.0, 0.9) -> 0	if EIC2202 == 1 then -> 1 (61, 0.921)  if EIC2202 == 0 then -> 0 (10, 0.833)  Default : -> 1
EIC3200	EIC2202 0 (10.0, 0.9) -> 0 EIC2202 1 EIC2100 1 (56.0, 0.964) -> 1 EIC2100 0 EIC2101 1 (2.0, 1.0) -> 0 EIC2101 0 (3.0, 0.667) -> 1	if EIC2202 == 1 then -> 1 (61, 0.905)  if EIC2100 == 0 and EIC2101 == 1 then -> 0 (5, 0.857)  if EIC2202 == 0 then -> 0 (10, 0.833)  if EIC2100 == 0 then -> 0 (12, 0.714)  Default : -> 1
EIC3204	EIC2204 1 (60.0, 0.95) -> 1 EIC2204 0 (11.0, 0.727) -> 0	if EIC2204 == 1 then -> 1 (60, 0.935)  if EIC2204 == 0 then -> 0 (11, 0.692)  Default : -> 1

Vejamos agora a matriz com a relação entre as disciplinas do segundo ano e as árvores de decisão das do terceiro, na tabela 4.15.

As disciplinas mais influentes são a EIC2100 (Física), que aparece em cinco casos, três dos quais na raiz, as disciplinas EIC2102 (Algoritmos e Estruturas de Dados I) e EIC2202 (Algoritmos e Estruturas de Dados II), que aparecem quatro vezes, três das quais na raiz, e a EIC2204 (Estatística Aplicada), que aparece três vezes, duas na raiz. As disciplinas que não aparecem em nenhuma árvore de decisão são as EIC2104 (Microprocessadores e Microcomputadores), EIC2105 (Laboratório de Circuitos e Sistemas) e EIC2205 (Laboratório de Arquitectura de Computadores).

Tabela 4.15 - Posição das disciplinas nas árvores de decisão

	EIC2100	EIC2101	EIC2102	EIC2103	EIC2104	EIC2105	EIC2200	EIC2201	EIC2202	EIC2203	EIC2204	EIC2205
EIC3100	-	-	1	-	-	-	-	2	-	-	-	-
EIC3101	2	-	1	-	-	-	-	-	-	-	-	-
EIC3102	1	-	-	-	-	-	2	-	-	-	-	-
EIC3103	1	-	-	-	-	-	-	-	-	-	-	-
EIC3104	-	-	1	2	-	-	-	-	-	-	3	-
EIC3105	-	-	-	-	-	-	-	-	1	-	-	-
EIC3200	2	3	-	-	-	-	-	-	1	-	-	-
EIC3201	-	2	-	-	-	-	-	-	1	-	-	-
EIC3202	1	-	-	-	-	-	-	2	-	-	-	-
EIC3203	-	-	2	-	-	-	-	-	-	-	1	-
EIC3204	-	-	-	-	-	-	-	-	-	-	1	-
EIC3205	-	-	-	-	-	-	-	2	3	1	-	-

Tal como já aconteceu no segundo ano, vamos agora ver, na tabela 4.16, as percentagens de acertos das árvores e regras de decisão usando o ano anterior, e o resultado da previsão do segundo semestre usando o primeiro.

No terceiro ano, as percentagens de acerto são melhores para a previsão do ano completo em cinco das seis disciplinas, sendo igual na outra. Pode-se pensar que relativamente ao terceiro ano será melhor gerar os modelos para o ano completo usando as disciplinas do segundo ano. No entanto, as percentagens não diferem muito, uma vez que a maior diferença se cifra em 4,23%. De referir ainda que as percentagens de acerto são iguais para as árvores e para as regras de decisão.

Tabela 4.16 - Percentagem de acertos para o terceiro ano

Disciplina	Árvore de decisão	Regras de decisão	Prev. 2.º semestre
EIC3100	90,14%	90,14%	-
EIC3101	88,73%	88,73%	-
EIC3102	87,32%	87,32%	-
EIC3103	87,32%	87,32%	-
EIC3104	91,55%	91,55%	-
EIC3105	88,73%	88,73%	-
EIC3200	85,92%	85,92%	83,10%
EIC3201	90,14%	90,14%	88,73%
EIC3202	88,73%	88,73%	88,73%
EIC3203	92,96%	92,96%	88,73%
EIC3204	88,73%	88,73%	84,51%
EIC3205	90,14%	90,14%	88,73%

#### 4.4. Quarto ano

O quarto ano é composto por treze disciplinas, sendo o primeiro semestre composto por seis, e o segundo por quatro obrigatórias e três opcionais, das quais se podem escolher duas.

Vamos começar por considerar a geração de modelos para o segundo semestre usando as disciplinas do primeiro. Considerando as áreas científicas, vamos tentar verificar a relação existente entre as disciplinas de diferentes semestres do quarto ano.

Na área científica de Bases de Dados e Sistemas de Tratamento de Informação, temos a disciplina EIC4105 (Laboratório de Sistemas de Gestão de Bases de Dados), do primeiro semestre, e a EIC4205 (Sistemas de Comunicação Integrada), do segundo semestre. Temos ainda, na área científica de Gestão e Métodos Quantitativos, as disciplinas EIC4104 (Contabilidade) e EIC4202 (Análise de Investimentos), na área científica de Inteligência Artificial e Robótica, as disciplinas EIC4103 (Sistemas Periciais e Robótica) e EIC4203 (Laboratório de Informação Avançada) e, na área científica de Informática Industrial, as disciplinas

EIC4102 (Informática Industrial) e EIC4204 (Concepção e Fabrico Assistido por Computador).

Na tabela 4.17 temos todas as árvore e regras de decisão.

**Tabela 4.17 - Modelos do C5.0 para algumas disciplinas do quarto ano**

Disciplina	Árvore de decisão	Regras de decisão
EIC4200	-> 1	Default : -> 1
EIC4201	-> 1	Default : -> 1
EIC4202	EIC4105 1 (34.0, 0.941) -> 1 EIC4105 0 (3.0, 0.667) -> 0	If EIC4105 == 1 then -> 1 (34, 0.917)  if EIC4105 == 0 then -> 0 (3, 0.6)  Default : -> 1
EIC4203	EIC4100 1 (33.0, 0.879) -> 1 EIC4100 0 (4.0, 0.75) -> 0	If EIC4100 == 1 then -> 1 (33, 0.857)  if EIC4100 == 0 then -> 0 (4, 0.667)  Default : -> 1
EIC4204	EIC4103 1 (30.0, 0.667) -> 1 EIC4103 0 (7.0, 1.0) -> 0	If EIC4103 == 1 then -> 1 (30, 0.656)  if EIC4103 == 0 then -> 0 (7, 0.889)  Default : -> 1
EIC4205	EIC4101 1 (32.0, 0.938) -> 1 EIC4101 0 (5.0, 1.0) -> 0	If EIC4101 == 1 then -> 1 (32, 0.912)  if EIC4101 == 0 then -> 0 (5, 0.857)  Default : -> 1
EIC4206	-> 0	Default : -> 0

Pela primeira vez aparecem disciplinas que não dependem de outras. São os casos das disciplinas EIC4200 (Computação Numérica e Simbólica) e EIC4201 (Aspectos Sociais da Informática), cuja classificação prevista é 1, e a disciplina EIC4206 (Sistemas Operativos Distribuídos), cuja classificação prevista é 0. Estes resultados vão de encontro ao que se poderia esperar, uma vez que estas disciplinas são as únicas que não têm qualquer outra da mesma área científica no primeiro semestre.

Relativamente às disciplinas da mesma área científica, não foi encontrada qualquer relação. Também não existe qualquer disciplina que seja determinante na previsão da classificação das disciplinas do segundo semestre, conforme se pode ver na tabela 4.18.

Tabela 4.18 - Posição das disciplinas nas árvores de decisão

	EIC4100	EIC4101	EIC4102	EIC4103	EIC4104	EIC4105
EIC4200	-	-	-	-	-	-
EIC4201	-	-	-	-	-	-
EIC4202	-	-	-	-	-	1
EIC4203	1	-	-	-	-	-
EIC4204	-	-	-	1	-	-
EIC4205	-	1	-	-	-	-
EIC4206	-	-	-	-	-	-

Neste caso não existe nenhuma disciplina a merecer especial destaque. Como já vimos, existem três disciplinas que não dependem de qualquer disciplina. As restantes quatro disciplinas dependem apenas de quatro disciplinas, todas diferentes: EIC4100 (Engenharia de Software), EIC4101 (Redes de Comunicação de Dados), EIC4103 (Sistemas Periciais e Robótica) e EIC4105 (Laboratório de Sistemas de Gestão de Bases de Dados). Não aparecem em nenhuma árvore as disciplinas EIC4102 (Informática Industrial) e EIC4104 (Contabilidade).

Vejamos ainda a percentagem de acertos dos modelos gerados, na tabela 4.19.

Tabela 4.19 - Percentagem de acertos para o quarto ano

Disciplina	Árvore de decisão	Regras de decisão
EIC4200	72,97%	72,97%
EIC4201	89,19%	89,19%
EIC4202	91,89%	91,89%
EIC4203	91,89%	91,89%
EIC4204	62,16%	62,16%
EIC4205	89,19%	89,19%
EIC4206	59,46%	59,46%

Neste caso não existem diferenças entre as percentagens de acertos das árvores e das regras de decisão. O valor mais baixo é o relativo à disciplina EIC4206,

com 59,46% de acertos, enquanto as mais altas são as EIC4202 e EIC4203, com 91,89%.

Vamos passar agora à previsão da classificação para o ano completo usando as disciplinas do ano anterior. Vamos considerar as disciplinas dos terceiro e quarto anos da mesma área científica.

Assim, temos na área científica de Bases de Dados e Sistemas de Tratamento de Informação, as disciplinas EIC3200 (Tecnologia de Sistemas de Gestão de Bases de Dados), do terceiro ano, e as EIC4105 (Laboratório de Sistemas de Gestão de Bases de Dados), do quarto. Na área de Gestão e Métodos Quantitativos, temos as disciplinas EIC3104 (Investigação Operacional) e EIC3204 (Gestão de Empresas), do terceiro ano, e as EIC4104 (Contabilidade) e EIC4202 (Análise de Investimentos), do quarto. Na área científica de Inteligência Artificial e Robótica, temos a disciplina EIC3203 (Introdução à Inteligência Artificial), do terceiro ano, e as EIC4103 (Sistemas Periciais e Robótica) e EIC4203 (Laboratório de Informática Avançada), do quarto. Na área científica de Metodologia e Engenharia de Software, temos a disciplina EIC3100 (Novas Metodologias de Programação), do terceiro ano, e a EIC4100 (Engenharia de Software), do quarto. Finalmente, na área científica de Sistemas Operativos, temos as disciplinas EIC3101 (Sistemas Operativos Clássicos) e EIC3205 (Laboratório de Sistemas Operativos), do terceiro ano, e a EIC4206 (Sistemas Operativos Distribuídos), do quarto.

Na tabela 4.20 temos as árvores e regras de decisão geradas. Nota-se que as disciplinas EIC4201 e EIC4206 continuam a não depender de nenhuma outra, sendo a classificação exactamente igual. Todas as restantes dependem apenas de uma disciplina.

Relativamente às áreas científicas, apenas na área de Gestão e Métodos Quantitativos se verifica a interdependência. Ambas as disciplinas do quarto ano dependem da EIC3104 (Investigação Operacional). Em todos os outros casos as disciplinas do quarto ano não dependem de disciplinas da mesma área do terceiro ano.

Vejamos, de seguida, na tabela 4.21, a matriz com as disciplinas do quarto ano e o local nas árvores de decisão em que aparecem.



Tabela 4.20 - Modelos do C5.0 para as disciplinas do quarto ano

Disciplina	Árvore de decisão	Regras de decisão
EIC4100	EIC3102 1 (34.0, 0.971) -> 1 EIC3102 0 (3.0, 1.0) -> 0	if EIC3102 == 1 then -> 1 (34, 0.944)  if EIC3102 == 0 then -> 0 (3, 0.8)  Default : -> 1
EIC4101	EIC3102 1 (34.0, 0.941) -> 1 EIC3102 0 (3.0, 1.0) -> 0	if EIC3102 == 1 then -> 1 (34, 0.917)  if EIC3102 == 0 then -> 0 (3, 0.8)  Default : -> 1
EIC4102	EIC3102 1 (34.0, 0.971) -> 1 EIC3102 0 (3.0, 1.0) -> 0	if EIC3102 == 1 then -> 1 (34, 0.944)  if EIC3102 == 0 then -> 0 (3, 0.8)  Default : -> 1
EIC4103	EIC3102 1 (34.0, 0.882) -> 1 EIC3102 0 (3.0, 1.0) -> 0	if EIC3102 == 1 then -> 1 (34, 0.861)  if EIC3102 == 0 then -> 0 (3, 0.8)  Default : -> 1
EIC4104	EIC3104 1 (35.0, 0.857) -> 1 EIC3104 0 (2.0, 1.0) -> 0	if EIC3104 == 1 then -> 1 (35, 0.838)  if EIC3104 == 0 then -> 0 (2, 0.75)  Default : -> 1
EIC4105	EIC3104 1 (35.0, 0.971) -> 1 EIC3104 0 (2.0, 1.0) -> 0	if EIC3104 == 1 then -> 1 (35, 0.946)  if EIC3104 == 0 then -> 0 (2, 0.75)  Default : -> 1
EIC4200	EIC3102 1 (34.0, 0.824) -> 1 EIC3102 0 (3.0, 0.667) -> 0	if EIC3102 == 1 then -> 1 (34, 0.806)  if EIC3102 == 0 then -> 0 (3, 0.6)  Default : -> 1
EIC4201	-> 1	Default : -> 1
EIC4202	EIC3104 1 (35.0, 0.943) -> 1 EIC3104 0 (2.0, 1.0) -> 0	if EIC3104 == 1 then -> 1 (35, 0.919)  if EIC3104 == 0 then -> 0 (2, 0.75)  Default : -> 1
EIC4203	EIC3102 1 (34.0, 0.882) -> 1 EIC3102 0 (3.0, 1.0) -> 0	if EIC3102 == 1 then -> 1 (34, 0.861)  if EIC3102 == 0 then -> 0 (3, 0.8)  Default : -> 1
EIC4204	EIC3102 1 (34.0, 0.588) -> 1 EIC3102 0 (3.0, 1.0) -> 0	if EIC3102 == 1 then -> 1 (34, 0.583)  if EIC3102 == 0 then -> 0 (3, 0.8)  Default : -> 1
EIC4205	EIC3102 1 (34.0, 0.882) -> 1 EIC3102 0 (3.0, 1.0) -> 0	if EIC3102 == 1 then -> 1 (34, 0.861)  if EIC3102 == 0 then -> 0 (3, 0.8)  Default : -> 1
EIC4206	-> 0	Default : -> 0

Tabela 4.21 - Posição das disciplinas nas árvores de decisão

	EIC3100	EIC3101	EIC3102	EIC3103	EIC3104	EIC3105	EIC3200	EIC3201	EIC3202	EIC3203	EIC3204	EIC3205
EIC4100	-	-	1	-	-	-	-	-	-	-	-	-
EIC4101	-	-	1	-	-	-	-	-	-	-	-	-
EIC4102	-	-	1	-	-	-	-	-	-	-	-	-
EIC4103	-	-	1	-	-	-	-	-	-	-	-	-
EIC4104	-	-	-	-	1	-	-	-	-	-	-	-
EIC4105	-	-	-	-	1	-	-	-	-	-	-	-
EIC4200	-	-	1	-	-	-	-	-	-	-	-	-
EIC4201	-	-	-	-	-	-	-	-	-	-	-	-
EIC4202	-	-	-	-	1	-	-	-	-	-	-	-
EIC4203	-	-	1	-	-	-	-	-	-	-	-	-
EIC4204	-	-	1	-	-	-	-	-	-	-	-	-
EIC4205	-	-	1	-	-	-	-	-	-	-	-	-
EIC4206	-	-	-	-	-	-	-	-	-	-	-	-

Contrariamente ao que acontecia com as a previsão da classificação usando as disciplinas do primeiro semestre, neste caso existem duas disciplinas que são muito importantes, sendo as únicas que aparecem nas árvores de decisão, sempre nas raízes: a EIC3102 (Teoria da Computação I), que aparece oito vezes, e a EIC3104 (Investigação Operacional), que aparece três vezes. Esta última, como já vimos, aparece duas vezes em disciplinas da mesma área científica.

Refira-se que este caso é bastante curioso, uma vez que estas disciplinas, sendo do primeiro semestre do terceiro ano, não aparecem em nenhuma árvore de decisão do segundo semestre do terceiro ano. Depreende-se que estas duas disciplinas são mais importantes para o ano seguinte, do que para o próprio ano.

Uma outra curiosidade a ter em conta é o facto de a disciplina EIC3102 (Teoria da Computação I) ter um peso tão acentuado nas disciplinas do quarto ano, e não ter qualquer influência na EIC3202 (Teoria da Computação II), o que seria uma interdependência bastante previsível.

Vejamos ainda, na tabela 4.22, as percentagens de acertos das árvores e regras de decisão geradas pelo C5.0 para a previsão da classificação para o ano completo, bem como a relativa à previsão da classificação para apenas o segundo semestre.

*Tabela 4.22 - Percentagem de acertos para o quarto ano*

Disciplina	Árvore de decisão	Regras de decisão	Prev. 2.º semestre
EIC4100	91,89%	91,89%	-
EIC4101	83,78%	83,78%	-
EIC4102	91,89%	91,89%	-
EIC4103	91,89%	91,89%	-
EIC4104	81,08%	81,08%	-
EIC4105	89,19%	89,19%	-
EIC4200	78,38%	78,38%	72,97%
EIC4201	89,19%	89,19%	89,19%
EIC4202	91,89%	91,89%	91,89%
EIC4203	89,19%	89,19%	91,89%
EIC4204	54,05%	54,05%	62,16%
EIC4205	89,19%	89,19%	89,19%
EIC4206	59,46%	59,46%	59,46%

Verifica-se que a disciplina com menor percentagem de acertos é a EIC4204 com 54,05% de acertos, seguido da EIC4206, com 59,46%. Estas duas estão bastante abaixo das restantes, uma vez que a terceira mais baixa tem 83,78% de acertos.

Relativamente à comparação das duas previsões de classificação, do ano completo e do segundo semestre, não se manteve a tendência que aconteceu no terceiro ano. Em quatro casos houve um empate, em dois casos a previsão da classificação do segundo semestre é melhor e apenas uma vez a previsão da classificação do ano completo é melhor.

#### 4.5. Quinto ano

O quinto ano é composto por nove disciplinas. O primeiro semestre é composto por sete disciplinas, três das quais opcionais, das quais são escolhidas duas. O segundo semestre é composto por duas disciplinas opcionais, das quais se escolhe uma.

Como já foi referido, para o quinto ano só se realizou a previsão da classificação do ano completo, tendo em conta o quarto ano. A razão para esta escolha prendeu-se com o facto de o segundo semestre ter apenas uma disciplina a escolher entre duas opcionais: a EIC5200 (Projecto) e a EIC5201 (Estágio). Estas disciplinas são geralmente frequentadas por alunos que se encontram no fim de curso, têm habitualmente uma percentagem de aprovações próxima dos 100% e não é habitual dependerem da aprovação a outras disciplinas, embora pressuponham o uso dos conhecimentos de todas as disciplinas do curso.

Refira-se como curiosidade o facto de nenhum aluno se ter inscrito na disciplina de Projecto, optando todos pelo Estágio. De referir ainda que todos os alunos foram aprovados, tal como se esperava, sendo quinze a pior nota atribuída.

Relativamente às disciplinas da mesma área científica, todas as seis disciplinas que compõem a área de Informática Industrial, são dos dois últimos anos. Assim, consideramos as disciplinas EIC4102 (Informática Industrial) e EIC4204 (Concepção e Fabrico Assistido por Computador), do quarto ano, e as EIC5103 (Logística), EIC5104 (Planeamento e Gestão da Produção) e a EIC5201 (Estágio). Não é considerada a EIC5200 (Projecto) dado ninguém ter escolhido esta opção.

Relativamente às restantes áreas científicas temos, na área científica de Aspectos Sociais e Profissionais, a disciplina EIC4201 (Aspectos Sociais da Informática), do quarto ano, e a EIC5102 (Seminários), do quinto, na área científica de Gestão e Métodos Quantitativos, as disciplinas EIC4104 (Contabilidade) e EIC4202 (Análise de Investimentos), do quarto ano, e a EIC5101 (Sistemas de Apoio à Decisão), do quinto e, finalmente, na área científica de Metodologia e Engenharia de Software, a disciplina EIC4100 (Engenharia de Software), do quarto ano, e a EIC5100 (Gestão de Projectos de Software), do quinto.

Vejamos as árvores e regras de decisão, que se encontram descritas na tabela 4.23.

Tabela 4.23 - Modelos do C5.0 para as disciplinas do quinto ano

Disciplina	Árvore de decisão	Regras de decisão
EIC5100	-> 1	Default : -> 1
EIC5101	-> 1	Default : -> 1
EIC5102	-> 1	Default : -> 1
EIC5103	EIC4104 1 (15.0, 1.0) -> 1 EIC4104 0 (2.0, 1.0) -> 0	if EIC4104 == 1 then -> 1 (15, 0.941)  if EIC4104 == 0 then -> 0 (2, 0.75)  Default : -> 1
EIC5104	EIC4104 0 (2.0, 1.0) -> 0 EIC4104 1  EIC4200 1 (13.0, 0.846) -> 1 EIC4200 0 (2.0, 1.0) -> 0	if EIC4104 == 1 and EIC4200 == 1 then -> 1 (13, 0.8)  if EIC4200 == 0 then -> 0 (2, 0.75)  if EIC4104 == 0 then -> 0 (2, 0.75)  Default : -> 1
EIC5105	EIC4104 1 (15.0, 0.8) -> 0 EIC4104 0 (2.0, 1.0) -> 1	if EIC4104 == 0 then -> 1 (2, 0.75)  if EIC4104 == 1 then -> 0 (15, 0.765)  Default : -> 0
EIC5106	-> 1	Default : -> 1
EIC5201	-> 1	Default : -> 1

Como se pode ver, existem cinco disciplinas que não dependem de nenhuma outra. São elas a EIC5100 (Gestão de Projectos de Software), a EIC5101 (Sistemas de Apoio a Decisão), a EIC5102 (Seminários), a EIC5106 (Sistemas Multimédia) e a já referida EIC5201 (Estágio), todas elas com a classificação 1, independentemente das outras disciplinas.

Em relação às outras disciplinas, as árvores são mais pequenas do que o habitual para a previsão da classificação do ano completo. Tal facto pode ficar a dever-se ao número de registos usados no treino e no teste serem inferiores, e também pelo facto de os alunos inscritos neste último ano terem um nível de aproveitamento geralmente bom.

Não se verifica qualquer interdependência entre disciplinas da mesma área científica. Parece evidente que as disciplinas do quinto não dependem muito dos anos anteriores. A hipótese que parece ser mais aceitável é que, sendo o último ano, existe uma motivação maior para o término do curso, o que leva os alunos a terem um esforço suplementar para colmatar algumas lacunas que possam ter.

Vejam agora na tabela 4.24 as posições que as disciplinas do quarto ano ocupam nas árvores de decisão das disciplinas do quinto.

Tabela 4.24 - Posição das disciplinas nas árvores de decisão

	EIC4100	EIC4101	EIC4102	EIC4103	EIC4104	EIC4105	EIC4200	EIC4201	EIC4202	EIC4203	EIC4204	EIC4205	EIC4206
EIC5100	-	-	-	-	-	-	-	-	-	-	-	-	-
EIC5101	-	-	-	-	-	-	-	-	-	-	-	-	-
EIC5102	-	-	-	-	-	-	-	-	-	-	-	-	-
EIC5103	-	-	-	-	1	-	-	-	-	-	-	-	-
EIC5104	-	-	-	-	1	-	2	-	-	-	-	-	-
EIC5105	-	-	-	-	1	-	-	-	-	-	-	-	-
EIC5106	-	-	-	-	-	-	-	-	-	-	-	-	-
EIC5200	-	-	-	-	-	-	-	-	-	-	-	-	-
EIC5201	-	-	-	-	-	-	-	-	-	-	-	-	-

Apenas duas disciplinas aparecem nas árvores de decisão. A mais importante é a EIC4104 (Contabilidade) que aparece três vezes, sempre na raiz. A outra disciplina só aparece uma vez, a EIC4200 (Computação Numérica e Simbólica), e apenas no nível dois de profundidade.

Estes resultados parecem no entanto ser mero acaso. Em primeiro lugar, o número de casos (dezassete) é bastante reduzido. Para além disso, como já foi referido, a motivação para terminar o curso parece ser mais importante neste último ano. Não parece muito crível que uma disciplina de Contabilidade possa ter uma influência muito grande no último ano da Licenciatura de Engenharia Informática e Computação.

Vejam ainda a percentagem de acertos dos modelos gerados, na tabela 4.25.

A disciplina com pior percentagem de acertos é a EIC5104 (Planeamento e Gestão da Produção), com 56,25%. As restantes disciplinas estão todas acima dos 75% de acertos. Como já se esperava, a disciplina EIC5201 (Estágio), obteve 100% de acertos.

Tabela 4.25 - Percentagens de acertos para o quinto ano

Disciplina	Árvore de decisão	Regras de decisão
EIC5100	75,00%	75,00%
EIC5101	75,00%	75,00%
EIC5102	75,00%	75,00%
EIC5103	81,25%	81,25%
EIC5104	56,25%	56,25%
EIC5105	87,50%	87,50%
EIC5106	81,25%	81,25%
EIC5201	100,00%	100,00%

Não parece que no quinto ano se possam tirar quaisquer resultados conclusivos. O número total de tuplos (dezassete para treino e dezasseis para teste) podem não ser suficientes.

#### 4.6. Perfil dos alunos

Uma última análise realizada consistiu em verificar se é possível, no primeiro ano de um aluno, prever qual a situação em que este estará ao fim de cinco anos. Por outras palavras, pretende-se verificar se um determinado aluno será um aluno que acabará o curso em cinco anos ou mais.

Na tabela 4.26, podemos ver um resumo dos resultados ao fim de cinco anos.

Tabela 4.26 – Resultados ao fim de cinco anos

PERFIL	N	F	C
F	16	3	2
S	3	5	11
B	0	0	6
MB	0	0	6

Como se pode ver, todos os alunos com o perfil B (Bom) e MB (Muito Bom), acabaram a licenciatura ao fim de cinco anos. De realçar também o facto de apenas

dois dos vinte e um alunos com o perfil F (Fraco) terem conseguido acabar o curso em cinco anos.

Tendo em conta que estes alunos se referem apenas ao primeiro ano de funcionamento da licenciatura, analisou-se também a situação dos alunos ao fim de três anos, usando os alunos que se inscreveram pela primeira vez no curso entre os anos lectivos de 1994/95 e 1996/97. Na tabela 4.27 podemos ver os resultados obtidos.

**Tabela 4.27 - Resultados ao fim de três anos**

PERFIL	N3	F31	F32	C3
F	48	4	6	1
S	4	8	21	23
B	0	0	1	26
MB	0	0	0	16

Os dados que se obtiveram anteriormente, são confirmados aqui. Contando apenas a situação ao fim de três anos e com mais alunos, continua-se a verificar que os alunos com perfil B e MB têm o terceiro ano concluído, com excepção de um, com perfil B.

Apenas um aluno com perfil F consegue ter o terceiro ano concluído ao fim de três anos, conseguindo também acabar o curso em cinco anos.

Pegando apenas nos alunos do ano lectivo de 1994/95, resolvemos comparar com o resultado ao fim de cinco anos, e ver as flutuações, considerando as relações entre N e N3, entre F31, F32 e F, e entre C3 e C. Os resultados encontram-se na tabela 4.28.

**Tabela 4.28 - Resultados ao fim de três anos**

PERFIL	N3 (N)	F31+ F32 (F)	C3 (C)
F	15 (16) ←1	2 + 3 (3)	1 (2) →1
S	2 (3) ←1	2 + 8 (5)	7 (11) →4
B	0 (0)	0 + 0 (0)	6 (6)
MB	0 (0)	0 + 0 (0)	6 (6)



Como se pode verificar, os alunos com perfil B e MB mantêm-se estáveis até ao fim do curso. De resto, existem duas flutuações de F31 para N e cinco de F32 para C, todas elas em alunos com perfil F e S.

Tendo em conta estes factos, tentou-se averiguar se, de entre as disciplinas do primeiro ano, existe alguma que seja determinante para a conclusão da licenciatura em cinco anos, ou para um bom desempenho ao fim de três anos.

Na tabelas 4.29 e 4.30, pode-se ver os modelos gerados pelo C5.0, para cada um dos casos (três anos, cinco anos e alunos de 1994/95 ao fim de três anos).

Tabela 4.29 - Árvores e regras de decisão relativos ao perfil dos alunos

Estudo	Árvore de decisão	Regras de decisão
Cinco anos	EIC1202 0 (4.0, 1.0) -> N EIC1202 1 (25.0, 0.52) -> C EIC1202 3 (15.0, 0.8) -> C EIC1202 2 EIC1101 [0 3] (0.0, 1.0) -> F EIC1101 2 (2.0, 1.0) -> F EIC1101 1 EIC1205 [0 3] (0.0, 1.0) -> N EIC1205 1 (2.0, 1.0) -> N EIC1205 2 (4.0, 0.75) -> F	<pre>                     if EIC1202 == 0                     then -&gt; N (4, 0.833)                      if EIC1101 == 1                     and EIC1202 == 2                     and EIC1205 == 1                     then -&gt; N (2, 0.75)                      if EIC1101 == 2                     and EIC1202 == 2                     then -&gt; F (2, 0.75)                      if EIC1202 == 2                     and EIC1205 == 2                     then -&gt; F (4, 0.667)                      if EIC1202 == 3                     then -&gt; C (15, 0.765)                      if EIC1202 == 1                     then -&gt; C (25, 0.519)                      Default : -&gt; C                 </pre>
Três anos (1994/95)	EIC1200 0 (4.0, 1.0) -> N3 EIC1200 2 (5.0, 1.0) -> C3 EIC1200 3 (4.0, 1.0) -> C3 EIC1200 1 EIC1204 0 (2.0, 1.0) -> N3 EIC1204 2 (2.0, 1.0) -> C3 EIC1204 3 (2.0, 0.5) -> F32 EIC1204 1 EIC1205 0 (0.0, 1.0) -> N3 EIC1205 1 (19.0, 0.526) -> N3 EIC1205 3 (4.0, 0.5) -> F32 EIC1205 2 EIC1102 0 (0.0, 1.0) -> C3 EIC1102 1 (4.0, 0.5) -> F31 EIC1102 2 (3.0, 1.0) -> F32 EIC1102 3 (3.0, 0.667) -> C3	<pre>                     if EIC1200 == 0                     then -&gt; N3 (4, 0.833)                      if EIC1200 == 1                     and EIC1204 == 1                     and EIC1205 == 1                     then -&gt; N3 (19, 0.524)                      if EIC1200 == 1                     then -&gt; N3 (39, 0.341)                      if EIC1102 == 1                     and EIC1200 == 1                     and EIC1205 == 2                     then -&gt; F31 (4, 0.5)                      if EIC1102 == 2                     and EIC1205 == 2                     then -&gt; F32 (3, 0.8)                      if EIC1204 == 2                     then -&gt; C3 (5, 0.857)                      if EIC1200 == 2                     then -&gt; C3 (5, 0.857)                      if EIC1102 == 3                     then -&gt; C3 (17, 0.684)                      Default : -&gt; C3                 </pre>

Tabela 4.30 - Árvores e regras de decisão relativos ao perfil dos alunos

Estudo	Árvore de decisão	Regras de decisão
Três anos	<pre> EIC1200 0 (34.0, 0.912) -&gt; N3 EIC1200 2 (15.0, 0.933) -&gt; C3 EIC1200 3 (12.0, 1.0) -&gt; C3 EIC1200 1   EIC1204 0 (9.0, 1.0) -&gt; N3   EIC1204 2 (14.0, 0.857) -&gt; C3   EIC1204 3     EIC1203 [0 2] (0.0, 1.0) -&gt; C3     EIC1203 1 (3.0, 0.667) -&gt; F32     EIC1203 3 (3.0, 1.0) -&gt; C3   EIC1204 1     EIC1206 0 (1.0, 1.0) -&gt; F32     EIC1206 2 (9.0, 0.778) -&gt; C3     EIC1206 3 (3.0, 0.333) -&gt; N3   EIC1206 1     EIC1101 0 (1.0, 1.0) -&gt; N3     EIC1101 2 (8.0, 0.75) -&gt; F32     EIC1101 3 (1.0, 1.0) -&gt; C3   EIC1101 1     EIC1104 0 (0.0, 1.0) -&gt; F32     EIC1104 2 (6.0, 0.5) -&gt; C3     EIC1104 3 (3.0, 1.0) -&gt; C3   EIC1104 1     EIC1203 0 (4.0, 0.5) -&gt; N3     EIC1203 1 (25.0, 0.36) -&gt; F32     EIC1203 2 (7.0, 0.429) -&gt; F31     EIC1203 3 (1.0, 1.0) -&gt; C3 </pre>	<pre> if EIC1204 == 0 then -&gt; N3 (29, 0.968)  if EIC1101 == 0 then -&gt; N3 (23, 0.92)  if EIC1203 == 0 then -&gt; N3 (27, 0.897)  if EIC1200 == 0 then -&gt; N3 (34, 0.889)  if EIC1101 == 1 and EIC1104 == 1 and EIC1200 == 1 and EIC1203 == 2 and EIC1204 == 1 then -&gt; F31 (7, 0.444)  if EIC1101 == 2 and EIC1200 == 1 and EIC1204 == 1 and EIC1206 == 1 then -&gt; F32 (8, 0.7)  if EIC1200 == 1 and EIC1204 == 1 and EIC1206 == 0 then -&gt; F32 (1, 0.667)  if EIC1203 == 1 and EIC1204 == 3 then -&gt; F32 (3, 0.6)  if EIC1200 == 1 and EIC1203 == 1 and EIC1204 == 1 and EIC1206 == 1 then -&gt; F32 (36, 0.395)  if EIC1200 == 3 then -&gt; C3 (12, 0.929)  if EIC1203 == 3 and EIC1204 == 3 then -&gt; C3 (7, 0.889)  if EIC1200 == 2 then -&gt; C3 (15, 0.882)  if EIC1200 == 1 and EIC1206 == 2 then -&gt; C3 (16, 0.833)  if EIC1104 == 3 and EIC1204 == 1 and EIC1206 == 1 then -&gt; C3 (4, 0.833)  if EIC1204 == 2 then -&gt; C3 (21, 0.826)  if EIC1101 == 1 and EIC1104 == 1 and EIC1203 == 3 then -&gt; C3 (3, 0.8)  if EIC1104 == 2 and EIC1200 == 1 then -&gt; C3 (20, 0.636)  Default : -&gt; F32 </pre>

Existem muitas ilações que se podem tirar daqui. Se compararmos os três casos, vemos que as disciplinas EIC1200 (Análise Matemática II) e EIC1204 (Arquitectura de Computadores) revelam-se essenciais para o sucesso ao fim de três anos.

A dependência da disciplina EIC1200 pode-se ficar a dever ao facto de, nos três primeiros anos, haver uma grande incidência das disciplinas das áreas científicas de Matemática e também de Modelação Física, num total de onze.

A dependência da disciplina EIC1204 já parece estar mais relacionada com o facto de se tratar de uma disciplina que aborda a estrutura interna de um computador, o que parece ser de grande importância para esta licenciatura.

No entanto a importância destas disciplinas acabam por se esbater ao fim do quinto ano, onde ganha maior consistência a dependência da disciplina EIC1202 (Introdução à Programação II). Este facto pode ter a ver com a motivação que um aluno que entra nesta licenciatura ganha ao ter um bom desempenho nesta disciplina. A não aprovação a esta disciplina pode ser um factor de desmotivação, uma vez que parece indicar que o aluno não terá feito a escolha acertada relativamente à licenciatura.

#### **4.7. Discussão dos resultados obtidos**

Expostos os resultados obtidos, vamos agora debruçar-nos sobre a informação que foi possível obter. O estudo de árvores e regras de decisão baseou-se em três tipos diferentes de casos:

- previsão da classificação das disciplinas do segundo semestre de um determinado ano, usando as disciplinas do primeiro semestre do mesmo ano, do primeiro ao quarto ano;
- previsão da classificação das disciplinas de um determinado ano completo, usando as disciplinas do ano anterior, do segundo ao quinto ano;
- previsão da classificação da situação de um aluno ao fim de três e cinco anos, usando o perfil baseado nas notas do primeiro ano.

Vamos começar por comparar os dois primeiros casos. Como já se viu, nos primeiro e quinto anos, não pode haver qualquer tipo de comparação, uma vez que só se realiza a análise de uma das formas.

Comparando os restantes três anos, colocámos a seguinte questão: alguma das previsões supera inequivocamente a outra? Os resultados parecem negar esta hipótese. Existem algumas disciplinas que parecem estar mais dependentes de disciplinas do mesmo ano, ao passo que outras parecem estar mais dependentes de disciplinas do ano anterior.

Um exemplo de disciplinas que dependem de outras do mesmo ano são as disciplinas complementares (Análise Matemática I e II, Introdução à Programação I e II, etc...). De notar que houve uma excepção, no terceiro ano, onde não foi determinada nenhuma interdependência entre as disciplinas Teoria da Computação I e II. Apesar disso, a disciplina Teoria da Computação I acabou por se revelar bastante importante na classificação das disciplinas do quarto ano.

Se tivermos em conta as percentagens de acertos na aplicação dos modelos aos conjuntos de teste, podemos considerar que a previsão do ano completo consegue valores ligeiramente superiores em alguns casos, especialmente no terceiro ano. Nos restantes anos constata-se algum equilíbrio.

As áreas científicas parecem não ser determinantes na aprovação das disciplinas. Tirando os casos das disciplinas complementares (mesmo assim com a excepção referida) não existe, de um modo geral, uma grande interdependência entre disciplinas da mesma área científica. Note-se que o facto de duas disciplinas serem da mesma área científica não implica que os programas das disciplinas tenham alguma relação entre si.

Relativamente às disciplinas mais relevantes ao longo dos diferentes anos, constata-se que existem duas fases distintas. Nos primeiros anos verifica-se que as disciplinas mais importantes são as relacionadas com informática e computadores: Introdução à Programação I, no segundo semestre do primeiro ano, Algoritmos e Estruturas de Dados I, e Sinais e Telecomunicações, no segundo semestre do segundo ano, Arquitectura de Computadores, Circuitos e Sistemas II, e Electrónica Digital, no segundo ano completo, Novas Metodologias de Programação, Laboratório de Computação Gráfica e Interfaces, e Programação Paralela, no segundo semestre do terceiro ano.

Nas previsões dos terceiro e quarto anos completos, começam a revelar-se mais importantes as disciplinas de outras áreas científicas: Física e Estatística Aplicada, no terceiro ano, e Teoria da Computação I e Investigação Operacional, no quarto ano.

A que se poderá dever esta mudança? A explicação que nos parece mais lógica terá a ver com dois factores importantes. Em primeiro lugar a motivação. Quando um aluno entra numa Licenciatura em Engenharia Informática e Computação tem, em princípio, maior motivação para as disciplinas relacionadas com informática e computadores. O segundo factor é a diversificação. Após a fase inicial de assimilação e reforço dos conhecimentos obtidos, existe uma necessidade de diversificação, aplicando estes conhecimentos a novos casos, mais complexos e de natureza diversa.

Vejamos agora a comparação entre estas duas análises e a última, relativa ao perfil dos alunos, baseado nas notas do primeiro ano. Existe alguma coerência entre as duas análises. Assim, existem três disciplinas do primeiro ano que são determinantes para um bom desempenho ao fim de três anos e para a conclusão da licenciatura em cinco anos: Análise Matemática II, Introdução à Programação II e Arquitectura de Computadores.

A disciplina de Arquitectura de Computadores é a disciplina mais relevante para classificação das disciplinas do segundo ano completo. Por outro lado a disciplina Introdução à Programação II é complementar de Introdução à Programação I, que é uma disciplina determinante para as disciplinas do segundo semestre do primeiro ano.

A disciplina de Análise Matemática II aparece em evidência na análise do perfil, relativamente aos três primeiros anos, mas não aparece com tanta evidência nas outras análises. Pode-se pensar que esta disciplina não tem tanta influência em nenhuma disciplina particular, mas sim na globalidade. Como já foi referido, existe um total de onze disciplinas de Matemática e Modelação Física num total de trinta e seis nos três primeiros anos, o que pode explicar, em grande parte, o facto de aparecer em evidência na análise feita para o terceiro ano e não acontecer o mesmo para o quinto ano.

# Capítulo 5

## Análise das regras de associação

Neste capítulo faz-se a análise das regras de associação. Esta análise é realizada por ano curricular, usando os mesmos critérios do que aconteceu com as árvores e regras de decisão. Neste caso usaram-se os dados todos, ao contrário da divisão que se fez para o C5.0.

Assim, da secção 5.1 até à 5.5 são descritos os resultados do primeiro ao quinto ano, respectivamente. Na secção 5.6 é feita uma análise tendo em conta o perfil dos alunos que entraram no ano lectivo de 1994/95. Finalmente, na secção 5.7, é feita uma discussão dos resultados obtidos.

### 5.1. Primeiro ano

No primeiro ano só se vai usar as disciplinas relativas ao primeiro semestre como entrada e as do segundo como saída. Como já vimos, todas as regras foram geradas com um suporte mínimo de 20% e uma confiança mínima de 80%. Restringiu-se o número de elementos que podem aparecer nas premissas a quatro. Neste caso só foram considerados os casos de aprovação na disciplina.

Na maior parte dos casos, todas as combinações possíveis de regras com uma, duas, três ou quatro disciplinas na premissa foram geradas. Isto acontece no primeiro ano onde, para cada disciplina, todas as combinações foram geradas, num total de cinquenta e seis casos.

A aplicação do filtro que foi desenvolvido durante a presente dissertação permitiu uma redução automática significativa do número de casos. Na tabela 5.1 podemos ver o número de regras de associação geradas pelo *apriori* e o número após a aplicação do filtro.

Tabela 5.1 - Número de regras geradas para o primeiro ano

Disciplina	Regras geradas	Regras após filtro
EIC1100	56	10
EIC1102	56	10
EIC1103	56	11
EIC1104	56	9
EIC1105	56	7
EIC1106	56	16

Nas regras de associação vamos ter em conta, para cada disciplina, os seguintes pontos (entre parêntesis está a forma como serão referenciados): regras com maior confiança (C), regras com maior suporte (S), regras correspondentes a um caminho na árvore de decisão correspondente (A), regras correspondentes a uma regra de decisão correspondente (R) e regras com apenas uma disciplina na premissa de maior confiança (1).

Vejamos os resultados de cada disciplina, assinalando as regras mais relevantes desta forma.

Para a disciplina EIC1200 foram consideradas as dez regras de associação, que estão na figura 5.1 (entre parêntesis estão o suporte e a confiança, respectivamente).

(C)	EIC1200 $\leq$ EIC1101 & EIC1102 & EIC1103 & EIC1100 (71,2%; 98,0%)
(C)	EIC1200 $\leq$ EIC1102 & EIC1103 & EIC1100 (71,2%; 98,0%)
(S)	EIC1200 $\leq$ EIC1106 (93,5%; 83,5%)
	EIC1200 $\leq$ EIC1103 & EIC1100 (73,7%; 97,6%)
	EIC1200 $\leq$ EIC1104 (86,3%; 85,8%)
	EIC1200 $\leq$ EIC1102 & EIC1100 (74,5%; 97,1%)
	EIC1200 $\leq$ EIC1101 (84,5%; 91,5%)
	EIC1200 $\leq$ EIC1101 & EIC1100 (78,4%; 96,8%)
	EIC1200 $\leq$ EIC1106 & EIC1101 (83,1%; 91,8%)
(AR1)	EIC1200 $\leq$ EIC1100 (79,5%; 96,4%)

Figura 5.1 - Regras de associação para a disciplina EIC1200

As duas primeiras regras são as que têm maior confiança ao passo que a terceira é a que tem maior suporte. Em relação às regras com maior confiança, que também têm o mesmo suporte, verifica-se que a primeira tem mais uma disciplina na premissa: a EIC1101. Isto significa que sempre que um aluno obtém aprovação às outras disciplinas (EIC1100, EIC1102 e EIC1103), também obtém aprovação à disciplina EIC1101.

Se olharmos para a sétima regra, verificamos que esta disciplina também aparece sozinha na premissa, com um suporte maior e uma confiança menor. Logo, poderemos considerar que, entre as duas regras, a que inclui esta disciplina transmite mais informação que a outra.

Outra característica que foi tida em conta foi verificar se as regras de associação, de alguma forma, coincidem com o que foi encontrado nas árvores e regras de decisão. Verifica-se que a última regra, corresponde a um caminho, isto é, uma sequência da raiz até uma folha, da árvore de decisão da disciplina EIC1200 e também a uma das regras de decisão. Para além disso, de entre as regras de associação que só têm uma disciplina na premissa, é a que detém melhor confiança (96,4%).

Uma chamada de atenção para o facto de só existir, no máximo, um caminho na árvore de decisão em que todos os ramos correspondem à classe 1 (o mesmo acontece para a classe 0). Assim, uma regra de associação só pode corresponder a um caminho na árvore de decisão, no máximo. Em relação às regras de decisão, uma vez que são simplificadas, pode haver mais do que uma regra.

Tendo em conta a relação entre as disciplinas da mesma área científica, as disciplinas EIC1100 e EIC1101, tal como EIC1200 da área científica de Matemática, aparecem em duas regras isoladamente, e numa outra em conjunto, para além de ambas fazerem parte da regra com maior confiança.

Para a disciplina EIC1202 consideraram-se também dez regras, que se encontram na figura 5.2.

As duas primeiras regras são as que têm maior confiança e suporte, respectivamente. Note-se que as três regras que aparecem na premissa da regra com maior confiança também aparecem isoladas noutras regras.



(C)	$EIC1202 \leq EIC1104 \ \& \ EIC1101 \ \& \ EIC1102$ (75,9%; 99,1%)
(S)	$EIC1202 \leq EIC1106$ (93,5%; 90,8%)
	$EIC1202 \leq EIC1101 \ \& \ EIC1102$ (78,1%, 98,6%)
	$EIC1202 \leq EIC1104$ (86,3%, 94,2%)
	$EIC1202 \leq EIC1106 \ \& \ EIC1104 \ \& \ EIC1102$ (78,4%; 98,2%)
	$EIC1202 \leq EIC1101$ (84,5%; 94,9%)
	$EIC1202 \leq EIC1104 \ \& \ EIC1102$ (79,5%; 97,7%)
	$EIC1202 \leq EIC1106 \ \& \ EIC1104$ (84,2%; 95,7%)
	$EIC1202 \leq EIC1106 \ \& \ EIC1102$ (82,0%; 97,4%)
(AR1)	$EIC1202 \leq EIC1102$ (83,5%; 97,0%)

*Figura 5.2 - Regras de associação para a disciplina EIC1202*

Tal como tinha acontecido na disciplina EIC1200, também aqui temos uma regra, a última, que corresponde a um caminho da árvore de decisão e a uma regra de decisão, sendo também, de entre as regras com apenas uma disciplina na premissa, aquela que tem maior confiança. A disciplina EIC1102, que aparece na premissa, é da mesma área científica da EIC1202, Linguagens de Programação.

De referir que em nenhum lado aparece a regra de associação:

$EIC1202 \leq EIC1101 \ \& \ EIC1104$

correspondente a uma regra de decisão. Em compensação estas duas disciplinas aparecem isoladamente nas premissas das quarta e sexta regras, e aparecem acompanhados da EIC1102, que corresponde a outra das regras de decisão.

Em relação à disciplina EIC1203, temos onze regras a considerar, que se encontram na tabela 5.3.

De referir que a disciplina EIC1103 é da área científica de Modelação Física, a mesma da EIC1203. Esta disciplina aparece na regra de maior confiança bem como na regra com apenas uma disciplina na premissa com maior confiança.

Passando à disciplina EIC1204, vamos considerar nove regras, constantes da figura 5.4.

(C)	$EIC1203 \Leftarrow EIC1101 \ \& \ EIC1103 \ \& \ EIC1100$ (73,4%; 97,1%)
(S)	$EIC1203 \Leftarrow EIC1106$ (93,5%; 81,9%) $EIC1203 \Leftarrow EIC1103 \ \& \ EIC1100$ (73,7%; 96,6%)
(R)	$EIC1203 \Leftarrow EIC1104$ (86,3%; 87,1%) $EIC1203 \Leftarrow EIC1104 \ \& \ EIC1101 \ \& \ EIC1103$ (75,5%; 94,8%) $EIC1203 \Leftarrow EIC1101$ (84,5%; 88,1%)
(A)	$EIC1203 \Leftarrow EIC1101 \ \& \ EIC1103$ (77,7%; 94,4%) $EIC1203 \Leftarrow EIC1106 \ \& \ EIC1104$ (84,2%; 88,5%) $EIC1203 \Leftarrow EIC1104 \ \& \ EIC1103$ (79,1%; 92,7%) $EIC1203 \Leftarrow EIC1106 \ \& \ EIC1101$ (83,1%; 89,2%)
(R1)	$EIC1203 \Leftarrow EIC1103$ (82,4%; 92,1%)

*Figura 5.3 - Regras de associação para a disciplina EIC1203*

(C)	$EIC1204 \Leftarrow EIC1104 \ \& \ EIC1101 \ \& \ EIC1102$ (75,9%; 96,2%)
(S)	$EIC1204 \Leftarrow EIC1106$ (93,5%; 88,5%) $EIC1204 \Leftarrow EIC1106 \ \& \ EIC1104 \ \& \ EIC1102$ (78,4%; 95,4%) $EIC1204 \Leftarrow EIC1101 \ \& \ EIC1100$ (78,4%; 95,4%) $EIC1204 \Leftarrow EIC1104$ (86,3%; 91,7%)
(AR)	$EIC1204 \Leftarrow EIC1104 \ \& \ EIC1102$ (79,5%; 95,0%)
(I)	$EIC1204 \Leftarrow EIC1101$ (84,5%; 93,2%) $EIC1204 \Leftarrow EIC1104 \ \& \ EIC1101$ (80,2%; 94,6%) $EIC1204 \Leftarrow EIC1106 \ \& \ EIC1101$ (83,1%; 93,9%)

*Figura 5.4 - Regras de associação para a disciplina EIC1204*

A sexta regra corresponde a um caminho na árvore de decisão e a uma regra de decisão. Também aqui não aparece a regra de associação:

$$EIC1204 \Leftarrow EIC1100$$

correspondente a uma regra de decisão. A única vez que esta disciplina aparece na premissa de uma regra de associação é na quarta regra, acompanhada da disciplina EIC1101. Esta última, apesar de não aparecer nas árvores e regras de decisão, é a disciplina relativa à regra de associação com apenas uma disciplina na premissa que tem maior confiança.

Na disciplina EIC1205 temos a considerar sete regras, que se encontram na figura 5.5.

(C)	$EIC1205 \leq EIC1104 \ \& \ EIC1101 \ \& \ EIC1102$ (75,9%; 98,6%)
(S)	$EIC1205 \leq EIC1106$ (93,5%; 93,8%)
(R1)	$EIC1205 \leq EIC1100$ (79,5%; 98,2%)
	$EIC1205 \leq EIC1104$ (86,3%; 95,0%)
	$EIC1205 \leq EIC1104 \ \& \ EIC1101$ (80,2%; 97,8%)
	$EIC1205 \leq EIC1101$ (84,5%; 97,0%)
	$EIC1205 \leq EIC1106 \ \& \ EIC1101$ (83,1%; 97,4%)

*Figura 5.5 - Regras de associação para a disciplina EIC1205*

A terceira regra corresponde a uma regra de decisão, sendo também, de entre as regras de associação com apenas uma disciplina na premissa, aquela que tem maior confiança, estando bastante próximo da que tem maior confiança.

Neste caso não aparece a regra de associação:

$EIC1205 \leq EIC1102$

que corresponde a um caminho numa árvore de decisão e a uma regra de decisão. A disciplina EIC1102 só aparece uma vez, na primeira regra de associação e acompanhado de outras duas.

Relativamente à disciplina EIC1104, da mesma área científica que a EIC1205, Sistemas Electrónicos Digitais, note-se que aparece na regra com maior confiança, aparecendo ainda numa outra isoladamente com um confiança muito alta (95,0%).

Finalmente, na figura 5.6, vamos considerar a disciplina EIC1206, com dezasseis regras de associação.

A disciplina EIC1100, apesar de estar na premissa da regra com maior confiança de entre as que têm apenas uma disciplina na premissa, não entra nem na árvore nem em nenhuma das regras de decisão geradas pelo C5.0 para a disciplina EIC1206.

Também aqui falta não aparece uma regra de associação:

$EIC1206 \leq EIC1102 \ \& \ EIC1103$

correspondente a uma regra de decisão. Este facto ainda é mais curioso quando notamos que a disciplina EIC1102 não aparece em nenhuma regra de associação para esta disciplina.

(C)	$EIC1206 \Leftarrow EIC1104 \ \& \ EIC1101 \ \& \ EIC1103 \ \& \ EIC1100$ (71,6%; 96,5%)
(C)	$EIC1206 \Leftarrow EIC1104 \ \& \ EIC1103 \ \& \ EIC1100$ (71,6%; 96,5%)
(S)	$EIC1206 \Leftarrow EIC1106$ (93,5%; 87,7%)
	$EIC1206 \Leftarrow EIC1106 \ \& \ EIC1104 \ \& \ EIC1101 \ \& \ EIC1100$ (74,1%; 96,1%)
	$EIC1206 \Leftarrow EIC1106 \ \& \ EIC1104 \ \& \ EIC1100$ (74,1%; 96,1%)
	$EIC1206 \Leftarrow EIC1104$ (86,3%; 91,3%)
	$EIC1206 \Leftarrow EIC1104 \ \& \ EIC1100$ (74,8%; 95,7%)
	$EIC1206 \Leftarrow EIC1104 \ \& \ EIC1101 \ \& \ EIC1100$ (74,8%; 95,7%)
(AR)	$EIC1206 \Leftarrow EIC1101$ (84,5%; 92,3%)
	$EIC1206 \Leftarrow EIC1104 \ \& \ EIC1101 \ \& \ EIC1103$ (75,5%; 95,2%)
	$EIC1206 \Leftarrow EIC1106 \ \& \ EIC1104$ (84,2%; 92,7%)
	$EIC1206 \Leftarrow EIC1106 \ \& \ EIC1101 \ \& \ EIC1100$ (77,7%; 94,9%)
	$EIC1206 \Leftarrow EIC1106 \ \& \ EIC1101$ (83,1%; 93,1%)
	$EIC1206 \Leftarrow EIC1106 \ \& \ EIC1104 \ \& \ EIC1101$ (79,1%; 94,5%)
	$EIC1206 \Leftarrow EIC1104 \ \& \ EIC1101$ (80,2%; 93,7%)
(1)	$EIC1206 \Leftarrow EIC1100$ (79,5%; 94,1%)

*Figura 5.6 - Regras de associação para a disciplina EIC1206*

Outra curiosidade prende-se com o facto de aparecem alguns pares de regras com o mesmo suporte e confiança, e cuja diferença assenta no facto de a disciplina EIC1101 estar ou não presente na premissa. Isto acontece nas duas primeiras regras, nas quarta e quinta, e nas sétima e oitava. Como já vimos, esta disciplina aparece isolada na premissa da nona regra, com 92,3% de confiança. Assim, podemos considerar que, das regras referidas, as que contêm a disciplina EIC1101, dão mais informação.

Considerando a área científica da disciplina EIC1206, Modelação Física, vamos considerar a disciplina EIC1103, da mesma área. Esta aparece apenas nas duas regras de maior confiança, não aparecendo em nenhuma isoladamente.

Para resumir os resultados temos, na tabela 5.2, uma matriz com as disciplinas do primeiro semestre nas colunas, as do segundo nas linhas, e os valores representam as relações entre as disciplinas em relação às regras com maior confiança (C) e suporte (S), e a regras com apenas uma disciplina na premissa com maior confiança (1). Não se referem as relações com árvores e regras de decisão, uma vez que estas relações já foram analisadas no capítulo anterior.

Tabela 5.2 - Relação entre as disciplinas do primeiro ano

	EIC1100	EIC1101	EIC1102	EIC1103	EIC1104	EIC1106
EIC1200	C 1	C	C	C	-	S
EIC1202	-	C	C 1	-	C	S
EIC1203	C	C	-	C 1	-	S
EIC1204	-	C 1	C	-	C	S
EIC1205	1	C	C	-	C	S
EIC1206	C 1	C	-	C	C	S

Note-se que, para todas as disciplinas do primeiro ano, a regra com maior suporte é uma regra em que na premissa aparece apenas a disciplina EIC1106 (Técnicas de Comunicação e Inglês Funcional). A razão para esta coincidência prende-se com o facto de esta disciplina ser a que tem maior percentagem de aprovações ao longo dos anos. Isto mostra, até certo ponto, que a confiança é mais importante para a determinação da importância de uma regra de associação.

Outra referência para a disciplina EIC1101 (Álgebra) que aparece em todas as regras de associação com maior confiança. No entanto, nas regras com apenas uma disciplina na premissa, é a disciplina EIC1100 (Análise Matemática I) que aparece mais vezes, num total de três. Pode-se supor que as disciplinas da área científica de Matemática tenham uma grande influência nas disciplinas do segundo semestre do primeiro ano.

## 5.2. Segundo ano

No segundo ano consideraram-se as duas situações que haviam sido consideradas para a classificação. No primeiro caso vamos usar as disciplinas do primeiro semestre do segundo ano como disciplinas de entrada, e as do segundo semestre do segundo ano como saída. No segundo caso vamos usar as disciplinas do primeiro ano como disciplinas de entrada, e as do segundo ano como saída. As disciplinas do segundo semestre serão comparadas em termos de suporte e confiança.

Começando pela primeira forma, temos novamente cinquenta e seis regras de associação para todas as disciplinas, antes da aplicação do filtro. Na tabela 5.3 temos o número de regras de associação, por disciplina, antes e após a aplicação do filtro.

Tabela 5.3 - Número de regras geradas para o segundo ano

Disciplina	Regras geradas	Regras após filtro
EIC2200	56	15
EIC2201	56	10
EIC2202	56	15
EIC2203	56	7
EIC2204	56	17
EIC2205	56	5

Começando pela disciplina EIC2200, após a aplicação do filtro, consideramos quinze regras, que se podem ver na figura 5.7.

(C)	EIC2200 <= EIC2101 & EIC2102 & EIC2100 (61,7%; 99,2%)
(C)	EIC2200 <= EIC2105 & EIC2101 & EIC2102 & EIC2100 (61,7%; 99,2%)
(SAR)	EIC2200 <= EIC2105 (83,6%; 90,5%)
	EIC2200 <= EIC2105 & EIC2102 & EIC2100 (63,6%; 98,5%)
	EIC2200 <= EIC2105 & EIC2104 (78,5%; 92,3%)
	EIC2200 <= EIC2102 & EIC2100 (64,0%; 97,8%)
(R)	EIC2200 <= EIC2103 (75,2%; 92,5%)
	EIC2200 <= EIC2105 & EIC2104 & EIC2101 & EIC2102 (65,4%; 97,1%)
	EIC2200 <= EIC2104 & EIC2101 & EIC2102 (65,4%; 97,1%)
(I)	EIC2200 <= EIC2102 (74,3%; 93,7%)
	EIC2200 <= EIC2101 & EIC2102 (66,8%; 96,5%)
	EIC2200 <= EIC2105 & EIC2101 & EIC2102 (66,8%; 96,5%)
	EIC2200 <= EIC2105 & EIC2102 (71,0%; 94,7%)
	EIC2200 <= EIC2103 & EIC2102 (68,2%; 95,9%)
	EIC2200 <= EIC2105 & EIC2104 & EIC2102 (69,6%; 95,3%)

Figura 5.7 - Regras de associação para a disciplina EIC2200

As duas primeiras regras são as com maior confiança, ao passo que a terceira é a que tem maior suporte. Para além das duas primeiras com o mesmo suporte e

confiança, também as oitava e nona, e as décima primeira e décima segunda, têm a mesma característica. A diferença entre cada par é a presença nas premissas de mais uma disciplina: a EIC2105.

Esta disciplina aparece na premissa da regra com maior suporte e tem uma confiança bastante alta (90,5%). Para além disso, esta regra de associação corresponde a um caminho na árvore de decisão e a uma das regras de decisão. Assim, pode-se concluir que as regras mencionadas que incluem esta disciplina serão mais importantes que as outras.

Existe ainda outra regra de associação correspondente a uma regra de decisão: a sétima. Apesar disto, a regra de associação de entre as com apenas uma disciplina na premissa com maior confiança é a décima, que não corresponde a nenhum dos modelos do C5.0.

Vamos agora considerar a disciplina EIC2201, onde teremos um total de dez regras de associação, conforme se pode ver na figura 5.8.

(C)	EIC2201 $\Leftarrow$ EIC2101 & EIC2102 & EIC2100 (61,7%; 99,2%)
(C)	EIC2201 $\Leftarrow$ EIC2105 & EIC2101 & EIC2102 & EIC2100 (61,7%; 99,2%)
(S)	EIC2201 $\Leftarrow$ EIC2105 (83,6%; 93,3%)
	EIC2201 $\Leftarrow$ EIC2105 & EIC2101 & EIC2100 (65,9%; 98,6%)
	EIC2201 $\Leftarrow$ EIC2101 & EIC2100 (65,9%; 98,6%)
(R)	EIC2201 $\Leftarrow$ EIC2104 (81,3%; 93,7%)
	EIC2201 $\Leftarrow$ EIC2105 & EIC2100 (68,7%; 98,0%)
	EIC2201 $\Leftarrow$ EIC2105 & EIC2104 (78,5%; 94,6%)
(AR1)	EIC2201 $\Leftarrow$ EIC2101 (74,3%; 96,9%)
	EIC2201 $\Leftarrow$ EIC2103 (75,2%; 95,7%)

*Figura 5.8 - Regras de associação para a disciplina EIC2201*

As duas primeiras regras são as que têm maior confiança, enquanto a terceira é a que tem maior suporte. Tanto as duas primeiras como as quarta e quinta têm em comum o facto de terem o mesmo suporte e a mesma confiança, e de a diferença entre as duas ser a presença, ou não, da disciplina EIC2105, como já aconteceu na disciplina EIC2200.

Tal como aconteceu nesta última disciplina, também aqui existe uma regra de associação em que a disciplina EIC2105 aparece isoladamente na premissa, sendo esta a regra com maior suporte. Apesar de não aparecer associada a nenhum modelo

do C5.0, pode-se considerar que as regras de associação contendo aquela disciplina são mais importantes que as outras.

Vamos passar à disciplina EIC2202, onde vamos considerar quinze casos, como se pode ver na figura 5.9.

(C)	EIC2202 <= EIC2101 & EIC2102 & EIC2100 (61,7%; 98,5%)
(C)	EIC2202 <= EIC2105 & EIC2101 & EIC2102 & EIC2100 (61,7%; 98,5%)
(S)	EIC2202 <= EIC2105 (83,6%; 82,1%)
	EIC2202 <= EIC2105 & EIC2102 & EIC2100 (63,6%; 97,8%)
	EIC2202 <= EIC2104 (81,3%; 84,5%)
	EIC2202 <= EIC2102 & EIC2100 (64,0%; 97,1%)
	EIC2202 <= EIC2105 & EIC2104 (78,5%; 86,3%)
	EIC2202 <= EIC2105 & EIC2104 & EIC2101 & EIC2102 (65,4%; 96,4%)
	EIC2202 <= EIC2104 & EIC2101 & EIC2102 (65,4%; 96,4%)
	EIC2202 <= EIC2103 (75,2%; 88,2%)
	EIC2202 <= EIC2101 & EIC2102 (66,8%; 95,8%)
	EIC2202 <= EIC2105 & EIC2101 & EIC2102 (66,8%; 95,8%)
(AR1)	EIC2202 <= EIC2102 (74,3%; 93,1%)
	EIC2202 <= EIC2104 & EIC2102 (70,6%; 95,4%)
	EIC2202 <= EIC2105 & EIC2102 (71,0%; 94,7%)

**Figura 5.9 - Regras de associação para a disciplina EIC2202**

As duas primeiras regras são as que têm maior confiança e a terceira é a que tem maior suporte. Mais uma vez temos pares de regras com o mesmo suporte e confiança, e cuja diferença reside na presença ou não da disciplina EIC2105 na premissa da regra. Isto acontece nas duas primeiras, nas oitava e nona, e nas décima primeira e décima segunda.

Tal como já acontecera nas duas disciplinas anteriores, esta disciplina aparece na premissa da regra com maior suporte, o que indica que as regras que contêm esta disciplina serão as mais importantes.

Relativamente aos modelos gerados pelo C5.0, a décima terceira regra corresponde a um caminho na árvore de decisão e a uma regra de decisão, para além de ser a regra, de entre as que só têm uma disciplina na premissa, com maior confiança. De referir ainda que a disciplina que aparece na premissa, a EIC2102, é da mesma área científica da EIC2202, Algoritmos e Estruturas de Dados. Estas disciplinas são complementares (Algoritmos e Estruturas de Dados I e II).



Passando para a disciplina EIC2203, consideramos sete regras, que se encontram na figura 5.10.

(C)	EIC2203 <= EIC2105 & EIC2102 & EIC2100 (63,6%; 100,0%)
(S)	EIC2203 <= EIC2105 (83,6%; 92,2%)
	EIC2203 <= EIC2105 & EIC2100 (68,7%; 99,3%)
	EIC2203 <= EIC2104 (81,3%; 92,5%)
	EIC2203 <= EIC2105 & EIC2103 (73,4%; 98,7%)
	EIC2203 <= EIC2105 & EIC2104 (78,5%; 94,0%)
(AR1)	EIC2203 <= EIC2103 (75,2%; 98,1%)

*Figura 5.10 - Regras de associação para a disciplina EIC2203*

Note-se que pela primeira vez se tem uma regra com confiança de 100%. Não aparece a regra de associação:

EIC2203 <= EIC2100

correspondente a uma regra de decisão. Este facto é compensado com a presença da disciplina EIC2100 na regra com maior confiança.

Vamos agora considerar a disciplina EIC2204, onde são consideradas dezassete regras de associação, que se encontram na figura 5.11.

Neste caso temos quatro pares de regras com os mesmos suporte e confiança, e cuja diferença é a presença ou não da disciplina EIC2105 na premissa. São elas as terceira e quarta, sexta e sétima, nona e décima, e décima segunda e décima terceira.

Uma vez mais a disciplina EIC2105 aparece isoladamente na premissa da regra com maior suporte, considerando-se assim que as regras consideradas que contêm esta disciplina serão as mais importantes. Não aparece em nenhum lado a regra de associação:

EIC2204 <= EIC2105 & EIC2102

que corresponde a um caminho na árvore de decisão e a uma regra de decisão. Além disso, a regra de associação, de entre as que só têm uma disciplina na premissa, com maior confiança é a décima quarta, e a disciplina que aparece na premissa, a EIC2101, não aparece nem na árvore de decisão nem em nenhuma das regras de decisão, aparecendo na regra de associação com maior confiança.

(C)	EIC2204 $\Leftarrow$ EIC2104 & EIC2101 & EIC2102 & EIC2100 (60,7%; 96,9%)
(S)	EIC2204 $\Leftarrow$ EIC2105 (83,6%; 82,7%)
	EIC2204 $\Leftarrow$ EIC2105 & EIC2104 & EIC2101 & EIC2100 (64,5%; 96,4%)
	EIC2204 $\Leftarrow$ EIC2104 & EIC2101 & EIC2100 (64,5%; 96,4%)
	EIC2204 $\Leftarrow$ EIC2104 (81,3%; 83,9%)
	EIC2204 $\Leftarrow$ EIC2101 & EIC2100 (65,9%; 95,7%)
	EIC2204 $\Leftarrow$ EIC2105 & EIC2101 & EIC2100 (65,9%; 95,7%)
	EIC2204 $\Leftarrow$ EIC2105 & EIC2104 (78,5%; 86,3%)
	EIC2204 $\Leftarrow$ EIC2105 & EIC2101 & EIC2102 (66,8%; 95,1%)
	EIC2204 $\Leftarrow$ EIC2101 & EIC2102 (66,8%; 95,1%)
	EIC2204 $\Leftarrow$ EIC2103 (75,2%; 88,8%)
	EIC2204 $\Leftarrow$ EIC2104 & EIC2103 & EIC2101 (67,8%; 94,5%)
	EIC2204 $\Leftarrow$ EIC2105 & EIC2104 & EIC2103 & EIC2101 (67,8%; 94,5%)
(I)	EIC2204 $\Leftarrow$ EIC2101 (74,3%; 91,2%)
	EIC2204 $\Leftarrow$ EIC2105 & EIC2104 & EIC2101 (70,6%; 94,0%)
	EIC2204 $\Leftarrow$ EIC2105 & EIC2101 (72,9%; 92,9%)
	EIC2204 $\Leftarrow$ EIC2104 & EIC2101 (71,0%; 93,4%)

**Figura 5.11 - Regras de associação para a disciplina EIC2204**

Para terminar, vamos considerar a disciplina EIC2205, com cinco regras de associação, que constam da figura 5.12.

(C)	EIC2205 $\Leftarrow$ EIC2104 & EIC2103 (72,0%; 99,4%)
(S)	EIC2205 $\Leftarrow$ EIC2105 (83,6%; 94,4%)
(ARI)	EIC2205 $\Leftarrow$ EIC2103 (75,2%; 98,8%)
(R)	EIC2205 $\Leftarrow$ EIC2104 (81,3%; 96,6%)
	EIC2205 $\Leftarrow$ EIC2105 & EIC2104 (78,5%; 97,6%)

**Figura 5.12 - Regras de associação para a disciplina EIC2205**

As duas primeiras regras são as que têm maior confiança e suporte, respectivamente. De entre as regra com apenas uma disciplina na premissa, a que tem maior confiança é a terceira, que corresponde a um caminho na árvore de decisão e a uma regra de decisão.

A quarta regra de associação também corresponde a uma regra de decisão e, para além disso, a disciplina na premissa, a EIC2104, também corresponde a uma disciplina da mesma área científica de Arquitectura de Computadores.

Na tabela 5.4 podemos ver as relações existentes entre as disciplinas dos dois semestres do segundo ano.

*Tabela 5.4 - Relação entre as disciplinas do segundo ano*

	EIC2100	EIC2101	EIC2102	EIC2103	EIC2104	EIC2105
EIC2200	C	C	C 1	-	-	C S
EIC2201	C	C 1	C	-	-	C S
EIC2202	C	C	C 1	-	-	C S
EIC2203	C	-	C	1	-	C S
EIC2204	C	C 1	C	-	C	S
EIC2205	-	-	-	C 1	C	S

De notar que, para todas as regras de associação, aparece sempre uma regra com a disciplina EIC2105 sempre com o mesmo suporte (83,6%), sendo este sempre o maior suporte. As razões poderiam ser as mesmas do que aconteceu com a disciplina EIC1106, no primeiro ano, isto é tratar-se uma disciplina com uma grande percentagem de aproveitamento.

No entanto, contrariamente ao que acontecia com a disciplina EIC1106, esta disciplina também aparece nas regras com maior confiança de quatro disciplinas. Assim, apesar de ser uma disciplina com um grande índice de aproveitamento, esta disciplina acaba por ter uma grande influência nas disciplinas do segundo semestre.

Relativamente às disciplinas que aparecem mais vezes nas regras com maior confiança são as EIC2100 e EIC2102, cinco vezes, seguidas da EIC2101 e da já referida EIC2105, ambas com quatro vezes. Relativamente às regras com apenas uma disciplina na premissa, existem três disciplinas que cobrem todos os casos, duas vezes cada: EIC2101, EIC2102 e EIC2103.

Olhando para os dados todos, nota-se que não existe nenhuma disciplina que assuma um papel preponderante para as disciplinas do segundo semestre. A razão para isto acontecer pode dever-se ao facto de as doze disciplinas que compõem os segundo ano pertencerem a oito áreas científicas diferentes, o que indica que não existam grandes relações entre as disciplinas, de um modo geral.

Vamos agora estudar as associação entre as disciplinas do primeiro ano e as do segundo ano completo. Neste caso vamos ter um número muito elevado de regras de associação. Na tabela 5.5 podemos ver o número de regras de associação geradas para cada disciplina e o número obtido após aplicação do filtro.

*Tabela 5.5 - Número de regras de associação para o segundo ano*

Disciplina	Regras geradas	Regras após filtro
EIC2100	759	41
EIC2101	792	57
EIC2102	789	75
EIC2103	791	31
EIC2104	793	37
EIC2105	793	19
EIC2200	793	40
EIC2201	793	56
EIC2202	765	126
EIC2203	793	18
EIC2204	778	77
EIC2205	793	30

Assim, vamos apenas apresentar as regras de associação mais relevantes, que são aquelas que temos destacado no estudo relativo ao segundo semestre: as regras com maiores suporte (referenciado com S) e confiança (C), as que correspondem a caminhos da árvore de decisão (A) ou a regras de decisão (R), e as que, de entre as regras com apenas uma disciplina na premissa da regra, tenham maior confiança (1).

Nas tabelas 5.6 e 5.7, temos as regras de associação mais relevantes, por disciplina, aparecendo sublinhadas as disciplinas da mesma área científica.

Constata-se que as disciplinas do segundo ano com disciplinas da mesma área científica no primeiro ano, têm sempre pelo menos uma destas nas regras com maior confiança.

Tal como já fizemos nos casos anteriores, podemos ver na tabela 5.8, de forma resumida, a relação entre as disciplinas do primeiro e segundo ano.

Tabela 5.6 - Regras de associação para as disciplinas do segundo ano

(C)	EIC2100 <= EIC1102 & <u>EIC1206</u> & EIC1204 & <u>EIC1203</u> (73,4%; 91,1%)
(S)	EIC2100 <= EIC1104 & EIC1101 (85,0%; 80,2%)
(I)	EIC2100 <= <u>EIC1203</u> (82,7%; 83,1%)
(C)	EIC2101 <= EIC1102 & EIC1206 & EIC1204 & EIC1200 (73,8%; 94,9%)
(C)	EIC2101 <= EIC1104 & EIC1206 & EIC1204 & EIC1200 (73,8%; 94,9%)
(S1)	EIC2101 <= EIC1205 (92,5%; 80,3%)
(C)	EIC2102 <= EIC1102 & EIC1204 & EIC1203 & EIC1100 (72,9%; 93,6%)
(C)	EIC2102 <= EIC1102 & EIC1204 & EIC1203 & EIC1200 (72,9%; 93,6%)
(S)	EIC2102 <= EIC1106 & EIC1104 (90,2%; 81,3%)
(AR)	EIC2102 <= EIC1102 & EIC1204 & EIC1203 (75,2%; 91,9%)
(I)	EIC2102 <= EIC1202 (89,3%; 82,2%)
(C)	EIC2103 <= EIC1102 & <u>EIC1206</u> & EIC1204 & <u>EIC1203</u> (73,4%; 96,2%)
(S)	EIC2103 <= EIC1104 (92,1%; 81,2%)
(AR)	EIC2103 <= EIC1104 & EIC1102 & <u>EIC1203</u> (78,0%; 94,0%)
(I)	EIC2103 <= <u>EIC1203</u> (82,7%; 89,8%)
(C)	EIC2104 <= EIC1102 & <u>EIC1204</u> & EIC1203 (75,2%; 96,3%)
(C)	EIC2104 <= EIC1104 & EIC1102 & <u>EIC1204</u> & EIC1203 (75,2%; 96,3%)
(C)	EIC2104 <= EIC1106 & EIC1102 & <u>EIC1204</u> & EIC1203 (75,2%; 96,3%)
(S)	EIC2104 <= EIC1106 (96,3%; 83,5%)
(I)	EIC2104 <= EIC1104 (92,1%; 86,8%)
(C)	EIC2105 <= EIC1104 & EIC1204 & <u>EIC1203</u> (76,2%; 98,2%)
(C)	EIC2105 <= EIC1106 & EIC1104 & EIC1204 & <u>EIC1203</u> (76,2%; 98,2%)
(S)	EIC2105 <= EIC1106 (96,3%; 86,4%)
(AR)	EIC2105 <= EIC1104 & <u>EIC1103</u> (86,9%; 93,5%)
(I)	EIC2105 <= EIC1104 (92,1%; 89,8%)
(C)	EIC2200 <= EIC1102 & EIC1103 & EIC1204 & EIC1203 (74,8%; 95,0%)
(C)	EIC2200 <= EIC1102 & EIC1202 & EIC1204 & EIC1203 (74,8%; 95,0%)
(C)	EIC2200 <= EIC1104 & EIC1202 & EIC1204 & EIC1203 (74,8%; 95,0%)
(C)	EIC2200 <= EIC1205 & EIC1102 & EIC1204 & EIC1203 (74,8%; 95,0%)
(S)	EIC2200 <= EIC1106 (96,3%; 82,0%)
(R)	EIC2200 <= EIC1103 & EIC1202 (84,1%; 90,0%)
(I)	EIC2200 <= EIC1202 (89,3%; 86,9%)
(C)	EIC2201 <= EIC1102 & EIC1101 & EIC1204 & EIC1200 (75,2%; 98,1%)
(C)	EIC2201 <= EIC1102 & EIC1204 & EIC1100 & EIC1200 (75,2%; 98,1%)
(S)	EIC2201 <= EIC1106 (96,3%; 84,0%)
(I)	EIC2201 <= EIC1102 (89,7%; 88,5%)
(I)	EIC2201 <= EIC1103 (89,7%; 88,5%)

Tabela 5.7 - Regras de associação para as disciplinas do segundo ano

(C)	EIC2202 ≤ EIC1102 & EIC1206 & EIC1204 & EIC1203 (73,4%; 90,4%)
(S)	EIC2202 ≤ EIC1106 & EIC1104 & EIC1102 (86,4%; 80,0%)
(A)	EIC2202 ≤ EIC1104 & EIC1204 & EIC1203 (76,2%; 87,7%)
(I)	EIC2202 ≤ EIC1204 (85,5%; 80,9%)
(C)	EIC2203 ≤ <u>EIC1104</u> & EIC1206 & EIC1204 & EIC1203 (74,3%; 98,7%)
(S)	EIC2203 ≤ EIC1106 (96,3%; 82,5%)
(AR)	EIC2203 ≤ <u>EIC1104</u> & EIC1206 & EIC1203 (77,6%; 98,2%)
(I)	EIC2203 ≤ EIC1203 (82,7%; 93,8%)
(C)	EIC2204 ≤ EIC1101 & EIC1206 & EIC1204 & EIC1203 (73,4%; 91,7%)
(C)	EIC2204 ≤ EIC1202 & EIC1204 & EIC1203 & EIC1100 (73,4%; 91,7%)
(C)	EIC2204 ≤ EIC1202 & EIC1204 & EIC1203 & EIC1200 (73,4%; 91,7%)
(C)	EIC2204 ≤ EIC1202 & EIC1206 & EIC1204 & EIC1203 (73,4%; 91,7%)
(S)	EIC2204 ≤ EIC1205 & EIC1202 (86,9%; 80,6%)
(I)	EIC2204 ≤ EIC1100 (81,8%; 85,1%)
(C)	EIC2205 ≤ EIC1102 & <u>EIC1204</u> & EIC1203 (75,2%; 98,1%)
(C)	EIC2205 ≤ EIC1104 & EIC1102 & <u>EIC1204</u> & EIC1203 (75,2%; 98,1%)
(C)	EIC2205 ≤ EIC1106 & EIC1102 & <u>EIC1204</u> & EIC1203 (75,2%; 98,1%)
(S)	EIC2205 ≤ EIC1106 (96,3%; 86,4%)
(R)	EIC2205 ≤ EIC1104 & EIC1203 (80,8%; 96,5%)
(I)	EIC2205 ≤ EIC1104 (92,1%; 88,3%)

Tabela 5.8 - Relação entre as disciplinas dos primeiro e segundo anos

	EIC1100	EIC1101	EIC1102	EIC1103	EIC1104	EIC1106	EIC1200	EIC1202	EIC1203	EIC1204	EIC1205	EIC1206
EIC2100	-	S	C	-	S	-	-	-	CS1	C	-	C
EIC2101	-	-	C	-	C	-	C	-	-	C	S1	C
EIC2102	C	-	C	-	S	S	C	1	C	C	-	-
EIC2103	-	-	C	-	S	-	-	-	C1	C	-	C
EIC2104	-	-	C	-	C1	CS	-	-	C	C	-	-
EIC2105	-	-	-	-	C1	CS	-	-	C	C	-	-
EIC2200	-	-	C	C	C	S	-	C1	C	C	C	-
EIC2201	C	C	C1	1	-	S	C	-	-	C	-	-
EIC2202	-	-	CS	-	S	S	-	-	C	C1	-	C
EIC2203	-	-	-	-	C	S	-	-	C1	C	-	C
EIC2204	C1	C	-	-	-	-	C	CS	C	C	S	C
EIC2205	-	-	C	-	C1	CS	-	-	C	C	-	-

Como se pode ver, a disciplina EIC1204 aparece em todas as regras de maior confiança. A disciplina que aparece mais vezes na regra com maior suporte é EIC1106, num total de sete. Relativamente às regras com apenas uma disciplina na premissa, as disciplinas EIC1104 e EIC1203 são as que aparecem mais vezes, três cada. Para além disso, a disciplina EIC1104 aparece ainda seis vezes nas regras com maior confiança e quatro nas com maior suporte, ao passo que a EIC1203 aparece dez vezes nas regras com maior confiança e uma vez nas com maior suporte.

Todas estas disciplinas são de grupos disciplinares diferentes, o que se explica pela heterogeneidade das áreas científicas do segundo ano.

Vamos agora fazer uma comparação das disciplinas do segundo semestre, considerando as regras com maior suporte e confiança. Esta comparação encontra-se na tabela 5.9.

*Tabela 5.9 - Comparação entre as regras de associação*

Disciplina	Primeiro e segundo semestre		Ano completo	
	Melhor confiança	Melhor suporte	Melhor confiança	Melhor suporte
EIC2200	(61,7%; 99,2%)	(83,6%; 90,5%)	(74,8%; 95,0%)	(96,3%; 82,0%)
EIC2201	(61,7%; 99,2%)	(83,6%; 93,3%)	(75,2%; 98,1%)	(96,3%; 84,0%)
EIC2202	(61,7%; 98,5%)	(83,6%; 82,1%)	(73,4%; 90,4%)	(86,4%; 80,0%)
EIC2203	(63,6%; 100,0%)	(83,6%; 92,2%)	(74,3%; 98,7%)	(96,3%; 82,5%)
EIC2204	(60,7%; 96,9%)	(83,6%; 82,7%)	(73,4%; 91,7%)	(86,9%; 80,6%)
EIC2205	(72,0%; 99,4%)	(83,6%; 94,4%)	(75,2%; 98,1%)	(96,3%; 86,4%)

Nota-se que o suporte é bastante melhor nas regras de associação geradas usando as disciplinas do primeiro ano como entrada e do segundo como saída, ao passo que a confiança é melhor nas regras geradas usando o primeiro semestre como entrada e o segundo como saída.

### 5.3. Terceiro ano

Vamos agora considerar o terceiro ano, começando por considerar as associações entre disciplinas do primeiro semestre e do segundo. Na tabela 5.10 temos o número de regras geradas pelo *apriori* e de obtidas após a aplicação do filtro.

Tabela 5.10 - Número de regras geradas para o terceiro ano

Disciplina	Regras geradas	Regras após filtro
EIC3200	56	14
EIC3201	56	14
EIC3202	56	12
EIC3203	56	24
EIC3204	56	8
EIC3205	56	10

Começando pela disciplina EIC3200, temos catorze regras de associação, que se podem ver na figura 5.13.

(C)	EIC3200 $\leq$ EIC3105 & EIC3102 (71,1%; 100%)
(C)	EIC3200 $\leq$ EIC3105 & EIC3103 & EIC3102 (71,1%; 100%)
(SAR)	EIC3200 $\leq$ EIC3105 (85,2%; 89,3%)
	EIC3200 $\leq$ EIC3105 & EIC3104 (71,8%; 99,0%)
	EIC3200 $\leq$ EIC3103 (83,1%; 90,7%)
	EIC3200 $\leq$ EIC3100 (83,1%; 90,7%)
	EIC3200 $\leq$ EIC3100 & EIC3101 (72,5%; 98,1%)
	EIC3200 $\leq$ EIC3103 & EIC3100 & EIC3101 (72,5%; 98,1%)
	EIC3200 $\leq$ EIC3105 & EIC3101 (72,5%; 98,1%)
	EIC3200 $\leq$ EIC3105 & EIC3100 (79,6%; 93,8%)
(I)	EIC3200 $\leq$ EIC3101 (73,9%; 97,1%)
	EIC3200 $\leq$ EIC3105 & EIC3103 (78,2%; 94,6%)
	EIC3200 $\leq$ EIC3103 & EIC3100 (78,2%; 94,6%)
	EIC3200 $\leq$ EIC3105 & EIC3103 & EIC3100 (76,1%; 96,3%)

Figura 5.13 - Regras de associação para a disciplina EIC3200



As duas primeiras têm uma confiança de 100%. A diferença entre as duas reside no facto de a segunda conter a disciplina EIC3103. Esta disciplina aparece isoladamente na premissa doutra regra de associação com maior suporte e menor confiança. Podemos assim considerar que a segunda será mais importante que a primeira.

Passando à disciplina EIC3201, vamos considerar catorze regras de associação, que constam da figura 5.14.

(C)	EIC3201 <= EIC3100 & EIC3102 & EIC3101 (68.3%; 100,0%)
(C)	EIC3201 <= EIC3103 & EIC3100 & EIC3102 & EIC3101 (68.3%; 100,0%)
(C)	EIC3201 <= EIC3105 & EIC3100 & EIC3102 & EIC3101 (68.3%; 100,0%)
(C)	EIC3201 <= EIC3105 & EIC3102 & EIC3101 (68.3%; 100,0%)
(C)	EIC3201 <= EIC3105 & EIC3103 & EIC3102 & EIC3101 (68.3%; 100,0%)
(S)	EIC3201 <= EIC3105 (85.2%; 87,6%)
	EIC3201 <= EIC3100 & EIC3104 (71.8%; 99,0%)
(AR)	EIC3201 <= EIC3100 (83.1%; 92,4%)
	EIC3201 <= EIC3100 & EIC3101 (72.5%; 98,1%)
	EIC3201 <= EIC3103 & EIC3100 & EIC3101 (72.5%; 98,1%)
	EIC3201 <= EIC3105 & EIC3100 (79.6%; 93,8%)
	EIC3201 <= EIC3103 & EIC3101 (73.2%; 97,1%)
	EIC3201 <= EIC3103 & EIC3100 (78.2%; 95,5%)
(I)	EIC3201 <= EIC3101 (73,9%; 96,2%)

*Figura 5.14 - Regras de associação para a disciplina EIC3201*

Pela primeira vez existem cinco regras de associação, as cinco primeiras, com 100% de confiança. De referir que as disciplinas EIC3101 e EIC3102 aparecem em todas estas cinco regras. A primeira é também a disciplina que aparece na premissa da regra com maior confiança de entre as que têm apenas uma disciplina na premissa. No entanto não é esta disciplina que aparece nos modelos do C5.0, mas sim a EIC3100, que também aparece em algumas das regras com maior confiança. Assim, parece não existir nenhuma disciplina determinante para a EIC3201.

Passando à disciplina EIC3202, na figura 5.15, onde vamos considerar doze regras de associação.

(C)	EIC3202 $\Leftarrow$ EIC3100 & EIC3104 & EIC3102 (67,6%; 97,9%)
(C)	EIC3202 $\Leftarrow$ EIC3103 & EIC3100 & EIC3104 & EIC3102 (67,6%; 97,9%)
(C)	EIC3202 $\Leftarrow$ EIC3105 & EIC3100 & EIC3104 & EIC3102 (67,6%; 97,9%)
(C)	EIC3202 $\Leftarrow$ EIC3105 & EIC3103 & EIC3104 & EIC3102 (67,6%; 97,9%)
(C)	EIC3202 $\Leftarrow$ EIC3105 & EIC3104 & EIC3102 (67,6%; 97,9%)
(S)	EIC3202 $\Leftarrow$ EIC3105 (85,2%; 84,3%)
	EIC3202 $\Leftarrow$ EIC3100 & EIC3104 (71,8%; 97,1%)
(AR)	EIC3202 $\Leftarrow$ EIC3103 (83,1%; 89,8%)
	EIC3202 $\Leftarrow$ EIC3103 & EIC3102 (73,2%; 96,2%)
	EIC3202 $\Leftarrow$ EIC3103 & EIC3100 (78,2%; 92,8%)
(I)	EIC3202 $\Leftarrow$ EIC3102 (73,9%; 95,2%)
	EIC3202 $\Leftarrow$ EIC3104 (74,6%; 93,4%)

**Figura 5.15 - Regras de associação para a disciplina EIC3202**

Comparando com a disciplina anterior, verificamos que as primeiras cinco regras (as com maior confiança) têm as premissas quase iguais. A única diferença é que, relativamente à EIC3201, nas regras da EIC3202 a disciplina EIC3101 é substituída pela EIC3104.

Relativamente à área científica, consideremos a disciplina EIC3102, da mesma área científica da EIC3202, Matemática. A disciplina EIC3102 aparece em todas as regras com maior confiança e ainda na regra com maior confiança entre as regras com apenas uma disciplina na premissa. Assim, contrariamente ao que acontecia nos modelos do C5.0, confirma-se a importância desta disciplina. De referir que as duas disciplinas são complementares (Teoria da Computação I e II).

Vejamos agora a disciplina EIC3203, com um total de vinte e quatro regras, na figura 5.16. De referir apenas que neste caso existem nove regras de associação com os mesmos suporte e confiança.

Passando à disciplina EIC3204, teremos a considerar oito regras de associação, conforme se pode ver na figura 5.17. As duas primeiras regras são as que têm maior confiança (100%). A diferença entre as duas primeiras regras reside na inclusão da disciplina EIC3103 na premissa da segunda. Apesar de não aparecer isoladamente na premissa de nenhuma regra, essa regra foi gerada inicialmente pelo *apriori*, tendo sido eliminada pelo filtro por ter confiança ligeiramente inferior à quinta regra, pelo que podemos considerar a regra que contém a disciplina EIC3103 como sendo a que fornece informação mais importante.

(C)	EIC3203 $\Leftarrow$ EIC3100 & EIC3104 & EIC3102 & EIC3101 (66,2%; 98,9%)
(C)	EIC3203 $\Leftarrow$ EIC3105 & EIC3104 & EIC3102 & EIC3101 (66,2%; 98,9%)
(S)	EIC3203 $\Leftarrow$ EIC3105 (85,2%; 84,3%)
	EIC3203 $\Leftarrow$ EIC3105 & EIC3103 & EIC3104 & EIC3101 (68,3%; 97,9%)
	EIC3203 $\Leftarrow$ EIC3105 & EIC3103 & EIC3102 & EIC3101 (68,3%; 97,9%)
	EIC3203 $\Leftarrow$ EIC3105 & EIC3102 & EIC3101 (68,3%; 97,9%)
	EIC3203 $\Leftarrow$ EIC3105 & EIC3100 & EIC3104 & EIC3101 (68,3%; 97,9%)
	EIC3203 $\Leftarrow$ EIC3105 & EIC3100 & EIC3102 & EIC3101 (68,3%; 97,9%)
	EIC3203 $\Leftarrow$ EIC3103 & EIC3100 & EIC3104 & EIC3101 (68,3%; 97,9%)
	EIC3203 $\Leftarrow$ EIC3103 & EIC3100 & EIC3102 & EIC3101 (68,3%; 97,9%)
	EIC3203 $\Leftarrow$ EIC3100 & EIC3104 & EIC3101 (68,3%; 97,9%)
	EIC3203 $\Leftarrow$ EIC3100 & EIC3102 & EIC3101 (68,3%; 97,9%)
	EIC3203 $\Leftarrow$ EIC3100 (83,1%; 87,3%)
	EIC3203 $\Leftarrow$ EIC3105 & EIC3100 & EIC3102 (70,4%; 97,0%)
	EIC3203 $\Leftarrow$ EIC3105 & EIC3103 & EIC3100 & EIC3102 (70,4%; 97,0%)
	EIC3203 $\Leftarrow$ EIC3105 & EIC3100 (79,6%; 89,4%)
	EIC3203 $\Leftarrow$ EIC3100 & EIC3104 (71,8%; 96,1%)
(AR)	EIC3203 $\Leftarrow$ EIC3103 & EIC3100 (78,2%; 91,0%)
	EIC3203 $\Leftarrow$ EIC3100 & EIC3101 (72,5%; 95,1%)
	EIC3203 $\Leftarrow$ EIC3103 & EIC3100 & EIC3101 (72,5%; 95,1%)
	EIC3203 $\Leftarrow$ EIC3105 & EIC3103 & EIC3100 (76,1%; 91,7%)
	EIC3203 $\Leftarrow$ EIC3103 & EIC3101 (73,2%; 94,2%)
	EIC3203 $\Leftarrow$ EIC3104 (74,6%; 92,5%)
(I)	EIC3203 $\Leftarrow$ EIC3101 (73,9%; 93,3%)

Figura 5.16 - Regras de associação para a disciplina EIC3203

(C)	EIC3204 $\Leftarrow$ EIC3100 & EIC3102 (71,1%; 100,0%)
(C)	EIC3204 $\Leftarrow$ EIC3103 & EIC3100 & EIC3102 (71,1%; 100,0%)
(S)	EIC3204 $\Leftarrow$ EIC3105 (85,2%; 91,7%)
	EIC3204 $\Leftarrow$ EIC3100 & EIC3104 (71,8%; 98,0%)
(AR1)	EIC3204 $\Leftarrow$ EIC3100 (83,1%; 93,2%)
	EIC3204 $\Leftarrow$ EIC3103 & EIC3102 (73,2%; 97,1%)
	EIC3204 $\Leftarrow$ EIC3105 & EIC3100 (79,6%; 94,7%)
	EIC3204 $\Leftarrow$ EIC3103 & EIC3100 (78,2%; 96,4%)

Figura 5.17 - Regras de associação para a disciplina EIC3204

Considerando a área científica, verifica-se que a disciplina EIC3104, da mesma área da EIC3204, aparece apenas numa das regras acompanhada pela

EIC3100. Tal como acontecia com os modelos gerados pelo C5.0, também aqui parece não haver uma grande interdependência entre as duas disciplinas.

Consideremos a disciplina EIC3205, na figura 5.18, com dez regras.

C)	$EIC3205 \leq EIC3105 \ \& \ EIC3102$ (71,1%; 100,0%)
(C)	$EIC3205 \leq EIC3105 \ \& \ EIC3103 \ \& \ EIC3102$ (71,1%; 100,0%)
(SAR)	$EIC3205 \leq EIC3105$ (85,2%; 91,7%)
	$EIC3205 \leq EIC3105 \ \& \ EIC3104$ (71,8%; 99,0%)
	$EIC3205 \leq EIC3105 \ \& \ EIC3100$ (79,6%; 94,7%)
	$EIC3205 \leq EIC3100 \ \& \ EIC3101$ (72,5%; 98,1%)
	$EIC3205 \leq EIC3103 \ \& \ EIC3100 \ \& \ EIC3101$ (72,5%; 98,1%)
	$EIC3205 \leq EIC3105 \ \& \ EIC3101$ (72,5%; 98,1%)
	$EIC3205 \leq EIC3105 \ \& \ EIC3103 \ \& \ EIC3100$ (76,1%; 96,3%)
(I)	$EIC3205 \leq EIC3101$ (73,9%; 97,1%)

*Figura 5.18 - Regras de associação para a disciplina EIC3205*

As duas primeiras regras são as que têm maior confiança. A diferença entre as duas primeiras regras reside no facto de a segunda incluir a disciplina EIC3103.

No entanto, esta disciplina não aparece destacada em nenhuma regra, aparecendo sempre acompanhada de duas outras disciplinas na premissa. Poderíamos ser levados a concluir que a primeira regra de associação seria a mais importante. No entanto é preciso não esquecer que, mais uma vez, esta regra foi suprimida no filtro mas foi originalmente gerada:

$$EIC3205 \leq EIC3103 \ (83,1\%; \ 89,8\%)$$

Assim, podemos concluir que a segunda regra será mais importante.

Tal como já foi feito nos anos anteriores vamos verificar, na tabela 5.11, as relações existentes entre as disciplinas.

A disciplina EIC3102 aparece nas regras de maior confiança em todas as disciplinas, seguido das EIC3103 e EIC3105, ambas com cinco disciplinas. Esta última disciplina aparece ainda em todas as regras de maior suporte. Relativamente às regras com apenas uma disciplina na premissa, a disciplina EIC3101 é a que aparece mais vezes, em quatro.

Tabela 5.11 - Relação entre as disciplinas do terceiro ano

	EIC3100	EIC3101	EIC3102	EIC3103	EIC3104	EIC3105
EIC3200	-	1	C	C	-	C S
EIC3201	C	C 1	C	C	-	C S
EIC3202	C	-	C 1	C	C	C S
EIC3203	C	C 1	C	-	C	C S
EIC3204	C 1	-	C	C	-	S
EIC3205	-	1	C	C	-	C S

Vamos agora estudar a associação entre as disciplinas do segundo ano e as do terceiro ano completo. Tal como anteriormente, vamos ter um número muito elevado de regras de associação. Na tabela 5.12 podemos ver o número de regras de associação geradas para cada disciplina e o número obtido após aplicação do filtro.

Tabela 5.12 - Número de regras de associação para o segundo ano

Disciplina	Regras geradas	Regras após filtro
EIC3100	793	28
EIC3101	787	29
EIC3102	787	65
EIC3103	793	25
EIC3104	788	51
EIC3105	793	16
EIC3200	793	38
EIC3201	793	61
EIC3202	789	45
EIC3203	783	97
EIC3204	793	32
EIC3205	793	31

Tal como fizemos para o segundo ano, vamos apenas apresentar as regras de associação mais relevantes, seguindo as mesmas referências, nas tabelas 5.13 e 5.14.

Tabela 5.13 - Regras de associação para o terceiro ano

(C)	EIC3100 $\leq$ EIC2201 & EIC2104 & EIC2102 & EIC2100 (74,6%; 99,1%)
(S)	EIC3100 $\leq$ EIC2205 (95,8%; 86,0%)
(AR)	EIC3100 $\leq$ EIC2201 & EIC2102 (83,1%; 95,8%)
(I)	EIC3100 $\leq$ EIC2102 (85,2%; 93,4%)
(C)	EIC3101 $\leq$ EIC2200 & EIC2202 & EIC2100 & EIC2204 (72,5%; 97,1%)
(S)	EIC3101 $\leq$ EIC2203 (92,3%; 80,2%)
(I)	EIC3101 $\leq$ EIC2102 (85,2%; 86,8%)
(C)	EIC3102 $\leq$ EIC2104 & EIC2200 & EIC2100 & EIC2204 (73,2%; 96,2%)
(S)	EIC3102 $\leq$ EIC2203 (92,3%; 80,2%)
(I)	EIC3102 $\leq$ EIC2100 (82,4%; 88,9%)
(C)	EIC3103 $\leq$ EIC2104 & <u>EIC2202</u> & EIC2100 (76,1%; 98,1%)
(S)	EIC3103 $\leq$ EIC2105 & EIC2104 & <u>EIC2202</u> & EIC2100 (76,1%; 98,1%)
(AR1)	EIC3103 $\leq$ EIC2100 (82,4%; 94,9%)
(C)	EIC3104 $\leq$ EIC2103 & EIC2102 & EIC2100 & <u>EIC2204</u> (72,5%; 96,1%)
(S)	EIC3104 $\leq$ EIC2203 (92,3%; 80,9%)
(AR)	EIC3104 $\leq$ EIC2103 & EIC2102 & <u>EIC2204</u> (76,1%; 94,4%)
(I)	EIC3104 $\leq$ <u>EIC2204</u> (82,4%; 88,9%)
(C)	EIC3105 $\leq$ <u>EIC2201</u> & EIC2103 & EIC2202 & EIC2100 (75,4%; 100,0%)
(S)	EIC3105 $\leq$ EIC2205 (95,8%; 88,2%)
(I)	EIC3105 $\leq$ EIC2102 (85,2%; 95,0%)
(C)	EIC3200 $\leq$ EIC2201 & EIC2102 & EIC2100 & EIC2204 (73,2%; 98,1%)
(S)	EIC3200 $\leq$ EIC2205 (95,8%; 82,4%)
(I)	EIC3200 $\leq$ EIC2201 (94,4%; 82,8%)
(I)	EIC3200 $\leq$ EIC2105 (94,4%; 82,8%)
(C)	EIC3201 $\leq$ EIC2201 & EIC2102 & EIC2100 & EIC2204 (73,2%; 99,0%)
(S)	EIC3201 $\leq$ EIC2205 (95,8%; 80,9%)
(I)	EIC3201 $\leq$ EIC2102 (85,2%; 89,3%)
(C)	EIC3202 $\leq$ EIC2201 & EIC2102 & EIC2100 & EIC2204 (73,2%; 95,2%)
(C)	EIC3202 $\leq$ EIC2201 & EIC2200 & EIC2100 & EIC2204 (73,2%; 95,2%)
(S)	EIC3202 $\leq$ EIC2205 & EIC2105 (93,0%; 80,3%)
(I)	EIC3202 $\leq$ EIC2100 (82,4%; 88,9%)
(C)	EIC3203 $\leq$ EIC2200 & EIC2202 & EIC2100 & <u>EIC2204</u> (72,5%; 97,1%)
(C)	EIC3203 $\leq$ EIC2201 & EIC2202 & EIC2100 & <u>EIC2204</u> (72,5%; 97,1%)
(S)	EIC3203 $\leq$ EIC2205 & EIC2201 (91,5%; 80,0%)
(AR)	EIC3203 $\leq$ EIC2102 & <u>EIC2204</u> (78,2%; 92,8%)
(I)	EIC3203 $\leq$ EIC2102 (85,2%; 86,0%)

Tabela 5.14 - Regras de associação para o terceiro ano

(C)	EIC3204 <= EIC2201 & EIC2104 & EIC2102 & EIC2204 (75,4%; 99,1%)
(C)	EIC3204 <= EIC2205 & EIC2201 & EIC2102 & EIC2100 (75,4%; 99,1%)
(S)	EIC3204 <= EIC2205 (95,8%; 86,8%)
(AR1)	EIC3204 <= EIC2204 (82,4%; 94,9%)
(C)	EIC3205 <= EIC2201 & EIC2102 & EIC2100 & EIC2204 (73,2%; 100,0%)
(S)	EIC3205 <= EIC2205 (95,8%; 82,4%)
(I)	EIC3205 <= EIC2102 (85,2%; 91,7%)

No segundo ano tínhamos verificado que as disciplinas da mesma área científica estavam sempre relacionadas na regra de maior confiança. No terceiro ano a tendência mantém-se, mas com uma exceção. Na área científica de Bases de Dados e Sistemas de Tratamento de Informação, as disciplinas EIC3200 e EIC2200 não estão relacionadas em nenhuma das regras apresentadas.

Vejamos agora na tabela 5.15 o resumo relativo às regras de associação mais relevantes.

Tabela 5.15 - Relação entre as disciplinas dos segundo e terceiro anos

	EIC2100	EIC2101	EIC2102	EIC2103	EIC2104	EIC2105	EIC2200	EIC2201	EIC2202	EIC2203	EIC2204	EIC2205
EIC3100	C	-	C 1	-	C	-	-	C	-	-	-	S
EIC3101	C	-	1	-	-	-	C	-	C	S	C	-
EIC3102	C 1	-	-	-	C	-	C	-	-	S	C	-
EIC3103	CS1	-	-	-	CS	S	-	-	CS	-	-	-
EIC3104	C	-	C	C	-	-	-	-	-	S	C 1	-
EIC3105	C	-	1	C	-	-	-	C	C	-	-	S
EIC3200	C	-	C	-	-	1	-	C 1	-	-	C	S
EIC3201	C	-	C 1	-	-	-	-	C	-	-	C	S
EIC3202	C 1	-	C	-	-	S	C	C	-	-	C	S
EIC3203	C	-	1	-	-	-	C	CS	C	-	C	S
EIC3204	C	-	C	-	C	-	-	C	-	-	C 1	CS
EIC3205	C	-	C 1	-	-	-	-	C	-	-	C	S

Como se pode ver, a disciplina EIC2100, aparece em todas as regras de melhor confiança. Relativamente ao suporte, a que aparece mais vezes é a EIC2205, num total de oito, e em relação às regras com apenas uma disciplina na premissa, a que aparece mais vezes é a EIC2102, num total de seis. Existe uma disciplina que não aparece em nenhuma das regras mais relevante, a EIC2101.

Vamos agora fazer uma comparação das disciplinas do segundo semestre, considerando as regras com maior suporte e confiança. Esta comparação encontra-se na tabela 5.16.

*Tabela 5.16 - Comparação entre as regras de associação*

Disciplina	Primeiro e segundo semestre		Ano completo	
	Melhor confiança	Melhor suporte	Melhor confiança	Melhor suporte
EIC3200	(71,1%; 100,0%)	(85,2%; 89,3%)	(73,2%; 98,1%)	(95,8%; 82,4%)
EIC3201	(68,3%; 100,0%)	(85,2%; 87,6%)	(73,2%; 99,0%)	(95,8%; 80,9%)
EIC3202	(67,6%; 97,9%)	(85,2%; 84,3%)	(73,2%; 95,2%)	(93,0%; 80,3%)
EIC3203	(66,2%; 98,9%)	(85,2%; 84,3%)	(72,5%; 97,1%)	(91,5%; 80,0%)
EIC3204	(71,1%; 100,0%)	(85,2%; 91,7%)	(75,4%; 99,1%)	(95,8%; 86,8%)
EIC3205	(71,1%; 100,0%)	(85,2%; 91,7%)	(73,2%; 100,0%)	(95,8%; 82,4%)

O que foi dito para o segundo ano também se aplica a este ano. A única nota de destaque é o facto de, para a disciplina EIC3205, a regra de associação com melhor confiança para o ano completo tem a mesma confiança e melhor suporte que a sua homóloga para o outro caso.

#### **5.4. Quarto ano**

Vamos começar a análise do quarto ano, usando as disciplinas do primeiro semestre como entrada e as do segundo como saída. Na tabela 5.17 encontram-se o número de regras geradas pelo *apriori* e após a aplicação do filtro.



Tabela 5.17 - Número de regras geradas para o terceiro ano

Disciplina	Regras geradas	Regras após filtro
EIC4200	55	26
EIC4201	56	3
EIC4202	56	5
EIC4203	56	16
EIC4204	0	0
EIC4205	56	14
EIC4206	0	0

Em primeiro lugar convém lembrar que existem três disciplinas opcionais: EIC4204, EIC4205 e EIC4206. Constatamos que duas destas disciplinas, a EIC4204 e a EIC4206, não originaram qualquer regra de associação.

Vamos olhar para as restantes, começando pela disciplina EIC4200. Neste caso não foram geradas as habituais cinquenta e seis, mas sim cinquenta e cinco. Após a aplicação do filtro, ficamos com vinte e seis regras de associação, que se podem ver na figura 5.19.

As três primeiras são as que têm maior confiança. Neste caso parece que não se pode considerar nenhuma das três regras iniciais como sendo a que tem melhor informação. As primeira e terceira regras obtêm-se da segunda adicionando à premissa uma disciplina: a EIC4100, na primeira, e a EIC4105, na terceira. Como estas disciplinas aparecem em várias regras seríamos levados a sugerir uma reunião das premissas (com cinco disciplinas, quando o máximo é quatro). Em caso de considerarmos um desempate, optaríamos pela primeira regra, uma vez que a disciplina EIC4100 aparece isoladamente na premissa de uma das regras com maior suporte.

Neste caso não existe qualquer relação com os modelos gerados pelo C5.0, uma vez que este classifica todos os casos como pertencendo à classe 1.

Vamos considerar, de seguida, a disciplina EIC4201, com apenas três regras de associação, como se pode constatar na figura 5.20.

(C)	EIC4200 $\Leftarrow$ EIC4100 & EIC4102 & EIC4103 & EIC4104 (66,2%; 89,8%)
(C)	EIC4200 $\Leftarrow$ EIC4102 & EIC4103 & EIC4104 (66,2%; 89,8%)
(C)	EIC4200 $\Leftarrow$ EIC4105 & EIC4102 & EIC4103 & EIC4104 (66,2%; 89,8%)
(S)	EIC4200 $\Leftarrow$ EIC4100 (87,8%; 83,1%)
(S)	EIC4200 $\Leftarrow$ EIC4102 (87,8%; 83,1%)
	EIC4200 $\Leftarrow$ EIC4105 & EIC4102 & EIC4104 (70,3%; 88,5%)
	EIC4200 $\Leftarrow$ EIC4105 & EIC4100 & EIC4102 & EIC4104 (70,3%; 88,5%)
	EIC4200 $\Leftarrow$ EIC4105 & EIC4102 (86,5%; 84,4%)
	EIC4200 $\Leftarrow$ EIC4105 & EIC4100 (86,5%; 84,4%)
	EIC4200 $\Leftarrow$ EIC4100 & EIC4102 (86,5%; 84,4%)
	EIC4200 $\Leftarrow$ EIC4100 & EIC4102 & EIC4103 & EIC4101 (78,4%; 87,9%)
	EIC4200 $\Leftarrow$ EIC4100 & EIC4103 & EIC4101 (78,4%; 87,9%)
	EIC4200 $\Leftarrow$ EIC4102 & EIC4103 & EIC4101 (78,4%; 87,9%)
	EIC4200 $\Leftarrow$ EIC4103 & EIC4101 (78,4%; 87,9%)
	EIC4200 $\Leftarrow$ EIC4105 & EIC4100 & EIC4103 & EIC4101 (78,4%; 87,9%)
	EIC4200 $\Leftarrow$ EIC4105 & EIC4102 & EIC4103 & EIC4101 (78,4%; 87,9%)
	EIC4200 $\Leftarrow$ EIC4105 & EIC4103 & EIC4101 (78,4%; 87,9%)
	EIC4200 $\Leftarrow$ EIC4105 & EIC4100 & EIC4102 (85,1%; 85,7%)
	EIC4200 $\Leftarrow$ EIC4105 & EIC4102 & EIC4101 (82,4%; 86,9%)
	EIC4200 $\Leftarrow$ EIC4105 & EIC4101 (82,4%; 86,9%)
	EIC4200 $\Leftarrow$ EIC4105 & EIC4100 & EIC4102 & EIC4101 (82,4%; 86,9%)
	EIC4200 $\Leftarrow$ EIC4105 & EIC4100 & EIC4101 (82,4%; 86,9%)
	EIC4200 $\Leftarrow$ EIC4102 & EIC4101 (82,4%; 86,9%)
(I)	EIC4200 $\Leftarrow$ EIC4101 (82,4%; 86,9%)
	EIC4200 $\Leftarrow$ EIC4100 & EIC4102 & EIC4101 (82,4%; 86,9%)
	EIC4200 $\Leftarrow$ EIC4100 & EIC4101 (82,4%; 86,9%)

**Figura 5.19 - Regras de associação para a disciplina EIC4200**

(C1)	EIC4201 $\Leftarrow$ EIC4100 (87,8%; 96,9%)
(C1)	EIC4201 $\Leftarrow$ EIC4102 (87,8%; 96,9%)
(S)	EIC4201 $\Leftarrow$ EIC4105 (93,2%; 94,2%)

**Figura 5.20 - Regras de associação para a disciplina EIC4201**

Neste caso só temos a considerar três regras de associação, todas elas com apenas uma disciplina na premissa, sendo as duas primeiras as que têm maior confiança e a última a que tem maior suporte. Não existe nenhuma relação com os

modelos gerados pelo C5.0, uma vez que este classifica todos os casos como sendo da classe 1.

Passamos à disciplina EIC4202, com cinco regras de associação, que se encontram na figura 5.21.

(C)	EIC4202 $\leq$ EIC4100 & EIC4102 & EIC4104 (71,6%; 98,1%)
(C)	EIC4202 $\leq$ EIC4102 & EIC4104 (71,6%; 98,1%)
(SAR)	EIC4202 $\leq$ EIC4105 (93,2%; 92,8%)
(1)	EIC4202 $\leq$ EIC4104 (75,7%; 96,4%)
	EIC4202 $\leq$ EIC4102 (87,8%; 95,4%)

*Figura 5.21 - Regras de associação para a disciplina EIC4202*

As duas primeiras regras são as que têm maior confiança. A diferença entre as duas primeiras reside na inclusão, na primeira regra, da disciplina EIC4100. Esta disciplina não aparece em nenhuma destas regras de associação, obtidas após a aplicação do filtro, mas aparece no conjunto inicial:

EIC4202  $\leq$  EIC4100 (87,8%, 93,8%)

Note-se que esta regra tem o mesmo suporte que a quinta, tendo sido eliminada pelo filtro por ter uma confiança ligeiramente mais baixa. Assim, podemos considerar que a primeira regra dá-nos informação mais importante que a segunda.

Relativamente à área científica, a disciplina EIC4104 pertence à mesma área científica da EIC4202, de Gestão e Métodos Quantitativos. Neste caso existe uma forte interdependência, uma vez que a EIC4104 aparece nas duas regras com maior confiança, aparecendo ainda isoladamente na regra com maior confiança entre as que têm apenas uma disciplina na premissa.

De referir ainda que a regra de maior suporte, a terceira, corresponde a um caminho na árvore de decisão e a uma regra de associação.

Consideremos agora a disciplina EIC4203, com dezasseis regras de associação, como se pode ver na figura 5.22.

(C)	EIC4203 <= EIC4100 & EIC4103 & EIC4104 (67,6%; 96,0%)
(C)	EIC4203 <= EIC4103 & EIC4104 (67,6%; 96,0%)
(C)	EIC4203 <= EIC4105 & EIC4100 & EIC4103 & EIC4104 (67,6%; 96,0%)
(C)	EIC4203 <= EIC4105 & EIC4103 & EIC4104 (67,6%; 96,0%)
(S)	EIC4203 <= EIC4105 (93,2%; 85,5%)
	EIC4203 <= EIC4105 & EIC4103 & EIC4101 (78,4%; 94,8%)
	EIC4203 <= EIC4105 & EIC4102 & EIC4103 & EIC4101 (78,4%; 94,8%)
	EIC4203 <= EIC4105 & EIC4100 & EIC4103 & EIC4101 (78,4%; 94,8%)
	EIC4203 <= EIC4103 & EIC4101 (78,4%; 94,8%)
	EIC4203 <= EIC4102 & EIC4103 & EIC4101 (78,4%; 94,8%)
	EIC4203 <= EIC4100 & EIC4103 & EIC4101 (78,4%; 94,8%)
	EIC4203 <= EIC4100 & EIC4102 & EIC4103 & EIC4101 (78,4%; 94,8%)
	EIC4203 <= EIC4102 (87,8%; 90,8%)
(AR)	EIC4203 <= EIC4100 (87,8%; 90,8%)
(I)	EIC4203 <= EIC4103 (83,8%; 93,5%)
	EIC4203 <= EIC4105 & EIC4103 (83,8%; 93,5%)

*Figura 5.22 - Regras de associação para a disciplina EIC4203*

Existem vários conjuntos de regras com o mesmo suporte e a mesma confiança. Para além dos quatro primeiros, estão na mesma situação todas as regras entre a sexta e a décima segunda, e os pares compostos pela décima terceira e décima quarta, e as duas últimas. Verifica-se facilmente que todas as disciplinas do primeiro semestre estão envolvidas nestas regras de associação, pelo que seria lógico concluir que todas elas são importantes para a disciplina EIC4203.

Relativamente à área científica, a disciplina EIC4103 pertence à mesma área científica da EIC4203, de Inteligência Artificial e Robótica. Neste caso existe uma forte interdependência, uma vez que a EIC4103 aparece nas quatro regras com maior confiança, aparecendo ainda isoladamente na regra com maior confiança entre as que têm apenas uma disciplina na premissa.

Como já foi referido, a disciplina EIC4204, não tem qualquer regra de associação. É uma situação que pode parecer estranha uma vez que existe uma árvore de decisão e uma regra de decisão a que se poderia associar a regra de associação:

EIC4204 <= EIC4103

No entanto, na própria árvore de decisão e na regra de decisão, a confiança que se tem na classificação é de apenas 66,7% o que colide com a confiança mínima das regras de associação (80%).

Vejamos a disciplina EIC4205, com catorze regras, na figura 5.23.

(C)	EIC4205 $\Leftarrow$ EIC4105 & EIC4100 & EIC4104 (71,6%; 96,2%)
(S)	EIC4205 $\Leftarrow$ EIC4105 (93,2%; 88,4%)
	EIC4205 $\Leftarrow$ EIC4105 & EIC4102 & EIC4101 (82,4%; 95,1%)
	EIC4205 $\Leftarrow$ EIC4105 & EIC4101 (82,4%; 95,1%)
	EIC4205 $\Leftarrow$ EIC4105 & EIC4100 & EIC4102 & EIC4101 (82,4%; 95,1%)
	EIC4205 $\Leftarrow$ EIC4105 & EIC4100 & EIC4101 (82,4%, 95,1%)
	EIC4205 $\Leftarrow$ EIC4102 & EIC4101 (82,4%, 95,1%)
(AR1)	EIC4205 $\Leftarrow$ EIC4101 (82,4%, 95,1%)
	EIC4205 $\Leftarrow$ EIC4100 & EIC4102 & EIC4101 (82,4%, 95,1%)
	EIC4205 $\Leftarrow$ EIC4100 & EIC4101 (82,4%, 95,1%)
	EIC4205 $\Leftarrow$ EIC4102 (87,8%; 92,3%)
	EIC4205 $\Leftarrow$ EIC4100 (87,8%; 92,3%)
	EIC4205 $\Leftarrow$ EIC4105 & EIC4100 (86,5%; 93,8%)
	EIC4205 $\Leftarrow$ EIC4105 & EIC4102 (86,5%; 93,8%)

**Figura 5.23 - Regras de associação para a disciplina EIC4205**

Existem também muitas regras com os mesmos suporte e confiança. Neste caso temos todas as disciplinas do primeiro semestre, com exceção da EIC4103, a intervir nas regras, pelo que se pode considerar que todas as disciplinas referidas são importantes para EIC4205.

A disciplina EIC4105, da mesma área científica da EIC4205, Bases de Dados e Sistemas de Tratamento de Informação, aparece em várias das regras geradas, em particular na regra de maior confiança e, isoladamente, na de maior suporte.

Uma última referência para a disciplina EIC4206, para a qual não foi gerada qualquer regra de associação. Aqui não há qualquer estranheza, uma vez que os modelos gerados pelo C5.0 classificam todos os casos como pertencendo à classe 0.

De referir que das disciplinas opcionais EIC4204, EIC4205 e EIC4206, a segunda foi a mais frequentada, o que pode explicar em grande parte o sucedido em relação à não geração de regras de associação para os outros dois casos.

Vejamos agora o resumo relativo às regras de associação, nos mesmos moldes dos anos anteriores, na tabela 5.18.

*Tabela 5.18 - Relação entre as disciplinas do quarto ano*

	EIC4100	EIC4101	EIC4102	EIC4103	EIC4104	EIC4105
EIC4200	C S	1	C S	C	C	C
EIC4201	C 1	-	C 1	-	-	S
EIC4202	-	C	C	-	C 1	S
EIC4203	C	-	-	C 1	C	C S
EIC4204	-	-	-	-	-	-
EIC4205	C	1	-	-	C	C S
EIC4206	-	-	-	-	-	-

Neste ano não existem disciplinas que apareçam em todas as regras de maior confiança ou maior suporte, mesmo não considerando as disciplinas EIC4204 e EIC4206, para as quais não foi gerada qualquer regras de associação.

Assim, as disciplinas EIC4100 e EIC4104 são as que aparecem mais vezes em regras com maior confiança, quatro vezes cada, enquanto a EIC4105 aparece o mesmo número de vezes nas com maior suporte. De referir ainda a EIC4101, que é a única a aparecer duas vezes nas regras com maior confiança entre as que só têm uma disciplina na premissa.

Vamos agora estudar as regras de associação geradas usando as disciplinas do terceiro ano como entrada e as do quarto ano completo como saída. Na tabela 5.19 podemos ver o número de regras de associação geradas para cada disciplina, assim como o número que se obtém após a aplicação do filtro.

Ao contrário do que temos feito, não vamos apresentar as regras mais relevantes, uma vez que existem bastantes com melhores suporte e confiança para cada disciplina. Vamos apenas considerar o resumo que temos feito em todos os anos anteriores, que se pode ver na tabela 5.20.



Neste caso existem três disciplinas que aparecem em todas as regras de maior confiança, com a exceção óbvia das duas disciplinas para as quais não foi gerada qualquer regra de associação. São elas a EIC3105, a EIC3200 e a EIC3205. De referir ainda que a primeira é ainda a única que aparece em todas as regras de maior suporte. Relativamente às regras com apenas uma disciplina na premissa, a EIC3102 é a que aparece mais vezes, num total de seis.

Vamos, de seguida, fazer uma comparação das disciplinas do segundo semestre, considerando as regras com maior suporte e confiança. Esta comparação encontra-se na tabela 5.21.

*Tabela 5.21 - Comparação entre as regras de associação*

Disciplina	Primeiro e segundo semestre		Ano completo	
	Melhor confiança	Melhor suporte	Melhor confiança	Melhor suporte
EIC4200	(66,2%; 89,8%)	(87,8%; 83,1%)	(87,8%; 86,2%)	(94,6%; 80,0%)
EIC4201	(87,8%; 96,9%)	(93,2%; 94,2%)	(91,9%; 97,1%)	(97,3%; 93,1%)
EIC4202	(71,6%; 98,1%)	(93,2%; 92,8%)	(85,1%; 98,4%)	(97,3%; 90,3%)
EIC4203	(67,6%; 96,0%)	(93,2%; 85,5%)	(86,5%; 92,2%)	(97,3%; 84,7%)
EIC4204	-	-	-	-
EIC4205	(71,6%; 96,2%)	(93,2%; 88,4%)	(87,8%; 92,3%)	(97,3%; 84,7%)
EIC4206	-	-	-	-

O que foi dito para os anos anteriores também se aplica a este ano, com a agravante de, em alguns casos, a confiança conseguir ser melhor no caso do ano completo, sendo o suporte bastante superior.

## 5.5. Quinto ano

Para o quinto ano vamos apenas estudar as regras de associação geradas usando as disciplinas do quarto ano como entrada e as do quinto ano completo como saída. Mais uma vez, vamos ter um número muito elevado de regras de associação, mesmo após a aplicação do filtro. Na tabela 5.22 podemos ver o número de regras de



associação geradas para cada disciplina, assim como o número que se obtém após a aplicação do filtro.

*Tabela 5.22 - Número de regras de associação para o quinto ano*

Disciplina	Regras geradas	Regras após filtro
EIC5100	895	83
EIC5101	858	115
EIC5102	895	83
EIC5103	794	78
EIC5104	21	20
EIC5105	0	0
EIC5106	1025	72
EIC5200	0	0
EIC5201	1025	123

Tal como aconteceu no quarto ano, não vamos poder apresentar as regras de associação mais relevantes. Vamos mostrar o resumo das relações entre as disciplinas dos quarto e quinto anos, na tabela 5.23.

*Tabela 5.23 - Relação entre as disciplinas dos quarto e quinto anos*

	EIC4100	EIC4101	EIC4102	EIC4103	EIC4104	EIC4105	EIC4200	EIC4201	EIC4202	EIC4203	EIC4204	EIC4205	EIC4206
EIC5100	C S	C I	C S	C S	-	C S	-	-	C	-	-	-	C
EIC5101	C S	C	C S	C S	C I	C S	-	C	C	C	-	C	C
EIC5102	C S	C I	C S	C S	-	C S	-	-	C	-	-	-	C
EIC5103	C S	C	C S	C S	C I	C S	-	-	C	-	-	-	C
EIC5104	S	C S	S	S	C S	S	C S	S	S	S	C S	S	-
EIC5105	-	-	-	-	-	-	-	-	-	-	-	-	-
EIC5106	C S	C	C S	C S	I	C S	-	-	C	-	-	-	C
EIC5200	-	-	-	-	-	-	-	-	-	-	-	-	-
EIC5201	C S	-	C S	C S	I	C S	-	C	C	C	C	C I	-

Como existem demasiadas regras com o mesmo suporte e a mesma confiança, mesmo entre as principais, é muito difícil tirar qualquer conclusão válida desta matriz. Um caso extremo é a disciplina EIC5104, onde as regras com maior suporte incluem todas as disciplinas do quarto ano excepto a EIC4206, que é opcional. Uma referência apenas para a EIC4200 que aparece apenas nas regras de maior confiança e suporte desta disciplina.

Vamos ainda mostrar, na tabela 5.24, os valores das regras com maior suporte e confiança.

*Tabela 5.24 - Melhores suporte e confiança das regras de associação*

Disciplina	Melhor confiança	Melhor suporte
EIC5100	(42,4%; 100,0%)	(100,0%; 87,9%)
EIC5101	(36,4%; 100,0%)	(100,0%; 84,8,%)
EIC5102	(42,4%; 100,0%)	(100,0%; 87,9,%)
EIC5103	(36,4%; 100,0%)	(100,0%; 81,8%)
EIC5104	(45,5%; 86,7%)	(48,5%; 81,2%)
EIC5105	-	-
EIC5106	(42,4%; 100,0%)	(100,0%; 90,9%)
EIC5200	-	-
EIC5201	(54,5%; 100,0%)	(100,0%; 93,9%)

Neste caso vemos que em todas as disciplinas, excepto a EIC5104 e as duas que não têm regras de associação, os melhores suportes e confiança são de 100%. Como já foi dito existem várias regras que têm simultaneamente os melhores suportes e confiança. As outras disciplinas são as opções menos frequentadas, pelo que pode vir daí a razão para os suporte e confiança serem mais baixos, num dos casos, e para a não geração de regras, nos outros dois casos.

## **5.6. Perfil de alunos**

Tal como se havia feito relativamente aos modelos do C5.0, também foram geradas regras de associação para a situação ao fim de três anos, para os anos lectivos

de 1994/95, 1995/96 e 1996/97, e ao fim cinco e três anos, para o ano lectivo de 1994/95, tendo em conta o perfil dos alunos, baseado nas notas do primeiro ano.

Assim, obtivemos cento e oitenta e três regras para o primeiro caso, oito após o filtro, cento e sessenta e nove para o segundo caso, seis após o filtro, e cento e sessenta e sete para o último, seis após o filtro. Estas regras estão na tabela 5.25.

**Tabela 5.25 - Regras de associação relativa às regras de associação**

Estudo	Regras de associação
Três anos	NDIS == N3 <= EIC1106 == 3 & EIC1204 == 0 (18,2%; 100,0%)
	NDIS == N3 <= EIC1204 == 0 (18,2%; 100,0%)
	NDIS == C3 <= EIC1106 == 3 & EIC1202 == 3 (25,8%; 80,5%)
	NDIS == C3 <= EIC1202 == 3 (25,8%; 80,5%)
	NDIS == N3 <= EIC1106 == 3 & EIC1100 == 0 (18,9%; 96,7%)
	NDIS == N3 <= EIC1100 == 0 (18,9%; 96,7%)
	NDIS == N3 <= EIC1200 == 0 (21,4%; 91,2%)
	NDIS == N3 <= EIC1106 == 3 & EIC1200 == 0 (21,4%; 91,2%)
Cinco anos	NDIS5 == C <= EIC1103 == 3 (21,2%; 100,0%)
	NDIS5 == C <= EIC1106 == 3 & EIC1103 == 3 (21,2%; 100,0%)
	NDIS5 == C <= EIC1106 == 3 & EIC1202 == 3 (28,8%; 80,0%)
	NDIS5 == C <= EIC1202 == 3 (28,8%; 80,0%)
	NDIS5 == C <= EIC1106 == 3 & EIC1102 == 3 & EIC1202 == 3 (25,0%; 84,6%)
	NDIS5 == C <= EIC1102 == 3 & EIC1202 == 3 (25,0%; 84,6%)
Três anos (1994/95)	NDIS3 == C3 <= EIC1103 == 3 (21,2%; 100,0%)
	NDIS3 == C3 <= EIC1106 == 3 & EIC1103 == 3 (21,2%; 100,0%)
	NDIS3 == C3 <= EIC1106 == 3 & EIC1202 == 3 (28,8%; 80,0%)
	NDIS3 == C3 <= EIC1202 == 3 (28,8%; 80,0%)
	NDIS3 == C3 <= EIC1106 == 3 & EIC1102 == 3 & EIC1202 == 3 (25,0%; 84,6%)
	NDIS3 == C3 <= EIC1102 == 3 & EIC1202 == 3 (25,0%; 84,6%)

Contrariamente ao que aconteceu nos modelos do C5.0, nas regras de associação obtemos, para os alunos que entraram no ano lectivo de 1994/95, após a aplicação do filtro, as regras com as mesmas premissas e valores de suporte e confiança, para o resultado ao fim de três e cinco anos. Assim, entre os alunos de 1994/95, as disciplinas que foram importantes para a aprovação a todas as disciplinas até ao terceiro ano continuaram a sê-lo para a conclusão da licenciatura em cinco anos.

No entanto, quando pegamos nos alunos que entraram em 1995/96 e 1996/97, os resultados são diferentes ao fim de três anos. A única disciplina que continua a aparecer como sendo importante é a EIC1202 (Introdução à Programação II). Este facto parece indicar que os alunos de 1994/95, o primeiro ano da licenciatura, tiveram um comportamento algo diferente, possivelmente motivado pela falta de conhecimento e de aconselhamento dos colegas mais avançados.

Note-se que, em todos os casos, temos pares de regras com os mesmos suporte e confiança, com a diferença do aparecimento da disciplina EIC1106

(Técnicas de Comunicação e Inglês Funcional) com o resultado 3 (Muito Bom). As regras que não contêm a condição “EIC1106 = = 3” serão as mais importantes, uma vez que a disciplina EIC1106 não aparece em nenhuma regra de associação isoladamente. Este facto pode-se dever à situação de a disciplina EIC1106 ser uma disciplina com uma percentagem de aprovações próxima dos 100%.

Assim, confirmam-se os resultados dos modelos do C5.0, com as disciplinas EIC1202 (Introdução à Programação II), EIC1200 (Análise Matemática II) e EIC1204 (Arquitetura de Computadores) a destacarem-se, assim como algumas das disciplinas das quais são complementares, como as EIC1102 (Introdução à Programação I) e EIC1100 (Análise Matemática II).

### **5.7. Discussão de resultados**

Tal como já aconteceu com as árvores e regras de decisão, vamos agora debruçar-nos sobre a informação que foi possível obter. O estudo de regras de associação também se baseou em três tipos diferentes de casos:

- análise de ligações das disciplinas do segundo semestre de um determinado ano, usando as disciplinas do primeiro semestre do mesmo ano, do primeiro ao quarto ano;
- análise de ligações das disciplinas de um determinado ano completo, usando as disciplinas do ano anterior, do segundo ao quinto ano;
- análise de ligações da situação de um aluno ao fim de três e cinco anos, usando o perfil baseado nas notas do primeiro ano.

Considerando os dois primeiro casos, e comparando os segundo, terceiro e quarto anos, tal como fizemos para os modelos do C5.0, colocámos a questão: alguma das formas supera inequivocamente a outra? A resposta que demos para os modelos do C5.0 mantém-se para as regras de associação. Existem disciplinas que parecem estar mais ligadas a disciplinas do mesmo ano, ao passo que outras estarão mais ligadas a disciplinas de anos diferentes.

De notar que a excepção dos modelos do C5.0, no terceiro ano, onde não foi determinada nenhuma interdependência entre as disciplinas Teoria da Computação I e II, acabou por não se manter aqui.

É preciso não esquecer que as regras de associação fornecem maior informação que as árvores e regras de decisão, uma vez que nos dão todas as associações possíveis, com determinado suporte e confiança, ao passo que os modelos do C5.0, em cada passo, dividem os dados tendo em conta os atributos onde se obtém maior ganho de informação ou o maior rácio de ganho de informação. Quando existem duas disciplinas muito próximas, só uma é escolhida, podendo ficar a outra esquecida no modelo.

Tendo em conta as regras de associação de maiores confiança e suporte para cada um dos anos, verifica-se que para o ano completo se obtém, geralmente, melhor suporte, ao passo que para o segundo semestre a confiança é melhor. No entanto, em alguns casos, a confiança para o ano completo consegue ser melhor.

Relativamente às áreas científicas, já existem mais algumas associações entre disciplinas da mesma área, mas esta continua a não ser determinante, uma vez que continuam a existir disciplinas de outras áreas que continuam a ser mais importantes, em alguns casos.

Vejamos agora a comparação entre estas duas análises e a última, relativa ao perfil dos alunos, baseado nas notas do primeiro ano. Tal como já aconteceu relativamente aos modelos do C5.0, existe alguma coerência entre as duas análises. Assim, são as mesmas disciplinas do primeiro ano que são determinantes para um bom desempenho na licenciatura.

# Capítulo 6

## Conclusões

Os objectivos da presente dissertação eram o estudo do processo de ECBD e a avaliação da adequação do emprego desta perante situações concretas, utilizando como estudo de caso a base de dados académica da FEUP. Apresentam-se neste capítulo as conclusões finais.

Assim, na próxima secção faz-se uma breve discussão sobre o trabalho realizado. Na secção 6.2 é feita uma comparação entre as árvores e regras de decisão e as regras de associação. Na secção 6.3 são discutidas algumas alternativas possíveis. Finalmente, na secção 6.4, são apresentadas algumas sugestões de trabalho futuro.

### **6.1. *Discussão do trabalho realizado***

A parte inicial do trabalho (pesquisa bibliográfica) foi a mais demorada, uma vez que, sendo a área de ECBD bastante recente, existem muitos livros que explicam o conceito de formas diferentes. Para além de muitas vezes se confundirem os conceitos de ECBD e de prospecção de dados, também se opta muitas vezes por explicá-los de forma mais dirigida a determinados segmentos (empresas, investigação, etc...).

A escolha da forma de abordar o tema que seria mais indicada para o trabalho em causa, apesar de ser um processo demorado, veio a revelar-se muito importante para o desenvolvimento do trabalho, especialmente na distinção entre o ponto de vista da prospecção de dados e da estatística.

Relativamente à preparação dos dados, foi a segunda fase mais demorada. Os dados encontravam-se com alguns erros de introdução e alguns campos

encontravam-se por preencher. A falta de experiência no uso da ferramenta Clementine, que também foi usada nesta fase, revelou-se um obstáculo que demorou a ultrapassar.

Tirando estas duas fases, o trabalho decorreu na maior normalidade. Os testes decorreram sem problemas, tirando o facto de o uso de bases de dados com muitos registos ter tido algum peso significativo em termos de desempenho.

No entanto, devido ao vasto leque de resultados que se podem obter, fica sempre a ideia que existe algo mais por descobrir. Sempre que uma hipótese é confirmada, ou um novo resultado é descoberto, novas hipóteses são geradas e novos resultados podem aparecer. Este ciclo contínuo de descoberta acaba por ser o grande motor da ECBD.

O filtro para as regras de associação, desenvolvido durante a presente dissertação, acabou por se revelar bastante útil, tendo permitido uma análise daquelas de uma forma mais eficiente, o que seria difícil de conseguir se tivéssemos que considerar o número elevado de regras geradas.

As soluções encontradas durante este trabalho acabaram por ir de encontro às pretensões iniciais. O uso de ECBD, aplicado a este caso, veio a revelar-se de grande utilidade, obtendo-se conclusões importantes como, por exemplo, sobre a relevância da área científica na interdependência entre as disciplinas e a detecção precoce de casos de insucesso escolar baseado no perfil dos alunos. Os resultados obtidos encorajam fortemente o uso de ECBD em bases de dados académicas.

## **6.2. Comparação entre os modelos gerados**

Os modelos gerados pelo C5.0 e pelo *apriori* são diferentes não apenas quanto à forma mas também quanto ao fim a que se destinam: classificação e análise de ligações.

Assim, os modelos gerados pelo C5.0, as árvores e regras de decisão, permitem de uma forma rápida classificar um determinado tuplo tendo em conta os valores dos seus atributos. Por outro lado, as regras de associação geradas pelo *apriori* permitem encontrar relações importantes entre as diversas disciplinas.

As árvores e regras de decisão resumem os dados de forma simples e legível, permitindo ao mesmo tempo classificar casos para os quais não se conhece a classe a

que pertencem. A aplicação dos modelos a novos dados, ou a um subconjunto dos dados seleccionado previamente para teste, permite-nos aferir sobre a sua real utilidade.

As regras de associação permitem-nos verificar as ligações existentes entre as disciplinas. O que pode acontecer é que o número de regras geradas pode ser demasiado elevado, o que torna a informação muitas vezes difícil de interpretar. Esta informação pode ser muitas vezes redundante. O filtro, desenvolvido durante a presente dissertação, permite reduzir em grande parte o número elevado de regras geradas.

Como já vimos, as regras de associação fornecem mais informação que as árvores e regras de decisão, uma vez que nos dão todas as associações possíveis, ao passo que os modelos do C5.0, em cada passo, dividem os dados tendo em conta os atributos onde se obtém maior ganho de informação ou o maior rácio de ganho de informação. Quando existem duas disciplinas muito próximas, só uma é escolhida, podendo ficar a outra esquecida no modelo.

As regras de associação têm ainda a particularidade de se basearem em dois factores, o suporte e a confiança, que num grande número de casos se deslocam em sentidos opostos. Enquanto nos modelos do C5.0, existe um caminho sequencial a seguir, nas regras de associação é necessário fazer uma análise tendo sempre em conta estes dois factores.

Na análise realizada pôde-se verificar que, em alguns casos, os modelos gerados para os mesmos anos mostravam conclusões diferentes relativamente às disciplinas que maior influência tinham no sucesso de outras. No entanto, as regras de associação correspondentes eram sempre geradas, embora num patamar de suporte e confiança relativamente inferiores às mais importantes.

Os testes realizados demonstram que a confiança parece ter um papel mais importante que o suporte, mas não podemos simplesmente esquecer este último. Uma regra pode ter uma confiança melhor mas, se tiver um suporte demasiado baixo, poderá não transmitir conhecimento útil.

Verificou-se que, que a confiança era melhor quando considerávamos um menor número de disciplinas, na análise do segundo semestre, ao passo que o suporte melhorava com o aumento do número de disciplinas. No entanto, quando o número de registos diminui, o que acontece no quarto ano, o maior número de disciplinas faz



com que a confiança aumente e o suporte se mantenha elevado, conseguindo, em muitos casos, ter ambos os valores comparativamente melhores.

De referir que, a partir das regras de associação, se pode construir um modelo de classificação, ordenando as regras para uma determinada disciplina por um determinado critério, em função do suporte e da confiança, e aplicar a primeira em que sejam satisfeitas todas as condições da premissa. Por omissão, poder-se-ia considerar a classe com maior percentagem de casos.

Tendo em conta os resultados obtidos nos modelos gerados, consegue-se verificar que existe concordância dos dois modelos nas conclusões principais:

- as áreas científicas, tirando algumas excepções, parecem não ser determinantes na aprovação das disciplinas;
- evidencia-se uma maior motivação de um aluno que se inscreve no primeiro ano de uma Licenciatura em Engenharia Informática e Computação para as disciplinas relacionadas com informática e computadores;
- o perfil dos alunos ao fim do primeiro ano do curso está directamente relacionado com o rendimento, principalmente nos alunos com os perfis “Bom” e “Muito Bom”;
- as disciplinas Análise Matemática I e II, da área científica de Matemática, são muito importantes para um bom desempenho ao fim dos três primeiros anos, podendo este facto estar relacionado com a existência de um total de onze disciplinas das áreas científicas de Matemática e Modelação Física, uma vez que as referidas disciplinas fornecem conhecimento de base para estas áreas científicas.

### **6.3. Discussão de alternativas**

Existem inúmeras alternativas que se podem considerar na ECBD de âmbito académico. A riqueza que se esconde por baixo dos dados pode ser explorada de múltiplas formas. Existem abordagens que diferem não só em relação aos tipos de algoritmos mas também relativamente à forma de tratar os dados.

Uma alternativa poderia ser o uso de outros algoritmos de prospecção de dados. Um exemplo, que chegou a ser equacionado, seria tentar encontrar

agrupamentos (*clusters*) de conjuntos de disciplinas que tenham determinadas características em comum. Por exemplo, algumas disciplinas podem ter uma grande percentagem de aprovações à primeira tentativa, outras só em recurso.

Outros agrupamentos podem também ser feitos tendo por base as médias das notas nas disciplinas, o que poderia permitir detectar o grau de exigência de algumas disciplinas relativamente a outras.

Existem também alternativas em termos de organização dos dados. O facto de se considerarem, de cada vez, apenas as disciplinas de cada ano lectivo, apesar de facilitar a compreensão dos modelos gerados, pode tornar-se limitativo. Por exemplo, podem existir disciplinas do quinto ano dependentes de outras do primeiro.

#### **6.4. Sugestões para trabalho futuro**

Existem algumas sugestões que poderão melhorar, em alguns aspectos, a pesquisa efectuada. Estes melhoramentos poder-se-ão verificar quer a nível dos dados quer a nível da ECBD.

Por exemplo, as bases de dados poderiam conter informação sobre as equivalências de disciplinas, tanto a nível de mudanças de curso como em mudanças nos planos de estudo. Esta informação permitiria estender o trabalho realizado, por exemplo, a cursos que tiveram alterações recentes no respectivo plano.

Uma sugestão a um nível mais global seria a integração da informação encontrada no processo de ECBD num sistema de apoio à decisão, que permitisse fazer previsões e correcções importantes para os anos seguintes.

# Referências

- [Agrawal et al., 1993] R. Agrawal, T. Imielinski, A. Swami. *Mining Association Rules between Sets of Items in Large Databases*. ACM SIGMOD Conference on Management of Data, 207-216. Washington, D. C. 1993.
- [Agrawal & Srikant, 1994] R. Agrawal, R. Srikant. *Fast Algorithms for Mining Association Rules*. VLDB Conference. Santiago, Chile. 1994.
- [Berry & Linoff, 1997] M. J. A. Berry, G. Linoff. *Data Mining Techniques – For Marketing, Sales and Customer Support*. John Wiley & Sons Ltd, 1997.
- [Breinman et al., 1984] L. Breinman, J. H. Friedman, R. A. Olshen, C. J. Stone. *Classification and Regression Trees*. Wadsworth, Belmont, California, 1984.
- [Cabena et al., 1998] P. Cabena, P. Hadjinian, R. Stadler, J. Verhees, A. Zanasi. *Discovering Data Mining – From Concept to Implementation*. Prentice Hall PTR. 1998.
- [Clementine, 1997] SPSS/ISL Clementine 5.0.4 User Guide. 1997.
- [Fayyad et al., 1996] U. Fayyad, G. Piatetsky-Shapiro, P. Smyth, R. Uthurusamy. *Advances in Knowledge Discovering and Data Mining*. MIT Press, Cambridge, MA, 1996.
- [Fayyad et al., 1996b] U. Fayyad, G. Piatetsky-Shapiro, P. Smyth. *The KDD Process for Extracting Useful Knowledge from Volumes of Data*. Communications of the ACM, 39 (11): 27-34, 1996.

- [Glymour et al., 1996] C. Glymour, D. Madigan, D. Pregibon, P. Smyth. *Statistical Inference and Data Mining*. Communications of the ACM, 39 (11): 35-41, 1996.
- [Michalski et al., 1998] R. S. Michalski, I. Bratko, M. Kubat. *Machine Learning and Data Mining – Methods and Applications*. John Wiley & Sons Ltd, 1998.
- [Mitchell, 1997] T. M. Mitchell. *Machine Learning*. McGraw-Hill International Editions, 1997.
- [Pyle, 1999] D. Pyle. *Data Preparation for Data Mining*. Morgan Kaufmann Publishers, Inc. 1999.
- [Quinlan, 1986] J. R. Quinlan. *Induction of Decision Trees*. Machine Learning 1, 81-106, 1986.
- [Quinlan, 1992] J. R. Quinlan. *C4.5: Programs for Machine Learning*. Morgan Kaufmann, San Francisco, 1992.
- [Westphal & Blaxton, 1998] C. Westphal, T. Blaxton. *Data Mining Solutions – Methods and Tools for Solving Real-World Problems*. John Wiley & Sons Ltd, 1998.

# **ANEXOS**

# **Anexo I**

## **Árvores e Regras de Decisão**

Tabela I.1 - Árvores e regras de decisão para o primeiro ano

Disciplina	Árvore de decisão	Regras de decisão
EIC1200	EIC1100 1 (110.0, 0.982) -> 1 EIC1100 0 EIC1101 0 (20.0, 0.95) -> 0 EIC1101 1 EIC1103 1 (7.0, 0.714) -> 0 EIC1103 0 (2.0, 1.0) -> 1	if EIC1100 == 1 then -> 1 (110, 0.973)  if EIC1100 == 0 then -> 0 (29, 0.806)  Default : -> 1
EIC1202	EIC1102 1 (115.0, 0.974) -> 1 EIC1102 0 EIC1101 0 (15.0, 0.867) -> 0 EIC1101 1 EIC1104 1 (6.0, 0.833) -> 1 EIC1104 0 (3.0, 1.0) -> 0	if EIC1101 == 1 and EIC1104 == 1 then -> 1 (112, 0.982)  if EIC1102 == 1 then -> 1 (115, 0.966)  if EIC1102 == 0 and EIC1104 == 0 then -> 0 (15, 0.824)  if EIC1101 == 0 and EIC1102 == 0 then -> 0 (15, 0.824)  Default : -> 1
EIC1203	EIC1103 1 EIC1101 1 (110.0, 0.936) -> 1 EIC1101 0 EIC1102 1 (3.0, 1.0) -> 0 EIC1102 0 (3.0, 0.667) -> 1 EIC1103 0 EIC1104 0 (15.0, 1.0) -> 0 EIC1104 1 EIC1102 1 (4.0, 1.0) -> 0 EIC1102 0 (4.0, 0.75) -> 1	if EIC1103 == 1 then -> 1 (116, 0.898)  if EIC1104 == 1 then -> 1 (118, 0.875)  if EIC1103 == 0 and EIC1104 == 0 then -> 0 (15, 0.941)  if EIC1102 == 1 and EIC1103 == 0 then -> 0 (7, 0.889)  Default : -> 1
EIC1204	EIC1102 1 EIC1104 1 (109.0, 0.972) -> 1 EIC1104 0 (6.0, 0.667) -> 0 EIC1102 0 EIC1100 1 (7.0, 0.857) -> 1 EIC1100 0 (17.0, 0.765) -> 0	if EIC1102 == 1 and EIC1104 == 1 then -> 1 (109, 0.964)  if EIC1100 == 1 then -> 1 (110, 0.955)  if EIC1100 == 0 and EIC1102 == 0 then -> 0 (17, 0.737)  if EIC1104 == 0 then -> 0 (21, 0.652)  Default : -> 1
EIC1205	EIC1102 1 (115.0, 0.965) -> 1 EIC1102 0 EIC1100 1 (7.0, 0.857) -> 1 EIC1100 0 (17.0, 0.706) -> 0	if EIC1100 == 1 then -> 1 (110, 0.964)  if EIC1102 == 1 then -> 1 (115, 0.957)  if EIC1100 == 0 and EIC1102 == 0 then -> 0 (17, 0.684)  Default : -> 1
EIC1206	EIC1101 1 (118.0, 0.915) -> 1 EIC1101 0 EIC1102 0 (15.0, 0.8) -> 0 EIC1102 1 EIC1103 1 (3.0, 1.0) -> 1 EIC1103 0 (3.0, 0.667) -> 0	if EIC1102 == 1 and EIC1103 == 1 then -> 1 (108, 0.927)  if EIC1101 == 1 then -> 1 (118, 0.908)  if EIC1101 == 0 and EIC1103 == 0 then -> 0 (15, 0.765)  if EIC1101 == 0 and EIC1102 == 0 then -> 0 (15, 0.765)  Default : -> 1

Tabela I.2 - Árvores e regras de decisão para o segundo ano (semestral)

Disciplina	Árvore de decisão	Regras de decisão
EIC2200	EIC2105 1 (86.0, 0.895) -> 1 EIC2105 0 EIC2103 1 (3.0, 0.667) -> 1 EIC2103 0 (18.0, 1.0) -> 0	if EIC2103 == 1 then -> 1 (79, 0.901)  if EIC2105 == 1 then -> 1 (86, 0.886)  if EIC2103 == 0 and EIC2105 == 0 then -> 0 (18, 0.95)  Default : -> 1
EIC2201	EIC2101 1 (75.0, 0.973) -> 1 EIC2101 0 EIC2104 1 (15.0, 0.733) -> 1 EIC2104 0 (17.0, 0.882) -> 0	if EIC2101 == 1 then -> 1 (75, 0.961)  if EIC2104 == 1 then -> 1 (85, 0.92)  if EIC2101 == 0 and EIC2104 == 0 then -> 0 (17, 0.842)  Default : -> 1
EIC2202	EIC2102 1 (76.0, 0.934) -> 1 EIC2102 0 (31.0, 0.935) -> 0	if EIC2102 == 1 then -> 1 (76, 0.923)  if EIC2102 == 0 then -> 0 (31, 0.909)  Default : -> 1
EIC2203	EIC2103 1 (79.0, 0.975) -> 1 EIC2103 0 EIC2100 1 (2.0, 1.0) -> 1 EIC2100 0 (26.0, 0.923) -> 0	if EIC2100 == 1 then -> 1 (70, 0.986)  if EIC2103 == 1 then -> 1 (79, 0.963)  if EIC2100 == 0 and EIC2103 == 0 then -> 0 (26, 0.893)  Default : -> 1
EIC2204	EIC2102 0 (31.0, 0.871) -> 0 EIC2102 1 EIC2105 1 (74.0, 0.905) -> 1 EIC2105 0 (2.0, 1.0) -> 0	if EIC2102 == 1 and EIC2105 == 1 then -> 1 (74, 0.895)  if EIC2105 == 0 then -> 0 (21, 0.87)  if EIC2102 == 0 then -> 0 (31, 0.848)  Default : -> 1
EIC2205	EIC2103 1 (79.0, 1.0) -> 1 EIC2103 0 EIC2104 1 (10.0, 0.8) -> 1 EIC2104 0 (18.0, 0.833) -> 0	if EIC2103 == 1 then -> 1 (79, 0.988)  if EIC2104 == 1 then -> 1 (85, 0.966)  if EIC2103 == 0 and EIC2104 == 0 then -> 0 (18, 0.8)  Default : -> 1



Tabela I.3- Árvores e regras de decisão para o segundo ano (anual)

Disciplina	Árvore de decisão	Regras de decisão
EIC2100	EIC1204 0 (21.0, 0.952) -> 0 EIC1204 1 EIC1203 1 (78.0, 0.872) -> 1 EIC1203 0 (8.0, 0.875) -> 0	if EIC1203 == 1 and EIC1204 == 1 then -> 1 (78, 0.862)  if EIC1204 == 0 then -> 0 (21, 0.913)  if EIC1203 == 0 then -> 0 (20, 0.909)  Default : -> 1
EIC2101	EIC1200 0 (25.0, 0.96) -> 0 EIC1200 1 EIC1104 1 (79.0, 0.937) -> 1 EIC1104 0 (3.0, 1.0) -> 0	if EIC1104 == 1 and EIC1200 == 1 then -> 1 (79, 0.926)  if EIC1200 == 0 then -> 0 (25, 0.926)  if EIC1104 == 0 then -> 0 (11, 0.923)  Default : -> 1
EIC2102	EIC1203 0 (20.0, 0.95) -> 0 EIC1203 1 EIC1102 0 (5.0, 1.0) -> 0 EIC1102 1 EIC1204 1 (77.0, 0.948) -> 1 EIC1204 0 (5.0, 0.6) -> 0	if EIC1102 == 1 and EIC1203 == 1 and EIC1204 == 1 then -> 1 (77, 0.937)  if EIC1102 == 0 then -> 0 (13, 0.933)  if EIC1203 == 0 then -> 0 (20, 0.909)  if EIC1204 == 0 then -> 0 (21, 0.87)  Default : -> 1
EIC2103	EIC1203 0 (20.0, 0.95) -> 0 EIC1203 1 EIC1104 0 (3.0, 1.0) -> 0 EIC1104 1 EIC1102 1 (81.0, 0.951) -> 1 EIC1102 0 (3.0, 0.667) -> 0	if EIC1102 == 1 and EIC1104 == 1 and EIC1203 == 1 then -> 1 (81, 0.94)  if EIC1104 == 0 then -> 0 (11, 0.923)  if EIC1203 == 0 then -> 0 (20, 0.909)  if EIC1102 == 0 then -> 0 (13, 0.867)  Default : -> 1
EIC2104	EIC1204 0 (21.0, 0.714) -> 0 EIC1204 1 EIC1104 1 (84.0, 0.94) -> 1 EIC1104 0 (2.0, 1.0) -> 0	if EIC1104 == 1 and EIC1204 == 1 then -> 1 (84, 0.93)  if EIC1104 == 0 then -> 0 (11, 0.769)  if EIC1204 == 0 then -> 0 (21, 0.696)  Default : -> 1

Tabela I.4– Árvores e regras de decisão para o segundo ano (anual)

Disciplina	Árvore de decisão	Regras de decisão
EIC2105	<pre> EIC1104 0   EIC1100 0 (7.0, 1.0) -&gt; 0   EIC1100 1     EIC1103 1 (2.0, 1.0) -&gt; 0     EIC1103 0 (2.0, 1.0) -&gt; 1 EIC1104 1   EIC1103 1 (88.0, 0.932) -&gt; 1   EIC1103 0     EIC1101 1 (3.0, 1.0) -&gt; 0     EIC1101 0       EIC1204 1 (2.0, 1.0) -&gt; 1       EIC1204 0 (3.0, 1.0) -&gt; 0                     </pre>	<pre> if EIC1103 == 1 and EIC1104 == 1 then -&gt; 1 (88, 0.922)  if EIC1104 == 1 and EIC1204 == 1 then -&gt; 1 (84, 0.919)  if EIC1100 == 1 and EIC1103 == 0 and EIC1104 == 0 then -&gt; 1 (2, 0.75)  if EIC1103 == 1 and EIC1104 == 0 then -&gt; 0 (4, 0.833)  if EIC1101 == 1 and EIC1103 == 0 and EIC1104 == 1 then -&gt; 0 (3, 0.8)  if EIC1103 == 0 then -&gt; 0 (15, 0.706)  Default : -&gt; 1                     </pre>
EIC2200	<pre> EIC1100 1   EIC1204 1 (79.0, 0.911) -&gt; 1   EIC1204 0     EIC1200 1 (4.0, 0.75) -&gt; 0     EIC1200 0 (3.0, 0.667) -&gt; 1 EIC1100 0   EIC1103 0 (11.0, 1.0) -&gt; 0   EIC1103 1     EIC1202 1 (4.0, 0.75) -&gt; 1     EIC1202 0 (6.0, 0.833) -&gt; 0                     </pre>	<pre> if EIC1103 == 1 and EIC1202 == 1 then -&gt; 1 (84, 0.872)  if EIC1100 == 1 then -&gt; 1 (86, 0.864)  if EIC1100 == 0 and EIC1103 == 0 then -&gt; 0 (11, 0.923)  if EIC1204 == 0 then -&gt; 0 (21, 0.783)  Default : -&gt; 1                     </pre>
EIC2201	<pre> EIC1206 0 (17.0, 0.765) -&gt; 0 EIC1206 1   EIC1102 1 (87.0, 0.931) -&gt; 1   EIC1102 0 (3.0, 0.667) -&gt; 0                     </pre>	<pre> if EIC1102 == 1 and EIC1206 == 1 then -&gt; 1 (87, 0.921)  if EIC1206 == 0 then -&gt; 0 (17, 0.737)  if EIC1102 == 0 then -&gt; 0 (13, 0.667)  Default : -&gt; 1                     </pre>
EIC2202	<pre> EIC1203 0 (20.0, 0.95) -&gt; 0 EIC1203 1   EIC1104 0 (3.0, 1.0) -&gt; 0   EIC1104 1     EIC1204 1 (77.0, 0.909) -&gt; 1     EIC1204 0       EIC1200 1 (3.0, 1.0) -&gt; 0       EIC1200 0         EIC1206 1 (2.0, 1.0) -&gt; 1         EIC1206 0 (2.0, 1.0) -&gt; 0                     </pre>	<pre> if EIC1203 == 1 then -&gt; 1 (87, 0.82)  if EIC1200 == 0 and EIC1203 == 1 and EIC1204 == 0 and EIC1206 == 1 then -&gt; 1 (2, 0.75)  if EIC1104 == 0 then -&gt; 0 (11, 0.923)  if EIC1203 == 0 then -&gt; 0 (20, 0.909)  if EIC1204 == 0 then -&gt; 0 (21, 0.87)  Default : -&gt; 1                     </pre>

Tabela I.5– Árvores e regras de decisão para o segundo ano (anual)

Disciplina	Árvore de decisão	Regras de decisão
EIC2203	<pre> EIC1206 0 (17.0, 0.941) -&gt; 0 EIC1206 1   EIC1104 0 (3.0, 1.0) -&gt; 0   EIC1104 1     EIC1203 1 (80.0, 0.975) -&gt; 1     EIC1203 0       EIC1100 1 (3.0, 0.667) -&gt; 1       EIC1100 0 (4.0, 1.0) -&gt; 0 </pre>	<pre> if EIC1104 == 1 and EIC1203 == 1 and EIC1206 == 1 then -&gt; 1 (80, 0.963)  if EIC1100 == 1 and EIC1104 == 1 and EIC1206 == 1 then -&gt; 1 (79, 0.951)  if EIC1100 == 0 and EIC1203 == 0 then -&gt; 0 (14, 0.938)  if EIC1104 == 0 then -&gt; 0 (11, 0.923)  if EIC1206 == 0 then -&gt; 0 (17, 0.895)  Default : -&gt; 1 </pre>
EIC2204	<pre> EIC1204 0 (21.0, 0.952) -&gt; 0 EIC1204 1   EIC1203 0 (8.0, 0.875) -&gt; 0   EIC1203 1     EIC1101 1 (76.0, 0.908) -&gt; 1     EIC1101 0 (2.0, 1.0) -&gt; 0 </pre>	<pre> if EIC1101 == 1 and EIC1203 == 1 and EIC1204 == 1 then -&gt; 1 (76, 0.897)  if EIC1204 == 0 then -&gt; 0 (21, 0.913)  if EIC1101 == 0 then -&gt; 0 (16, 0.889)  if EIC1203 == 0 then -&gt; 0 (20, 0.864)  Default : -&gt; 1 </pre>
EIC2205	<pre> EIC1204 1 (86.0, 0.965) -&gt; 1 EIC1204 0   EIC1203 0 (12.0, 0.833) -&gt; 0   EIC1203 1     EIC1104 1 (7.0, 0.714) -&gt; 1     EIC1104 0 (2.0, 1.0) -&gt; 0 </pre>	<pre> if EIC1104 == 1 and EIC1203 == 1 then -&gt; 1 (84, 0.965)  if EIC1204 == 1 then -&gt; 1 (86, 0.955)  if EIC1104 == 0 and EIC1203 == 1 then -&gt; 0 (3, 0.8)  if EIC1203 == 0 and EIC1204 == 0 then -&gt; 0 (12, 0.786)  Default : -&gt; 1 </pre>

Tabela I.6 – Árvores e regras de decisão para o terceiro ano (semestral)

Disciplina	Árvore de decisão	Regras de decisão
EIC3200	EIC3105 1 (58.0, 0.948) -> 1 EIC3105 0 (13.0, 0.846) -> 0	if EIC3105 == 1 then -> 1 (58, 0.933)  if EIC3105 == 0 then -> 0 (13, 0.8)  Default : -> 1
EIC3201	EIC3100 1 (57.0, 0.965) -> 1 EIC3100 0 (14.0, 1.0) -> 0	if EIC3100 == 1 then -> 1 (57, 0.949)  if EIC3100 == 0 then -> 0 (14, 0.938)  Default : -> 1
EIC3202	EIC3103 1 (59.0, 0.915) -> 1 EIC3103 0 (12.0, 1.0) -> 0	if EIC3103 == 1 then -> 1 (59, 0.902)  if EIC3103 == 0 then -> 0 (12, 0.929)  Default : -> 1
EIC3203	EIC3100 0 (14.0, 0.929) -> 0 EIC3100 1 EIC3103 1 (55.0, 0.927) -> 1 EIC3103 0 (2.0, 1.0) -> 0	if EIC3100 == 1 and EIC3103 == 1 then -> 1 (55, 0.912)  if EIC3100 == 0 then -> 0 (14, 0.875)  if EIC3103 == 0 then -> 0 (12, 0.857)  Default : -> 1
EIC3204	EIC3100 1 (57.0, 0.947) -> 1 EIC3100 0 EIC3103 1 (4.0, 1.0) -> 0 EIC3103 0 (10.0, 0.6) -> 1	if EIC3100 == 1 then -> 1 (57, 0.932)  if EIC3100 == 0 and EIC3103 == 0 then -> 1 (10, 0.583)  if EIC3100 == 0 and EIC3103 == 1 then -> 0 (4, 0.833)  Default : -> 1
EIC3205	EIC3105 1 (58.0, 0.966) -> 1 EIC3105 0 (13.0, 0.846) -> 0	if EIC3105 == 1 then -> 1 (58, 0.95)  if EIC3105 == 0 then -> 0 (13, 0.8)  Default : -> 1

Tabela I.7– Árvores e regras de decisão para o terceiro ano (anual)

Disciplina	Árvore de decisão	Regras de decisão
EIC3100	EIC2102 0 (9.0, 1.0) -> 0 EIC2102 1 EIC2201 1 (60.0, 0.95) -> 1 EIC2201 0 (2.0, 1.0) -> 0	if EIC2102 == 1 and EIC2201 == 1 then -> 1 (60, 0.935)  if EIC2102 == 0 then -> 0 (9, 0.909)  if EIC2201 == 0 then -> 0 (3, 0.8)  Default : -> 1
EIC3101	EIC2102 0 (9.0, 1.0) -> 0 EIC2102 1 EIC2100 1 (56.0, 0.964) -> 1 EIC2100 0 (6.0, 0.833) -> 0	if EIC2100 == 1 and EIC2102 == 1 then -> 1 (56, 0.948)  if EIC2102 == 0 then -> 0 (9, 0.909)  if EIC2100 == 0 then -> 0 (12, 0.857)  Default : -> 1
EIC3102	EIC2100 0 (12.0, 1.0) -> 0 EIC2100 1 EIC2200 1 (56.0, 0.929) -> 1 EIC2200 0 (3.0, 0.667) -> 0	if EIC2100 == 1 and EIC2200 == 1 then -> 1 (56, 0.914)  if EIC2100 == 0 then -> 0 (12, 0.929)  if EIC2200 == 0 then -> 0 (7, 0.778)  Default : -> 1
EIC3103	EIC2100 1 (59.0, 0.966) -> 1 EIC2100 0 (12.0, 0.833) -> 0	if EIC2100 == 1 then -> 1 (59, 0.951)  if EIC2100 == 0 then -> 0 (12, 0.786)  Default : -> 1
EIC3104	EIC2102 0 (9.0, 1.0) -> 0 EIC2102 1 EIC2103 0 (3.0, 1.0) -> 0 EIC2103 1 EIC2204 1 (54.0, 0.963) -> 1 EIC2204 0 (5.0, 0.6) -> 0	if EIC2102 == 1 and EIC2103 == 1 and EIC2204 == 1 then -> 1 (54, 0.946)  if EIC2102 == 0 then -> 0 (9, 0.909)  if EIC2103 == 0 then -> 0 (8, 0.9)  if EIC2204 == 0 then -> 0 (11, 0.769)  Default : -> 1

Tabela I.8– Árvores e regras de decisão para o terceiro ano (anual)

Disciplina	Árvore de decisão	Regras de decisão
EIC3105	EIC2202 1 (61.0, 0.934) -> 1 EIC2202 0 (10.0, 0.9) -> 0	if EIC2202 == 1 then -> 1 (61, 0.921)  if EIC2202 == 0 then -> 0 (10, 0.833)  Default : -> 1
EIC3200	EIC2202 0 (10.0, 0.9) -> 0 EIC2202 1 EIC2100 1 (56.0, 0.964) -> 1 EIC2100 0 EIC2101 1 (2.0, 1.0) -> 0 EIC2101 0 (3.0, 0.667) -> 1	if EIC2202 == 1 then -> 1 (61, 0.905)  if EIC2100 == 0 and EIC2101 == 1 then -> 0 (5, 0.857)  if EIC2202 == 0 then -> 0 (10, 0.833)  if EIC2100 == 0 then -> 0 (12, 0.714)  Default : -> 1
EIC3201	EIC2202 0 (10.0, 1.0) -> 0 EIC2202 1 EIC2101 1 (57.0, 0.947) -> 1 EIC2101 0 (4.0, 0.75) -> 0	if EIC2101 == 1 and EIC2202 == 1 then -> 1 (57, 0.932)  if EIC2202 == 0 then -> 0 (10, 0.917)  if EIC2101 == 0 then -> 0 (8, 0.8)  Default : -> 1
EIC3202	EIC2100 0 (12.0, 0.917) -> 0 EIC2100 1 EIC2201 1 (57.0, 0.93) -> 1 EIC2201 0 (2.0, 1.0) -> 0	if EIC2100 == 1 and EIC2201 == 1 then -> 1 (57, 0.915)  if EIC2100 == 0 then -> 0 (12, 0.857)  if EIC2201 == 0 then -> 0 (3, 0.8)  Default : -> 1
EIC3203	EIC2204 0 (11.0, 1.0) -> 0 EIC2204 1 EIC2102 1 (56.0, 0.929) -> 1 EIC2102 0 (4.0, 1.0) -> 0	if EIC2102 == 1 and EIC2204 == 1 then -> 1 (56, 0.914)  if EIC2204 == 0 then -> 0 (11, 0.923)  if EIC2102 == 0 then -> 0 (9, 0.909)  Default : -> 1
EIC3204	EIC2204 1 (60.0, 0.95) -> 1 EIC2204 0 (11.0, 0.727) -> 0	if EIC2204 == 1 then -> 1 (60, 0.935)  if EIC2204 == 0 then -> 0 (11, 0.692)  Default : -> 1
EIC3205	EIC2203 0 (6.0, 1.0) -> 0 EIC2203 1 EIC2201 0 (2.0, 1.0) -> 0 EIC2201 1 EIC2202 1 (59.0, 0.966) -> 1 EIC2202 0 (4.0, 0.75) -> 0	if EIC2201 == 1 and EIC2202 == 1 and EIC2203 == 1 then -> 1 (59, 0.951)  if EIC2203 == 0 then -> 0 (6, 0.875)  if EIC2202 == 0 then -> 0 (10, 0.833)  if EIC2201 == 0 then -> 0 (3, 0.8)  Default : -> 1

Tabela I.9– Árvores e regras de decisão para o quarto ano (semestral)

Disciplina	Árvore de decisão	Regras de decisão
EIC4200	-> 1	Default : -> 1
EIC4201	-> 1	Default : -> 1
EIC4202	EIC4105 1 (34.0, 0.941) -> 1 EIC4105 0 (3.0, 0.667) -> 0	If EIC4105 == 1 then -> 1 (34, 0.917)  if EIC4105 == 0 then -> 0 (3, 0.6)  Default : -> 1
EIC4203	EIC4100 1 (33.0, 0.879) -> 1 EIC4100 0 (4.0, 0.75) -> 0	If EIC4100 == 1 then -> 1 (33, 0.857)  if EIC4100 == 0 then -> 0 (4, 0.667)  Default : -> 1
EIC4204	EIC4103 1 (30.0, 0.667) -> 1 EIC4103 0 (7.0, 1.0) -> 0	If EIC4103 == 1 then -> 1 (30, 0.656)  if EIC4103 == 0 then -> 0 (7, 0.889)  Default : -> 1
EIC4205	EIC4101 1 (32.0, 0.938) -> 1 EIC4101 0 (5.0, 1.0) -> 0	If EIC4101 == 1 then -> 1 (32, 0.912)  if EIC4101 == 0 then -> 0 (5, 0.857)  Default : -> 1
EIC4206	-> 0	Default : -> 0

Tabela I.10– Árvores e regras de decisão para o quarto ano (anual)

Disciplina	Árvore de decisão	Regras de decisão
EIC4100	EIC3102 1 (34.0, 0.971) -> 1 EIC3102 0 (3.0, 1.0) -> 0	if EIC3102 == 1 then -> 1 (34, 0.944)  if EIC3102 == 0 then -> 0 (3, 0.8)  Default : -> 1
EIC4101	EIC3102 1 (34.0, 0.941) -> 1 EIC3102 0 (3.0, 1.0) -> 0	if EIC3102 == 1 then -> 1 (34, 0.917)  if EIC3102 == 0 then -> 0 (3, 0.8)  Default : -> 1
EIC4102	EIC3102 1 (34.0, 0.971) -> 1 EIC3102 0 (3.0, 1.0) -> 0	if EIC3102 == 1 then -> 1 (34, 0.944)  if EIC3102 == 0 then -> 0 (3, 0.8)  Default : -> 1
EIC4103	EIC3102 1 (34.0, 0.882) -> 1 EIC3102 0 (3.0, 1.0) -> 0	if EIC3102 == 1 then -> 1 (34, 0.861)  if EIC3102 == 0 then -> 0 (3, 0.8)  Default : -> 1
EIC4104	EIC3104 1 (35.0, 0.857) -> 1 EIC3104 0 (2.0, 1.0) -> 0	if EIC3104 == 1 then -> 1 (35, 0.838)  if EIC3104 == 0 then -> 0 (2, 0.75)  Default : -> 1
EIC4105	EIC3104 1 (35.0, 0.971) -> 1 EIC3104 0 (2.0, 1.0) -> 0	if EIC3104 == 1 then -> 1 (35, 0.946)  if EIC3104 == 0 then -> 0 (2, 0.75)  Default : -> 1
EIC4200	EIC3102 1 (34.0, 0.824) -> 1 EIC3102 0 (3.0, 0.667) -> 0	if EIC3102 == 1 then -> 1 (34, 0.806)  if EIC3102 == 0 then -> 0 (3, 0.6)  Default : -> 1
EIC4201	-> 1	Default : -> 1
EIC4202	EIC3104 1 (35.0, 0.943) -> 1 EIC3104 0 (2.0, 1.0) -> 0	if EIC3104 == 1 then -> 1 (35, 0.919)  if EIC3104 == 0 then -> 0 (2, 0.75)  Default : -> 1
EIC4203	EIC3102 1 (34.0, 0.882) -> 1 EIC3102 0 (3.0, 1.0) -> 0	if EIC3102 == 1 then -> 1 (34, 0.861)  if EIC3102 == 0 then -> 0 (3, 0.8)  Default : -> 1
EIC4204	EIC3102 1 (34.0, 0.588) -> 1 EIC3102 0 (3.0, 1.0) -> 0	if EIC3102 == 1 then -> 1 (34, 0.583)  if EIC3102 == 0 then -> 0 (3, 0.8)  Default : -> 1
EIC4205	EIC3102 1 (34.0, 0.882) -> 1 EIC3102 0 (3.0, 1.0) -> 0	if EIC3102 == 1 then -> 1 (34, 0.861)  if EIC3102 == 0 then -> 0 (3, 0.8)  Default : -> 1
EIC4206	-> 0	Default : -> 0



Tabela I.11– Árvores e regras de decisão para o quinto ano (anual)

Disciplina	Árvore de decisão	Regras de decisão
EIC5100	-> 1	Default : -> 1
EIC5101	-> 1	Default : -> 1
EIC5102	-> 1	Default : -> 1
EIC5103	EIC4104 1 (15.0, 1.0) -> 1 EIC4104 0 (2.0, 1.0) -> 0	if EIC4104 == 1 then -> 1 (15, 0.941)  if EIC4104 == 0 then -> 0 (2, 0.75)  Default : -> 1
EIC5104	EIC4104 0 (2.0, 1.0) -> 0 EIC4104 1 EIC4200 1 (13.0, 0.846) -> 1 EIC4200 0 (2.0, 1.0) -> 0	if EIC4104 == 1 and EIC4200 == 1 then -> 1 (13, 0.8)  if EIC4200 == 0 then -> 0 (2, 0.75)  if EIC4104 == 0 then -> 0 (2, 0.75)  Default : -> 1
EIC5105	EIC4104 1 (15.0, 0.8) -> 0 EIC4104 0 (2.0, 1.0) -> 1	if EIC4104 == 0 then -> 1 (2, 0.75)  if EIC4104 == 1 then -> 0 (15, 0.765)  Default : -> 0
EIC5106	-> 1	Default : -> 1
EIC5201	-> 1	Default : -> 1

**Anexo II**  
**Análise de Árvores e Regras de**  
**Decisão**

Tabela II.1– Análise das árvores e regras de decisão para o primeiro ano (sem.)

Disciplina	Análise das árvores de decisão	Análise das regras de decisão
<b>EIC1200</b>	Correct : 128 ( 92.09%) Wrong : 11 ( 7.91%) Total : 139  Coincidence Matrix \$C-EIC1200 0 1 0 23 9 1 2 105	Correct : 131 ( 94.24%) Wrong : 8 ( 5.76%) Total : 139  Coincidence Matrix \$C-EIC1200 0 1 0 26 6 1 2 105
<b>EIC1202</b>	Correct : 127 ( 91.37%) Wrong : 12 ( 8.63%) Total : 139  Coincidence Matrix \$C-EIC1202 0 1 0 11 7 1 5 116	Correct : 127 ( 91.37%) Wrong : 12 ( 8.63%) Total : 139  Coincidence Matrix \$C-EIC1202 0 1 0 11 7 1 5 116
<b>EIC1203</b>	Correct : 127 ( 91.37%) Wrong : 12 ( 8.63%) Total : 139  Coincidence Matrix \$C-EIC1203 0 1 0 22 9 1 3 105	Correct : 128 ( 92.09%) Wrong : 11 ( 7.91%) Total : 139  Coincidence Matrix \$C-EIC1203 0 1 0 21 10 1 1 107
<b>EIC1204</b>	Correct : 122 ( 87.77%) Wrong : 17 ( 12.23%) Total : 139  Coincidence Matrix \$C-EIC1204 0 1 0 13 10 1 7 109	Correct : 126 ( 90.65%) Wrong : 13 ( 9.35%) Total : 139  Coincidence Matrix \$C-EIC1204 0 1 0 13 10 1 3 113
<b>EIC1205</b>	Correct : 127 ( 91.37%) Wrong : 12 ( 8.63%) Total : 139  Coincidence Matrix \$C-EIC1205 0 1 0 7 4 1 8 120	Correct : 127 ( 91.37%) Wrong : 12 ( 8.63%) Total : 139  Coincidence Matrix \$C-EIC1205 0 1 0 7 4 1 8 120
<b>EIC1206</b>	Correct : 125 ( 89.93%) Wrong : 14 ( 10.07%) Total : 139  Coincidence Matrix \$C-EIC1206 0 1 0 14 9 1 5 111	Correct : 125 ( 89.93%) Wrong : 14 ( 10.07%) Total : 139  Coincidence Matrix \$C-EIC1206 0 1 0 14 9 1 5 111

Tabela II.2– Análise das árvores e regras de decisão para o segundo ano (sem.)

Disciplina	Análise das árvores de decisão	Análise das regras de decisão
EIC2200	Correct : 94 ( 87.85%) Wrong : 13 ( 12.15%) Total : 107  Coincidence Matrix \$C-EIC2200 0 1 0 8 8 1 5 86	Correct : 94 ( 87.85%) Wrong : 13 ( 12.15%) Total : 107  Coincidence Matrix \$C-EIC2200 0 1 0 8 8 1 5 86
EIC2201	Correct : 98 ( 91.59%) Wrong : 9 ( 8.41%) Total : 107  Coincidence Matrix \$C-EIC2201 0 1 0 12 5 1 4 86	Correct : 98 ( 91.59%) Wrong : 9 ( 8.41%) Total : 107  Coincidence Matrix \$C-EIC2201 0 1 0 12 5 1 4 86
EIC2202	Correct : 100 ( 93.46%) Wrong : 7 ( 6.54%) Total : 107  Coincidence Matrix \$C-EIC2202 0 1 0 23 6 1 1 77	Correct : 100 ( 93.46%) Wrong : 7 ( 6.54%) Total : 107  Coincidence Matrix \$C-EIC2202 0 1 0 23 6 1 1 77
EIC2203	Correct : 98 ( 91.59%) Wrong : 9 ( 8.41%) Total : 107  Coincidence Matrix \$C-EIC2203 0 1 0 15 2 1 7 83	Correct : 98 ( 91.59%) Wrong : 9 ( 8.41%) Total : 107  Coincidence Matrix \$C-EIC2203 0 1 0 15 2 1 7 83
EIC2204	Correct : 90 ( 84.11%) Wrong : 17 ( 15.89%) Total : 107  Coincidence Matrix \$C-EIC2204 0 1 0 19 7 1 10 71	Correct : 90 ( 84.11%) Wrong : 17 ( 15.89%) Total : 107  Coincidence Matrix \$C-EIC2204 0 1 0 19 7 1 10 71
EIC2205	Correct : 100 ( 93.46%) Wrong : 7 ( 6.54%) Total : 107  Coincidence Matrix \$C-EIC2205 0 1 0 13 5 1 2 87	Correct : 100 ( 93.46%) Wrong : 7 ( 6.54%) Total : 107  Coincidence Matrix \$C-EIC2205 0 1 0 13 5 1 2 87

Tabela II.3– Análise das árvores e regras de decisão para o segundo ano (anual)

Disciplina	Análise das árvores de decisão	Análise das regras de decisão
EIC2100	Correct : 95 ( 88.79%) Wrong : 12 ( 11.21%) Total : 107  Coincidence Matrix \$C-EIC2100 0 1 0 18 10 1 2 77	Correct : 95 ( 88.79%) Wrong : 12 ( 11.21%) Total : 107  Coincidence Matrix \$C-EIC2100 0 1 0 18 10 1 2 77
EIC2101	Correct : 94 ( 87.85%) Wrong : 13 ( 12.15%) Total : 107  Coincidence Matrix \$C-EIC2101 0 1 0 14 9 1 4 80	Correct : 94 ( 87.85%) Wrong : 13 ( 12.15%) Total : 107  Coincidence Matrix \$C-EIC2101 0 1 0 14 9 1 4 80
EIC2102	Correct : 90 ( 84.11%) Wrong : 17 ( 15.89%) Total : 107  Coincidence Matrix \$C-EIC2102 0 1 0 15 9 1 8 75	Correct : 90 ( 84.11%) Wrong : 17 ( 15.89%) Total : 107  Coincidence Matrix \$C-EIC2102 0 1 0 15 9 1 8 75
EIC2103	Correct : 99 ( 92.52%) Wrong : 8 ( 7.48%) Total : 107  Coincidence Matrix \$C-EIC2103 0 1 0 19 6 1 2 80	Correct : 99 ( 92.52%) Wrong : 8 ( 7.48%) Total : 107  Coincidence Matrix \$C-EIC2103 0 1 0 19 6 1 2 80
EIC2104	Correct : 97 ( 90.65%) Wrong : 10 ( 9.35%) Total : 107  Coincidence Matrix \$C-EIC2104 0 1 0 11 7 1 3 86	Correct : 97 ( 90.65%) Wrong : 10 ( 9.35%) Total : 107  Coincidence Matrix \$C-EIC2104 0 1 0 11 7 1 3 86
EIC2105	Correct : 101 ( 94.39%) Wrong : 6 ( 5.61%) Total : 107  Coincidence Matrix \$C-EIC2105 0 1 0 8 6 1 0 93	Correct : 101 ( 94.39%) Wrong : 6 ( 5.61%) Total : 107  Coincidence Matrix \$C-EIC2105 0 1 0 8 6 1 0 93
EIC2200	Correct : 97 ( 90.65%) Wrong : 10 ( 9.35%) Total : 107  Coincidence Matrix \$C-EIC2200 0 1 0 9 7 1 3 88	Correct : 94 ( 87.85%) Wrong : 13 ( 12.15%) Total : 107  Coincidence Matrix \$C-EIC2200 0 1 0 5 11 1 2 89
EIC2201	Correct : 96 ( 89.72%) Wrong : 11 ( 10.28%) Total : 107  Coincidence Matrix \$C-EIC2201 0 1 0 11 6 1 5 85	Correct : 96 ( 89.72%) Wrong : 11 ( 10.28%) Total : 107  Coincidence Matrix \$C-EIC2201 0 1 0 11 6 1 5 85

Tabela II.4– Análise das árvores e regras de decisão para o segundo ano (anual)

Disciplina	Análise das árvores de decisão	Análise das regras de decisão
EIC2202	Correct : 88 ( 82.24%) Wrong : 19 ( 17.76%) Total : 107  Coincidence Matrix \$C-EIC2202 0 1 0 15 14 1 5 73	Correct : 88 ( 82.24%) Wrong : 19 ( 17.76%) Total : 107  Coincidence Matrix \$C-EIC2202 0 1 0 15 14 1 5 73
EIC2203	Correct : 101 ( 94.39%) Wrong : 6 ( 5.61%) Total : 107  Coincidence Matrix \$C-EIC2203 0 1 0 15 2 1 4 86	Correct : 101 ( 94.39%) Wrong : 6 ( 5.61%) Total : 107  Coincidence Matrix \$C-EIC2203 0 1 0 15 2 1 4 86
EIC2204	Correct : 93 ( 86.92%) Wrong : 14 ( 13.08%) Total : 107  Coincidence Matrix \$C-EIC2204 0 1 0 17 9 1 5 76	Correct : 93 ( 86.92%) Wrong : 14 ( 13.08%) Total : 107  Coincidence Matrix \$C-EIC2204 0 1 0 17 9 1 5 76
EIC2205	Correct : 96 ( 89.72%) Wrong : 11 ( 10.28%) Total : 107  Coincidence Matrix \$C-EIC2205 0 1 0 7 11 1 0 89	Correct : 96 ( 89.72%) Wrong : 11 ( 10.28%) Total : 107  Coincidence Matrix \$C-EIC2205 0 1 0 7 11 1 0 89

Tabela II.5 – Análise das árvores e regras de decisão para o terceiro ano (sem.)

Disciplina	Análise das árvores de decisão	Análise das regras de decisão
EIC3200	Correct : 59 ( 83.10%) Wrong : 12 ( 16.90%) Total : 71  Coincidence Matrix \$C-EIC3200 0 1 0 6 10 1 2 53	Correct : 59 ( 83.10%) Wrong : 12 ( 16.90%) Total : 71  Coincidence Matrix \$C-EIC3200 0 1 0 6 10 1 2 53
EIC3201	Correct : 63 ( 88.73%) Wrong : 8 ( 11.27%) Total : 71  Coincidence Matrix \$C-EIC3201 0 1 0 9 7 1 1 54	Correct : 63 ( 88.73%) Wrong : 8 ( 11.27%) Total : 71  Coincidence Matrix \$C-EIC3201 0 1 0 9 7 1 1 54
EIC3202	Correct : 63 ( 88.73%) Wrong : 8 ( 11.27%) Total : 71  Coincidence Matrix \$C-EIC3202 0 1 0 11 7 1 1 52	Correct : 63 ( 88.73%) Wrong : 8 ( 11.27%) Total : 71  Coincidence Matrix \$C-EIC3202 0 1 0 11 7 1 1 52
EIC3203	Correct : 63 ( 88.73%) Wrong : 8 ( 11.27%) Total : 71  Coincidence Matrix \$C-EIC3203 0 1 0 13 6 1 2 50	Correct : 63 ( 88.73%) Wrong : 8 ( 11.27%) Total : 71  Coincidence Matrix \$C-EIC3203 0 1 0 13 6 1 2 50
EIC3204	Correct : 60 ( 84.51%) Wrong : 11 ( 15.49%) Total : 71  Coincidence Matrix \$C-EIC3204 0 1 0 2 10 1 1 58	Correct : 60 ( 84.51%) Wrong : 11 ( 15.49%) Total : 71  Coincidence Matrix \$C-EIC3204 0 1 0 2 10 1 1 58
EIC3205	Correct : 63 ( 88.73%) Wrong : 8 ( 11.27%) Total : 71  Coincidence Matrix \$C-EIC3205 0 1 0 8 8 1 0 55	Correct : 63 ( 88.73%) Wrong : 8 ( 11.27%) Total : 71  Coincidence Matrix \$C-EIC3205 0 1 0 8 8 1 0 55

Tabela II.6–Análise das árvores e regras de decisão para o terceiro ano (anual)

Disciplina	Análise das árvores de decisão	Análise das regras de decisão
EIC3100	Correct : 64 ( 90.14%) Wrong : 7 ( 9.86%) Total : 71  Coincidence Matrix \$C-EIC3100 0 1 0 8 2 1 5 56	Correct : 64 ( 90.14%) Wrong : 7 ( 9.86%) Total : 71  Coincidence Matrix \$C-EIC3100 0 1 0 8 2 1 5 56
EIC3101	Correct : 63 ( 88.73%) Wrong : 8 ( 11.27%) Total : 71  Coincidence Matrix \$C-EIC3101 0 1 0 15 6 1 2 48	Correct : 63 ( 88.73%) Wrong : 8 ( 11.27%) Total : 71  Coincidence Matrix \$C-EIC3101 0 1 0 15 6 1 2 48
EIC3102	Correct : 62 ( 87.32%) Wrong : 9 ( 12.68%) Total : 71  Coincidence Matrix \$C-EIC3102 0 1 0 13 6 1 3 49	Correct : 62 ( 87.32%) Wrong : 9 ( 12.68%) Total : 71  Coincidence Matrix \$C-EIC3102 0 1 0 13 6 1 3 49
EIC3103	Correct : 62 ( 87.32%) Wrong : 9 ( 12.68%) Total : 71  Coincidence Matrix \$C-EIC3103 0 1 0 8 4 1 5 54	Correct : 62 ( 87.32%) Wrong : 9 ( 12.68%) Total : 71  Coincidence Matrix \$C-EIC3103 0 1 0 8 4 1 5 54
EIC3104	Correct : 65 ( 91.55%) Wrong : 6 ( 8.45%) Total : 71  Coincidence Matrix \$C-EIC3104 0 1 0 15 4 1 2 50	Correct : 65 ( 91.55%) Wrong : 6 ( 8.45%) Total : 71  Coincidence Matrix \$C-EIC3104 0 1 0 15 4 1 2 50
EIC3105	Correct : 63 ( 88.73%) Wrong : 8 ( 11.27%) Total : 71  Coincidence Matrix \$C-EIC3105 0 1 0 7 1 1 7 56	Correct : 63 ( 88.73%) Wrong : 8 ( 11.27%) Total : 71  Coincidence Matrix \$C-EIC3105 0 1 0 7 1 1 7 56
EIC3200	Correct : 61 ( 85.92%) Wrong : 10 ( 14.08%) Total : 71  Coincidence Matrix \$C-EIC3200 0 1 0 11 5 1 5 50	Correct : 61 ( 85.92%) Wrong : 10 ( 14.08%) Total : 71  Coincidence Matrix \$C-EIC3200 0 1 0 11 5 1 5 50
EIC3201	Correct : 64 ( 90.14%) Wrong : 7 ( 9.86%) Total : 71  Coincidence Matrix \$C-EIC3201 0 1 0 13 3 1 4 51	Correct : 64 ( 90.14%) Wrong : 7 ( 9.86%) Total : 71  Coincidence Matrix \$C-EIC3201 0 1 0 13 3 1 4 51



Tabela II.7– Análise das árvores e regras de decisão para o terceiro ano (anual)

Disciplina	Análise das árvores de decisão	Análise das regras de decisão
EIC3202	Correct : 63 ( 88.73%) Wrong : 8 ( 11.27%) Total : 71  Coincidence Matrix \$C-EIC3202 0 1 0 12 6 1 2 51	Correct : 63 ( 88.73%) Wrong : 8 ( 11.27%) Total : 71  Coincidence Matrix \$C-EIC3202 0 1 0 12 6 1 2 51
EIC3203	Correct : 66 ( 92.96%) Wrong : 5 ( 7.04%) Total : 71  Coincidence Matrix \$C-EIC3203 0 1 0 15 4 1 1 51	Correct : 66 ( 92.96%) Wrong : 5 ( 7.04%) Total : 71  Coincidence Matrix \$C-EIC3203 0 1 0 15 4 1 1 51
EIC3204	Correct : 63 ( 88.73%) Wrong : 8 ( 11.27%) Total : 71  Coincidence Matrix \$C-EIC3204 0 1 0 9 3 1 5 54	Correct : 63 ( 88.73%) Wrong : 8 ( 11.27%) Total : 71  Coincidence Matrix \$C-EIC3204 0 1 0 9 3 1 5 54
EIC3205	Correct : 64 ( 90.14%) Wrong : 7 ( 9.86%) Total : 71  Coincidence Matrix \$C-EIC3205 0 1 0 12 4 1 3 52	Correct : 64 ( 90.14%) Wrong : 7 ( 9.86%) Total : 71  Coincidence Matrix \$C-EIC3205 0 1 0 12 4 1 3 52

Tabela II.8– Análise das árvores e regras de decisão para o quarto ano (sem.)

Disciplina	Análise das árvores de decisão	Análise das regras de decisão
EIC4200	Correct : 27 ( 72.97%) Wrong : 10 ( 27.03%) Total : 37  Coincidence Matrix \$C-EIC4200 0 1 0 10 1 0 27	Correct : 27 ( 72.97%) Wrong : 10 ( 27.03%) Total : 37  Coincidence Matrix \$C-EIC4200 0 1 0 10 1 0 27
EIC4201	Correct : 33 ( 89.19%) Wrong : 4 ( 10.81%) Total : 37  Coincidence Matrix \$C-EIC4201 0 1 0 4 1 0 33	Correct : 33 ( 89.19%) Wrong : 4 ( 10.81%) Total : 37  Coincidence Matrix \$C-EIC4201 0 1 0 4 1 0 33
EIC4202	Correct : 34 ( 91.89%) Wrong : 3 ( 8.11%) Total : 37  Coincidence Matrix \$C-EIC4202 0 1 0 2 3 1 0 32	Correct : 34 ( 91.89%) Wrong : 3 ( 8.11%) Total : 37  Coincidence Matrix \$C-EIC4202 0 1 0 2 3 1 0 32
EIC4203	Correct : 34 ( 91.89%) Wrong : 3 ( 8.11%) Total : 37  Coincidence Matrix \$C-EIC4203 0 1 0 4 2 1 1 30	Correct : 34 ( 91.89%) Wrong : 3 ( 8.11%) Total : 37  Coincidence Matrix \$C-EIC4203 0 1 0 4 2 1 1 30
EIC4204	Correct : 23 ( 62.16%) Wrong : 14 ( 37.84%) Total : 37  Coincidence Matrix \$C-EIC4204 0 1 0 5 14 1 0 18	Correct : 23 ( 62.16%) Wrong : 14 ( 37.84%) Total : 37  Coincidence Matrix \$C-EIC4204 0 1 0 5 14 1 0 18
EIC4205	Correct : 33 ( 89.19%) Wrong : 4 ( 10.81%) Total : 37  Coincidence Matrix \$C-EIC4205 0 1 0 5 1 1 3 28	Correct : 33 ( 89.19%) Wrong : 4 ( 10.81%) Total : 37  Coincidence Matrix \$C-EIC4205 0 1 0 5 1 1 3 28
EIC4206	Correct : 22 ( 59.46%) Wrong : 15 ( 40.54%) Total : 37  Coincidence Matrix \$C-EIC4206 0 1 0 22 0 1 15 0	Correct : 22 ( 59.46%) Wrong : 15 ( 40.54%) Total : 37  Coincidence Matrix \$C-EIC4206 0 1 0 22 0 1 15 0

Tabela II.9– Análise das árvores e regras de decisão para o quarto ano (anual)

Disciplina	Análise das árvores de decisão	Análise das regras de decisão
EIC4100	Correct : 34 ( 91.89%) Wrong : 3 ( 8.11%) Total : 37  Coincidence Matrix \$C-EIC4100 0 1 0 2 3 1 0 32	Correct : 34 ( 91.89%) Wrong : 3 ( 8.11%) Total : 37  Coincidence Matrix \$C-EIC4100 0 1 0 2 3 1 0 32
EIC4101	Correct : 31 ( 83.78%) Wrong : 6 ( 16.22%) Total : 37  Coincidence Matrix \$C-EIC4101 0 1 0 2 6 1 0 29	Correct : 31 ( 83.78%) Wrong : 6 ( 16.22%) Total : 37  Coincidence Matrix \$C-EIC4101 0 1 0 2 6 1 0 29
EIC4102	Correct : 34 ( 91.89%) Wrong : 3 ( 8.11%) Total : 37  Coincidence Matrix \$C-EIC4102 0 1 0 2 3 1 0 32	Correct : 34 ( 91.89%) Wrong : 3 ( 8.11%) Total : 37  Coincidence Matrix \$C-EIC4102 0 1 0 2 3 1 0 32
EIC4103	Correct : 34 ( 91.89%) Wrong : 3 ( 8.11%) Total : 37  Coincidence Matrix \$C-EIC4103 0 1 0 2 3 1 0 32	Correct : 34 ( 91.89%) Wrong : 3 ( 8.11%) Total : 37  Coincidence Matrix \$C-EIC4103 0 1 0 2 3 1 0 32
EIC4104	Correct : 30 ( 81.08%) Wrong : 7 ( 18.92%) Total : 37  Coincidence Matrix \$C-EIC4104 0 1 0 5 6 1 1 25	Correct : 30 ( 81.08%) Wrong : 7 ( 18.92%) Total : 37  Coincidence Matrix \$C-EIC4104 0 1 0 5 6 1 1 25
EIC4105	Correct : 33 ( 89.19%) Wrong : 4 ( 10.81%) Total : 37  Coincidence Matrix \$C-EIC4105 0 1 0 2 0 1 4 31	Correct : 33 ( 89.19%) Wrong : 4 ( 10.81%) Total : 37  Coincidence Matrix \$C-EIC4105 0 1 0 2 0 1 4 31
EIC4200	Correct : 29 ( 78.38%) Wrong : 8 ( 21.62%) Total : 37  Coincidence Matrix \$C-EIC4200 0 1 0 2 8 1 0 27	Correct : 29 ( 78.38%) Wrong : 8 ( 21.62%) Total : 37  Coincidence Matrix \$C-EIC4200 0 1 0 2 8 1 0 27
EIC4201	Correct : 33 ( 89.19%) Wrong : 4 ( 10.81%) Total : 37  Coincidence Matrix \$C-EIC4201 0 1 0 0 4 1 0 33	Correct : 33 ( 89.19%) Wrong : 4 ( 10.81%) Total : 37  Coincidence Matrix \$C-EIC4201 0 1 0 0 4 1 0 33
EIC4202	Correct : 34 ( 91.89%) Wrong : 3 ( 8.11%) Total : 37  Coincidence Matrix \$C-EIC4202 0 1 0 4 1 1 2 30	Correct : 34 ( 91.89%) Wrong : 3 ( 8.11%) Total : 37  Coincidence Matrix \$C-EIC4202 0 1 0 4 1 1 2 30

Tabela II.10– Análise das árvores e regras de decisão para o quarto ano (anual)

Disciplina	Análise das árvores de decisão	Análise das regras de decisão
EIC4203	Correct : 33 ( 89.19%) Wrong : 4 ( 10.81%) Total : 37  Coincidence Matrix \$C-EIC4203 0 1 0 2 4 1 0 31	Correct : 33 ( 89.19%) Wrong : 4 ( 10.81%) Total : 37  Coincidence Matrix \$C-EIC4203 0 1 0 2 4 1 0 31
EIC4204	Correct : 20 ( 54.05%) Wrong : 17 ( 45.95%) Total : 37  Coincidence Matrix \$C-EIC4204 0 1 0 2 17 1 0 18	Correct : 20 ( 54.05%) Wrong : 17 ( 45.95%) Total : 37  Coincidence Matrix \$C-EIC4204 0 1 0 2 17 1 0 18
EIC4205	Correct : 33 ( 89.19%) Wrong : 4 ( 10.81%) Total : 37  Coincidence Matrix \$C-EIC4205 0 1 0 2 4 1 0 31	Correct : 33 ( 89.19%) Wrong : 4 ( 10.81%) Total : 37  Coincidence Matrix \$C-EIC4205 0 1 0 2 4 1 0 31
EIC4206	Correct : 22 ( 59.46%) Wrong : 15 ( 40.54%) Total : 37  Coincidence Matrix \$C-EIC4206 0 1 0 22 0 1 15 0	Correct : 22 ( 59.46%) Wrong : 15 ( 40.54%) Total : 37  Coincidence Matrix \$C-EIC4206 0 1 0 22 0 1 15 0

Tabela II.11 – Análise das árvores e regras de decisão para o quinto ano (anual)

Disciplina	Análise das árvores de decisão	Análise das regras de decisão
EIC5100	Correct : 12 ( 75.00%) Wrong : 4 ( 25.00%) Total : 16  Coincidence Matrix \$C-EIC5100 0 1 0 4 1 0 12	Correct : 12 ( 75.00%) Wrong : 4 ( 25.00%) Total : 16  Coincidence Matrix \$C-EIC5100 0 1 0 4 1 0 12
EIC5101	Correct : 12 ( 75.00%) Wrong : 4 ( 25.00%) Total : 16  Coincidence Matrix \$C-EIC5101 0 1 0 4 1 0 12	Correct : 12 ( 75.00%) Wrong : 4 ( 25.00%) Total : 16  Coincidence Matrix \$C-EIC5101 0 1 0 4 1 0 12
EIC5102	Correct : 12 ( 75.00%) Wrong : 4 ( 25.00%) Total : 16  Coincidence Matrix \$C-EIC5102 0 1 0 4 1 0 12	Correct : 12 ( 75.00%) Wrong : 4 ( 25.00%) Total : 16  Coincidence Matrix \$C-EIC5102 0 1 0 4 1 0 12
EIC5103	Correct : 13 ( 81.25%) Wrong : 3 ( 18.75%) Total : 16  Coincidence Matrix \$C-EIC5103 0 1 0 1 3 1 0 12	Correct : 13 ( 81.25%) Wrong : 3 ( 18.75%) Total : 16  Coincidence Matrix \$C-EIC5103 0 1 0 1 3 1 0 12
EIC5104	Correct : 9 ( 56.25%) Wrong : 7 ( 43.75%) Total : 16  Coincidence Matrix \$C-EIC5104 0 1 0 1 4 1 3 8	Correct : 9 ( 56.25%) Wrong : 7 ( 43.75%) Total : 16  Coincidence Matrix \$C-EIC5104 0 1 0 1 4 1 3 8
EIC5105	Correct : 14 ( 87.50%) Wrong : 2 ( 12.50%) Total : 16  Coincidence Matrix \$C-EIC5105 0 1 0 14 1 1 1 0	Correct : 14 ( 87.50%) Wrong : 2 ( 12.50%) Total : 16  Coincidence Matrix \$C-EIC5105 0 1 0 14 1 1 1 0
EIC5106	Correct : 13 ( 81.25%) Wrong : 3 ( 18.75%) Total : 16  Coincidence Matrix \$C-EIC5106 0 1 0 0 3 1 0 13	Correct : 13 ( 81.25%) Wrong : 3 ( 18.75%) Total : 16  Coincidence Matrix \$C-EIC5106 0 1 0 0 3 1 0 13
EIC5201	Correct : 16 (100.00%) Wrong : 0 ( 0.00%) Total : 16  Coincidence Matrix \$C-EIC5201 1 1 16	Correct : 16 (100.00%) Wrong : 0 ( 0.00%) Total : 16  Coincidence Matrix \$C-EIC5201 1 1 16

# **Anexo III**

## **Regras de Associação**

EIC1200	<=	EIC1100	(221:79.5%, 0.964)
EIC1200	<=	EIC1101 & EIC1100	(218:78.4%, 0.968)
EIC1200	<=	EIC1101 & EIC1102 & EIC1100	(206:74.1%, 0.971)
EIC1200	<=	EIC1101 & EIC1102 & EIC1103 & EIC1100	(198:71.2%, 0.98)
EIC1200	<=	EIC1101 & EIC1102 & EIC1103	(207:74.5%, 0.942)
EIC1200	<=	EIC1101 & EIC1102	(217:78.1%, 0.931)
EIC1200	<=	EIC1101 & EIC1103 & EIC1100	(204:73.4%, 0.975)
EIC1200	<=	EIC1101 & EIC1103	(216:77.7%, 0.931)
EIC1200	<=	EIC1101	(235:84.5%, 0.915)
EIC1200	<=	EIC1102 & EIC1100	(207:74.5%, 0.971)
EIC1200	<=	EIC1102 & EIC1103 & EIC1100	(198:71.2%, 0.98)
EIC1200	<=	EIC1102 & EIC1103	(213:76.6%, 0.925)
EIC1200	<=	EIC1102	(232:83.5%, 0.884)
EIC1200	<=	EIC1103 & EIC1100	(205:73.7%, 0.976)
EIC1200	<=	EIC1103	(229:82.4%, 0.891)
EIC1200	<=	EIC1104 & EIC1100	(208:74.8%, 0.966)
EIC1200	<=	EIC1104 & EIC1101 & EIC1100	(208:74.8%, 0.966)
EIC1200	<=	EIC1104 & EIC1101 & EIC1102 & EIC1100	(201:72.3%, 0.97)
EIC1200	<=	EIC1104 & EIC1101 & EIC1102 & EIC1103	(203:73.0%, 0.946)
EIC1200	<=	EIC1104 & EIC1101 & EIC1102	(211:75.9%, 0.934)
EIC1200	<=	EIC1104 & EIC1101 & EIC1103 & EIC1100	(199:71.6%, 0.975)
EIC1200	<=	EIC1104 & EIC1101 & EIC1103	(210:75.5%, 0.933)
EIC1200	<=	EIC1104 & EIC1101	(223:80.2%, 0.915)
EIC1200	<=	EIC1104 & EIC1102 & EIC1100	(201:72.3%, 0.97)
EIC1200	<=	EIC1104 & EIC1102 & EIC1103 & EIC1100	(195:70.1%, 0.979)
EIC1200	<=	EIC1104 & EIC1102 & EIC1103	(209:75.2%, 0.928)
EIC1200	<=	EIC1104 & EIC1102	(221:79.5%, 0.9)
EIC1200	<=	EIC1104 & EIC1103 & EIC1100	(199:71.6%, 0.975)
EIC1200	<=	EIC1104 & EIC1103	(220:79.1%, 0.9)
EIC1200	<=	EIC1104	(240:86.3%, 0.858)
EIC1200	<=	EIC1106 & EIC1100	(219:78.8%, 0.963)
EIC1200	<=	EIC1106 & EIC1101 & EIC1100	(216:77.7%, 0.968)
EIC1200	<=	EIC1106 & EIC1101 & EIC1102 & EIC1100	(204:73.4%, 0.971)
EIC1200	<=	EIC1106 & EIC1101 & EIC1102 & EIC1103	(206:74.1%, 0.942)
EIC1200	<=	EIC1106 & EIC1101 & EIC1102	(215:77.3%, 0.93)
EIC1200	<=	EIC1106 & EIC1101 & EIC1103 & EIC1100	(203:73.0%, 0.975)
EIC1200	<=	EIC1106 & EIC1101 & EIC1103	(215:77.3%, 0.93)
EIC1200	<=	EIC1106 & EIC1101	(231:83.1%, 0.918)
EIC1200	<=	EIC1106 & EIC1102 & EIC1100	(205:73.7%, 0.971)
EIC1200	<=	EIC1106 & EIC1102 & EIC1103 & EIC1100	(197:70.9%, 0.98)
EIC1200	<=	EIC1106 & EIC1102 & EIC1103	(212:76.3%, 0.925)
EIC1200	<=	EIC1106 & EIC1102	(228:82.0%, 0.89)
EIC1200	<=	EIC1106 & EIC1103 & EIC1100	(204:73.4%, 0.975)
EIC1200	<=	EIC1106 & EIC1103	(226:81.3%, 0.898)
EIC1200	<=	EIC1106 & EIC1104 & EIC1100	(206:74.1%, 0.966)
EIC1200	<=	EIC1106 & EIC1104 & EIC1101 & EIC1100	(206:74.1%, 0.966)
EIC1200	<=	EIC1106 & EIC1104 & EIC1101 & EIC1102	(209:75.2%, 0.933)
EIC1200	<=	EIC1106 & EIC1104 & EIC1101 & EIC1103	(209:75.2%, 0.933)
EIC1200	<=	EIC1106 & EIC1104 & EIC1101	(220:79.1%, 0.918)
EIC1200	<=	EIC1106 & EIC1104 & EIC1102 & EIC1100	(199:71.6%, 0.97)
EIC1200	<=	EIC1106 & EIC1104 & EIC1102 & EIC1103	(208:74.8%, 0.928)
EIC1200	<=	EIC1106 & EIC1104 & EIC1102	(218:78.4%, 0.904)
EIC1200	<=	EIC1106 & EIC1104 & EIC1103 & EIC1100	(198:71.2%, 0.975)
EIC1200	<=	EIC1106 & EIC1104 & EIC1103	(218:78.4%, 0.904)
EIC1200	<=	EIC1106 & EIC1104	(234:84.2%, 0.872)
EIC1200	<=	EIC1106	(260:93.5%, 0.835)

Figura III.1 - Regras de associação: disciplina EIC1200

Melhor suporte: 260:93.5% (conf=83.5%)  
 Melhor confiança: 98.0% (sup=198:71.2%)

EIC1200	<=	EIC1101 & EIC1102 & EIC1103 & EIC1100	(198:71.2%, 0.98)
EIC1200	<=	EIC1102 & EIC1103 & EIC1100	(198:71.2%, 0.98)
EIC1200	<=	EIC1106	(260:93.5%, 0.835)
EIC1200	<=	EIC1103 & EIC1100	(205:73.7%, 0.976)
EIC1200	<=	EIC1104	(240:86.3%, 0.858)
EIC1200	<=	EIC1102 & EIC1100	(207:74.5%, 0.971)
EIC1200	<=	EIC1101	(235:84.5%, 0.915)
EIC1200	<=	EIC1101 & EIC1100	(218:78.4%, 0.968)
EIC1200	<=	EIC1106 & EIC1101	(231:83.1%, 0.918)
EIC1200	<=	EIC1100	(221:79.5%, 0.964)

Figura III.2 - Regras de associação para a disciplina EIC1200 após o filtro

EIC1202	<=	EIC1100	(221:79.5%, 0.964)
EIC1202	<=	EIC1101 & EIC1100	(218:78.4%, 0.968)
EIC1202	<=	EIC1101 & EIC1102 & EIC1100	(206:74.1%, 0.99)
EIC1202	<=	EIC1101 & EIC1102 & EIC1103 & EIC1100	(198:71.2%, 0.99)
EIC1202	<=	EIC1101 & EIC1102 & EIC1103	(207:74.5%, 0.986)
EIC1202	<=	EIC1101 & EIC1102	(217:78.1%, 0.986)
EIC1202	<=	EIC1101 & EIC1103 & EIC1100	(204:73.4%, 0.98)
EIC1202	<=	EIC1101 & EIC1103	(216:77.7%, 0.972)
EIC1202	<=	EIC1101	(235:84.5%, 0.949)
EIC1202	<=	EIC1102 & EIC1100	(207:74.5%, 0.99)
EIC1202	<=	EIC1102 & EIC1103 & EIC1100	(198:71.2%, 0.99)
EIC1202	<=	EIC1102 & EIC1103	(213:76.6%, 0.981)
EIC1202	<=	EIC1102	(232:83.5%, 0.97)
EIC1202	<=	EIC1103 & EIC1100	(205:73.7%, 0.976)
EIC1202	<=	EIC1103	(229:82.4%, 0.943)
EIC1202	<=	EIC1104 & EIC1100	(208:74.8%, 0.986)
EIC1202	<=	EIC1104 & EIC1101 & EIC1100	(208:74.8%, 0.986)
EIC1202	<=	EIC1104 & EIC1101 & EIC1102 & EIC1100	(201:72.3%, 0.99)
EIC1202	<=	EIC1104 & EIC1101 & EIC1102 & EIC1103	(203:73.0%, 0.99)
EIC1202	<=	EIC1104 & EIC1101 & EIC1102	(211:75.9%, 0.991)
EIC1202	<=	EIC1104 & EIC1101 & EIC1103 & EIC1100	(199:71.6%, 0.985)
EIC1202	<=	EIC1104 & EIC1101 & EIC1103	(210:75.5%, 0.981)
EIC1202	<=	EIC1104 & EIC1101	(223:80.2%, 0.973)
EIC1202	<=	EIC1104 & EIC1102 & EIC1100	(201:72.3%, 0.99)
EIC1202	<=	EIC1104 & EIC1102 & EIC1103 & EIC1100	(195:70.1%, 0.99)
EIC1202	<=	EIC1104 & EIC1102 & EIC1103	(209:75.2%, 0.986)
EIC1202	<=	EIC1104 & EIC1102	(221:79.5%, 0.977)
EIC1202	<=	EIC1104 & EIC1103 & EIC1100	(199:71.6%, 0.985)
EIC1202	<=	EIC1104 & EIC1103	(220:79.1%, 0.964)
EIC1202	<=	EIC1104	(240:86.3%, 0.942)
EIC1202	<=	EIC1106 & EIC1100	(219:78.8%, 0.963)
EIC1202	<=	EIC1106 & EIC1101 & EIC1100	(216:77.7%, 0.968)
EIC1202	<=	EIC1106 & EIC1101 & EIC1102 & EIC1100	(204:73.4%, 0.99)
EIC1202	<=	EIC1106 & EIC1101 & EIC1102 & EIC1103	(206:74.1%, 0.985)
EIC1202	<=	EIC1106 & EIC1101 & EIC1102	(215:77.3%, 0.986)
EIC1202	<=	EIC1106 & EIC1101 & EIC1103 & EIC1100	(203:73.0%, 0.98)
EIC1202	<=	EIC1106 & EIC1101 & EIC1103	(215:77.3%, 0.972)
EIC1202	<=	EIC1106 & EIC1101	(231:83.1%, 0.957)
EIC1202	<=	EIC1106 & EIC1102 & EIC1100	(205:73.7%, 0.99)
EIC1202	<=	EIC1106 & EIC1102 & EIC1103 & EIC1100	(197:70.9%, 0.99)
EIC1202	<=	EIC1106 & EIC1102 & EIC1103	(212:76.3%, 0.981)
EIC1202	<=	EIC1106 & EIC1102	(228:82.0%, 0.974)
EIC1202	<=	EIC1106 & EIC1103 & EIC1100	(204:73.4%, 0.975)
EIC1202	<=	EIC1106 & EIC1103	(226:81.3%, 0.951)
EIC1202	<=	EIC1106 & EIC1104 & EIC1100	(206:74.1%, 0.985)
EIC1202	<=	EIC1106 & EIC1104 & EIC1101 & EIC1100	(206:74.1%, 0.985)
EIC1202	<=	EIC1106 & EIC1104 & EIC1101 & EIC1102	(209:75.2%, 0.99)
EIC1202	<=	EIC1106 & EIC1104 & EIC1101 & EIC1103	(209:75.2%, 0.981)
EIC1202	<=	EIC1106 & EIC1104 & EIC1101	(220:79.1%, 0.977)
EIC1202	<=	EIC1106 & EIC1104 & EIC1102 & EIC1100	(199:71.6%, 0.99)
EIC1202	<=	EIC1106 & EIC1104 & EIC1102 & EIC1103	(208:74.8%, 0.986)
EIC1202	<=	EIC1106 & EIC1104 & EIC1102	(218:78.4%, 0.982)
EIC1202	<=	EIC1106 & EIC1104 & EIC1103 & EIC1100	(198:71.2%, 0.985)
EIC1202	<=	EIC1106 & EIC1104 & EIC1103	(218:78.4%, 0.968)
EIC1202	<=	EIC1106 & EIC1104	(234:84.2%, 0.957)
EIC1202	<=	EIC1106	(260:93.5%, 0.908)

Figura III.3 - Regras de associação: disciplina EIC1202

Melhor suporte:	260:93.5%	(conf=90.8%)	
Melhor confiança:	99.1%	(sup=211:75.9%)	
EIC1202	<=	EIC1104 & EIC1101 & EIC1102	(211:75.9%, 0.991)
EIC1202	<=	EIC1106	(260:93.5%, 0.908)
EIC1202	<=	EIC1101 & EIC1102	(217:78.1%, 0.986)
EIC1202	<=	EIC1104	(240:86.3%, 0.942)
EIC1202	<=	EIC1106 & EIC1104 & EIC1102	(218:78.4%, 0.982)
EIC1202	<=	EIC1101	(235:84.5%, 0.949)
EIC1202	<=	EIC1104 & EIC1102	(221:79.5%, 0.977)
EIC1202	<=	EIC1106 & EIC1104	(234:84.2%, 0.957)
EIC1202	<=	EIC1106 & EIC1102	(228:82.0%, 0.974)
EIC1202	<=	EIC1102	(232:83.5%, 0.97)

Figura III.4 - Regras de associação: disciplina EIC1202 após o filtro



EIC1203	<=	EIC1100	(221:79.5%, 0.905)
EIC1203	<=	EIC1101 & EIC1100	(218:78.4%, 0.917)
EIC1203	<=	EIC1101 & EIC1102 & EIC1100	(206:74.1%, 0.932)
EIC1203	<=	EIC1101 & EIC1102 & EIC1103 & EIC1100	(198:71.2%, 0.97)
EIC1203	<=	EIC1101 & EIC1102 & EIC1103	(207:74.5%, 0.947)
EIC1203	<=	EIC1101 & EIC1102	(217:78.1%, 0.903)
EIC1203	<=	EIC1101 & EIC1103 & EIC1100	(204:73.4%, 0.971)
EIC1203	<=	EIC1101 & EIC1103	(216:77.7%, 0.944)
EIC1203	<=	EIC1101	(235:84.5%, 0.881)
EIC1203	<=	EIC1102 & EIC1100	(207:74.5%, 0.928)
EIC1203	<=	EIC1102 & EIC1103 & EIC1100	(198:71.2%, 0.97)
EIC1203	<=	EIC1102 & EIC1103	(213:76.6%, 0.93)
EIC1203	<=	EIC1102	(232:83.5%, 0.858)
EIC1203	<=	EIC1103 & EIC1100	(205:73.7%, 0.966)
EIC1203	<=	EIC1103	(229:82.4%, 0.921)
EIC1203	<=	EIC1104 & EIC1100	(208:74.8%, 0.938)
EIC1203	<=	EIC1104 & EIC1101 & EIC1100	(208:74.8%, 0.938)
EIC1203	<=	EIC1104 & EIC1101 & EIC1102 & EIC1100	(201:72.3%, 0.94)
EIC1203	<=	EIC1104 & EIC1101 & EIC1102 & EIC1103	(203:73.0%, 0.951)
EIC1203	<=	EIC1104 & EIC1101 & EIC1102	(211:75.9%, 0.915)
EIC1203	<=	EIC1104 & EIC1101 & EIC1103 & EIC1100	(199:71.6%, 0.97)
EIC1203	<=	EIC1104 & EIC1101 & EIC1103	(210:75.5%, 0.948)
EIC1203	<=	EIC1104 & EIC1101	(223:80.2%, 0.906)
EIC1203	<=	EIC1104 & EIC1102 & EIC1100	(201:72.3%, 0.94)
EIC1203	<=	EIC1104 & EIC1102 & EIC1103 & EIC1100	(195:70.1%, 0.969)
EIC1203	<=	EIC1104 & EIC1102 & EIC1103	(209:75.2%, 0.933)
EIC1203	<=	EIC1104 & EIC1102	(221:79.5%, 0.887)
EIC1203	<=	EIC1104 & EIC1103 & EIC1100	(199:71.6%, 0.97)
EIC1203	<=	EIC1104 & EIC1103	(220:79.1%, 0.927)
EIC1203	<=	EIC1104	(240:86.3%, 0.871)
EIC1203	<=	EIC1106 & EIC1100	(219:78.8%, 0.909)
EIC1203	<=	EIC1106 & EIC1101 & EIC1100	(216:77.7%, 0.921)
EIC1203	<=	EIC1106 & EIC1101 & EIC1102 & EIC1100	(204:73.4%, 0.936)
EIC1203	<=	EIC1106 & EIC1101 & EIC1102 & EIC1103	(206:74.1%, 0.947)
EIC1203	<=	EIC1106 & EIC1101 & EIC1102	(215:77.3%, 0.907)
EIC1203	<=	EIC1106 & EIC1101 & EIC1103 & EIC1100	(203:73.0%, 0.97)
EIC1203	<=	EIC1106 & EIC1101 & EIC1103	(215:77.3%, 0.944)
EIC1203	<=	EIC1106 & EIC1101	(231:83.1%, 0.892)
EIC1203	<=	EIC1106 & EIC1102 & EIC1100	(205:73.7%, 0.932)
EIC1203	<=	EIC1106 & EIC1102 & EIC1103 & EIC1100	(197:70.9%, 0.97)
EIC1203	<=	EIC1106 & EIC1102 & EIC1103	(212:76.3%, 0.929)
EIC1203	<=	EIC1106 & EIC1102	(228:82.0%, 0.868)
EIC1203	<=	EIC1106 & EIC1103 & EIC1100	(204:73.4%, 0.966)
EIC1203	<=	EIC1106 & EIC1103	(226:81.3%, 0.92)
EIC1203	<=	EIC1106 & EIC1104 & EIC1100	(206:74.1%, 0.942)
EIC1203	<=	EIC1106 & EIC1104 & EIC1101 & EIC1100	(206:74.1%, 0.942)
EIC1203	<=	EIC1106 & EIC1104 & EIC1101 & EIC1102	(209:75.2%, 0.919)
EIC1203	<=	EIC1106 & EIC1104 & EIC1101 & EIC1103	(209:75.2%, 0.947)
EIC1203	<=	EIC1106 & EIC1104 & EIC1101	(220:79.1%, 0.914)
EIC1203	<=	EIC1106 & EIC1104 & EIC1102 & EIC1100	(199:71.6%, 0.945)
EIC1203	<=	EIC1106 & EIC1104 & EIC1102 & EIC1103	(208:74.8%, 0.933)
EIC1203	<=	EIC1106 & EIC1104 & EIC1102	(218:78.4%, 0.894)
EIC1203	<=	EIC1106 & EIC1104 & EIC1103 & EIC1100	(198:71.2%, 0.97)
EIC1203	<=	EIC1106 & EIC1104 & EIC1103	(218:78.4%, 0.927)
EIC1203	<=	EIC1106 & EIC1104	(234:84.2%, 0.885)
EIC1203	<=	EIC1106	(260:93.5%, 0.819)

Figura III.5 - Regras de associação: disciplina EIC1203

Melhor suporte:	260:93.5%	(conf=81.9%)	
Melhor confiança:	97.1%	(sup=204:73.4%)	
EIC1203	<=	EIC1101 & EIC1103 & EIC1100	(204:73.4%, 0.971)
EIC1203	<=	EIC1106	(260:93.5%, 0.819)
EIC1203	<=	EIC1103 & EIC1100	(205:73.7%, 0.966)
EIC1203	<=	EIC1104	(240:86.3%, 0.871)
EIC1203	<=	EIC1104 & EIC1101 & EIC1103	(210:75.5%, 0.948)
EIC1203	<=	EIC1101	(235:84.5%, 0.881)
EIC1203	<=	EIC1101 & EIC1103	(216:77.7%, 0.944)
EIC1203	<=	EIC1106 & EIC1104	(234:84.2%, 0.885)
EIC1203	<=	EIC1104 & EIC1103	(220:79.1%, 0.927)
EIC1203	<=	EIC1106 & EIC1101	(231:83.1%, 0.892)
EIC1203	<=	EIC1103	(229:82.4%, 0.921)

Figura III.6 - Regras de associação: disciplina EIC1203 após o filtro

EIC1204	<=	EIC1100	(221:79.5%, 0.946)
EIC1204	<=	EIC1101 & EIC1100	(218:78.4%, 0.954)
EIC1204	<=	EIC1101 & EIC1102 & EIC1100	(206:74.1%, 0.956)
EIC1204	<=	EIC1101 & EIC1102 & EIC1103 & EIC1100	(198:71.2%, 0.955)
EIC1204	<=	EIC1101 & EIC1102 & EIC1103	(207:74.5%, 0.952)
EIC1204	<=	EIC1101 & EIC1102	(217:78.1%, 0.954)
EIC1204	<=	EIC1101 & EIC1103 & EIC1100	(204:73.4%, 0.956)
EIC1204	<=	EIC1101 & EIC1103	(216:77.7%, 0.94)
EIC1204	<=	EIC1101	(235:84.5%, 0.932)
EIC1204	<=	EIC1102 & EIC1100	(207:74.5%, 0.957)
EIC1204	<=	EIC1102 & EIC1103 & EIC1100	(198:71.2%, 0.955)
EIC1204	<=	EIC1102 & EIC1103	(213:76.6%, 0.948)
EIC1204	<=	EIC1102	(232:83.5%, 0.931)
EIC1204	<=	EIC1103 & EIC1100	(205:73.7%, 0.951)
EIC1204	<=	EIC1103	(229:82.4%, 0.917)
EIC1204	<=	EIC1104 & EIC1100	(208:74.8%, 0.962)
EIC1204	<=	EIC1104 & EIC1101 & EIC1100	(208:74.8%, 0.962)
EIC1204	<=	EIC1104 & EIC1101 & EIC1102 & EIC1100	(201:72.3%, 0.96)
EIC1204	<=	EIC1104 & EIC1101 & EIC1102 & EIC1103	(203:73.0%, 0.961)
EIC1204	<=	EIC1104 & EIC1101 & EIC1102	(211:75.9%, 0.962)
EIC1204	<=	EIC1104 & EIC1101 & EIC1103 & EIC1100	(199:71.6%, 0.96)
EIC1204	<=	EIC1104 & EIC1101 & EIC1103	(210:75.5%, 0.948)
EIC1204	<=	EIC1104 & EIC1101	(223:80.2%, 0.946)
EIC1204	<=	EIC1104 & EIC1102 & EIC1100	(201:72.3%, 0.96)
EIC1204	<=	EIC1104 & EIC1102 & EIC1103 & EIC1100	(195:70.1%, 0.959)
EIC1204	<=	EIC1104 & EIC1102 & EIC1103	(209:75.2%, 0.957)
EIC1204	<=	EIC1104 & EIC1102	(221:79.5%, 0.95)
EIC1204	<=	EIC1104 & EIC1103 & EIC1100	(199:71.6%, 0.96)
EIC1204	<=	EIC1104 & EIC1103	(220:79.1%, 0.932)
EIC1204	<=	EIC1104	(240:86.3%, 0.917)
EIC1204	<=	EIC1106 & EIC1100	(219:78.8%, 0.945)
EIC1204	<=	EIC1106 & EIC1101 & EIC1100	(216:77.7%, 0.954)
EIC1204	<=	EIC1106 & EIC1101 & EIC1102 & EIC1100	(204:73.4%, 0.956)
EIC1204	<=	EIC1106 & EIC1101 & EIC1102 & EIC1103	(206:74.1%, 0.951)
EIC1204	<=	EIC1106 & EIC1101 & EIC1102	(215:77.3%, 0.953)
EIC1204	<=	EIC1106 & EIC1101 & EIC1103 & EIC1100	(203:73.0%, 0.956)
EIC1204	<=	EIC1106 & EIC1101 & EIC1103	(215:77.3%, 0.94)
EIC1204	<=	EIC1106 & EIC1101	(231:83.1%, 0.939)
EIC1204	<=	EIC1106 & EIC1102 & EIC1100	(205:73.7%, 0.956)
EIC1204	<=	EIC1106 & EIC1102 & EIC1103 & EIC1100	(197:70.9%, 0.954)
EIC1204	<=	EIC1106 & EIC1102 & EIC1103	(212:76.3%, 0.948)
EIC1204	<=	EIC1106 & EIC1102	(228:82.0%, 0.934)
EIC1204	<=	EIC1106 & EIC1103 & EIC1100	(204:73.4%, 0.951)
EIC1204	<=	EIC1106 & EIC1103	(226:81.3%, 0.92)
EIC1204	<=	EIC1106 & EIC1104 & EIC1100	(206:74.1%, 0.961)
EIC1204	<=	EIC1106 & EIC1104 & EIC1101 & EIC1100	(206:74.1%, 0.961)
EIC1204	<=	EIC1106 & EIC1104 & EIC1101 & EIC1102	(209:75.2%, 0.962)
EIC1204	<=	EIC1106 & EIC1104 & EIC1101 & EIC1103	(209:75.2%, 0.947)
EIC1204	<=	EIC1106 & EIC1104 & EIC1101	(220:79.1%, 0.95)
EIC1204	<=	EIC1106 & EIC1104 & EIC1102 & EIC1100	(199:71.6%, 0.96)
EIC1204	<=	EIC1106 & EIC1104 & EIC1102 & EIC1103	(208:74.8%, 0.957)
EIC1204	<=	EIC1106 & EIC1104 & EIC1102	(218:78.4%, 0.954)
EIC1204	<=	EIC1106 & EIC1104 & EIC1103 & EIC1100	(198:71.2%, 0.96)
EIC1204	<=	EIC1106 & EIC1104 & EIC1103	(218:78.4%, 0.936)
EIC1204	<=	EIC1106 & EIC1104	(234:84.2%, 0.932)
EIC1204	<=	EIC1106	(260:93.5%, 0.885)

Figura III.7 - Regras de associação: disciplina EIC1204

Melhor suporte:	260:93.5%	(conf=88.5%)	
Melhor confiança:	96.2%	(sup=211:75.9%)	
EIC1204	<=	EIC1104 & EIC1101 & EIC1102	(211:75.9%, 0.962)
EIC1204	<=	EIC1106	(260:93.5%, 0.885)
EIC1204	<=	EIC1106 & EIC1104 & EIC1102	(218:78.4%, 0.954)
EIC1204	<=	EIC1101 & EIC1100	(218:78.4%, 0.954)
EIC1204	<=	EIC1104	(240:86.3%, 0.917)
EIC1204	<=	EIC1104 & EIC1102	(221:79.5%, 0.95)
EIC1204	<=	EIC1101	(235:84.5%, 0.932)
EIC1204	<=	EIC1104 & EIC1101	(223:80.2%, 0.946)
EIC1204	<=	EIC1106 & EIC1101	(231:83.1%, 0.939)

Figura III.8 - Regras de associação: disciplina EIC1204 após o filtro

EIC1205	<=	EIC1100	(221:79.5%, 0.982)
EIC1205	<=	EIC1101 & EIC1100	(218:78.4%, 0.982)
EIC1205	<=	EIC1101 & EIC1102 & EIC1100	(206:74.1%, 0.985)
EIC1205	<=	EIC1101 & EIC1102 & EIC1103 & EIC1100	(198:71.2%, 0.985)
EIC1205	<=	EIC1101 & EIC1102 & EIC1103	(207:74.5%, 0.981)
EIC1205	<=	EIC1101 & EIC1102	(217:78.1%, 0.982)
EIC1205	<=	EIC1101 & EIC1103 & EIC1100	(204:73.4%, 0.985)
EIC1205	<=	EIC1101 & EIC1103	(216:77.7%, 0.977)
EIC1205	<=	EIC1101	(235:84.5%, 0.97)
EIC1205	<=	EIC1102 & EIC1100	(207:74.5%, 0.986)
EIC1205	<=	EIC1102 & EIC1103 & EIC1100	(198:71.2%, 0.985)
EIC1205	<=	EIC1102 & EIC1103	(213:76.6%, 0.977)
EIC1205	<=	EIC1102	(232:83.5%, 0.966)
EIC1205	<=	EIC1103 & EIC1100	(205:73.7%, 0.985)
EIC1205	<=	EIC1103	(229:82.4%, 0.965)
EIC1205	<=	EIC1104 & EIC1100	(208:74.8%, 0.981)
EIC1205	<=	EIC1104 & EIC1101 & EIC1100	(208:74.8%, 0.981)
EIC1205	<=	EIC1104 & EIC1101 & EIC1102 & EIC1100	(201:72.3%, 0.985)
EIC1205	<=	EIC1104 & EIC1101 & EIC1102 & EIC1103	(203:73.0%, 0.985)
EIC1205	<=	EIC1104 & EIC1101 & EIC1102	(211:75.9%, 0.986)
EIC1205	<=	EIC1104 & EIC1101 & EIC1103 & EIC1100	(199:71.6%, 0.985)
EIC1205	<=	EIC1104 & EIC1101 & EIC1103	(210:75.5%, 0.981)
EIC1205	<=	EIC1104 & EIC1101	(223:80.2%, 0.978)
EIC1205	<=	EIC1104 & EIC1102 & EIC1100	(201:72.3%, 0.985)
EIC1205	<=	EIC1104 & EIC1102 & EIC1103 & EIC1100	(195:70.1%, 0.985)
EIC1205	<=	EIC1104 & EIC1102 & EIC1103	(209:75.2%, 0.981)
EIC1205	<=	EIC1104 & EIC1102	(221:79.5%, 0.968)
EIC1205	<=	EIC1104 & EIC1103 & EIC1100	(199:71.6%, 0.985)
EIC1205	<=	EIC1104 & EIC1103	(220:79.1%, 0.968)
EIC1205	<=	EIC1104	(240:86.3%, 0.95)
EIC1205	<=	EIC1106 & EIC1100	(219:78.8%, 0.982)
EIC1205	<=	EIC1106 & EIC1101 & EIC1100	(216:77.7%, 0.981)
EIC1205	<=	EIC1106 & EIC1101 & EIC1102 & EIC1100	(204:73.4%, 0.985)
EIC1205	<=	EIC1106 & EIC1101 & EIC1102 & EIC1103	(206:74.1%, 0.981)
EIC1205	<=	EIC1106 & EIC1101 & EIC1102	(215:77.3%, 0.981)
EIC1205	<=	EIC1106 & EIC1101 & EIC1103 & EIC1100	(203:73.0%, 0.985)
EIC1205	<=	EIC1106 & EIC1101 & EIC1103	(215:77.3%, 0.977)
EIC1205	<=	EIC1106 & EIC1101	(231:83.1%, 0.974)
EIC1205	<=	EIC1106 & EIC1102 & EIC1100	(205:73.7%, 0.985)
EIC1205	<=	EIC1106 & EIC1102 & EIC1103 & EIC1100	(197:70.9%, 0.985)
EIC1205	<=	EIC1106 & EIC1102 & EIC1103	(212:76.3%, 0.976)
EIC1205	<=	EIC1106 & EIC1102	(228:82.0%, 0.969)
EIC1205	<=	EIC1106 & EIC1103 & EIC1100	(204:73.4%, 0.985)
EIC1205	<=	EIC1106 & EIC1103	(226:81.3%, 0.969)
EIC1205	<=	EIC1106 & EIC1104 & EIC1100	(206:74.1%, 0.981)
EIC1205	<=	EIC1106 & EIC1104 & EIC1101 & EIC1100	(206:74.1%, 0.981)
EIC1205	<=	EIC1106 & EIC1104 & EIC1101 & EIC1102	(209:75.2%, 0.986)
EIC1205	<=	EIC1106 & EIC1104 & EIC1101 & EIC1103	(209:75.2%, 0.981)
EIC1205	<=	EIC1106 & EIC1104 & EIC1101	(220:79.1%, 0.977)
EIC1205	<=	EIC1106 & EIC1104 & EIC1102 & EIC1100	(199:71.6%, 0.985)
EIC1205	<=	EIC1106 & EIC1104 & EIC1102 & EIC1103	(208:74.8%, 0.981)
EIC1205	<=	EIC1106 & EIC1104 & EIC1102	(218:78.4%, 0.972)
EIC1205	<=	EIC1106 & EIC1104 & EIC1103 & EIC1100	(198:71.2%, 0.985)
EIC1205	<=	EIC1106 & EIC1104 & EIC1103	(218:78.4%, 0.972)
EIC1205	<=	EIC1106 & EIC1104	(234:84.2%, 0.962)
EIC1205	<=	EIC1106	(260:93.5%, 0.938)

Figura III.9 - Regras de associação: disciplina EIC1205

Melhor suporte:	260:93.5%	{conf=93.8%}	
Melhor confiança:	98.6%	{sup=211:75.9%}	
EIC1205	<=	EIC1104 & EIC1101 & EIC1102	(211:75.9%, 0.986)
EIC1205	<=	EIC1106	(260:93.5%, 0.938)
EIC1205	<=	EIC1100	(221:79.5%, 0.982)
EIC1205	<=	EIC1104	(240:86.3%, 0.95)
EIC1205	<=	EIC1104 & EIC1101	(223:80.2%, 0.978)
EIC1205	<=	EIC1101	(235:84.5%, 0.97)
EIC1205	<=	EIC1106 & EIC1101	(231:83.1%, 0.974)

Figura III.10 - Regras de associação: disciplina EIC1205 após o filtro

EIC1206	<=	EIC1100	(221:79.5%, 0.941)
EIC1206	<=	EIC1101 & EIC1100	(218:78.4%, 0.945)
EIC1206	<=	EIC1101 & EIC1102 & EIC1100	(206:74.1%, 0.947)
EIC1206	<=	EIC1101 & EIC1102 & EIC1103 & EIC1100	(198:71.2%, 0.96)
EIC1206	<=	EIC1101 & EIC1102 & EIC1103	(207:74.5%, 0.947)
EIC1206	<=	EIC1101 & EIC1102	(217:78.1%, 0.931)
EIC1206	<=	EIC1101 & EIC1103 & EIC1100	(204:73.4%, 0.956)
EIC1206	<=	EIC1101 & EIC1103	(216:77.7%, 0.94)
EIC1206	<=	EIC1101	(235:84.5%, 0.923)
EIC1206	<=	EIC1102 & EIC1100	(207:74.5%, 0.947)
EIC1206	<=	EIC1102 & EIC1103 & EIC1100	(198:71.2%, 0.96)
EIC1206	<=	EIC1102 & EIC1103	(213:76.6%, 0.944)
EIC1206	<=	EIC1102	(232:83.5%, 0.909)
EIC1206	<=	EIC1103 & EIC1100	(205:73.7%, 0.956)
EIC1206	<=	EIC1103	(229:82.4%, 0.917)
EIC1206	<=	EIC1104 & EIC1100	(208:74.8%, 0.957)
EIC1206	<=	EIC1104 & EIC1101 & EIC1100	(208:74.8%, 0.957)
EIC1206	<=	EIC1104 & EIC1101 & EIC1102 & EIC1100	(201:72.3%, 0.955)
EIC1206	<=	EIC1104 & EIC1101 & EIC1102 & EIC1103	(203:73.0%, 0.956)
EIC1206	<=	EIC1104 & EIC1101 & EIC1102	(211:75.9%, 0.943)
EIC1206	<=	EIC1104 & EIC1101 & EIC1103 & EIC1100	(199:71.6%, 0.965)
EIC1206	<=	EIC1104 & EIC1101 & EIC1103	(210:75.5%, 0.952)
EIC1206	<=	EIC1104 & EIC1101	(223:80.2%, 0.937)
EIC1206	<=	EIC1104 & EIC1102 & EIC1100	(201:72.3%, 0.955)
EIC1206	<=	EIC1104 & EIC1102 & EIC1103 & EIC1100	(195:70.1%, 0.964)
EIC1206	<=	EIC1104 & EIC1102 & EIC1103	(209:75.2%, 0.952)
EIC1206	<=	EIC1104	(221:79.5%, 0.932)
EIC1206	<=	EIC1104 & EIC1103 & EIC1100	(199:71.6%, 0.965)
EIC1206	<=	EIC1104 & EIC1103	(220:79.1%, 0.936)
EIC1206	<=	EIC1104	(240:86.3%, 0.913)
EIC1206	<=	EIC1106 & EIC1100	(219:78.8%, 0.945)
EIC1206	<=	EIC1106 & EIC1101 & EIC1100	(216:77.7%, 0.949)
EIC1206	<=	EIC1106 & EIC1101 & EIC1102 & EIC1100	(204:73.4%, 0.951)
EIC1206	<=	EIC1106 & EIC1101 & EIC1102 & EIC1103	(206:74.1%, 0.947)
EIC1206	<=	EIC1106 & EIC1101 & EIC1102	(215:77.3%, 0.935)
EIC1206	<=	EIC1106 & EIC1101 & EIC1103 & EIC1100	(203:73.0%, 0.956)
EIC1206	<=	EIC1106 & EIC1101 & EIC1103	(215:77.3%, 0.94)
EIC1206	<=	EIC1106 & EIC1101	(231:83.1%, 0.931)
EIC1206	<=	EIC1106 & EIC1102 & EIC1100	(205:73.7%, 0.951)
EIC1206	<=	EIC1106 & EIC1102 & EIC1103 & EIC1100	(197:70.9%, 0.959)
EIC1206	<=	EIC1106 & EIC1102 & EIC1103	(212:76.3%, 0.943)
EIC1206	<=	EIC1106 & EIC1102	(228:82.0%, 0.921)
EIC1206	<=	EIC1106 & EIC1103 & EIC1100	(204:73.4%, 0.956)
EIC1206	<=	EIC1106 & EIC1103	(226:81.3%, 0.92)
EIC1206	<=	EIC1106 & EIC1104 & EIC1100	(206:74.1%, 0.961)
EIC1206	<=	EIC1106 & EIC1104 & EIC1101 & EIC1100	(206:74.1%, 0.961)
EIC1206	<=	EIC1106 & EIC1104 & EIC1101 & EIC1102	(209:75.2%, 0.947)
EIC1206	<=	EIC1106 & EIC1104 & EIC1101 & EIC1103	(209:75.2%, 0.952)
EIC1206	<=	EIC1106 & EIC1104 & EIC1101	(220:79.1%, 0.945)
EIC1206	<=	EIC1106 & EIC1104 & EIC1102 & EIC1100	(199:71.6%, 0.96)
EIC1206	<=	EIC1106 & EIC1104 & EIC1102 & EIC1103	(208:74.8%, 0.952)
EIC1206	<=	EIC1106 & EIC1104 & EIC1102	(218:78.4%, 0.94)
EIC1206	<=	EIC1106 & EIC1104 & EIC1103 & EIC1100	(198:71.2%, 0.965)
EIC1206	<=	EIC1106 & EIC1104 & EIC1103	(218:78.4%, 0.936)
EIC1206	<=	EIC1106 & EIC1104	(234:84.2%, 0.927)
EIC1206	<=	EIC1106	(260:93.5%, 0.877)

Figura III.11 - Regras de associação: disciplina EIC1206

Melhor suporte: 260:93.5% (conf=87.7%)			
Melhor confiança: 96.5% (sup=199:71.6%)			
EIC1206	<=	EIC1104 & EIC1101 & EIC1103 & EIC1100	(199:71.6%, 0.965)
EIC1206	<=	EIC1104 & EIC1103 & EIC1100	(199:71.6%, 0.965)
EIC1206	<=	EIC1106	(260:93.5%, 0.877)
EIC1206	<=	EIC1106 & EIC1104 & EIC1101 & EIC1100	(206:74.1%, 0.961)
EIC1206	<=	EIC1106 & EIC1104 & EIC1100	(206:74.1%, 0.961)
EIC1206	<=	EIC1104	(240:86.3%, 0.913)
EIC1206	<=	EIC1104 & EIC1100	(208:74.8%, 0.957)
EIC1206	<=	EIC1104 & EIC1101 & EIC1100	(208:74.8%, 0.957)
EIC1206	<=	EIC1101	(235:84.5%, 0.923)
EIC1206	<=	EIC1104 & EIC1101 & EIC1103	(210:75.5%, 0.952)
EIC1206	<=	EIC1106 & EIC1104	(234:84.2%, 0.927)
EIC1206	<=	EIC1106 & EIC1101 & EIC1100	(216:77.7%, 0.949)
EIC1206	<=	EIC1106 & EIC1101	(231:83.1%, 0.931)
EIC1206	<=	EIC1106 & EIC1104 & EIC1101	(220:79.1%, 0.945)
EIC1206	<=	EIC1104 & EIC1101	(223:80.2%, 0.937)
EIC1206	<=	EIC1100	(221:79.5%, 0.941)

Figura III.12 - Regras de associação: disciplina EIC1206 após o filtro

EIC2200	<=	EIC2100	(149:69.6%, 0.946)
EIC2200	<=	EIC2101 & EIC2100	(141:65.9%, 0.965)
EIC2200	<=	EIC2101 & EIC2102 & EIC2100	(132:61.7%, 0.992)
EIC2200	<=	EIC2101 & EIC2102	(143:66.8%, 0.965)
EIC2200	<=	EIC2101	(159:74.3%, 0.918)
EIC2200	<=	EIC2102 & EIC2100	(137:64.0%, 0.978)
EIC2200	<=	EIC2102	(159:74.3%, 0.937)
EIC2200	<=	EIC2103 & EIC2100	(144:67.3%, 0.951)
EIC2200	<=	EIC2103 & EIC2101 & EIC2100	(138:64.5%, 0.964)
EIC2200	<=	EIC2103 & EIC2101 & EIC2102 & EIC2100	(130:60.7%, 0.992)
EIC2200	<=	EIC2103 & EIC2101 & EIC2102	(140:65.4%, 0.964)
EIC2200	<=	EIC2103 & EIC2101	(151:70.6%, 0.934)
EIC2200	<=	EIC2103 & EIC2102 & EIC2100	(133:62.1%, 0.985)
EIC2200	<=	EIC2103 & EIC2102	(146:68.2%, 0.959)
EIC2200	<=	EIC2103	(161:75.2%, 0.925)
EIC2200	<=	EIC2104 & EIC2100	(145:67.8%, 0.952)
EIC2200	<=	EIC2104 & EIC2101 & EIC2100	(138:64.5%, 0.964)
EIC2200	<=	EIC2104 & EIC2101 & EIC2102 & EIC2100	(130:60.7%, 0.992)
EIC2200	<=	EIC2104 & EIC2101 & EIC2102	(140:65.4%, 0.971)
EIC2200	<=	EIC2104 & EIC2101	(152:71.0%, 0.934)
EIC2200	<=	EIC2104 & EIC2102 & EIC2100	(134:62.6%, 0.985)
EIC2200	<=	EIC2104 & EIC2102	(151:70.6%, 0.947)
EIC2200	<=	EIC2104 & EIC2103 & EIC2100	(141:65.9%, 0.95)
EIC2200	<=	EIC2104 & EIC2103 & EIC2101 & EIC2100	(135:63.1%, 0.963)
EIC2200	<=	EIC2104 & EIC2103 & EIC2101 & EIC2102	(137:64.0%, 0.971)
EIC2200	<=	EIC2104 & EIC2103 & EIC2101	(145:67.8%, 0.945)
EIC2200	<=	EIC2104 & EIC2103 & EIC2102 & EIC2100	(131:61.2%, 0.985)
EIC2200	<=	EIC2104 & EIC2103 & EIC2102	(142:66.4%, 0.965)
EIC2200	<=	EIC2104 & EIC2103	(154:72.0%, 0.935)
EIC2200	<=	EIC2104	(174:81.3%, 0.902)
EIC2200	<=	EIC2105 & EIC2100	(147:68.7%, 0.952)
EIC2200	<=	EIC2105 & EIC2101 & EIC2100	(141:65.9%, 0.965)
EIC2200	<=	EIC2105 & EIC2101 & EIC2102 & EIC2100	(132:61.7%, 0.992)
EIC2200	<=	EIC2105 & EIC2101 & EIC2102	(143:66.8%, 0.965)
EIC2200	<=	EIC2105 & EIC2101	(156:72.9%, 0.936)
EIC2200	<=	EIC2105 & EIC2102 & EIC2100	(136:63.6%, 0.985)
EIC2200	<=	EIC2105 & EIC2102	(152:71.0%, 0.947)
EIC2200	<=	EIC2105 & EIC2103 & EIC2100	(143:66.8%, 0.951)
EIC2200	<=	EIC2105 & EIC2103 & EIC2101 & EIC2100	(138:64.5%, 0.964)
EIC2200	<=	EIC2105 & EIC2103 & EIC2101 & EIC2102	(140:65.4%, 0.964)
EIC2200	<=	EIC2105 & EIC2103 & EIC2101	(150:70.1%, 0.94)
EIC2200	<=	EIC2105 & EIC2103 & EIC2102 & EIC2100	(133:62.1%, 0.985)
EIC2200	<=	EIC2105 & EIC2103 & EIC2102	(144:67.3%, 0.958)
EIC2200	<=	EIC2105 & EIC2103	(157:73.4%, 0.93)
EIC2200	<=	EIC2105 & EIC2104 & EIC2100	(144:67.3%, 0.951)
EIC2200	<=	EIC2105 & EIC2104 & EIC2101 & EIC2100	(138:64.5%, 0.964)
EIC2200	<=	EIC2105 & EIC2104 & EIC2101 & EIC2102	(140:65.4%, 0.971)
EIC2200	<=	EIC2105 & EIC2104 & EIC2101	(151:70.6%, 0.94)
EIC2200	<=	EIC2105 & EIC2104 & EIC2102 & EIC2100	(134:62.6%, 0.985)
EIC2200	<=	EIC2105 & EIC2104 & EIC2102	(149:69.6%, 0.953)
EIC2200	<=	EIC2105 & EIC2104 & EIC2103 & EIC2100	(140:65.4%, 0.95)
EIC2200	<=	EIC2105 & EIC2104 & EIC2103 & EIC2101	(145:67.8%, 0.945)
EIC2200	<=	EIC2105 & EIC2104 & EIC2103 & EIC2102	(141:65.9%, 0.965)
EIC2200	<=	EIC2105 & EIC2104 & EIC2103	(152:71.0%, 0.934)
EIC2200	<=	EIC2105 & EIC2104	(168:78.5%, 0.923)
EIC2200	<=	EIC2105	(179:83.6%, 0.905)

Figura III.13 - Regras de associação: disciplina EIC2200 (sem.)

Melhor suporte:	179:83.6% (conf=90.5%)		
Melhor confiança:	99.2% (sup=132:61.7%)		
EIC2200	<=	EIC2101 & EIC2102 & EIC2100	(132:61.7%, 0.992)
EIC2200	<=	EIC2105 & EIC2101 & EIC2102 & EIC2100	(132:61.7%, 0.992)
EIC2200	<=	EIC2105	(179:83.6%, 0.905)
EIC2200	<=	EIC2105 & EIC2102 & EIC2100	(136:63.6%, 0.985)
EIC2200	<=	EIC2105 & EIC2104	(168:78.5%, 0.923)
EIC2200	<=	EIC2102 & EIC2100	(137:64.0%, 0.978)
EIC2200	<=	EIC2103	(161:75.2%, 0.925)
EIC2200	<=	EIC2105 & EIC2104 & EIC2101 & EIC2102	(140:65.4%, 0.971)
EIC2200	<=	EIC2104 & EIC2101 & EIC2102	(140:65.4%, 0.971)
EIC2200	<=	EIC2102	(159:74.3%, 0.937)
EIC2200	<=	EIC2101 & EIC2102	(143:66.8%, 0.965)
EIC2200	<=	EIC2105 & EIC2101 & EIC2102	(143:66.8%, 0.965)
EIC2200	<=	EIC2105 & EIC2102	(152:71.0%, 0.947)
EIC2200	<=	EIC2103 & EIC2102	(146:68.2%, 0.959)
EIC2200	<=	EIC2105 & EIC2104 & EIC2102	(149:69.6%, 0.953)

Figura III.14 - Regras de associação: disciplina EIC2200 (sem.) após o filtro

EIC2201	<=	EIC2100	(149:69.6%, 0.966)
EIC2201	<=	EIC2101 & EIC2100	(141:65.9%, 0.986)
EIC2201	<=	EIC2101 & EIC2102 & EIC2100	(132:61.7%, 0.992)
EIC2201	<=	EIC2101 & EIC2102	(143:66.8%, 0.979)
EIC2201	<=	EIC2101	(159:74.3%, 0.969)
EIC2201	<=	EIC2102 & EIC2100	(137:64.0%, 0.978)
EIC2201	<=	EIC2102	(159:74.3%, 0.943)
EIC2201	<=	EIC2103 & EIC2100	(144:67.3%, 0.972)
EIC2201	<=	EIC2103 & EIC2101 & EIC2100	(138:64.5%, 0.986)
EIC2201	<=	EIC2103 & EIC2101 & EIC2102 & EIC2100	(130:60.7%, 0.992)
EIC2201	<=	EIC2103 & EIC2101 & EIC2102	(140:65.4%, 0.979)
EIC2201	<=	EIC2103 & EIC2101	(151:70.6%, 0.967)
EIC2201	<=	EIC2103 & EIC2102 & EIC2100	(133:62.1%, 0.985)
EIC2201	<=	EIC2103 & EIC2102	(146:68.2%, 0.973)
EIC2201	<=	EIC2103	(161:75.2%, 0.957)
EIC2201	<=	EIC2104 & EIC2100	(145:67.8%, 0.972)
EIC2201	<=	EIC2104 & EIC2101 & EIC2100	(138:64.5%, 0.986)
EIC2201	<=	EIC2104 & EIC2101 & EIC2102 & EIC2100	(130:60.7%, 0.992)
EIC2201	<=	EIC2104 & EIC2101 & EIC2102	(140:65.4%, 0.979)
EIC2201	<=	EIC2104 & EIC2101	(152:71.0%, 0.967)
EIC2201	<=	EIC2104 & EIC2102 & EIC2100	(134:62.6%, 0.985)
EIC2201	<=	EIC2104 & EIC2102	(151:70.6%, 0.967)
EIC2201	<=	EIC2104 & EIC2103 & EIC2100	(141:65.9%, 0.972)
EIC2201	<=	EIC2104 & EIC2103 & EIC2101 & EIC2100	(135:63.1%, 0.985)
EIC2201	<=	EIC2104 & EIC2103 & EIC2101 & EIC2102	(137:64.0%, 0.978)
EIC2201	<=	EIC2104 & EIC2103 & EIC2101	(145:67.8%, 0.966)
EIC2201	<=	EIC2104 & EIC2103 & EIC2102 & EIC2100	(131:61.2%, 0.985)
EIC2201	<=	EIC2104 & EIC2103 & EIC2102	(142:66.4%, 0.972)
EIC2201	<=	EIC2104 & EIC2103	(154:72.0%, 0.955)
EIC2201	<=	EIC2104	(174:81.3%, 0.937)
EIC2201	<=	EIC2105 & EIC2100	(147:68.7%, 0.98)
EIC2201	<=	EIC2105 & EIC2101 & EIC2100	(141:65.9%, 0.986)
EIC2201	<=	EIC2105 & EIC2101 & EIC2102 & EIC2100	(132:61.7%, 0.992)
EIC2201	<=	EIC2105 & EIC2101 & EIC2102	(143:66.8%, 0.979)
EIC2201	<=	EIC2105 & EIC2101	(156:72.9%, 0.968)
EIC2201	<=	EIC2105 & EIC2102 & EIC2100	(136:63.6%, 0.985)
EIC2201	<=	EIC2105 & EIC2102	(152:71.0%, 0.967)
EIC2201	<=	EIC2105 & EIC2103 & EIC2100	(143:66.8%, 0.979)
EIC2201	<=	EIC2105 & EIC2103 & EIC2101 & EIC2100	(138:64.5%, 0.986)
EIC2201	<=	EIC2105 & EIC2103 & EIC2101 & EIC2102	(140:65.4%, 0.979)
EIC2201	<=	EIC2105 & EIC2103 & EIC2101	(150:70.1%, 0.967)
EIC2201	<=	EIC2105 & EIC2103 & EIC2102 & EIC2100	(133:62.1%, 0.985)
EIC2201	<=	EIC2105 & EIC2103 & EIC2102	(144:67.3%, 0.972)
EIC2201	<=	EIC2105 & EIC2103	(157:73.4%, 0.962)
EIC2201	<=	EIC2105 & EIC2104 & EIC2100	(144:67.3%, 0.979)
EIC2201	<=	EIC2105 & EIC2104 & EIC2101 & EIC2100	(138:64.5%, 0.986)
EIC2201	<=	EIC2105 & EIC2104 & EIC2101 & EIC2102	(140:65.4%, 0.979)
EIC2201	<=	EIC2105 & EIC2104 & EIC2101	(151:70.6%, 0.967)
EIC2201	<=	EIC2105 & EIC2104 & EIC2102 & EIC2100	(134:62.6%, 0.985)
EIC2201	<=	EIC2105 & EIC2104 & EIC2102	(149:69.6%, 0.966)
EIC2201	<=	EIC2105 & EIC2104 & EIC2103 & EIC2100	(140:65.4%, 0.979)
EIC2201	<=	EIC2105 & EIC2104 & EIC2103 & EIC2101	(145:67.8%, 0.966)
EIC2201	<=	EIC2105 & EIC2104 & EIC2103 & EIC2102	(141:65.9%, 0.972)
EIC2201	<=	EIC2105 & EIC2104 & EIC2103	(152:71.0%, 0.961)
EIC2201	<=	EIC2105 & EIC2104	(168:78.5%, 0.946)
EIC2201	<=	EIC2105	(179:83.6%, 0.933)

**Figura III.15 - Regras de associação: disciplina EIC2201 (sem.)**

Melhor suporte:	179:83.6% (conf=93.3%)		
Melhor confiança:	99.2% (sup=132:61.7%)		
EIC2201	<=	EIC2101 & EIC2102 & EIC2100	(132:61.7%, 0.992)
EIC2201	<=	EIC2105 & EIC2101 & EIC2102 & EIC2100	(132:61.7%, 0.992)
EIC2201	<=	EIC2105	(179:83.6%, 0.933)
EIC2201	<=	EIC2105 & EIC2101 & EIC2100	(141:65.9%, 0.986)
EIC2201	<=	EIC2101 & EIC2100	(141:65.9%, 0.986)
EIC2201	<=	EIC2104	(174:81.3%, 0.937)
EIC2201	<=	EIC2105 & EIC2100	(147:68.7%, 0.98)
EIC2201	<=	EIC2105 & EIC2104	(168:78.5%, 0.946)
EIC2201	<=	EIC2101	(159:74.3%, 0.969)
EIC2201	<=	EIC2103	(161:75.2%, 0.957)

**Figura III.16 - Regras de associação: disciplina EIC2201 (sem.) após o filtro**

EIC2202	<=	EIC2100	(149:69.6%, 0.913)
EIC2202	<=	EIC2101 & EIC2100	(141:65.9%, 0.929)
EIC2202	<=	EIC2101 & EIC2102 & EIC2100	(132:61.7%, 0.985)
EIC2202	<=	EIC2101 & EIC2102	(143:66.8%, 0.958)
EIC2202	<=	EIC2101	(159:74.3%, 0.868)
EIC2202	<=	EIC2102 & EIC2100	(137:64.0%, 0.971)
EIC2202	<=	EIC2102	(159:74.3%, 0.931)
EIC2202	<=	EIC2103 & EIC2100	(144:67.3%, 0.924)
EIC2202	<=	EIC2103 & EIC2101 & EIC2100	(138:64.5%, 0.935)
EIC2202	<=	EIC2103 & EIC2101 & EIC2102 & EIC2100	(130:60.7%, 0.985)
EIC2202	<=	EIC2103 & EIC2101 & EIC2102	(140:65.4%, 0.957)
EIC2202	<=	EIC2103 & EIC2101	(151:70.6%, 0.894)
EIC2202	<=	EIC2103 & EIC2102 & EIC2100	(133:62.1%, 0.977)
EIC2202	<=	EIC2103 & EIC2102	(146:68.2%, 0.952)
EIC2202	<=	EIC2103	(161:75.2%, 0.882)
EIC2202	<=	EIC2104 & EIC2100	(145:67.8%, 0.924)
EIC2202	<=	EIC2104 & EIC2101 & EIC2100	(138:64.5%, 0.935)
EIC2202	<=	EIC2104 & EIC2101 & EIC2102 & EIC2100	(130:60.7%, 0.985)
EIC2202	<=	EIC2104 & EIC2101 & EIC2102	(140:65.4%, 0.964)
EIC2202	<=	EIC2104 & EIC2101	(152:71.0%, 0.895)
EIC2202	<=	EIC2104 & EIC2102 & EIC2100	(134:62.6%, 0.978)
EIC2202	<=	EIC2104 & EIC2102	(151:70.6%, 0.954)
EIC2202	<=	EIC2104 & EIC2103 & EIC2100	(141:65.9%, 0.929)
EIC2202	<=	EIC2104 & EIC2103 & EIC2101 & EIC2100	(135:63.1%, 0.941)
EIC2202	<=	EIC2104 & EIC2103 & EIC2101 & EIC2102	(137:64.0%, 0.964)
EIC2202	<=	EIC2104 & EIC2103 & EIC2101	(145:67.8%, 0.917)
EIC2202	<=	EIC2104 & EIC2103 & EIC2102 & EIC2100	(131:61.2%, 0.977)
EIC2202	<=	EIC2104 & EIC2103 & EIC2102	(142:66.4%, 0.958)
EIC2202	<=	EIC2104 & EIC2103	(154:72.0%, 0.903)
EIC2202	<=	EIC2104	(174:81.3%, 0.845)
EIC2202	<=	EIC2105 & EIC2100	(147:68.7%, 0.925)
EIC2202	<=	EIC2105 & EIC2101 & EIC2100	(141:65.9%, 0.929)
EIC2202	<=	EIC2105 & EIC2101 & EIC2102 & EIC2100	(132:61.7%, 0.985)
EIC2202	<=	EIC2105 & EIC2101 & EIC2102	(143:66.8%, 0.958)
EIC2202	<=	EIC2105 & EIC2101	(156:72.9%, 0.885)
EIC2202	<=	EIC2105 & EIC2102 & EIC2100	(136:63.6%, 0.978)
EIC2202	<=	EIC2105 & EIC2102	(152:71.0%, 0.947)
EIC2202	<=	EIC2105 & EIC2103 & EIC2100	(143:66.8%, 0.93)
EIC2202	<=	EIC2105 & EIC2103 & EIC2101 & EIC2100	(138:64.5%, 0.935)
EIC2202	<=	EIC2105 & EIC2103 & EIC2101 & EIC2102	(140:65.4%, 0.957)
EIC2202	<=	EIC2105 & EIC2103 & EIC2101	(150:70.1%, 0.9)
EIC2202	<=	EIC2105 & EIC2103 & EIC2102 & EIC2100	(133:62.1%, 0.977)
EIC2202	<=	EIC2105 & EIC2103 & EIC2102	(144:67.3%, 0.951)
EIC2202	<=	EIC2105 & EIC2103	(157:73.4%, 0.892)
EIC2202	<=	EIC2105 & EIC2104 & EIC2100	(144:67.3%, 0.931)
EIC2202	<=	EIC2105 & EIC2104 & EIC2101 & EIC2100	(138:64.5%, 0.935)
EIC2202	<=	EIC2105 & EIC2104 & EIC2101 & EIC2102	(140:65.4%, 0.964)
EIC2202	<=	EIC2105 & EIC2104 & EIC2101	(151:70.6%, 0.901)
EIC2202	<=	EIC2105 & EIC2104 & EIC2102 & EIC2100	(134:62.6%, 0.978)
EIC2202	<=	EIC2105 & EIC2104 & EIC2102	(149:69.6%, 0.953)
EIC2202	<=	EIC2105 & EIC2104 & EIC2103 & EIC2100	(140:65.4%, 0.936)
EIC2202	<=	EIC2105 & EIC2104 & EIC2103 & EIC2101	(145:67.8%, 0.917)
EIC2202	<=	EIC2105 & EIC2104 & EIC2103 & EIC2102	(141:65.9%, 0.957)
EIC2202	<=	EIC2105 & EIC2104 & EIC2103	(152:71.0%, 0.908)
EIC2202	<=	EIC2105 & EIC2104	(168:78.5%, 0.863)
EIC2202	<=	EIC2105	(179:83.6%, 0.821)

Figura III.17 - Regras de associação: disciplina EIC2202 (sem.)

Melhor suporte:	179:83.6%	(conf=82.1%)	
Melhor confiança:	98.5%	(sup=132:61.7%)	
EIC2202	<=	EIC2101 & EIC2102 & EIC2100	(132:61.7%, 0.985)
EIC2202	<=	EIC2105 & EIC2101 & EIC2102 & EIC2100	(132:61.7%, 0.985)
EIC2202	<=	EIC2105	(179:83.6%, 0.821)
EIC2202	<=	EIC2105 & EIC2102 & EIC2100	(136:63.6%, 0.978)
EIC2202	<=	EIC2104	(174:81.3%, 0.845)
EIC2202	<=	EIC2102 & EIC2100	(137:64.0%, 0.971)
EIC2202	<=	EIC2105 & EIC2104	(168:78.5%, 0.863)
EIC2202	<=	EIC2105 & EIC2104 & EIC2101 & EIC2102	(140:65.4%, 0.964)
EIC2202	<=	EIC2104 & EIC2101 & EIC2102	(140:65.4%, 0.964)
EIC2202	<=	EIC2103	(161:75.2%, 0.882)
EIC2202	<=	EIC2101 & EIC2102	(143:66.8%, 0.958)
EIC2202	<=	EIC2105 & EIC2101 & EIC2102	(143:66.8%, 0.958)
EIC2202	<=	EIC2102	(159:74.3%, 0.931)
EIC2202	<=	EIC2104 & EIC2102	(151:70.6%, 0.954)
EIC2202	<=	EIC2105 & EIC2102	(152:71.0%, 0.947)

Figura III.18 - Regras de associação: disciplina EIC2202 (sem.) após o filtro

EIC2203	<=	EIC2100	{149:69.6%, 0.987}
EIC2203	<=	EIC2101 & EIC2100	{141:65.9%, 0.993}
EIC2203	<=	EIC2101 & EIC2102 & EIC2100	{132:61.7%, 1.0}
EIC2203	<=	EIC2101 & EIC2102	{143:66.8%, 0.993}
EIC2203	<=	EIC2101	{159:74.3%, 0.969}
EIC2203	<=	EIC2102 & EIC2100	{137:64.0%, 0.993}
EIC2203	<=	EIC2102	{159:74.3%, 0.95}
EIC2203	<=	EIC2103 & EIC2100	{144:67.3%, 0.993}
EIC2203	<=	EIC2103 & EIC2101 & EIC2100	{138:64.5%, 0.993}
EIC2203	<=	EIC2103 & EIC2101 & EIC2102 & EIC2100	{130:60.7%, 1.0}
EIC2203	<=	EIC2103 & EIC2101 & EIC2102	{140:65.4%, 0.993}
EIC2203	<=	EIC2103 & EIC2101	{151:70.6%, 0.98}
EIC2203	<=	EIC2103 & EIC2102 & EIC2100	{133:62.1%, 1.0}
EIC2203	<=	EIC2103 & EIC2102	{146:68.2%, 0.993}
EIC2203	<=	EIC2103	{161:75.2%, 0.981}
EIC2203	<=	EIC2104 & EIC2100	{145:67.8%, 0.993}
EIC2203	<=	EIC2104 & EIC2101 & EIC2100	{138:64.5%, 0.993}
EIC2203	<=	EIC2104 & EIC2101 & EIC2102 & EIC2100	{130:60.7%, 1.0}
EIC2203	<=	EIC2104 & EIC2101 & EIC2102	{140:65.4%, 0.993}
EIC2203	<=	EIC2104 & EIC2101	{152:71.0%, 0.98}
EIC2203	<=	EIC2104 & EIC2102 & EIC2100	{134:62.6%, 1.0}
EIC2203	<=	EIC2104 & EIC2102	{151:70.6%, 0.974}
EIC2203	<=	EIC2104 & EIC2103 & EIC2100	{141:65.9%, 0.993}
EIC2203	<=	EIC2104 & EIC2103 & EIC2101 & EIC2100	{135:63.1%, 0.993}
EIC2203	<=	EIC2104 & EIC2103 & EIC2101 & EIC2102	{137:64.0%, 0.993}
EIC2203	<=	EIC2104 & EIC2103 & EIC2101	{145:67.8%, 0.986}
EIC2203	<=	EIC2104 & EIC2103 & EIC2102 & EIC2100	{131:61.2%, 1.0}
EIC2203	<=	EIC2104 & EIC2103 & EIC2102	{142:66.4%, 0.993}
EIC2203	<=	EIC2104 & EIC2103	{154:72.0%, 0.987}
EIC2203	<=	EIC2104	{174:81.3%, 0.925}
EIC2203	<=	EIC2105 & EIC2100	{147:68.7%, 0.993}
EIC2203	<=	EIC2105 & EIC2101 & EIC2100	{141:65.9%, 0.993}
EIC2203	<=	EIC2105 & EIC2101 & EIC2102 & EIC2100	{132:61.7%, 1.0}
EIC2203	<=	EIC2105 & EIC2101 & EIC2102	{143:66.8%, 0.993}
EIC2203	<=	EIC2105 & EIC2101	{156:72.9%, 0.987}
EIC2203	<=	EIC2105 & EIC2102 & EIC2100	{136:63.6%, 1.0}
EIC2203	<=	EIC2105 & EIC2102	{152:71.0%, 0.974}
EIC2203	<=	EIC2105 & EIC2103 & EIC2100	{143:66.8%, 0.993}
EIC2203	<=	EIC2105 & EIC2103 & EIC2101 & EIC2100	{138:64.5%, 0.993}
EIC2203	<=	EIC2105 & EIC2103 & EIC2101 & EIC2102	{140:65.4%, 0.993}
EIC2203	<=	EIC2105 & EIC2103 & EIC2101	{150:70.1%, 0.987}
EIC2203	<=	EIC2105 & EIC2103 & EIC2102 & EIC2100	{133:62.1%, 1.0}
EIC2203	<=	EIC2105 & EIC2103 & EIC2102	{144:67.3%, 0.993}
EIC2203	<=	EIC2105 & EIC2103	{157:73.4%, 0.987}
EIC2203	<=	EIC2105 & EIC2104 & EIC2100	{144:67.3%, 0.993}
EIC2203	<=	EIC2105 & EIC2104 & EIC2101 & EIC2100	{138:64.5%, 0.993}
EIC2203	<=	EIC2105 & EIC2104 & EIC2101 & EIC2102	{140:65.4%, 0.993}
EIC2203	<=	EIC2105 & EIC2104 & EIC2101	{151:70.6%, 0.987}
EIC2203	<=	EIC2105 & EIC2104 & EIC2102 & EIC2100	{134:62.6%, 1.0}
EIC2203	<=	EIC2105 & EIC2104 & EIC2102	{149:69.6%, 0.973}
EIC2203	<=	EIC2105 & EIC2104 & EIC2103 & EIC2100	{140:65.4%, 0.993}
EIC2203	<=	EIC2105 & EIC2104 & EIC2103 & EIC2101	{145:67.8%, 0.986}
EIC2203	<=	EIC2105 & EIC2104 & EIC2103 & EIC2102	{141:65.9%, 0.993}
EIC2203	<=	EIC2105 & EIC2104 & EIC2103	{152:71.0%, 0.987}
EIC2203	<=	EIC2105 & EIC2104	{168:78.5%, 0.94}
EIC2203	<=	EIC2105	{179:83.6%, 0.922}

Figura III.19 - Regras de associação: disciplina EIC2203 (sem.)

Melhor suporte:	179:83.6% (conf=92.2%)		
Melhor confiança:	100.0% (sup=136:63.6%)		
EIC2203	<=	EIC2105 & EIC2102 & EIC2100	{136:63.6%, 1.0}
EIC2203	<=	EIC2105	{179:83.6%, 0.922}
EIC2203	<=	EIC2105 & EIC2100	{147:68.7%, 0.993}
EIC2203	<=	EIC2104	{174:81.3%, 0.925}
EIC2203	<=	EIC2105 & EIC2103	{157:73.4%, 0.987}
EIC2203	<=	EIC2105 & EIC2104	{168:78.5%, 0.94}
EIC2203	<=	EIC2103	{161:75.2%, 0.981}

Figura III.20 - Regras de associação: disciplina EIC2203 (sem.) após o filtro



EIC2204	<=	EIC2100	(149:69.6%, 0.926)
EIC2204	<=	EIC2101 & EIC2100	(141:65.9%, 0.957)
EIC2204	<=	EIC2101 & EIC2102 & EIC2100	(132:61.7%, 0.962)
EIC2204	<=	EIC2101 & EIC2102	(143:66.8%, 0.951)
EIC2204	<=	EIC2101	(159:74.3%, 0.912)
EIC2204	<=	EIC2102 & EIC2100	(137:64.0%, 0.942)
EIC2204	<=	EIC2102	(159:74.3%, 0.874)
EIC2204	<=	EIC2103 & EIC2100	(144:67.3%, 0.931)
EIC2204	<=	EIC2103 & EIC2101 & EIC2100	(138:64.5%, 0.957)
EIC2204	<=	EIC2103 & EIC2101 & EIC2102 & EIC2100	(130:60.7%, 0.962)
EIC2204	<=	EIC2103 & EIC2101 & EIC2102	(140:65.4%, 0.95)
EIC2204	<=	EIC2103 & EIC2101	(151:70.6%, 0.927)
EIC2204	<=	EIC2103 & EIC2102 & EIC2100	(133:62.1%, 0.947)
EIC2204	<=	EIC2103 & EIC2102	(146:68.2%, 0.925)
EIC2204	<=	EIC2103	(161:75.2%, 0.888)
EIC2204	<=	EIC2104 & EIC2100	(145:67.8%, 0.938)
EIC2204	<=	EIC2104 & EIC2101 & EIC2100	(138:64.5%, 0.964)
EIC2204	<=	EIC2104 & EIC2101 & EIC2102 & EIC2100	(130:60.7%, 0.969)
EIC2204	<=	EIC2104 & EIC2101 & EIC2102	(140:65.4%, 0.957)
EIC2204	<=	EIC2104 & EIC2101	(152:71.0%, 0.934)
EIC2204	<=	EIC2104 & EIC2102 & EIC2100	(134:62.6%, 0.955)
EIC2204	<=	EIC2104 & EIC2102	(151:70.6%, 0.901)
EIC2204	<=	EIC2104 & EIC2103 & EIC2100	(141:65.9%, 0.936)
EIC2204	<=	EIC2104 & EIC2103 & EIC2101 & EIC2100	(135:63.1%, 0.963)
EIC2204	<=	EIC2104 & EIC2103 & EIC2101 & EIC2102	(137:64.0%, 0.956)
EIC2204	<=	EIC2104 & EIC2103 & EIC2101	(145:67.8%, 0.945)
EIC2204	<=	EIC2104 & EIC2103 & EIC2102 & EIC2100	(131:61.2%, 0.954)
EIC2204	<=	EIC2104 & EIC2103 & EIC2102	(142:66.4%, 0.93)
EIC2204	<=	EIC2104 & EIC2103	(154:72.0%, 0.903)
EIC2204	<=	EIC2104	(174:81.3%, 0.839)
EIC2204	<=	EIC2105 & EIC2100	(147:68.7%, 0.932)
EIC2204	<=	EIC2105 & EIC2101 & EIC2100	(141:65.9%, 0.957)
EIC2204	<=	EIC2105 & EIC2101 & EIC2102 & EIC2100	(132:61.7%, 0.962)
EIC2204	<=	EIC2105 & EIC2101 & EIC2102	(143:66.8%, 0.951)
EIC2204	<=	EIC2105 & EIC2101	(156:72.9%, 0.929)
EIC2204	<=	EIC2105 & EIC2102 & EIC2100	(136:63.6%, 0.949)
EIC2204	<=	EIC2105 & EIC2102	(152:71.0%, 0.908)
EIC2204	<=	EIC2105 & EIC2103 & EIC2100	(143:66.8%, 0.93)
EIC2204	<=	EIC2105 & EIC2103 & EIC2101 & EIC2100	(138:64.5%, 0.957)
EIC2204	<=	EIC2105 & EIC2103 & EIC2101 & EIC2102	(140:65.4%, 0.95)
EIC2204	<=	EIC2105 & EIC2103 & EIC2101	(150:70.1%, 0.933)
EIC2204	<=	EIC2105 & EIC2103 & EIC2102 & EIC2100	(133:62.1%, 0.947)
EIC2204	<=	EIC2105 & EIC2103 & EIC2102	(144:67.3%, 0.931)
EIC2204	<=	EIC2105 & EIC2103	(157:73.4%, 0.898)
EIC2204	<=	EIC2105 & EIC2104 & EIC2100	(144:67.3%, 0.938)
EIC2204	<=	EIC2105 & EIC2104 & EIC2101 & EIC2100	(138:64.5%, 0.964)
EIC2204	<=	EIC2105 & EIC2104 & EIC2101 & EIC2102	(140:65.4%, 0.957)
EIC2204	<=	EIC2105 & EIC2104 & EIC2101	(151:70.6%, 0.94)
EIC2204	<=	EIC2105 & EIC2104 & EIC2102 & EIC2100	(134:62.6%, 0.955)
EIC2204	<=	EIC2105 & EIC2104 & EIC2102	(149:69.6%, 0.913)
EIC2204	<=	EIC2105 & EIC2104 & EIC2103 & EIC2100	(140:65.4%, 0.936)
EIC2204	<=	EIC2105 & EIC2104 & EIC2103 & EIC2101	(145:67.8%, 0.945)
EIC2204	<=	EIC2105 & EIC2104 & EIC2103 & EIC2102	(141:65.9%, 0.936)
EIC2204	<=	EIC2105 & EIC2104 & EIC2103	(152:71.0%, 0.908)
EIC2204	<=	EIC2105 & EIC2104	(168:78.5%, 0.863)
EIC2204	<=	EIC2105	(179:83.6%, 0.827)

Figura III.21 - Regras de associação: disciplina EIC2204 (sem.)

Melhor suporte:	179:83.6% (conf=82.7%)		
Melhor confiança:	96.9% (sup=130:60.7%)		
EIC2204	<=	EIC2104 & EIC2101 & EIC2102 & EIC2100	(130:60.7%, 0.969)
EIC2204	<=	EIC2105	(179:83.6%, 0.827)
EIC2204	<=	EIC2105 & EIC2104 & EIC2101 & EIC2100	(138:64.5%, 0.964)
EIC2204	<=	EIC2104 & EIC2101 & EIC2100	(138:64.5%, 0.964)
EIC2204	<=	EIC2104	(174:81.3%, 0.839)
EIC2204	<=	EIC2101 & EIC2100	(141:65.9%, 0.957)
EIC2204	<=	EIC2105 & EIC2101 & EIC2100	(141:65.9%, 0.957)
EIC2204	<=	EIC2105 & EIC2104	(168:78.5%, 0.863)
EIC2204	<=	EIC2105 & EIC2101 & EIC2102	(143:66.8%, 0.951)
EIC2204	<=	EIC2101 & EIC2102	(143:66.8%, 0.951)
EIC2204	<=	EIC2103	(161:75.2%, 0.888)
EIC2204	<=	EIC2104 & EIC2103 & EIC2101	(145:67.8%, 0.945)
EIC2204	<=	EIC2105 & EIC2104 & EIC2103 & EIC2101	(145:67.8%, 0.945)
EIC2204	<=	EIC2101	(159:74.3%, 0.912)
EIC2204	<=	EIC2105 & EIC2104 & EIC2101	(151:70.6%, 0.94)
EIC2204	<=	EIC2105 & EIC2101	(156:72.9%, 0.929)
EIC2204	<=	EIC2104 & EIC2101	(152:71.0%, 0.934)

Figura III.22 - Regras de associação: disciplina EIC2204 (sem.) após o filtro

EIC2205	<=	EIC2100	{149:69.6%, 0.98}
EIC2205	<=	EIC2101 & EIC2100	{141:65.9%, 0.986}
EIC2205	<=	EIC2101 & EIC2102 & EIC2100	{132:61.7%, 0.992}
EIC2205	<=	EIC2101 & EIC2102	{143:66.8%, 0.993}
EIC2205	<=	EIC2101	{159:74.3%, 0.981}
EIC2205	<=	EIC2102 & EIC2100	{137:64.0%, 0.985}
EIC2205	<=	EIC2102	{159:74.3%, 0.956}
EIC2205	<=	EIC2103 & EIC2100	{144:67.3%, 0.986}
EIC2205	<=	EIC2103 & EIC2101 & EIC2100	{138:64.5%, 0.986}
EIC2205	<=	EIC2103 & EIC2101 & EIC2102 & EIC2100	{130:60.7%, 0.992}
EIC2205	<=	EIC2103 & EIC2101 & EIC2102	{140:65.4%, 0.993}
EIC2205	<=	EIC2103 & EIC2101	{151:70.6%, 0.987}
EIC2205	<=	EIC2103 & EIC2102 & EIC2100	{133:62.1%, 0.992}
EIC2205	<=	EIC2103 & EIC2102	{146:68.2%, 0.993}
EIC2205	<=	EIC2103	{161:75.2%, 0.988}
EIC2205	<=	EIC2104 & EIC2100	{145:67.8%, 0.993}
EIC2205	<=	EIC2104 & EIC2101 & EIC2100	{138:64.5%, 0.993}
EIC2205	<=	EIC2104 & EIC2101 & EIC2102 & EIC2100	{130:60.7%, 0.992}
EIC2205	<=	EIC2104 & EIC2101 & EIC2102	{140:65.4%, 0.993}
EIC2205	<=	EIC2104 & EIC2101	{152:71.0%, 0.987}
EIC2205	<=	EIC2104 & EIC2102 & EIC2100	{134:62.6%, 0.993}
EIC2205	<=	EIC2104 & EIC2102	{151:70.6%, 0.98}
EIC2205	<=	EIC2104 & EIC2103 & EIC2100	{141:65.9%, 0.993}
EIC2205	<=	EIC2104 & EIC2103 & EIC2101 & EIC2100	{135:63.1%, 0.993}
EIC2205	<=	EIC2104 & EIC2103 & EIC2101 & EIC2102	{137:64.0%, 0.993}
EIC2205	<=	EIC2104 & EIC2103 & EIC2101	{145:67.8%, 0.993}
EIC2205	<=	EIC2104 & EIC2103 & EIC2102 & EIC2100	{131:61.2%, 0.992}
EIC2205	<=	EIC2104 & EIC2103 & EIC2102	{142:66.4%, 0.993}
EIC2205	<=	EIC2104 & EIC2103	{154:72.0%, 0.994}
EIC2205	<=	EIC2104	{174:81.3%, 0.966}
EIC2205	<=	EIC2105 & EIC2100	{147:68.7%, 0.986}
EIC2205	<=	EIC2105 & EIC2101 & EIC2100	{141:65.9%, 0.986}
EIC2205	<=	EIC2105 & EIC2101 & EIC2102 & EIC2100	{132:61.7%, 0.992}
EIC2205	<=	EIC2105 & EIC2101 & EIC2102	{143:66.8%, 0.993}
EIC2205	<=	EIC2105 & EIC2101	{156:72.9%, 0.987}
EIC2205	<=	EIC2105 & EIC2102 & EIC2100	{136:63.6%, 0.993}
EIC2205	<=	EIC2105 & EIC2102	{152:71.0%, 0.98}
EIC2205	<=	EIC2105 & EIC2103 & EIC2100	{143:66.8%, 0.986}
EIC2205	<=	EIC2105 & EIC2103 & EIC2101 & EIC2100	{138:64.5%, 0.986}
EIC2205	<=	EIC2105 & EIC2103 & EIC2101 & EIC2102	{140:65.4%, 0.993}
EIC2205	<=	EIC2105 & EIC2103 & EIC2101	{150:70.1%, 0.987}
EIC2205	<=	EIC2105 & EIC2103 & EIC2102 & EIC2100	{133:62.1%, 0.992}
EIC2205	<=	EIC2105 & EIC2103 & EIC2102	{144:67.3%, 0.993}
EIC2205	<=	EIC2105 & EIC2103	{157:73.4%, 0.987}
EIC2205	<=	EIC2105 & EIC2104 & EIC2100	{144:67.3%, 0.993}
EIC2205	<=	EIC2105 & EIC2104 & EIC2101 & EIC2100	{138:64.5%, 0.993}
EIC2205	<=	EIC2105 & EIC2104 & EIC2101 & EIC2102	{140:65.4%, 0.993}
EIC2205	<=	EIC2105 & EIC2104 & EIC2101	{151:70.6%, 0.993}
EIC2205	<=	EIC2105 & EIC2104 & EIC2102 & EIC2100	{134:62.6%, 0.993}
EIC2205	<=	EIC2105 & EIC2104 & EIC2102	{149:69.6%, 0.98}
EIC2205	<=	EIC2105 & EIC2104 & EIC2103 & EIC2100	{140:65.4%, 0.993}
EIC2205	<=	EIC2105 & EIC2104 & EIC2103 & EIC2101	{145:67.8%, 0.993}
EIC2205	<=	EIC2105 & EIC2104 & EIC2103 & EIC2102	{141:65.9%, 0.993}
EIC2205	<=	EIC2105 & EIC2104 & EIC2103	{152:71.0%, 0.993}
EIC2205	<=	EIC2105 & EIC2104	{168:78.5%, 0.976}
EIC2205	<=	EIC2105	{179:83.6%, 0.944}

**Figura III.23 - Regras de associação: disciplina EIC2205 (sem.)**

Melhor suporte:	179:83.6% (conf=94.4%)		
Melhor confiança:	99.4% (sup=154:72.0%)		
EIC2205	<=	EIC2104 & EIC2103	{154:72.0%, 0.994}
EIC2205	<=	EIC2105	{179:83.6%, 0.944}
EIC2205	<=	EIC2103	{161:75.2%, 0.988}
EIC2205	<=	EIC2104	{174:81.3%, 0.966}
EIC2205	<=	EIC2105 & EIC2104	{168:78.5%, 0.976}

**Figura III.24 - Regras de associação: disciplina EIC2205 (sem.) após o filtro**

Melhor suporte: 182:85.0% (conf=80.2%)  
 Melhor confiança: 91.1% (sup=157:73.4%)

EIC2100 <= EIC1102 & EIC1206 & EIC1204 & EIC1203 (157:73.4%, 0.911)  
 EIC2100 <= EIC1104 & EIC1101 (182:85.0%, 0.802)  
 EIC2100 <= EIC1104 & EIC1206 & EIC1204 & EIC1203 (159:74.3%, 0.906)  
 EIC2100 <= EIC1104 & EIC1206 (181:84.6%, 0.818)  
 EIC2100 <= EIC1106 & EIC1206 & EIC1204 & EIC1203 (160:74.8%, 0.9)  
 EIC2100 <= EIC1206 & EIC1204 & EIC1203 (160:74.8%, 0.9)  
 EIC2100 <= EIC1102 & EIC1206 (178:83.2%, 0.82)  
 EIC2100 <= EIC1106 & EIC1102 & EIC1206 (178:83.2%, 0.82)  
 EIC2100 <= EIC1106 & EIC1102 & EIC1204 & EIC1203 (161:75.2%, 0.894)  
 EIC2100 <= EIC1104 & EIC1102 & EIC1204 & EIC1203 (161:75.2%, 0.894)  
 EIC2100 <= EIC1102 & EIC1204 & EIC1203 (161:75.2%, 0.894)  
 EIC2100 <= EIC1203 (177:82.7%, 0.831)  
 EIC2100 <= EIC1104 & EIC1204 & EIC1203 (163:76.2%, 0.89)  
 EIC2100 <= EIC1106 & EIC1104 & EIC1204 & EIC1203 (163:76.2%, 0.89)  
 EIC2100 <= EIC1104 & EIC1102 & EIC1202 & EIC1101 (175:81.8%, 0.834)  
 EIC2100 <= EIC1104 & EIC1102 & EIC1206 (175:81.8%, 0.834)  
 EIC2100 <= EIC1106 & EIC1104 & EIC1102 & EIC1206 (175:81.8%, 0.834)  
 EIC2100 <= EIC1106 & EIC1203 (175:81.8%, 0.834)  
 EIC2100 <= EIC1106 & EIC1204 & EIC1203 (164:76.6%, 0.884)  
 EIC2100 <= EIC1104 & EIC1203 (173:80.8%, 0.85)  
 EIC2100 <= EIC1104 & EIC1206 & EIC1203 (166:77.6%, 0.88)  
 EIC2100 <= EIC1102 & EIC1202 & EIC1101 & EIC1206 (170:79.4%, 0.853)  
 EIC2100 <= EIC1104 & EIC1100 (170:79.4%, 0.853)  
 EIC2100 <= EIC1104 & EIC1101 & EIC1100 (170:79.4%, 0.853)  
 EIC2100 <= EIC1104 & EIC1202 & EIC1101 & EIC1206 (170:79.4%, 0.853)  
 EIC2100 <= EIC1104 & EIC1206 & EIC1204 (170:79.4%, 0.853)  
 EIC2100 <= EIC1106 & EIC1104 & EIC1100 (170:79.4%, 0.853)  
 EIC2100 <= EIC1106 & EIC1104 & EIC1101 & EIC1100 (170:79.4%, 0.853)  
 EIC2100 <= EIC1106 & EIC1104 & EIC1206 & EIC1204 (170:79.4%, 0.853)  
 EIC2100 <= EIC1106 & EIC1202 & EIC1101 & EIC1100 (170:79.4%, 0.853)  
 EIC2100 <= EIC1106 & EIC1205 & EIC1104 & EIC1203 (170:79.4%, 0.853)  
 EIC2100 <= EIC1202 & EIC1101 & EIC1100 (170:79.4%, 0.853)  
 EIC2100 <= EIC1205 & EIC1104 & EIC1203 (170:79.4%, 0.853)  
 EIC2100 <= EIC1206 & EIC1203 (168:78.5%, 0.869)  
 EIC2100 <= EIC1106 & EIC1104 & EIC1102 & EIC1100 (169:79.0%, 0.858)  
 EIC2100 <= EIC1106 & EIC1102 & EIC1202 & EIC1100 (169:79.0%, 0.858)  
 EIC2100 <= EIC1104 & EIC1102 & EIC1101 & EIC1206 (169:79.0%, 0.858)  
 EIC2100 <= EIC1104 & EIC1102 & EIC1101 & EIC1100 (169:79.0%, 0.858)  
 EIC2100 <= EIC1104 & EIC1102 & EIC1100 (169:79.0%, 0.858)  
 EIC2100 <= EIC1102 & EIC1202 & EIC1101 & EIC1100 (169:79.0%, 0.858)  
 EIC2100 <= EIC1102 & EIC1202 & EIC1100 (169:79.0%, 0.858)

**Figura III.25 - Regras de associação: disciplina EIC2100 (anual) após o filtro**

Melhor suporte: 198:92.5% (conf=80.3%)  
 Melhor confiança: 94.9% (sup=158:73.8%)

EIC2101	<=	EIC1102	&	EIC1206	&	EIC1204	&	EIC1200	(158:73.8%, 0.949)
EIC2101	<=	EIC1104	&	EIC1206	&	EIC1204	&	EIC1200	(158:73.8%, 0.949)
EIC2101	<=	EIC1205	(198:92.5%, 0.803)						
EIC2101	<=	EIC1104	&	EIC1102	&	EIC1206	&	EIC1200	(161:75.2%, 0.944)
EIC2101	<=	EIC1106	&	EIC1205	(195:91.1%, 0.81)				
EIC2101	<=	EIC1104	&	EIC1102	&	EIC1203	&	EIC1100	(162:75.7%, 0.938)
EIC2101	<=	EIC1104	&	EIC1206	&	EIC1100	&	EIC1200	(162:75.7%, 0.938)
EIC2101	<=	EIC1205	&	EIC1102	&	EIC1206	&	EIC1200	(162:75.7%, 0.938)
EIC2101	<=	EIC1205	&	EIC1104	&	EIC1206	&	EIC1200	(162:75.7%, 0.938)
EIC2101	<=	EIC1106	&	EIC1104	(193:90.2%, 0.819)				
EIC2101	<=	EIC1205	&	EIC1104	&	EIC1102	&	EIC1200	(165:77.1%, 0.933)
EIC2101	<=	EIC1205	&	EIC1104	&	EIC1100	&	EIC1200	(165:77.1%, 0.933)
EIC2101	<=	EIC1104	&	EIC1102	&	EIC1101	&	EIC1203	(165:77.1%, 0.933)
EIC2101	<=	EIC1104	&	EIC1102	&	EIC1101	&	EIC1200	(165:77.1%, 0.933)
EIC2101	<=	EIC1104	&	EIC1102	&	EIC1100	&	EIC1200	(165:77.1%, 0.933)
EIC2101	<=	EIC1106	&	EIC1102	(190:88.8%, 0.826)				
EIC2101	<=	EIC1205	&	EIC1104	&	EIC1102	&	EIC1100	(167:78.0%, 0.928)
EIC2101	<=	EIC1205	&	EIC1104	(187:87.4%, 0.845)				
EIC2101	<=	EIC1205	&	EIC1104	&	EIC1101	&	EIC1100	(168:78.5%, 0.923)
EIC2101	<=	EIC1205	&	EIC1104	&	EIC1100	(168:78.5%, 0.923)		
EIC2101	<=	EIC1106	&	EIC1205	&	EIC1104	&	EIC1100	(168:78.5%, 0.923)
EIC2101	<=	EIC1106	&	EIC1205	&	EIC1104	(186:86.9%, 0.849)		
EIC2101	<=	EIC1104	&	EIC1102	&	EIC1100	(169:79.0%, 0.917)		
EIC2101	<=	EIC1104	&	EIC1102	&	EIC1101	&	EIC1100	(169:79.0%, 0.917)
EIC2101	<=	EIC1106	&	EIC1104	&	EIC1102	&	EIC1100	(169:79.0%, 0.917)
EIC2101	<=	EIC1106	&	EIC1205	&	EIC1102	&	EIC1100	(169:79.0%, 0.917)
EIC2101	<=	EIC1205	&	EIC1102	&	EIC1100	(169:79.0%, 0.917)		
EIC2101	<=	EIC1205	&	EIC1102	&	EIC1101	&	EIC1100	(169:79.0%, 0.917)
EIC2101	<=	EIC1106	&	EIC1205	&	EIC1102	(184:86.0%, 0.853)		
EIC2101	<=	EIC1205	&	EIC1104	&	EIC1203	(170:79.4%, 0.912)		
EIC2101	<=	EIC1106	&	EIC1205	&	EIC1104	&	EIC1203	(170:79.4%, 0.912)
EIC2101	<=	EIC1106	&	EIC1104	&	EIC1101	&	EIC1100	(170:79.4%, 0.912)
EIC2101	<=	EIC1106	&	EIC1104	&	EIC1100	(170:79.4%, 0.912)		
EIC2101	<=	EIC1104	&	EIC1101	&	EIC1100	(170:79.4%, 0.912)		
EIC2101	<=	EIC1104	&	EIC1100	(170:79.4%, 0.912)				
EIC2101	<=	EIC1205	&	EIC1101	(182:85.0%, 0.863)				
EIC2101	<=	EIC1104	&	EIC1101	(182:85.0%, 0.863)				
EIC2101	<=	EIC1102	&	EIC1100	(171:79.9%, 0.906)				
EIC2101	<=	EIC1102	&	EIC1101	&	EIC1100	(171:79.9%, 0.906)		
EIC2101	<=	EIC1106	&	EIC1102	&	EIC1100	(171:79.9%, 0.906)		
EIC2101	<=	EIC1106	&	EIC1102	&	EIC1101	&	EIC1100	(171:79.9%, 0.906)
EIC2101	<=	EIC1106	&	EIC1205	&	EIC1101	&	EIC1100	(171:79.9%, 0.906)
EIC2101	<=	EIC1205	&	EIC1101	&	EIC1100	(171:79.9%, 0.906)		
EIC2101	<=	EIC1106	&	EIC1104	&	EIC1101	(181:84.6%, 0.867)		
EIC2101	<=	EIC1106	&	EIC1205	&	EIC1101	(181:84.6%, 0.867)		
EIC2101	<=	EIC1205	&	EIC1102	&	EIC1103	&	EIC1101	(173:80.8%, 0.902)
EIC2101	<=	EIC1205	&	EIC1104	&	EIC1102	(180:84.1%, 0.872)		
EIC2101	<=	EIC1106	&	EIC1205	&	EIC1104	&	EIC1102	(180:84.1%, 0.872)
EIC2101	<=	EIC1106	&	EIC1102	&	EIC1101	(180:84.1%, 0.872)		
EIC2101	<=	EIC1102	&	EIC1101	(180:84.1%, 0.872)				
EIC2101	<=	EIC1205	&	EIC1104	&	EIC1102	&	EIC1101	(175:81.8%, 0.897)
EIC2101	<=	EIC1205	&	EIC1104	&	EIC1101	(179:83.6%, 0.877)		
EIC2101	<=	EIC1205	&	EIC1102	&	EIC1101	(177:82.7%, 0.887)		
EIC2101	<=	EIC1106	&	EIC1205	&	EIC1102	&	EIC1101	(177:82.7%, 0.887)
EIC2101	<=	EIC1106	&	EIC1104	&	EIC1102	&	EIC1101	(177:82.7%, 0.887)
EIC2101	<=	EIC1104	&	EIC1102	&	EIC1101	(177:82.7%, 0.887)		
EIC2101	<=	EIC1106	&	EIC1205	&	EIC1104	&	EIC1101	(178:83.2%, 0.882)

**Figura III.26 - Regras de associação: disciplina EIC2101 (anual) após o filtro**

Melhor suporte: 193:90.2% (conf=81.3%)  
 Melhor confiança: 93.6% (sup=156:72.9%)

EIC2102	<=	EIC1102	&	EIC1204	&	EIC1203	&	EIC1100	(156:72.9%, 0.936)
EIC2102	<=	EIC1102	&	EIC1204	&	EIC1203	&	EIC1200	(156:72.9%, 0.936)
EIC2102	<=	EIC1106	&	EIC1104	(193:90.2%, 0.813)				
EIC2102	<=	EIC1102	&	EIC1103	&	EIC1204	&	EIC1100	(159:74.3%, 0.931)
EIC2102	<=	EIC1102	(192:89.7%, 0.818)						
EIC2102	<=	EIC1102	&	EIC1103	&	EIC1204	&	EIC1200	(160:74.8%, 0.925)
EIC2102	<=	EIC1103	&	EIC1202	&	EIC1204	&	EIC1100	(160:74.8%, 0.925)
EIC2102	<=	EIC1104	&	EIC1103	&	EIC1204	&	EIC1100	(160:74.8%, 0.925)
EIC2102	<=	EIC1202	(191:89.3%, 0.822)						
EIC2102	<=	EIC1205	&	EIC1103	&	EIC1204	&	EIC1100	(161:75.2%, 0.919)
EIC2102	<=	EIC1106	&	EIC1103	&	EIC1204	&	EIC1100	(161:75.2%, 0.919)
EIC2102	<=	EIC1106	&	EIC1102	&	EIC1204	&	EIC1203	(161:75.2%, 0.919)
EIC2102	<=	EIC1104	&	EIC1204	&	EIC1100	&	EIC1200	(161:75.2%, 0.919)
EIC2102	<=	EIC1104	&	EIC1202	&	EIC1204	&	EIC1200	(161:75.2%, 0.919)
EIC2102	<=	EIC1104	&	EIC1202	&	EIC1204	&	EIC1100	(161:75.2%, 0.919)
EIC2102	<=	EIC1104	&	EIC1202	&	EIC1203	&	EIC1100	(161:75.2%, 0.919)
EIC2102	<=	EIC1104	&	EIC1103	&	EIC1204	&	EIC1200	(161:75.2%, 0.919)
EIC2102	<=	EIC1104	&	EIC1102	&	EIC1204	&	EIC1203	(161:75.2%, 0.919)
EIC2102	<=	EIC1104	&	EIC1102	&	EIC1204	&	EIC1200	(161:75.2%, 0.919)
EIC2102	<=	EIC1104	&	EIC1102	&	EIC1204	&	EIC1100	(161:75.2%, 0.919)
EIC2102	<=	EIC1104	&	EIC1101	&	EIC1204	&	EIC1200	(161:75.2%, 0.919)
EIC2102	<=	EIC1103	&	EIC1204	&	EIC1100	(161:75.2%, 0.919)		
EIC2102	<=	EIC1103	&	EIC1204	&	EIC1100	&	EIC1200	(161:75.2%, 0.919)
EIC2102	<=	EIC1103	&	EIC1202	&	EIC1204	&	EIC1200	(161:75.2%, 0.919)
EIC2102	<=	EIC1103	&	EIC1101	&	EIC1204	&	EIC1200	(161:75.2%, 0.919)
EIC2102	<=	EIC1103	&	EIC1101	&	EIC1204	&	EIC1100	(161:75.2%, 0.919)
EIC2102	<=	EIC1102	&	EIC1204	&	EIC1203	(161:75.2%, 0.919)		
EIC2102	<=	EIC1102	&	EIC1204	&	EIC1100	&	EIC1200	(161:75.2%, 0.919)
EIC2102	<=	EIC1102	&	EIC1101	&	EIC1204	&	EIC1200	(161:75.2%, 0.919)
EIC2102	<=	EIC1106	&	EIC1202	(189:88.3%, 0.825)				
EIC2102	<=	EIC1104	&	EIC1103	&	EIC1202	&	EIC1100	(165:77.1%, 0.915)
EIC2102	<=	EIC1104	&	EIC1102	(186:86.9%, 0.839)				
EIC2102	<=	EIC1102	&	EIC1103	&	EIC1202	&	EIC1100	(166:77.6%, 0.91)
EIC2102	<=	EIC1106	&	EIC1104	&	EIC1102	(185:86.4%, 0.843)		
EIC2102	<=	EIC1102	&	EIC1202	(185:86.4%, 0.843)				
EIC2102	<=	EIC1103	&	EIC1202	&	EIC1100	(167:78.0%, 0.904)		
EIC2102	<=	EIC1103	&	EIC1202	&	EIC1101	&	EIC1100	(167:78.0%, 0.904)
EIC2102	<=	EIC1104	&	EIC1102	&	EIC1103	&	EIC1100	(167:78.0%, 0.904)
EIC2102	<=	EIC1104	&	EIC1102	&	EIC1202	&	EIC1100	(167:78.0%, 0.904)
EIC2102	<=	EIC1104	&	EIC1202	&	EIC1101	&	EIC1100	(167:78.0%, 0.904)
EIC2102	<=	EIC1106	&	EIC1103	&	EIC1202	&	EIC1100	(167:78.0%, 0.904)
EIC2102	<=	EIC1106	&	EIC1104	&	EIC1202	&	EIC1100	(167:78.0%, 0.904)
EIC2102	<=	EIC1104	&	EIC1202	(184:86.0%, 0.848)				
EIC2102	<=	EIC1106	&	EIC1104	&	EIC1202	(184:86.0%, 0.848)		
EIC2102	<=	EIC1106	&	EIC1104	&	EIC1103	&	EIC1100	(168:78.5%, 0.899)
EIC2102	<=	EIC1106	&	EIC1102	&	EIC1103	&	EIC1100	(168:78.5%, 0.899)
EIC2102	<=	EIC1104	&	EIC1103	&	EIC1101	&	EIC1100	(168:78.5%, 0.899)
EIC2102	<=	EIC1104	&	EIC1103	&	EIC1100	(168:78.5%, 0.899)		
EIC2102	<=	EIC1102	&	EIC1103	&	EIC1101	&	EIC1100	(168:78.5%, 0.899)
EIC2102	<=	EIC1102	&	EIC1103	&	EIC1100	(168:78.5%, 0.899)		
EIC2102	<=	EIC1106	&	EIC1104	&	EIC1102	&	EIC1202	(181:84.6%, 0.856)
EIC2102	<=	EIC1104	&	EIC1102	&	EIC1202	(181:84.6%, 0.856)		
EIC2102	<=	EIC1102	&	EIC1202	&	EIC1100	(169:79.0%, 0.893)		
EIC2102	<=	EIC1102	&	EIC1202	&	EIC1101	&	EIC1100	(169:79.0%, 0.893)
EIC2102	<=	EIC1104	&	EIC1102	&	EIC1100	(169:79.0%, 0.893)		
EIC2102	<=	EIC1104	&	EIC1102	&	EIC1101	&	EIC1100	(169:79.0%, 0.893)
EIC2102	<=	EIC1106	&	EIC1102	&	EIC1202	&	EIC1100	(169:79.0%, 0.893)
EIC2102	<=	EIC1106	&	EIC1104	&	EIC1102	&	EIC1100	(169:79.0%, 0.893)
EIC2102	<=	EIC1103	&	EIC1202	(180:84.1%, 0.861)				
EIC2102	<=	EIC1106	&	EIC1103	&	EIC1202	(180:84.1%, 0.861)		
EIC2102	<=	EIC1202	&	EIC1101	&	EIC1100	(170:79.4%, 0.888)		
EIC2102	<=	EIC1106	&	EIC1202	&	EIC1101	&	EIC1100	(170:79.4%, 0.888)
EIC2102	<=	EIC1106	&	EIC1104	&	EIC1101	&	EIC1100	(170:79.4%, 0.888)
EIC2102	<=	EIC1106	&	EIC1104	&	EIC1100	(170:79.4%, 0.888)		
EIC2102	<=	EIC1106	&	EIC1103	&	EIC1101	&	EIC1100	(170:79.4%, 0.888)
EIC2102	<=	EIC1104	&	EIC1101	&	EIC1100	(170:79.4%, 0.888)		
EIC2102	<=	EIC1104	&	EIC1100	(170:79.4%, 0.888)				
EIC2102	<=	EIC1103	&	EIC1101	&	EIC1100	(170:79.4%, 0.888)		
EIC2102	<=	EIC1106	&	EIC1104	&	EIC1103	&	EIC1202	(178:83.2%, 0.871)
EIC2102	<=	EIC1104	&	EIC1103	&	EIC1202	(178:83.2%, 0.871)		
EIC2102	<=	EIC1104	&	EIC1103	&	EIC1202	&	EIC1101	(174:81.3%, 0.885)
EIC2102	<=	EIC1103	&	EIC1202	&	EIC1101	(176:82.2%, 0.875)		
EIC2102	<=	EIC1104	&	EIC1102	&	EIC1103	&	EIC1202	(176:82.2%, 0.875)
EIC2102	<=	EIC1106	&	EIC1103	&	EIC1202	&	EIC1101	(176:82.2%, 0.875)

Figura III.27 - Regras de associação: disciplina EIC2102 (anual) após o filtro

Melhor suporte: 197:92.1% (conf=81.2%)  
 Melhor confiança: 96.2% (sup=157:73.4%)

EIC2103 <= EIC1102 & EIC1206 & EIC1204 & EIC1203 (157:73.4%, 0.962)  
 EIC2103 <= EIC1104 (197:92.1%, 0.812)  
 EIC2103 <= EIC1104 & EIC1206 & EIC1204 & EIC1203 (159:74.3%, 0.956)  
 EIC2103 <= EIC1106 & EIC1104 (193:90.2%, 0.824)  
 EIC2103 <= EIC1104 & EIC1102 & EIC1206 & EIC1203 (162:75.7%, 0.951)  
 EIC2103 <= EIC1106 & EIC1102 (190:88.8%, 0.832)  
 EIC2103 <= EIC1106 & EIC1104 & EIC1204 & EIC1203 (163:76.2%, 0.945)  
 EIC2103 <= EIC1106 & EIC1102 & EIC1206 & EIC1203 (163:76.2%, 0.945)  
 EIC2103 <= EIC1104 & EIC1204 & EIC1203 (163:76.2%, 0.945)  
 EIC2103 <= EIC1102 & EIC1206 & EIC1203 (163:76.2%, 0.945)  
 EIC2103 <= EIC1205 & EIC1104 (187:87.4%, 0.845)  
 EIC2103 <= EIC1104 & EIC1102 & EIC1203 (167:78.0%, 0.94)  
 EIC2103 <= EIC1106 & EIC1104 & EIC1102 & EIC1203 (167:78.0%, 0.94)  
 EIC2103 <= EIC1104 & EIC1102 (186:86.9%, 0.849)  
 EIC2103 <= EIC1106 & EIC1205 & EIC1104 (186:86.9%, 0.849)  
 EIC2103 <= EIC1106 & EIC1102 & EIC1203 (168:78.5%, 0.935)  
 EIC2103 <= EIC1102 & EIC1203 (168:78.5%, 0.935)  
 EIC2103 <= EIC1106 & EIC1104 & EIC1102 (185:86.4%, 0.854)  
 EIC2103 <= EIC1106 & EIC1205 & EIC1104 & EIC1203 (170:79.4%, 0.924)  
 EIC2103 <= EIC1205 & EIC1104 & EIC1203 (170:79.4%, 0.924)  
 EIC2103 <= EIC1104 & EIC1101 (182:85.0%, 0.857)  
 EIC2103 <= EIC1205 & EIC1101 (182:85.0%, 0.857)  
 EIC2103 <= EIC1205 & EIC1102 & EIC1202 (182:85.0%, 0.857)  
 EIC2103 <= EIC1104 & EIC1203 (173:80.8%, 0.919)  
 EIC2103 <= EIC1104 & EIC1206 (181:84.6%, 0.867)  
 EIC2103 <= EIC1106 & EIC1203 (175:81.8%, 0.903)  
 EIC2103 <= EIC1106 & EIC1205 & EIC1104 & EIC1102 (180:84.1%, 0.872)  
 EIC2103 <= EIC1205 & EIC1104 & EIC1102 (180:84.1%, 0.872)  
 EIC2103 <= EIC1203 (177:82.7%, 0.898)  
 EIC2103 <= EIC1205 & EIC1104 & EIC1102 & EIC1202 (178:83.2%, 0.876)  
 EIC2103 <= EIC1106 & EIC1205 & EIC1104 & EIC1101 (178:83.2%, 0.876)

**Figura III.28 - Regras de associação: disciplina EIC2103 (anual) após o filtro**

Melhor suporte: 206:96.3% (conf=83.5%)  
 Melhor confiança: 96.3% (sup=161:75.2%)

EIC2104 <= EIC1102 & EIC1204 & EIC1203 (161:75.2%, 0.963)  
 EIC2104 <= EIC1104 & EIC1102 & EIC1204 & EIC1203 (161:75.2%, 0.963)  
 EIC2104 <= EIC1106 & EIC1102 & EIC1204 & EIC1203 (161:75.2%, 0.963)  
 EIC2104 <= EIC1106 (206:96.3%, 0.835)  
 EIC2104 <= EIC1102 & EIC1103 & EIC1101 & EIC1204 (166:77.6%, 0.958)  
 EIC2104 <= EIC1205 (198:92.5%, 0.864)  
 EIC2104 <= EIC1102 & EIC1103 & EIC1202 & EIC1204 (169:79.0%, 0.953)  
 EIC2104 <= EIC1104 & EIC1103 & EIC1202 & EIC1204 (169:79.0%, 0.953)  
 EIC2104 <= EIC1104 (197:92.1%, 0.868)  
 EIC2104 <= EIC1205 & EIC1103 & EIC1202 & EIC1204 (170:79.4%, 0.947)  
 EIC2104 <= EIC1205 & EIC1102 & EIC1103 & EIC1204 (170:79.4%, 0.947)  
 EIC2104 <= EIC1106 & EIC1103 & EIC1202 & EIC1204 (170:79.4%, 0.947)  
 EIC2104 <= EIC1106 & EIC1102 & EIC1103 & EIC1204 (170:79.4%, 0.947)  
 EIC2104 <= EIC1104 & EIC1102 & EIC1103 & EIC1204 (170:79.4%, 0.947)  
 EIC2104 <= EIC1103 & EIC1202 & EIC1204 (170:79.4%, 0.947)  
 EIC2104 <= EIC1102 & EIC1103 & EIC1204 (170:79.4%, 0.947)  
 EIC2104 <= EIC1106 & EIC1205 (195:91.1%, 0.872)  
 EIC2104 <= EIC1205 & EIC1102 & EIC1103 & EIC1202 (175:81.8%, 0.943)  
 EIC2104 <= EIC1205 & EIC1104 & EIC1103 & EIC1202 (175:81.8%, 0.943)  
 EIC2104 <= EIC1106 & EIC1104 (193:90.2%, 0.881)  
 EIC2104 <= EIC1205 & EIC1104 & EIC1102 & EIC1202 (178:83.2%, 0.938)  
 EIC2104 <= EIC1106 & EIC1102 (190:88.8%, 0.884)  
 EIC2104 <= EIC1106 & EIC1205 & EIC1104 & EIC1102 (180:84.1%, 0.928)  
 EIC2104 <= EIC1106 & EIC1205 & EIC1104 & EIC1202 (180:84.1%, 0.928)  
 EIC2104 <= EIC1205 & EIC1104 & EIC1102 (180:84.1%, 0.928)  
 EIC2104 <= EIC1205 & EIC1104 & EIC1202 (180:84.1%, 0.928)  
 EIC2104 <= EIC1106 & EIC1202 (189:88.3%, 0.889)  
 EIC2104 <= EIC1106 & EIC1205 & EIC1102 & EIC1202 (181:84.6%, 0.923)  
 EIC2104 <= EIC1106 & EIC1104 & EIC1102 & EIC1202 (181:84.6%, 0.923)  
 EIC2104 <= EIC1104 & EIC1102 & EIC1202 (181:84.6%, 0.923)  
 EIC2104 <= EIC1205 & EIC1104 (187:87.4%, 0.898)  
 EIC2104 <= EIC1205 & EIC1102 & EIC1202 (182:85.0%, 0.918)  
 EIC2104 <= EIC1104 & EIC1102 (186:86.9%, 0.903)  
 EIC2104 <= EIC1106 & EIC1205 & EIC1104 (186:86.9%, 0.903)  
 EIC2104 <= EIC1205 & EIC1202 (186:86.9%, 0.903)  
 EIC2104 <= EIC1106 & EIC1205 & EIC1202 (185:86.4%, 0.908)  
 EIC2104 <= EIC1106 & EIC1104 & EIC1102 (185:86.4%, 0.908)

**Figura III.29 - Regras de associação: disciplina EIC2104 (anual) após o filtro**

Melhor suporte: 206:96.3% (conf=86.4%)  
 Melhor confiança: 98.2% (sup=163:76.2%)

EIC2105 <= EIC1104 & EIC1204 & EIC1203 (163:76.2%, 0.982)  
 EIC2105 <= EIC1106 & EIC1104 & EIC1204 & EIC1203 (163:76.2%, 0.982)  
 EIC2105 <= EIC1106 (206:96.3%, 0.864)  
 EIC2105 <= EIC1104 & EIC1103 & EIC1101 & EIC1204 (167:78.0%, 0.976)  
 EIC2105 <= EIC1205 (198:92.5%, 0.889)  
 EIC2105 <= EIC1104 & EIC1103 & EIC1204 (171:79.9%, 0.971)  
 EIC2105 <= EIC1106 & EIC1104 & EIC1103 & EIC1204 (171:79.9%, 0.971)  
 EIC2105 <= EIC1205 & EIC1104 & EIC1103 & EIC1204 (171:79.9%, 0.971)  
 EIC2105 <= EIC1104 (197:92.1%, 0.898)  
 EIC2105 <= EIC1104 & EIC1203 (173:80.8%, 0.965)  
 EIC2105 <= EIC1106 & EIC1205 (195:91.1%, 0.903)  
 EIC2105 <= EIC1205 & EIC1104 & EIC1102 & EIC1103 (176:82.2%, 0.96)  
 EIC2105 <= EIC1106 & EIC1104 (193:90.2%, 0.912)  
 EIC2105 <= EIC1205 & EIC1104 & EIC1103 (180:84.1%, 0.956)  
 EIC2105 <= EIC1106 & EIC1205 & EIC1104 & EIC1103 (180:84.1%, 0.956)  
 EIC2105 <= EIC1205 & EIC1104 (187:87.4%, 0.93)  
 EIC2105 <= EIC1106 & EIC1204 (181:84.6%, 0.939)  
 EIC2105 <= EIC1104 & EIC1103 (186:86.9%, 0.935)  
 EIC2105 <= EIC1106 & EIC1205 & EIC1104 (186:86.9%, 0.935)

**Figura III.30 - Regras de associação: disciplina EIC2105 (anual) após o filtro**

Melhor suporte: 206:96.3% (conf=82.0%)  
 Melhor confiança: 95.0% (sup=160:74.8%)

EIC2200 <= EIC1102 & EIC1103 & EIC1204 & EIC1203 (160:74.8%, 0.95)  
 EIC2200 <= EIC1102 & EIC1202 & EIC1204 & EIC1203 (160:74.8%, 0.95)  
 EIC2200 <= EIC1104 & EIC1202 & EIC1204 & EIC1203 (160:74.8%, 0.95)  
 EIC2200 <= EIC1205 & EIC1102 & EIC1204 & EIC1203 (160:74.8%, 0.95)  
 EIC2200 <= EIC1106 (206:96.3%, 0.82)  
 EIC2200 <= EIC1102 & EIC1103 & EIC1206 & EIC1204 (164:76.6%, 0.945)  
 EIC2200 <= EIC1205 (198:92.5%, 0.838)  
 EIC2200 <= EIC1102 & EIC1103 & EIC1101 & EIC1204 (166:77.6%, 0.94)  
 EIC2200 <= EIC1104 (197:92.1%, 0.843)  
 EIC2200 <= EIC1205 & EIC1102 & EIC1103 & EIC1204 (170:79.4%, 0.935)  
 EIC2200 <= EIC1106 & EIC1102 & EIC1103 & EIC1204 (170:79.4%, 0.935)  
 EIC2200 <= EIC1104 & EIC1102 & EIC1103 & EIC1204 (170:79.4%, 0.935)  
 EIC2200 <= EIC1102 & EIC1103 & EIC1204 (170:79.4%, 0.935)  
 EIC2200 <= EIC1106 & EIC1205 (195:91.1%, 0.846)  
 EIC2200 <= EIC1104 & EIC1103 & EIC1204 (171:79.9%, 0.93)  
 EIC2200 <= EIC1106 & EIC1104 & EIC1103 & EIC1204 (171:79.9%, 0.93)  
 EIC2200 <= EIC1205 & EIC1104 & EIC1103 & EIC1204 (171:79.9%, 0.93)  
 EIC2200 <= EIC1106 & EIC1104 (193:90.2%, 0.86)  
 EIC2200 <= EIC1106 & EIC1205 & EIC1103 & EIC1204 (172:80.4%, 0.924)  
 EIC2200 <= EIC1106 & EIC1103 & EIC1204 (172:80.4%, 0.924)  
 EIC2200 <= EIC1202 (191:89.3%, 0.869)  
 EIC2200 <= EIC1205 & EIC1102 & EIC1204 (176:82.2%, 0.92)  
 EIC2200 <= EIC1106 & EIC1202 (189:88.3%, 0.873)  
 EIC2200 <= EIC1102 & EIC1204 (177:82.7%, 0.915)  
 EIC2200 <= EIC1104 & EIC1103 (186:86.9%, 0.887)  
 EIC2200 <= EIC1104 & EIC1103 & EIC1202 (178:83.2%, 0.91)  
 EIC2200 <= EIC1106 & EIC1104 & EIC1103 & EIC1202 (178:83.2%, 0.91)  
 EIC2200 <= EIC1106 & EIC1104 & EIC1103 (185:86.4%, 0.892)  
 EIC2200 <= EIC1106 & EIC1104 & EIC1102 & EIC1103 (179:83.6%, 0.905)  
 EIC2200 <= EIC1104 & EIC1102 & EIC1103 (179:83.6%, 0.905)  
 EIC2200 <= EIC1106 & EIC1204 (181:84.6%, 0.895)  
 EIC2200 <= EIC1106 & EIC1102 & EIC1103 (181:84.6%, 0.895)  
 EIC2200 <= EIC1102 & EIC1103 (181:84.6%, 0.895)  
 EIC2200 <= EIC1103 & EIC1202 (180:84.1%, 0.9)  
 EIC2200 <= EIC1106 & EIC1103 & EIC1202 (180:84.1%, 0.9)  
 EIC2200 <= EIC1106 & EIC1205 & EIC1104 & EIC1102 (180:84.1%, 0.9)  
 EIC2200 <= EIC1106 & EIC1205 & EIC1104 & EIC1103 (180:84.1%, 0.9)  
 EIC2200 <= EIC1205 & EIC1104 & EIC1102 (180:84.1%, 0.9)  
 EIC2200 <= EIC1205 & EIC1104 & EIC1103 (180:84.1%, 0.9)  
 EIC2200 <= EIC1205 & EIC1204 (180:84.1%, 0.9)

**Figura III.31 - Regras de associação: disciplina EIC2200 (anual) após o filtro**

Melhor suporte: 206:96.3% (conf=84.0%)	
Melhor confiança: 98.1% (sup=161:75.2%)	
EIC2201	<= EIC1102 & EIC1101 & EIC1204 & EIC1200 (161:75.2%, 0.981)
EIC2201	<= EIC1102 & EIC1204 & EIC1100 & EIC1200 (161:75.2%, 0.981)
EIC2201	<= EIC1106 (206:96.3%, 0.84)
EIC2201	<= EIC1205 & EIC1102 & EIC1204 & EIC1200 (162:75.7%, 0.975)
EIC2201	<= EIC1205 & EIC1102 & EIC1204 & EIC1100 (162:75.7%, 0.975)
EIC2201	<= EIC1202 & EIC1204 & EIC1100 & EIC1200 (162:75.7%, 0.975)
EIC2201	<= EIC1202 & EIC1101 & EIC1204 & EIC1200 (162:75.7%, 0.975)
EIC2201	<= EIC1106 & EIC1102 & EIC1204 & EIC1200 (162:75.7%, 0.975)
EIC2201	<= EIC1106 & EIC1102 & EIC1204 & EIC1100 (162:75.7%, 0.975)
EIC2201	<= EIC1104 & EIC1102 & EIC1206 & EIC1203 (162:75.7%, 0.975)
EIC2201	<= EIC1102 & EIC1204 & EIC1200 (162:75.7%, 0.975)
EIC2201	<= EIC1102 & EIC1204 & EIC1100 (162:75.7%, 0.975)
EIC2201	<= EIC1102 & EIC1202 & EIC1204 & EIC1200 (162:75.7%, 0.975)
EIC2201	<= EIC1102 & EIC1202 & EIC1204 & EIC1100 (162:75.7%, 0.975)
EIC2201	<= EIC1102 & EIC1101 & EIC1204 & EIC1100 (162:75.7%, 0.975)
EIC2201	<= EIC1205 (198:92.5%, 0.864)
EIC2201	<= EIC1102 & EIC1103 & EIC1101 & EIC1204 (166:77.6%, 0.97)
EIC2201	<= EIC1104 (197:92.1%, 0.873)
EIC2201	<= EIC1205 & EIC1104 & EIC1102 & EIC1100 (167:78.0%, 0.964)
EIC2201	<= EIC1106 & EIC1104 & EIC1102 & EIC1203 (167:78.0%, 0.964)
EIC2201	<= EIC1104 & EIC1103 & EIC1101 & EIC1204 (167:78.0%, 0.964)
EIC2201	<= EIC1104 & EIC1102 & EIC1203 (167:78.0%, 0.964)
EIC2201	<= EIC1103 & EIC1202 & EIC1101 & EIC1204 (167:78.0%, 0.964)
EIC2201	<= EIC1106 & EIC1104 (193:90.2%, 0.881)
EIC2201	<= EIC1102 & EIC1101 & EIC1204 (170:79.4%, 0.959)
EIC2201	<= EIC1102 & EIC1103 & EIC1204 (170:79.4%, 0.959)
EIC2201	<= EIC1102 & EIC1202 & EIC1101 & EIC1204 (170:79.4%, 0.959)
EIC2201	<= EIC1104 & EIC1102 & EIC1103 & EIC1204 (170:79.4%, 0.959)
EIC2201	<= EIC1106 & EIC1102 & EIC1101 & EIC1204 (170:79.4%, 0.959)
EIC2201	<= EIC1106 & EIC1102 & EIC1101 & EIC1204 (170:79.4%, 0.959)
EIC2201	<= EIC1205 & EIC1102 & EIC1101 & EIC1204 (170:79.4%, 0.959)
EIC2201	<= EIC1205 & EIC1102 & EIC1101 & EIC1206 (170:79.4%, 0.959)
EIC2201	<= EIC1205 & EIC1102 & EIC1103 & EIC1204 (170:79.4%, 0.959)
EIC2201	<= EIC1205 & EIC1102 & EIC1103 & EIC1206 (170:79.4%, 0.959)
EIC2201	<= EIC1102 (192:89.7%, 0.885)
EIC2201	<= EIC1103 (192:89.7%, 0.885)
EIC2201	<= EIC1205 & EIC1102 & EIC1103 & EIC1101 (173:80.8%, 0.954)
EIC2201	<= EIC1106 & EIC1102 (190:88.8%, 0.895)
EIC2201	<= EIC1205 & EIC1104 & EIC1102 & EIC1103 (176:82.2%, 0.949)
EIC2201	<= EIC1205 & EIC1104 (187:87.4%, 0.898)
EIC2201	<= EIC1205 & EIC1102 & EIC1103 (177:82.7%, 0.944)
EIC2201	<= EIC1205 & EIC1102 & EIC1101 (177:82.7%, 0.944)
EIC2201	<= EIC1106 & EIC1205 & EIC1102 & EIC1103 (177:82.7%, 0.944)
EIC2201	<= EIC1106 & EIC1205 & EIC1102 & EIC1101 (177:82.7%, 0.944)
EIC2201	<= EIC1104 & EIC1103 (186:86.9%, 0.909)
EIC2201	<= EIC1104 & EIC1102 (186:86.9%, 0.909)
EIC2201	<= EIC1106 & EIC1205 & EIC1104 & EIC1102 (180:84.1%, 0.933)
EIC2201	<= EIC1205 & EIC1104 & EIC1102 (180:84.1%, 0.933)
EIC2201	<= EIC1106 & EIC1104 & EIC1102 (185:86.4%, 0.914)
EIC2201	<= EIC1205 & EIC1102 (185:86.4%, 0.914)
EIC2201	<= EIC1106 & EIC1205 & EIC1102 & EIC1202 (181:84.6%, 0.923)
EIC2201	<= EIC1106 & EIC1205 & EIC1101 (181:84.6%, 0.923)
EIC2201	<= EIC1106 & EIC1204 (181:84.6%, 0.923)
EIC2201	<= EIC1106 & EIC1102 & EIC1103 (181:84.6%, 0.923)
EIC2201	<= EIC1102 & EIC1103 (181:84.6%, 0.923)
EIC2201	<= EIC1106 & EIC1205 & EIC1102 (184:86.0%, 0.918)

**Figura III.32 - Regras de associação: disciplina EIC2201 (anual) após o filtro**





EIC2202	<=	EIC1102	&	EIC1103	&	EIC1101	&	EIC1204	(166:77.6%, 0.873)
EIC2202	<=	EIC1103	&	EIC1202	&	EIC1101			(176:82.2%, 0.835)
EIC2202	<=	EIC1106	&	EIC1102	&	EIC1204			(176:82.2%, 0.835)
EIC2202	<=	EIC1106	&	EIC1103	&	EIC1202	&	EIC1101	(176:82.2%, 0.835)
EIC2202	<=	EIC1104	&	EIC1103	&	EIC1101	&	EIC1204	(167:78.0%, 0.868)
EIC2202	<=	EIC1103	&	EIC1202	&	EIC1101	&	EIC1204	(167:78.0%, 0.868)
EIC2202	<=	EIC1106	&	EIC1104	&	EIC1102	&	EIC1204	(175:81.8%, 0.84)
EIC2202	<=	EIC1104	&	EIC1102	&	EIC1204			(175:81.8%, 0.84)
EIC2202	<=	EIC1103	&	EIC1101	&	EIC1204			(168:78.5%, 0.863)
EIC2202	<=	EIC1104	&	EIC1102	&	EIC1206	&	EIC1204	(168:78.5%, 0.863)
EIC2202	<=	EIC1106	&	EIC1103	&	EIC1101	&	EIC1204	(168:78.5%, 0.863)
EIC2202	<=	EIC1205	&	EIC1103	&	EIC1101	&	EIC1204	(168:78.5%, 0.863)
EIC2202	<=	EIC1104	&	EIC1103	&	EIC1202	&	EIC1101	(174:81.3%, 0.845)
EIC2202	<=	EIC1106	&	EIC1102	&	EIC1206	&	EIC1204	(169:79.0%, 0.858)
EIC2202	<=	EIC1104	&	EIC1202	&	EIC1101	&	EIC1204	(169:79.0%, 0.858)
EIC2202	<=	EIC1104	&	EIC1103	&	EIC1202	&	EIC1204	(169:79.0%, 0.858)
EIC2202	<=	EIC1104	&	EIC1102	&	EIC1101	&	EIC1204	(169:79.0%, 0.858)
EIC2202	<=	EIC1102	&	EIC1206	&	EIC1204			(169:79.0%, 0.858)
EIC2202	<=	EIC1102	&	EIC1103	&	EIC1202	&	EIC1204	(169:79.0%, 0.858)
EIC2202	<=	EIC1205	&	EIC1202	&	EIC1101	&	EIC1204	(171:79.9%, 0.848)
EIC2202	<=	EIC1205	&	EIC1104	&	EIC1103	&	EIC1204	(171:79.9%, 0.848)
EIC2202	<=	EIC1202	&	EIC1101	&	EIC1204			(171:79.9%, 0.848)
EIC2202	<=	EIC1106	&	EIC1202	&	EIC1101	&	EIC1204	(171:79.9%, 0.848)
EIC2202	<=	EIC1106	&	EIC1104	&	EIC1103	&	EIC1204	(171:79.9%, 0.848)
EIC2202	<=	EIC1104	&	EIC1103	&	EIC1204			(171:79.9%, 0.848)
EIC2202	<=	EIC1104	&	EIC1103	&	EIC1202	&	EIC1206	(171:79.9%, 0.848)
EIC2202	<=	EIC1102	&	EIC1101	&	EIC1204			(170:79.4%, 0.853)
EIC2202	<=	EIC1102	&	EIC1103	&	EIC1204			(170:79.4%, 0.853)
EIC2202	<=	EIC1102	&	EIC1202	&	EIC1101	&	EIC1204	(170:79.4%, 0.853)
EIC2202	<=	EIC1103	&	EIC1202	&	EIC1101	&	EIC1206	(170:79.4%, 0.853)
EIC2202	<=	EIC1103	&	EIC1202	&	EIC1204			(170:79.4%, 0.853)
EIC2202	<=	EIC1104	&	EIC1101	&	EIC1204			(170:79.4%, 0.853)
EIC2202	<=	EIC1104	&	EIC1102	&	EIC1103	&	EIC1204	(170:79.4%, 0.853)
EIC2202	<=	EIC1104	&	EIC1103	&	EIC1101	&	EIC1206	(170:79.4%, 0.853)
EIC2202	<=	EIC1104	&	EIC1202	&	EIC1101	&	EIC1206	(170:79.4%, 0.853)
EIC2202	<=	EIC1104	&	EIC1206	&	EIC1204			(170:79.4%, 0.853)
EIC2202	<=	EIC1106	&	EIC1102	&	EIC1101	&	EIC1204	(170:79.4%, 0.853)
EIC2202	<=	EIC1106	&	EIC1102	&	EIC1103	&	EIC1204	(170:79.4%, 0.853)
EIC2202	<=	EIC1106	&	EIC1103	&	EIC1202	&	EIC1204	(170:79.4%, 0.853)
EIC2202	<=	EIC1106	&	EIC1104	&	EIC1101	&	EIC1204	(170:79.4%, 0.853)
EIC2202	<=	EIC1106	&	EIC1104	&	EIC1206	&	EIC1204	(170:79.4%, 0.853)
EIC2202	<=	EIC1205	&	EIC1102	&	EIC1101	&	EIC1204	(170:79.4%, 0.853)
EIC2202	<=	EIC1205	&	EIC1102	&	EIC1103	&	EIC1204	(170:79.4%, 0.853)
EIC2202	<=	EIC1205	&	EIC1103	&	EIC1202	&	EIC1204	(170:79.4%, 0.853)
EIC2202	<=	EIC1205	&	EIC1104	&	EIC1101	&	EIC1204	(170:79.4%, 0.853)

**Figura III.34 – Regras de associação: disciplina EIC2202 (anual) após o filtro (cont.)**

Melhor suporte: 206:96.3% (conf=82.5%)									
Melhor confiança: 98.7% (sup=159:74.3%)									
EIC2203	<=	EIC1104	&	EIC1206	&	EIC1204	&	EIC1203	(159:74.3%, 0.987)
EIC2203	<=	EIC1106							(206:96.3%, 0.825)
EIC2203	<=	EIC1104	&	EIC1206	&	EIC1203			(166:77.6%, 0.982)
EIC2203	<=	EIC1205							(198:92.5%, 0.848)
EIC2203	<=	EIC1206	&	EIC1203					(168:78.5%, 0.97)
EIC2203	<=	EIC1104							(197:92.1%, 0.863)
EIC2203	<=	EIC1205	&	EIC1104	&	EIC1203			(170:79.4%, 0.965)
EIC2203	<=	EIC1106	&	EIC1205	&	EIC1104	&	EIC1203	(170:79.4%, 0.965)
EIC2203	<=	EIC1106	&	EIC1104					(193:90.2%, 0.876)
EIC2203	<=	EIC1104	&	EIC1203					(173:80.8%, 0.96)
EIC2203	<=	EIC1205	&	EIC1104					(187:87.4%, 0.898)
EIC2203	<=	EIC1106	&	EIC1203					(175:81.8%, 0.943)
EIC2203	<=	EIC1106	&	EIC1205	&	EIC1104			(186:86.9%, 0.903)
EIC2203	<=	EIC1203							(177:82.7%, 0.938)
EIC2203	<=	EIC1104	&	EIC1101					(182:85.0%, 0.907)
EIC2203	<=	EIC1205	&	EIC1101					(182:85.0%, 0.907)
EIC2203	<=	EIC1106	&	EIC1205	&	EIC1104	&	EIC1101	(178:83.2%, 0.927)
EIC2203	<=	EIC1104	&	EIC1206					(181:84.6%, 0.923)

**Figura III.35 - Regras de associação: disciplina EIC2203 (anual) após o filtro**

Melhor suporte: 186:86.9% (conf=80.6%)  
 Melhor confiança: 91.7% (sup=157:73.4%)

EIC2204	<=	EIC1101	&	EIC1206	&	EIC1204	&	EIC1203	(157:73.4%, 0.917)
EIC2204	<=	EIC1202	&	EIC1204	&	EIC1203	&	EIC1100	(157:73.4%, 0.917)
EIC2204	<=	EIC1202	&	EIC1204	&	EIC1203	&	EIC1200	(157:73.4%, 0.917)
EIC2204	<=	EIC1202	&	EIC1206	&	EIC1204	&	EIC1203	(157:73.4%, 0.917)
EIC2204	<=	EIC1205	&	EIC1202	(186:86.9%, 0.806)				
EIC2204	<=	EIC1205	&	EIC1206	&	EIC1204	&	EIC1203	(159:74.3%, 0.912)
EIC2204	<=	EIC1106	&	EIC1205	&	EIC1202	(185:86.4%, 0.811)		
EIC2204	<=	EIC1103	&	EIC1202	&	EIC1204	&	EIC1100	(160:74.8%, 0.906)
EIC2204	<=	EIC1106	&	EIC1206	&	EIC1204	&	EIC1203	(160:74.8%, 0.906)
EIC2204	<=	EIC1202	&	EIC1101	&	EIC1204	&	EIC1203	(160:74.8%, 0.906)
EIC2204	<=	EIC1202	&	EIC1206	&	EIC1203	&	EIC1100	(160:74.8%, 0.906)
EIC2204	<=	EIC1202	&	EIC1206	&	EIC1204	&	EIC1100	(160:74.8%, 0.906)
EIC2204	<=	EIC1206	&	EIC1204	&	EIC1203	(160:74.8%, 0.906)		
EIC2204	<=	EIC1204	(183:85.5%, 0.814)						
EIC2204	<=	EIC1202	&	EIC1101	&	EIC1206	&	EIC1203	(162:75.7%, 0.901)
EIC2204	<=	EIC1205	&	EIC1101	(182:85.0%, 0.819)				
EIC2204	<=	EIC1103	&	EIC1202	&	EIC1203	&	EIC1100	(163:76.2%, 0.896)
EIC2204	<=	EIC1103	&	EIC1202	&	EIC1206	&	EIC1100	(163:76.2%, 0.896)
EIC2204	<=	EIC1103	&	EIC1202	&	EIC1206	&	EIC1203	(163:76.2%, 0.896)
EIC2204	<=	EIC1104	&	EIC1202	&	EIC1206	&	EIC1100	(163:76.2%, 0.896)
EIC2204	<=	EIC1106	&	EIC1202	&	EIC1203	&	EIC1100	(163:76.2%, 0.896)
EIC2204	<=	EIC1106	&	EIC1202	&	EIC1204	&	EIC1100	(163:76.2%, 0.896)
EIC2204	<=	EIC1106	&	EIC1202	&	EIC1206	&	EIC1203	(163:76.2%, 0.896)
EIC2204	<=	EIC1106	&	EIC1205	&	EIC1204	&	EIC1203	(163:76.2%, 0.896)
EIC2204	<=	EIC1202	&	EIC1101	&	EIC1203	&	EIC1100	(163:76.2%, 0.896)
EIC2204	<=	EIC1202	&	EIC1101	&	EIC1204	&	EIC1100	(163:76.2%, 0.896)
EIC2204	<=	EIC1202	&	EIC1203	&	EIC1100	(163:76.2%, 0.896)		
EIC2204	<=	EIC1202	&	EIC1204	&	EIC1100	(163:76.2%, 0.896)		
EIC2204	<=	EIC1202	&	EIC1206	&	EIC1203	(163:76.2%, 0.896)		
EIC2204	<=	EIC1205	&	EIC1202	&	EIC1203	&	EIC1100	(163:76.2%, 0.896)
EIC2204	<=	EIC1205	&	EIC1202	&	EIC1204	&	EIC1100	(163:76.2%, 0.896)
EIC2204	<=	EIC1205	&	EIC1202	&	EIC1206	&	EIC1203	(163:76.2%, 0.896)
EIC2204	<=	EIC1106	&	EIC1204	(181:84.6%, 0.823)				
EIC2204	<=	EIC1106	&	EIC1205	&	EIC1101	(181:84.6%, 0.823)		
EIC2204	<=	EIC1106	&	EIC1205	&	EIC1206	(181:84.6%, 0.823)		
EIC2204	<=	EIC1205	&	EIC1206	(181:84.6%, 0.823)				
EIC2204	<=	EIC1205	&	EIC1202	&	EIC1206	&	EIC1100	(165:77.1%, 0.891)
EIC2204	<=	EIC1202	&	EIC1101	(180:84.1%, 0.828)				
EIC2204	<=	EIC1106	&	EIC1202	&	EIC1101	(180:84.1%, 0.828)		
EIC2204	<=	EIC1106	&	EIC1202	&	EIC1206	&	EIC1100	(166:77.6%, 0.886)
EIC2204	<=	EIC1106	&	EIC1205	&	EIC1206	&	EIC1203	(166:77.6%, 0.886)
EIC2204	<=	EIC1202	&	EIC1101	&	EIC1206	&	EIC1100	(166:77.6%, 0.886)
EIC2204	<=	EIC1202	&	EIC1206	&	EIC1100	(166:77.6%, 0.886)		
EIC2204	<=	EIC1205	&	EIC1103	&	EIC1202	&	EIC1100	(166:77.6%, 0.886)
EIC2204	<=	EIC1205	&	EIC1104	&	EIC1202	&	EIC1100	(166:77.6%, 0.886)
EIC2204	<=	EIC1205	&	EIC1206	&	EIC1203	(166:77.6%, 0.886)		
EIC2204	<=	EIC1106	&	EIC1205	&	EIC1202	&	EIC1101	(178:83.2%, 0.837)
EIC2204	<=	EIC1205	&	EIC1202	&	EIC1101	(178:83.2%, 0.837)		
EIC2204	<=	EIC1205	&	EIC1202	&	EIC1101	&	EIC1203	(167:78.0%, 0.88)
EIC2204	<=	EIC1205	&	EIC1104	&	EIC1102	&	EIC1100	(167:78.0%, 0.88)
EIC2204	<=	EIC1205	&	EIC1101	&	EIC1206	&	EIC1100	(167:78.0%, 0.88)
EIC2204	<=	EIC1202	&	EIC1101	&	EIC1203	(167:78.0%, 0.88)		
EIC2204	<=	EIC1106	&	EIC1206	&	EIC1203	(167:78.0%, 0.88)		
EIC2204	<=	EIC1106	&	EIC1202	&	EIC1101	&	EIC1203	(167:78.0%, 0.88)
EIC2204	<=	EIC1106	&	EIC1104	&	EIC1202	&	EIC1100	(167:78.0%, 0.88)
EIC2204	<=	EIC1106	&	EIC1103	&	EIC1202	&	EIC1100	(167:78.0%, 0.88)
EIC2204	<=	EIC1104	&	EIC1202	&	EIC1101	&	EIC1100	(167:78.0%, 0.88)
EIC2204	<=	EIC1104	&	EIC1202	&	EIC1100	(167:78.0%, 0.88)		
EIC2204	<=	EIC1104	&	EIC1102	&	EIC1202	&	EIC1100	(167:78.0%, 0.88)
EIC2204	<=	EIC1103	&	EIC1202	&	EIC1101	&	EIC1203	(167:78.0%, 0.88)
EIC2204	<=	EIC1103	&	EIC1202	&	EIC1101	&	EIC1100	(167:78.0%, 0.88)
EIC2204	<=	EIC1103	&	EIC1202	&	EIC1100	(167:78.0%, 0.88)		
EIC2204	<=	EIC1205	&	EIC1202	&	EIC1206	(176:82.2%, 0.841)		
EIC2204	<=	EIC1205	&	EIC1104	&	EIC1206	(176:82.2%, 0.841)		
EIC2204	<=	EIC1205	&	EIC1102	&	EIC1202	&	EIC1101	(176:82.2%, 0.841)
EIC2204	<=	EIC1106	&	EIC1205	&	EIC1104	&	EIC1206	(176:82.2%, 0.841)
EIC2204	<=	EIC1106	&	EIC1103	&	EIC1202	&	EIC1101	(176:82.2%, 0.841)
EIC2204	<=	EIC1103	&	EIC1202	&	EIC1101	(176:82.2%, 0.841)		
EIC2204	<=	EIC1106	&	EIC1205	&	EIC1202	&	EIC1100	(170:79.4%, 0.876)
EIC2204	<=	EIC1205	&	EIC1202	&	EIC1100	(170:79.4%, 0.876)		
EIC2204	<=	EIC1100	(175:81.8%, 0.851)						
EIC2204	<=	EIC1106	&	EIC1100	(175:81.8%, 0.851)				
EIC2204	<=	EIC1202	&	EIC1100	(171:79.9%, 0.871)				
EIC2204	<=	EIC1106	&	EIC1202	&	EIC1100	(171:79.9%, 0.871)		
EIC2204	<=	EIC1205	&	EIC1100	(173:80.8%, 0.861)				
EIC2204	<=	EIC1106	&	EIC1205	&	EIC1100	(173:80.8%, 0.861)		

Figura III.36 - Regras de associação: disciplina EIC2204 (anual) após o filtro

Melhor suporte: 206:96.3% (conf=86.4%)  
 Melhor confiança: 98.1% (sup=161:75.2%)

EIC2205 <= EIC1102 & EIC1204 & EIC1203 (161:75.2%, 0.981)  
 EIC2205 <= EIC1104 & EIC1102 & EIC1204 & EIC1203 (161:75.2%, 0.981)  
 EIC2205 <= EIC1106 & EIC1102 & EIC1204 & EIC1203 (161:75.2%, 0.981)  
 EIC2205 <= EIC1106 (206:96.3%, 0.864)  
 EIC2205 <= EIC1106 & EIC1104 & EIC1102 & EIC1203 (167:78.0%, 0.976)  
 EIC2205 <= EIC1104 & EIC1102 & EIC1203 (167:78.0%, 0.976)  
 EIC2205 <= EIC1205 (198:92.5%, 0.879)  
 EIC2205 <= EIC1102 & EIC1203 (168:78.5%, 0.97)  
 EIC2205 <= EIC1106 & EIC1102 & EIC1203 (168:78.5%, 0.97)  
 EIC2205 <= EIC1104 (197:92.1%, 0.883)  
 EIC2205 <= EIC1104 & EIC1203 (173:80.8%, 0.965)  
 EIC2205 <= EIC1106 & EIC1205 (195:91.1%, 0.892)  
 EIC2205 <= EIC1205 & EIC1104 & EIC1102 & EIC1101 (175:81.8%, 0.96)  
 EIC2205 <= EIC1205 & EIC1104 & EIC1202 & EIC1101 (175:81.8%, 0.96)  
 EIC2205 <= EIC1106 & EIC1104 (193:90.2%, 0.896)  
 EIC2205 <= EIC1205 & EIC1102 & EIC1101 (177:82.7%, 0.955)  
 EIC2205 <= EIC1106 & EIC1205 & EIC1102 & EIC1101 (177:82.7%, 0.955)  
 EIC2205 <= EIC1106 & EIC1102 (190:88.8%, 0.905)  
 EIC2205 <= EIC1106 & EIC1205 & EIC1104 & EIC1101 (178:83.2%, 0.949)  
 EIC2205 <= EIC1106 & EIC1205 & EIC1202 & EIC1101 (178:83.2%, 0.949)  
 EIC2205 <= EIC1205 & EIC1202 & EIC1101 (178:83.2%, 0.949)  
 EIC2205 <= EIC1205 & EIC1104 (187:87.4%, 0.914)  
 EIC2205 <= EIC1205 & EIC1104 & EIC1101 (179:83.6%, 0.944)  
 EIC2205 <= EIC1205 & EIC1202 (186:86.9%, 0.919)  
 EIC2205 <= EIC1106 & EIC1205 & EIC1104 (186:86.9%, 0.919)  
 EIC2205 <= EIC1106 & EIC1205 & EIC1101 (181:84.6%, 0.939)  
 EIC2205 <= EIC1106 & EIC1205 & EIC1202 (185:86.4%, 0.924)  
 EIC2205 <= EIC1205 & EIC1102 (185:86.4%, 0.924)  
 EIC2205 <= EIC1205 & EIC1101 (182:85.0%, 0.934)  
 EIC2205 <= EIC1106 & EIC1205 & EIC1102 (184:86.0%, 0.929)

**Figura III.37 - Regras de associação: disciplina EIC2205 (anual) após o filtro**

EIC3200	<=	EIC3100	&	EIC3101	{103:72.5%, 0.981}				
EIC3200	<=	EIC3100	&	EIC3102	&	EIC3101	{97:68.3%, 1.0}		
EIC3200	<=	EIC3100	&	EIC3102	{101:71.1%, 0.99}				
EIC3200	<=	EIC3100	&	EIC3104	&	EIC3101	{97:68.3%, 0.99}		
EIC3200	<=	EIC3100	&	EIC3104	&	EIC3102	&	EIC3101	{94:66.2%, 1.0}
EIC3200	<=	EIC3100	&	EIC3104	&	EIC3102	{96:67.6%, 1.0}		
EIC3200	<=	EIC3100	&	EIC3104	{102:71.8%, 0.98}				
EIC3200	<=	EIC3100	{118:83.1%, 0.907}						
EIC3200	<=	EIC3101	{105:73.9%, 0.971}						
EIC3200	<=	EIC3102	&	EIC3101	{98:69.0%, 0.99}				
EIC3200	<=	EIC3102	{105:73.9%, 0.962}						
EIC3200	<=	EIC3103	&	EIC3100	&	EIC3101	{103:72.5%, 0.981}		
EIC3200	<=	EIC3103	&	EIC3100	&	EIC3102	&	EIC3101	{97:68.3%, 1.0}
EIC3200	<=	EIC3103	&	EIC3100	&	EIC3102	{101:71.1%, 0.99}		
EIC3200	<=	EIC3103	&	EIC3100	&	EIC3104	&	EIC3101	{97:68.3%, 0.99}
EIC3200	<=	EIC3103	&	EIC3100	&	EIC3104	&	EIC3102	{96:67.6%, 1.0}
EIC3200	<=	EIC3103	&	EIC3100	&	EIC3104	{101:71.1%, 0.98}		
EIC3200	<=	EIC3103	&	EIC3100	{111:78.2%, 0.946}				
EIC3200	<=	EIC3103	&	EIC3101	{104:73.2%, 0.971}				
EIC3200	<=	EIC3103	&	EIC3102	&	EIC3101	{98:69.0%, 0.99}		
EIC3200	<=	EIC3103	&	EIC3102	{104:73.2%, 0.971}				
EIC3200	<=	EIC3103	&	EIC3104	&	EIC3101	{98:69.0%, 0.98}		
EIC3200	<=	EIC3103	&	EIC3104	&	EIC3102	&	EIC3101	{95:66.9%, 0.989}
EIC3200	<=	EIC3103	&	EIC3104	&	EIC3102	{97:68.3%, 0.99}		
EIC3200	<=	EIC3103	&	EIC3104	{102:71.8%, 0.971}				
EIC3200	<=	EIC3103	{118:83.1%, 0.907}						
EIC3200	<=	EIC3104	&	EIC3101	{99:69.7%, 0.98}				
EIC3200	<=	EIC3104	&	EIC3102	&	EIC3101	{95:66.9%, 0.989}		
EIC3200	<=	EIC3104	&	EIC3102	{98:69.0%, 0.98}				
EIC3200	<=	EIC3104	{106:74.6%, 0.953}						
EIC3200	<=	EIC3105	&	EIC3100	&	EIC3101	{102:71.8%, 0.98}		
EIC3200	<=	EIC3105	&	EIC3100	&	EIC3102	&	EIC3101	{97:68.3%, 1.0}
EIC3200	<=	EIC3105	&	EIC3100	&	EIC3102	{100:70.4%, 1.0}		
EIC3200	<=	EIC3105	&	EIC3100	&	EIC3104	&	EIC3101	{97:68.3%, 0.99}
EIC3200	<=	EIC3105	&	EIC3100	&	EIC3104	&	EIC3102	{96:67.6%, 1.0}
EIC3200	<=	EIC3105	&	EIC3100	&	EIC3104	{101:71.1%, 0.99}		
EIC3200	<=	EIC3105	&	EIC3100	{113:79.6%, 0.938}				
EIC3200	<=	EIC3105	&	EIC3101	{103:72.5%, 0.981}				
EIC3200	<=	EIC3105	&	EIC3102	&	EIC3101	{97:68.3%, 1.0}		
EIC3200	<=	EIC3105	&	EIC3102	{101:71.1%, 1.0}				
EIC3200	<=	EIC3105	&	EIC3103	&	EIC3100	&	EIC3101	{102:71.8%, 0.98}
EIC3200	<=	EIC3105	&	EIC3103	&	EIC3100	&	EIC3102	{100:70.4%, 1.0}
EIC3200	<=	EIC3105	&	EIC3103	&	EIC3100	&	EIC3104	{100:70.4%, 0.99}
EIC3200	<=	EIC3105	&	EIC3103	&	EIC3100	{108:76.1%, 0.963}		
EIC3200	<=	EIC3105	&	EIC3103	&	EIC3101	{102:71.8%, 0.98}		
EIC3200	<=	EIC3105	&	EIC3103	&	EIC3102	&	EIC3101	{97:68.3%, 1.0}
EIC3200	<=	EIC3105	&	EIC3103	&	EIC3102	{101:71.1%, 1.0}		
EIC3200	<=	EIC3105	&	EIC3103	&	EIC3104	&	EIC3101	{97:68.3%, 0.99}
EIC3200	<=	EIC3105	&	EIC3103	&	EIC3104	&	EIC3102	{96:67.6%, 1.0}
EIC3200	<=	EIC3105	&	EIC3103	&	EIC3104	{100:70.4%, 0.99}		
EIC3200	<=	EIC3105	&	EIC3103	{111:78.2%, 0.946}				
EIC3200	<=	EIC3105	&	EIC3104	&	EIC3101	{98:69.0%, 0.99}		
EIC3200	<=	EIC3105	&	EIC3104	&	EIC3102	&	EIC3101	{94:66.2%, 1.0}
EIC3200	<=	EIC3105	&	EIC3104	&	EIC3102	{96:67.6%, 1.0}		
EIC3200	<=	EIC3105	&	EIC3104	{102:71.8%, 0.99}				
EIC3200	<=	EIC3105	{121:85.2%, 0.893}						

Figura III.38 - Regras de associação: disciplina EIC3200 (sem.)

Melhor suporte:	121:85.2%	(conf=89.3%)					
Melhor confiança:	100.0%	(sup=101:71.1%)					
EIC3200	<=	EIC3105	&	EIC3102	{101:71.1%, 1.0}		
EIC3200	<=	EIC3105	&	EIC3103	&	EIC3102	{101:71.1%, 1.0}
EIC3200	<=	EIC3105	{121:85.2%, 0.893}				
EIC3200	<=	EIC3105	&	EIC3104	{102:71.8%, 0.99}		
EIC3200	<=	EIC3103	{118:83.1%, 0.907}				
EIC3200	<=	EIC3100	{118:83.1%, 0.907}				
EIC3200	<=	EIC3100	&	EIC3101	{103:72.5%, 0.981}		
EIC3200	<=	EIC3103	&	EIC3100	&	EIC3101	{103:72.5%, 0.981}
EIC3200	<=	EIC3105	&	EIC3101	{103:72.5%, 0.981}		
EIC3200	<=	EIC3105	&	EIC3100	{113:79.6%, 0.938}		
EIC3200	<=	EIC3101	{105:73.9%, 0.971}				
EIC3200	<=	EIC3105	&	EIC3103	{111:78.2%, 0.946}		
EIC3200	<=	EIC3103	&	EIC3100	{111:78.2%, 0.946}		
EIC3200	<=	EIC3105	&	EIC3103	&	EIC3100	{108:76.1%, 0.963}

Figura III.39 - Regras de associação: disciplina EIC3200(sem.) após o filtro

EIC3201	<=	EIC3100	&	EIC3101	{	103:72.5%	,	0.981	}				
EIC3201	<=	EIC3100	&	EIC3102	&	EIC3101	{	97:68.3%	,	1.0	}		
EIC3201	<=	EIC3100	&	EIC3102	{	101:71.1%	,	0.99	}				
EIC3201	<=	EIC3100	&	EIC3104	&	EIC3101	{	97:68.3%	,	0.99	}		
EIC3201	<=	EIC3100	&	EIC3104	&	EIC3102	&	EIC3101	{	94:66.2%	,	1.0	}
EIC3201	<=	EIC3100	&	EIC3104	&	EIC3102	{	96:67.6%	,	1.0	}		
EIC3201	<=	EIC3100	&	EIC3104	{	102:71.8%	,	0.99	}				
EIC3201	<=	EIC3100	{	118:83.1%	,	0.924	}						
EIC3201	<=	EIC3101	{	105:73.9%	,	0.962	}						
EIC3201	<=	EIC3102	&	EIC3101	{	98:69.0%	,	0.99	}				
EIC3201	<=	EIC3102	{	105:73.9%	,	0.952	}						
EIC3201	<=	EIC3103	&	EIC3100	&	EIC3101	{	103:72.5%	,	0.981	}		
EIC3201	<=	EIC3103	&	EIC3100	&	EIC3102	&	EIC3101	{	97:68.3%	,	1.0	}
EIC3201	<=	EIC3103	&	EIC3100	&	EIC3102	{	101:71.1%	,	0.99	}		
EIC3201	<=	EIC3103	&	EIC3100	&	EIC3104	&	EIC3101	{	97:68.3%	,	0.99	}
EIC3201	<=	EIC3103	&	EIC3100	&	EIC3104	&	EIC3102	{	96:67.6%	,	1.0	}
EIC3201	<=	EIC3103	&	EIC3100	&	EIC3104	{	101:71.1%	,	0.99	}		
EIC3201	<=	EIC3103	&	EIC3100	{	111:78.2%	,	0.955	}				
EIC3201	<=	EIC3103	&	EIC3101	{	104:73.2%	,	0.971	}				
EIC3201	<=	EIC3103	&	EIC3102	&	EIC3101	{	98:69.0%	,	0.99	}		
EIC3201	<=	EIC3103	&	EIC3102	{	104:73.2%	,	0.962	}				
EIC3201	<=	EIC3103	&	EIC3104	&	EIC3101	{	98:69.0%	,	0.98	}		
EIC3201	<=	EIC3103	&	EIC3104	&	EIC3102	&	EIC3101	{	95:66.9%	,	0.989	}
EIC3201	<=	EIC3103	&	EIC3104	&	EIC3102	{	97:68.3%	,	0.99	}		
EIC3201	<=	EIC3103	&	EIC3104	{	102:71.8%	,	0.98	}				
EIC3201	<=	EIC3103	{	118:83.1%	,	0.907	}						
EIC3201	<=	EIC3104	&	EIC3101	{	99:69.7%	,	0.97	}				
EIC3201	<=	EIC3104	&	EIC3102	&	EIC3101	{	95:66.9%	,	0.989	}		
EIC3201	<=	EIC3104	&	EIC3102	{	98:69.0%	,	0.98	}				
EIC3201	<=	EIC3104	{	106:74.6%	,	0.953	}						
EIC3201	<=	EIC3105	&	EIC3100	&	EIC3101	{	102:71.8%	,	0.98	}		
EIC3201	<=	EIC3105	&	EIC3100	&	EIC3102	&	EIC3101	{	97:68.3%	,	1.0	}
EIC3201	<=	EIC3105	&	EIC3100	&	EIC3102	{	100:70.4%	,	0.99	}		
EIC3201	<=	EIC3105	&	EIC3100	&	EIC3104	&	EIC3101	{	97:68.3%	,	0.99	}
EIC3201	<=	EIC3105	&	EIC3100	&	EIC3104	&	EIC3102	{	96:67.6%	,	1.0	}
EIC3201	<=	EIC3105	&	EIC3100	&	EIC3104	{	101:71.1%	,	0.99	}		
EIC3201	<=	EIC3105	&	EIC3100	{	113:79.6%	,	0.938	}				
EIC3201	<=	EIC3105	&	EIC3101	{	103:72.5%	,	0.971	}				
EIC3201	<=	EIC3105	&	EIC3102	&	EIC3101	{	97:68.3%	,	1.0	}		
EIC3201	<=	EIC3105	&	EIC3102	{	101:71.1%	,	0.98	}				
EIC3201	<=	EIC3105	&	EIC3103	&	EIC3100	&	EIC3101	{	102:71.8%	,	0.98	}
EIC3201	<=	EIC3105	&	EIC3103	&	EIC3100	&	EIC3102	{	100:70.4%	,	0.99	}
EIC3201	<=	EIC3105	&	EIC3103	&	EIC3100	&	EIC3104	{	100:70.4%	,	0.99	}
EIC3201	<=	EIC3105	&	EIC3103	&	EIC3100	{	108:76.1%	,	0.954	}		
EIC3201	<=	EIC3105	&	EIC3103	&	EIC3101	{	102:71.8%	,	0.98	}		
EIC3201	<=	EIC3105	&	EIC3103	&	EIC3102	&	EIC3101	{	97:68.3%	,	1.0	}
EIC3201	<=	EIC3105	&	EIC3103	&	EIC3102	{	101:71.1%	,	0.98	}		
EIC3201	<=	EIC3105	&	EIC3103	&	EIC3104	&	EIC3101	{	97:68.3%	,	0.99	}
EIC3201	<=	EIC3105	&	EIC3103	&	EIC3104	&	EIC3102	{	96:67.6%	,	1.0	}
EIC3201	<=	EIC3105	&	EIC3103	&	EIC3104	{	100:70.4%	,	0.99	}		
EIC3201	<=	EIC3105	&	EIC3103	{	111:78.2%	,	0.928	}				
EIC3201	<=	EIC3105	&	EIC3104	&	EIC3101	{	98:69.0%	,	0.98	}		
EIC3201	<=	EIC3105	&	EIC3104	&	EIC3102	&	EIC3101	{	94:66.2%	,	1.0	}
EIC3201	<=	EIC3105	&	EIC3104	&	EIC3102	{	96:67.6%	,	1.0	}		
EIC3201	<=	EIC3105	&	EIC3104	{	102:71.8%	,	0.98	}				
EIC3201	<=	EIC3105	{	121:85.2%	,	0.876	}						

Figura III.40 - Regras de associação: disciplina EIC3201 (sem.)

Melhor suporte: 121:85.2% (conf=87.6%)													
Melhor confiança: 100.0% (sup=97:68.3%)													
EIC3201	<=	EIC3100	&	EIC3102	&	EIC3101	{	97:68.3%	,	1.0	}		
EIC3201	<=	EIC3103	&	EIC3100	&	EIC3102	&	EIC3101	{	97:68.3%	,	1.0	}
EIC3201	<=	EIC3105	&	EIC3100	&	EIC3102	&	EIC3101	{	97:68.3%	,	1.0	}
EIC3201	<=	EIC3105	&	EIC3102	&	EIC3101	{	97:68.3%	,	1.0	}		
EIC3201	<=	EIC3105	&	EIC3103	&	EIC3102	&	EIC3101	{	97:68.3%	,	1.0	}
EIC3201	<=	EIC3105	{	121:85.2%	,	0.876	}						
EIC3201	<=	EIC3100	&	EIC3104	{	102:71.8%	,	0.99	}				
EIC3201	<=	EIC3100	{	118:83.1%	,	0.924	}						
EIC3201	<=	EIC3100	&	EIC3101	{	103:72.5%	,	0.981	}				
EIC3201	<=	EIC3103	&	EIC3100	&	EIC3101	{	103:72.5%	,	0.981	}		
EIC3201	<=	EIC3105	&	EIC3100	{	113:79.6%	,	0.938	}				
EIC3201	<=	EIC3103	&	EIC3101	{	104:73.2%	,	0.971	}				
EIC3201	<=	EIC3103	&	EIC3100	{	111:78.2%	,	0.955	}				
EIC3201	<=	EIC3101	{	105:73.9%	,	0.962	}						

Figura III.41 - Regras de associação: disciplina EIC3201 (sem.) após o filtro

EIC3202	<=	EIC3100	&	EIC3101	(103:72.5%, 0.942)				
EIC3202	<=	EIC3100	&	EIC3102	&	EIC3101	(97:68.3%, 0.969)		
EIC3202	<=	EIC3100	&	EIC3102	(101:71.1%, 0.97)				
EIC3202	<=	EIC3100	&	EIC3104	&	EIC3101	(97:68.3%, 0.969)		
EIC3202	<=	EIC3100	&	EIC3104	&	EIC3102	&	EIC3101	(94:66.2%, 0.979)
EIC3202	<=	EIC3100	&	EIC3104	&	EIC3102	(96:67.6%, 0.979)		
EIC3202	<=	EIC3100	&	EIC3104	(102:71.8%, 0.971)				
EIC3202	<=	EIC3100	(118:83.1%, 0.881)						
EIC3202	<=	EIC3101	(105:73.9%, 0.924)						
EIC3202	<=	EIC3102	&	EIC3101	(98:69.0%, 0.959)				
EIC3202	<=	EIC3102	(105:73.9%, 0.952)						
EIC3202	<=	EIC3103	&	EIC3100	&	EIC3101	(103:72.5%, 0.942)		
EIC3202	<=	EIC3103	&	EIC3100	&	EIC3102	&	EIC3101	(97:68.3%, 0.969)
EIC3202	<=	EIC3103	&	EIC3100	&	EIC3102	(101:71.1%, 0.97)		
EIC3202	<=	EIC3103	&	EIC3100	&	EIC3104	&	EIC3101	(97:68.3%, 0.969)
EIC3202	<=	EIC3103	&	EIC3100	&	EIC3104	&	EIC3102	(96:67.6%, 0.979)
EIC3202	<=	EIC3103	&	EIC3100	&	EIC3104	(101:71.1%, 0.97)		
EIC3202	<=	EIC3103	&	EIC3100	(111:78.2%, 0.928)				
EIC3202	<=	EIC3103	&	EIC3101	(104:73.2%, 0.933)				
EIC3202	<=	EIC3103	&	EIC3102	&	EIC3101	(98:69.0%, 0.959)		
EIC3202	<=	EIC3103	&	EIC3102	(104:73.2%, 0.962)				
EIC3202	<=	EIC3103	&	EIC3104	&	EIC3101	(98:69.0%, 0.959)		
EIC3202	<=	EIC3103	&	EIC3104	&	EIC3102	&	EIC3101	(95:66.9%, 0.968)
EIC3202	<=	EIC3103	&	EIC3104	&	EIC3102	(97:68.3%, 0.969)		
EIC3202	<=	EIC3103	&	EIC3104	(102:71.8%, 0.961)				
EIC3202	<=	EIC3103	(118:83.1%, 0.898)						
EIC3202	<=	EIC3104	&	EIC3101	(99:69.7%, 0.949)				
EIC3202	<=	EIC3104	&	EIC3102	&	EIC3101	(95:66.9%, 0.968)		
EIC3202	<=	EIC3104	&	EIC3102	(98:69.0%, 0.959)				
EIC3202	<=	EIC3104	(106:74.6%, 0.934)						
EIC3202	<=	EIC3105	&	EIC3100	&	EIC3101	(102:71.8%, 0.941)		
EIC3202	<=	EIC3105	&	EIC3100	&	EIC3102	&	EIC3101	(97:68.3%, 0.969)
EIC3202	<=	EIC3105	&	EIC3100	&	EIC3102	(100:70.4%, 0.97)		
EIC3202	<=	EIC3105	&	EIC3100	&	EIC3104	&	EIC3101	(97:68.3%, 0.969)
EIC3202	<=	EIC3105	&	EIC3100	&	EIC3104	&	EIC3102	(96:67.6%, 0.979)
EIC3202	<=	EIC3105	&	EIC3100	&	EIC3104	(101:71.1%, 0.97)		
EIC3202	<=	EIC3105	&	EIC3100	(113:79.6%, 0.894)				
EIC3202	<=	EIC3105	&	EIC3101	(103:72.5%, 0.932)				
EIC3202	<=	EIC3105	&	EIC3102	&	EIC3101	(97:68.3%, 0.969)		
EIC3202	<=	EIC3105	&	EIC3102	(101:71.1%, 0.97)				
EIC3202	<=	EIC3105	&	EIC3103	&	EIC3100	&	EIC3101	(102:71.8%, 0.941)
EIC3202	<=	EIC3105	&	EIC3103	&	EIC3100	&	EIC3102	(100:70.4%, 0.97)
EIC3202	<=	EIC3105	&	EIC3103	&	EIC3100	&	EIC3104	(100:70.4%, 0.97)
EIC3202	<=	EIC3105	&	EIC3103	&	EIC3100	(108:76.1%, 0.926)		
EIC3202	<=	EIC3105	&	EIC3103	&	EIC3101	(102:71.8%, 0.941)		
EIC3202	<=	EIC3105	&	EIC3103	&	EIC3102	&	EIC3101	(97:68.3%, 0.969)
EIC3202	<=	EIC3105	&	EIC3103	&	EIC3102	(101:71.1%, 0.97)		
EIC3202	<=	EIC3105	&	EIC3103	&	EIC3104	&	EIC3101	(97:68.3%, 0.969)
EIC3202	<=	EIC3105	&	EIC3103	&	EIC3104	&	EIC3102	(96:67.6%, 0.979)
EIC3202	<=	EIC3105	&	EIC3103	&	EIC3104	(100:70.4%, 0.97)		
EIC3202	<=	EIC3105	&	EIC3103	(111:78.2%, 0.91)				
EIC3202	<=	EIC3105	&	EIC3104	&	EIC3101	(98:69.0%, 0.959)		
EIC3202	<=	EIC3105	&	EIC3104	&	EIC3102	&	EIC3101	(94:66.2%, 0.979)
EIC3202	<=	EIC3105	&	EIC3104	&	EIC3102	(96:67.6%, 0.979)		
EIC3202	<=	EIC3105	&	EIC3104	(102:71.8%, 0.961)				
EIC3202	<=	EIC3105	(121:85.2%, 0.843)						

Figura III.42 - Regras de associação: disciplina EIC3202 (sem.)

Melhor suporte:	121:85.2%	(conf=84.3%)							
Melhor confiança:	97.9%	(sup=96:67.6%)							
EIC3202	<=	EIC3100	&	EIC3104	&	EIC3102	(96:67.6%, 0.979)		
EIC3202	<=	EIC3103	&	EIC3100	&	EIC3104	&	EIC3102	(96:67.6%, 0.979)
EIC3202	<=	EIC3105	&	EIC3100	&	EIC3104	&	EIC3102	(96:67.6%, 0.979)
EIC3202	<=	EIC3105	&	EIC3103	&	EIC3104	&	EIC3102	(96:67.6%, 0.979)
EIC3202	<=	EIC3105	&	EIC3104	&	EIC3102	(96:67.6%, 0.979)		
EIC3202	<=	EIC3105	(121:85.2%, 0.843)						
EIC3202	<=	EIC3100	&	EIC3104	(102:71.8%, 0.971)				
EIC3202	<=	EIC3103	(118:83.1%, 0.898)						
EIC3202	<=	EIC3103	&	EIC3102	(104:73.2%, 0.962)				
EIC3202	<=	EIC3103	&	EIC3100	(111:78.2%, 0.928)				
EIC3202	<=	EIC3102	(105:73.9%, 0.952)						
EIC3202	<=	EIC3104	(106:74.6%, 0.934)						

Figura III.43 - Regras de associação: disciplina EIC3202 (sem.) após o filtro

EIC3203	<=	EIC3100	&	EIC3101	(103:72.5%, 0.951)				
EIC3203	<=	EIC3100	&	EIC3102	&	EIC3101	(97:68.3%, 0.979)		
EIC3203	<=	EIC3100	&	EIC3102	(101:71.1%, 0.96)				
EIC3203	<=	EIC3100	&	EIC3104	&	EIC3101	(97:68.3%, 0.979)		
EIC3203	<=	EIC3100	&	EIC3104	&	EIC3102	&	EIC3101	(94:66.2%, 0.989)
EIC3203	<=	EIC3100	&	EIC3104	&	EIC3102	(96:67.6%, 0.979)		
EIC3203	<=	EIC3100	&	EIC3104	(102:71.8%, 0.961)				
EIC3203	<=	EIC3100	(118:83.1%, 0.873)						
EIC3203	<=	EIC3101	(105:73.9%, 0.933)						
EIC3203	<=	EIC3102	&	EIC3101	(98:69.0%, 0.969)				
EIC3203	<=	EIC3102	(105:73.9%, 0.924)						
EIC3203	<=	EIC3103	&	EIC3100	&	EIC3101	(103:72.5%, 0.951)		
EIC3203	<=	EIC3103	&	EIC3100	&	EIC3102	&	EIC3101	(97:68.3%, 0.979)
EIC3203	<=	EIC3103	&	EIC3100	&	EIC3102	(101:71.1%, 0.96)		
EIC3203	<=	EIC3103	&	EIC3100	&	EIC3104	&	EIC3101	(97:68.3%, 0.979)
EIC3203	<=	EIC3103	&	EIC3100	&	EIC3104	&	EIC3102	(96:67.6%, 0.979)
EIC3203	<=	EIC3103	&	EIC3100	&	EIC3104	(101:71.1%, 0.96)		
EIC3203	<=	EIC3103	&	EIC3100	(111:78.2%, 0.91)				
EIC3203	<=	EIC3103	&	EIC3101	(104:73.2%, 0.942)				
EIC3203	<=	EIC3103	&	EIC3102	&	EIC3101	(98:69.0%, 0.969)		
EIC3203	<=	EIC3103	&	EIC3102	(104:73.2%, 0.933)				
EIC3203	<=	EIC3103	&	EIC3104	&	EIC3101	(98:69.0%, 0.969)		
EIC3203	<=	EIC3103	&	EIC3104	&	EIC3102	&	EIC3101	(95:66.9%, 0.979)
EIC3203	<=	EIC3103	&	EIC3104	&	EIC3102	(97:68.3%, 0.969)		
EIC3203	<=	EIC3103	&	EIC3104	(102:71.8%, 0.951)				
EIC3203	<=	EIC3103	(118:83.1%, 0.856)						
EIC3203	<=	EIC3104	&	EIC3101	(99:69.7%, 0.96)				
EIC3203	<=	EIC3104	&	EIC3102	&	EIC3101	(95:66.9%, 0.979)		
EIC3203	<=	EIC3104	&	EIC3102	(98:69.0%, 0.959)				
EIC3203	<=	EIC3104	(106:74.6%, 0.925)						
EIC3203	<=	EIC3105	&	EIC3100	&	EIC3101	(102:71.8%, 0.951)		
EIC3203	<=	EIC3105	&	EIC3100	&	EIC3102	&	EIC3101	(97:68.3%, 0.979)
EIC3203	<=	EIC3105	&	EIC3100	&	EIC3102	(100:70.4%, 0.97)		
EIC3203	<=	EIC3105	&	EIC3100	&	EIC3104	&	EIC3101	(97:68.3%, 0.979)
EIC3203	<=	EIC3105	&	EIC3100	&	EIC3104	&	EIC3102	(96:67.6%, 0.979)
EIC3203	<=	EIC3105	&	EIC3100	&	EIC3104	(101:71.1%, 0.96)		
EIC3203	<=	EIC3105	&	EIC3100	(113:79.6%, 0.894)				
EIC3203	<=	EIC3105	&	EIC3101	(103:72.5%, 0.942)				
EIC3203	<=	EIC3105	&	EIC3102	&	EIC3101	(97:68.3%, 0.979)		
EIC3203	<=	EIC3105	&	EIC3102	(101:71.1%, 0.96)				
EIC3203	<=	EIC3105	&	EIC3103	&	EIC3100	&	EIC3101	(102:71.8%, 0.951)
EIC3203	<=	EIC3105	&	EIC3103	&	EIC3102	&	EIC3101	(97:68.3%, 0.979)
EIC3203	<=	EIC3105	&	EIC3103	&	EIC3102	(101:71.1%, 0.96)		
EIC3203	<=	EIC3105	&	EIC3103	&	EIC3104	&	EIC3101	(97:68.3%, 0.979)
EIC3203	<=	EIC3105	&	EIC3103	&	EIC3104	&	EIC3102	(96:67.6%, 0.979)
EIC3203	<=	EIC3105	&	EIC3103	&	EIC3104	(100:70.4%, 0.96)		
EIC3203	<=	EIC3105	&	EIC3103	(111:78.2%, 0.892)				
EIC3203	<=	EIC3105	&	EIC3104	&	EIC3101	(98:69.0%, 0.969)		
EIC3203	<=	EIC3105	&	EIC3104	&	EIC3102	&	EIC3101	(94:66.2%, 0.989)
EIC3203	<=	EIC3105	&	EIC3104	&	EIC3102	(96:67.6%, 0.979)		
EIC3203	<=	EIC3105	&	EIC3104	(102:71.8%, 0.951)				
EIC3203	<=	EIC3105	(121:85.2%, 0.843)						

Figura III.44 - Regras de associação: disciplina EIC3203 (sem.)

Melhor suporte:	121:85.2%	(conf=84.3%)							
Melhor confiança:	98.9%	(sup=94:66.2%)							
EIC3203	<=	EIC3100	&	EIC3104	&	EIC3102	&	EIC3101	(94:66.2%, 0.989)
EIC3203	<=	EIC3105	&	EIC3104	&	EIC3102	&	EIC3101	(94:66.2%, 0.989)
EIC3203	<=	EIC3105	(121:85.2%, 0.843)						
EIC3203	<=	EIC3105	&	EIC3103	&	EIC3104	&	EIC3101	(97:68.3%, 0.979)
EIC3203	<=	EIC3105	&	EIC3103	&	EIC3102	&	EIC3101	(97:68.3%, 0.979)
EIC3203	<=	EIC3105	&	EIC3102	&	EIC3101	(97:68.3%, 0.979)		
EIC3203	<=	EIC3105	&	EIC3100	&	EIC3104	&	EIC3101	(97:68.3%, 0.979)
EIC3203	<=	EIC3105	&	EIC3100	&	EIC3102	&	EIC3101	(97:68.3%, 0.979)
EIC3203	<=	EIC3103	&	EIC3100	&	EIC3104	&	EIC3101	(97:68.3%, 0.979)
EIC3203	<=	EIC3103	&	EIC3100	&	EIC3102	&	EIC3101	(97:68.3%, 0.979)
EIC3203	<=	EIC3100	&	EIC3104	&	EIC3101	(97:68.3%, 0.979)		
EIC3203	<=	EIC3100	&	EIC3102	&	EIC3101	(97:68.3%, 0.979)		
EIC3203	<=	EIC3100	(118:83.1%, 0.873)						
EIC3203	<=	EIC3105	&	EIC3100	&	EIC3102	(100:70.4%, 0.97)		
EIC3203	<=	EIC3105	&	EIC3103	&	EIC3100	&	EIC3102	(100:70.4%, 0.97)
EIC3203	<=	EIC3105	&	EIC3100	(113:79.6%, 0.894)				
EIC3203	<=	EIC3100	&	EIC3104	(102:71.8%, 0.961)				
EIC3203	<=	EIC3103	&	EIC3100	(111:78.2%, 0.91)				
EIC3203	<=	EIC3100	&	EIC3101	(103:72.5%, 0.951)				
EIC3203	<=	EIC3103	&	EIC3100	&	EIC3101	(103:72.5%, 0.951)		
EIC3203	<=	EIC3105	&	EIC3103	&	EIC3100	(108:76.1%, 0.917)		
EIC3203	<=	EIC3103	&	EIC3101	(104:73.2%, 0.942)				
EIC3203	<=	EIC3104	(106:74.6%, 0.925)						
EIC3203	<=	EIC3101	(105:73.9%, 0.933)						

Figura III.45 - Regras de associação: disciplina EIC3203 (sem.) após o filtro



EIC3204	<=	EIC3100	&	EIC3101	(103:72.5%, 0.971)				
EIC3204	<=	EIC3100	&	EIC3102	&	EIC3101	(97:68.3%, 1.0)		
EIC3204	<=	EIC3100	&	EIC3102	(101:71.1%, 1.0)				
EIC3204	<=	EIC3100	&	EIC3104	&	EIC3101	(97:68.3%, 0.979)		
EIC3204	<=	EIC3100	&	EIC3104	&	EIC3102	&	EIC3101	(94:66.2%, 1.0)
EIC3204	<=	EIC3100	&	EIC3104	(96:67.6%, 1.0)				
EIC3204	<=	EIC3100	&	EIC3104	(102:71.8%, 0.98)				
EIC3204	<=	EIC3100	(118:83.1%, 0.932)						
EIC3204	<=	EIC3101	(105:73.9%, 0.952)						
EIC3204	<=	EIC3102	&	EIC3101	(98:69.0%, 0.99)				
EIC3204	<=	EIC3102	(105:73.9%, 0.962)						
EIC3204	<=	EIC3103	&	EIC3100	&	EIC3101	(103:72.5%, 0.971)		
EIC3204	<=	EIC3103	&	EIC3100	&	EIC3102	&	EIC3101	(97:68.3%, 1.0)
EIC3204	<=	EIC3103	&	EIC3100	(101:71.1%, 1.0)				
EIC3204	<=	EIC3103	&	EIC3100	&	EIC3104	&	EIC3101	(97:68.3%, 0.979)
EIC3204	<=	EIC3103	&	EIC3100	&	EIC3104	(96:67.6%, 1.0)		
EIC3204	<=	EIC3103	&	EIC3100	&	EIC3104	(101:71.1%, 0.98)		
EIC3204	<=	EIC3103	&	EIC3100	(111:78.2%, 0.964)				
EIC3204	<=	EIC3103	&	EIC3101	(104:73.2%, 0.962)				
EIC3204	<=	EIC3103	&	EIC3102	&	EIC3101	(98:69.0%, 0.99)		
EIC3204	<=	EIC3103	&	EIC3102	(104:73.2%, 0.971)				
EIC3204	<=	EIC3103	&	EIC3104	&	EIC3101	(98:69.0%, 0.969)		
EIC3204	<=	EIC3103	&	EIC3104	&	EIC3102	&	EIC3101	(95:66.9%, 0.989)
EIC3204	<=	EIC3103	&	EIC3104	(97:68.3%, 0.99)				
EIC3204	<=	EIC3103	&	EIC3104	(102:71.8%, 0.971)				
EIC3204	<=	EIC3103	(118:83.1%, 0.915)						
EIC3204	<=	EIC3104	&	EIC3101	(99:69.7%, 0.96)				
EIC3204	<=	EIC3104	&	EIC3102	&	EIC3101	(95:66.9%, 0.989)		
EIC3204	<=	EIC3104	&	EIC3102	(98:69.0%, 0.98)				
EIC3204	<=	EIC3104	(106:74.6%, 0.943)						
EIC3204	<=	EIC3105	&	EIC3100	&	EIC3101	(102:71.8%, 0.971)		
EIC3204	<=	EIC3105	&	EIC3100	&	EIC3102	&	EIC3101	(97:68.3%, 1.0)
EIC3204	<=	EIC3105	&	EIC3100	&	EIC3102	(100:70.4%, 1.0)		
EIC3204	<=	EIC3105	&	EIC3100	&	EIC3104	&	EIC3101	(97:68.3%, 0.979)
EIC3204	<=	EIC3105	&	EIC3100	&	EIC3104	&	EIC3102	(96:67.6%, 1.0)
EIC3204	<=	EIC3105	&	EIC3100	&	EIC3104	(101:71.1%, 0.98)		
EIC3204	<=	EIC3105	&	EIC3100	(113:79.6%, 0.947)				
EIC3204	<=	EIC3105	&	EIC3101	(103:72.5%, 0.961)				
EIC3204	<=	EIC3105	&	EIC3102	&	EIC3101	(97:68.3%, 1.0)		
EIC3204	<=	EIC3105	&	EIC3102	(101:71.1%, 0.99)				
EIC3204	<=	EIC3105	&	EIC3103	&	EIC3100	&	EIC3101	(102:71.8%, 0.971)
EIC3204	<=	EIC3105	&	EIC3103	&	EIC3100	(100:70.4%, 1.0)		
EIC3204	<=	EIC3105	&	EIC3103	&	EIC3100	(100:70.4%, 0.98)		
EIC3204	<=	EIC3105	&	EIC3103	&	EIC3100	(108:76.1%, 0.963)		
EIC3204	<=	EIC3105	&	EIC3103	&	EIC3101	(102:71.8%, 0.971)		
EIC3204	<=	EIC3105	&	EIC3103	&	EIC3102	&	EIC3101	(97:68.3%, 1.0)
EIC3204	<=	EIC3105	&	EIC3103	&	EIC3102	(101:71.1%, 0.99)		
EIC3204	<=	EIC3105	&	EIC3103	&	EIC3104	&	EIC3101	(97:68.3%, 0.979)
EIC3204	<=	EIC3105	&	EIC3103	&	EIC3104	(96:67.6%, 1.0)		
EIC3204	<=	EIC3105	&	EIC3103	&	EIC3104	(100:70.4%, 0.98)		
EIC3204	<=	EIC3105	&	EIC3103	(111:78.2%, 0.946)				
EIC3204	<=	EIC3105	&	EIC3104	&	EIC3101	(98:69.0%, 0.969)		
EIC3204	<=	EIC3105	&	EIC3104	&	EIC3102	&	EIC3101	(94:66.2%, 1.0)
EIC3204	<=	EIC3105	&	EIC3104	&	EIC3102	(96:67.6%, 1.0)		
EIC3204	<=	EIC3105	&	EIC3104	(102:71.8%, 0.971)				
EIC3204	<=	EIC3105	(121:85.2%, 0.917)						

Figura III.46 - Regras de associação: disciplina EIC3204 (sem.)

Melhor suporte:	121:85.2%	(conf=91.7%)					
Melhor confiança:	100.0%	(sup=101:71.1%)					
EIC3204	<=	EIC3100	&	EIC3102	(101:71.1%, 1.0)		
EIC3204	<=	EIC3103	&	EIC3100	&	EIC3102	(101:71.1%, 1.0)
EIC3204	<=	EIC3105	(121:85.2%, 0.917)				
EIC3204	<=	EIC3100	&	EIC3104	(102:71.8%, 0.98)		
EIC3204	<=	EIC3100	(118:83.1%, 0.932)				
EIC3204	<=	EIC3103	&	EIC3102	(104:73.2%, 0.971)		
EIC3204	<=	EIC3105	&	EIC3100	(113:79.6%, 0.947)		
EIC3204	<=	EIC3103	&	EIC3100	(111:78.2%, 0.964)		

Figura III.47 - Regras de associação: disciplina EIC3204 (sem.) após o filtro

EIC3205	<=	EIC3100	&	EIC3101	(103:72.5%, 0.981)				
EIC3205	<=	EIC3100	&	EIC3102	&	EIC3101	(97:68.3%, 1.0)		
EIC3205	<=	EIC3100	&	EIC3102	(101:71.1%, 0.99)				
EIC3205	<=	EIC3100	&	EIC3104	&	EIC3101	(97:68.3%, 0.99)		
EIC3205	<=	EIC3100	&	EIC3104	&	EIC3102	&	EIC3101	(94:66.2%, 1.0)
EIC3205	<=	EIC3100	&	EIC3104	&	EIC3102	(96:67.6%, 1.0)		
EIC3205	<=	EIC3100	&	EIC3104	(102:71.8%, 0.98)				
EIC3205	<=	EIC3100	(118:83.1%, 0.915)						
EIC3205	<=	EIC3101	(105:73.9%, 0.971)						
EIC3205	<=	EIC3102	&	EIC3101	(98:69.0%, 0.99)				
EIC3205	<=	EIC3102	(105:73.9%, 0.962)						
EIC3205	<=	EIC3103	&	EIC3100	&	EIC3101	(103:72.5%, 0.981)		
EIC3205	<=	EIC3103	&	EIC3100	&	EIC3102	&	EIC3101	(97:68.3%, 1.0)
EIC3205	<=	EIC3103	&	EIC3100	&	EIC3102	(101:71.1%, 0.99)		
EIC3205	<=	EIC3103	&	EIC3100	&	EIC3104	&	EIC3101	(97:68.3%, 0.99)
EIC3205	<=	EIC3103	&	EIC3100	&	EIC3104	&	EIC3102	(96:67.6%, 1.0)
EIC3205	<=	EIC3103	&	EIC3100	&	EIC3104	(101:71.1%, 0.98)		
EIC3205	<=	EIC3103	&	EIC3100	(111:78.2%, 0.946)				
EIC3205	<=	EIC3103	&	EIC3101	(104:73.2%, 0.971)				
EIC3205	<=	EIC3103	&	EIC3102	&	EIC3101	(98:69.0%, 0.99)		
EIC3205	<=	EIC3103	&	EIC3102	(104:73.2%, 0.971)				
EIC3205	<=	EIC3103	&	EIC3104	&	EIC3101	(98:69.0%, 0.98)		
EIC3205	<=	EIC3103	&	EIC3104	&	EIC3102	&	EIC3101	(95:66.9%, 0.989)
EIC3205	<=	EIC3103	&	EIC3104	&	EIC3102	(97:68.3%, 0.99)		
EIC3205	<=	EIC3103	&	EIC3104	(102:71.8%, 0.971)				
EIC3205	<=	EIC3103	(118:83.1%, 0.898)						
EIC3205	<=	EIC3104	&	EIC3101	(99:69.7%, 0.98)				
EIC3205	<=	EIC3104	&	EIC3102	&	EIC3101	(95:66.9%, 0.989)		
EIC3205	<=	EIC3104	&	EIC3102	(98:69.0%, 0.98)				
EIC3205	<=	EIC3104	(106:74.6%, 0.953)						
EIC3205	<=	EIC3105	&	EIC3100	&	EIC3101	(102:71.8%, 0.98)		
EIC3205	<=	EIC3105	&	EIC3100	&	EIC3102	&	EIC3101	(97:68.3%, 1.0)
EIC3205	<=	EIC3105	&	EIC3100	&	EIC3102	(100:70.4%, 1.0)		
EIC3205	<=	EIC3105	&	EIC3100	&	EIC3104	&	EIC3101	(97:68.3%, 0.99)
EIC3205	<=	EIC3105	&	EIC3100	&	EIC3104	&	EIC3102	(96:67.6%, 1.0)
EIC3205	<=	EIC3105	&	EIC3100	&	EIC3104	(101:71.1%, 0.99)		
EIC3205	<=	EIC3105	&	EIC3100	(113:79.6%, 0.947)				
EIC3205	<=	EIC3105	&	EIC3101	(103:72.5%, 0.981)				
EIC3205	<=	EIC3105	&	EIC3102	&	EIC3101	(97:68.3%, 1.0)		
EIC3205	<=	EIC3105	&	EIC3102	(101:71.1%, 1.0)				
EIC3205	<=	EIC3105	&	EIC3103	&	EIC3100	&	EIC3101	(102:71.8%, 0.98)
EIC3205	<=	EIC3105	&	EIC3103	&	EIC3100	&	EIC3102	(100:70.4%, 1.0)
EIC3205	<=	EIC3105	&	EIC3103	&	EIC3100	&	EIC3104	(100:70.4%, 0.99)
EIC3205	<=	EIC3105	&	EIC3103	&	EIC3100	(108:76.1%, 0.963)		
EIC3205	<=	EIC3105	&	EIC3103	&	EIC3101	(102:71.8%, 0.98)		
EIC3205	<=	EIC3105	&	EIC3103	&	EIC3102	&	EIC3101	(97:68.3%, 1.0)
EIC3205	<=	EIC3105	&	EIC3103	&	EIC3102	(101:71.1%, 1.0)		
EIC3205	<=	EIC3105	&	EIC3103	&	EIC3104	&	EIC3101	(97:68.3%, 0.99)
EIC3205	<=	EIC3105	&	EIC3103	&	EIC3104	&	EIC3102	(96:67.6%, 1.0)
EIC3205	<=	EIC3105	&	EIC3103	&	EIC3104	(100:70.4%, 0.99)		
EIC3205	<=	EIC3105	&	EIC3103	(111:78.2%, 0.946)				
EIC3205	<=	EIC3105	&	EIC3104	&	EIC3101	(98:69.0%, 0.99)		
EIC3205	<=	EIC3105	&	EIC3104	&	EIC3102	&	EIC3101	(94:66.2%, 1.0)
EIC3205	<=	EIC3105	&	EIC3104	&	EIC3102	(96:67.6%, 1.0)		
EIC3205	<=	EIC3105	&	EIC3104	(102:71.8%, 0.99)				
EIC3205	<=	EIC3105	(121:85.2%, 0.917)						

Figura III.48 - Regras de associação: disciplina EIC3205 (sem.)

Melhor suporte:	121:85.2%	(conf=91.7%)							
Melhor confiança:	100.0%	(sup=101:71.1%)							
EIC3205	<=	EIC3105	&	EIC3102	(101:71.1%, 1.0)				
EIC3205	<=	EIC3105	&	EIC3103	&	EIC3102	(101:71.1%, 1.0)		
EIC3205	<=	EIC3105	(121:85.2%, 0.917)						
EIC3205	<=	EIC3105	&	EIC3104	(102:71.8%, 0.99)				
EIC3205	<=	EIC3105	&	EIC3100	(113:79.6%, 0.947)				
EIC3205	<=	EIC3100	&	EIC3101	(103:72.5%, 0.981)				
EIC3205	<=	EIC3103	&	EIC3100	&	EIC3101	(103:72.5%, 0.981)		
EIC3205	<=	EIC3105	&	EIC3103	&	EIC3100	(108:76.1%, 0.963)		
EIC3205	<=	EIC3105	&	EIC3103	&	EIC3101	(102:71.8%, 0.98)		
EIC3205	<=	EIC3105	&	EIC3103	&	EIC3102	&	EIC3101	(97:68.3%, 1.0)
EIC3205	<=	EIC3105	&	EIC3103	&	EIC3102	(101:71.1%, 1.0)		
EIC3205	<=	EIC3105	&	EIC3103	&	EIC3104	&	EIC3101	(97:68.3%, 0.99)
EIC3205	<=	EIC3105	&	EIC3103	&	EIC3104	&	EIC3102	(96:67.6%, 1.0)
EIC3205	<=	EIC3105	&	EIC3103	&	EIC3104	(100:70.4%, 0.99)		
EIC3205	<=	EIC3105	&	EIC3103	(111:78.2%, 0.946)				
EIC3205	<=	EIC3105	&	EIC3104	&	EIC3101	(98:69.0%, 0.99)		
EIC3205	<=	EIC3105	&	EIC3104	&	EIC3102	&	EIC3101	(94:66.2%, 1.0)
EIC3205	<=	EIC3105	&	EIC3104	&	EIC3102	(96:67.6%, 1.0)		
EIC3205	<=	EIC3105	&	EIC3104	(102:71.8%, 0.99)				
EIC3205	<=	EIC3105	(121:85.2%, 0.917)						

Figura III.49 - Regras de associação: disciplina EIC3205 (sem.) após o filtro

Melhor suporte: 136:95.8% (conf=86.0%)  
 Melhor confiança: 99.1% (sup=106:74.6%)

EIC3100 <= EIC2201 & EIC2104 & EIC2102 & EIC2100 (106:74.6%, 0.991)  
 EIC3100 <= EIC2205 (136:95.8%, 0.86)  
 EIC3100 <= EIC2201 & EIC2104 & EIC2103 & EIC2102 (110:77.5%, 0.982)  
 EIC3100 <= EIC2201 & EIC2104 & EIC2101 & EIC2102 (110:77.5%, 0.982)  
 EIC3100 <= EIC2201 & EIC2103 & EIC2101 & EIC2102 (110:77.5%, 0.982)  
 EIC3100 <= EIC2201 (134:94.4%, 0.873)  
 EIC3100 <= EIC2105 (134:94.4%, 0.873)  
 EIC3100 <= EIC2201 & EIC2104 & EIC2102 (114:80.3%, 0.974)  
 EIC3100 <= EIC2205 & EIC2105 (132:93.0%, 0.879)  
 EIC3100 <= EIC2105 & EIC2201 & EIC2102 (116:81.7%, 0.966)  
 EIC3100 <= EIC2203 (131:92.3%, 0.885)  
 EIC3100 <= EIC2201 & EIC2102 (118:83.1%, 0.958)  
 EIC3100 <= EIC2104 (130:91.5%, 0.892)  
 EIC3100 <= EIC2205 & EIC2201 (130:91.5%, 0.892)  
 EIC3100 <= EIC2203 & EIC2102 (119:83.8%, 0.941)  
 EIC3100 <= EIC2201 & EIC2203 & EIC2101 (119:83.8%, 0.941)  
 EIC3100 <= EIC2201 & EIC2104 & EIC2200 (119:83.8%, 0.941)  
 EIC3100 <= EIC2105 & EIC2201 & EIC2203 & EIC2101 (119:83.8%, 0.941)  
 EIC3100 <= EIC2105 & EIC2102 (119:83.8%, 0.941)  
 EIC3100 <= EIC2105 & EIC2201 (129:90.8%, 0.899)  
 EIC3100 <= EIC2102 (121:85.2%, 0.934)  
 EIC3100 <= EIC2201 & EIC2203 & EIC2104 (121:85.2%, 0.934)  
 EIC3100 <= EIC2201 & EIC2203 (127:89.4%, 0.906)  
 EIC3100 <= EIC2205 & EIC2105 & EIC2201 (127:89.4%, 0.906)  
 EIC3100 <= EIC2203 & EIC2101 (122:85.9%, 0.926)  
 EIC3100 <= EIC2105 & EIC2203 & EIC2101 (122:85.9%, 0.926)  
 EIC3100 <= EIC2104 & EIC2200 (122:85.9%, 0.926)  
 EIC3100 <= EIC2201 & EIC2104 (125:88.0%, 0.92)

**Figura III.50 - Regras de associação: disciplina EIC3100 (anual) após o filtro**

Melhor suporte: 131:92.3% (conf=80.2%)  
 Melhor confiança: 97.1% (sup=103:72.5%)

EIC3101 <= EIC2200 & EIC2202 & EIC2100 & EIC2204 (103:72.5%, 0.971)  
 EIC3101 <= EIC2203 (131:92.3%, 0.802)  
 EIC3101 <= EIC2200 & EIC2102 & EIC2202 & EIC2100 (105:73.9%, 0.962)  
 EIC3101 <= EIC2205 & EIC2203 (129:90.8%, 0.806)  
 EIC3101 <= EIC2105 & EIC2201 (129:90.8%, 0.806)  
 EIC3101 <= EIC2105 & EIC2104 (129:90.8%, 0.806)  
 EIC3101 <= EIC2200 & EIC2101 & EIC2102 & EIC2100 (107:75.4%, 0.953)  
 EIC3101 <= EIC2105 & EIC2203 (128:90.1%, 0.82)  
 EIC3101 <= EIC2200 (128:90.1%, 0.82)  
 EIC3101 <= EIC2203 & EIC2200 & EIC2101 & EIC2202 (109:76.8%, 0.945)  
 EIC3101 <= EIC2200 & EIC2101 & EIC2102 & EIC2202 (109:76.8%, 0.945)  
 EIC3101 <= EIC2105 & EIC2200 (125:88.0%, 0.84)  
 EIC3101 <= EIC2105 & EIC2200 & EIC2102 & EIC2202 (111:78.2%, 0.937)  
 EIC3101 <= EIC2105 & EIC2203 & EIC2200 & EIC2202 (111:78.2%, 0.937)  
 EIC3101 <= EIC2203 & EIC2200 (123:86.6%, 0.854)  
 EIC3101 <= EIC2203 & EIC2200 & EIC2102 & EIC2202 (112:78.9%, 0.929)  
 EIC3101 <= EIC2203 & EIC2200 & EIC2101 & EIC2102 (112:78.9%, 0.929)  
 EIC3101 <= EIC2200 & EIC2101 & EIC2102 (112:78.9%, 0.929)  
 EIC3101 <= EIC2105 & EIC2203 & EIC2102 & EIC2202 (112:78.9%, 0.929)  
 EIC3101 <= EIC2105 & EIC2200 & EIC2101 & EIC2102 (112:78.9%, 0.929)  
 EIC3101 <= EIC2102 (121:85.2%, 0.868)  
 EIC3101 <= EIC2105 & EIC2203 & EIC2200 & EIC2102 (114:80.3%, 0.921)  
 EIC3101 <= EIC2105 & EIC2203 & EIC2200 (120:84.5%, 0.875)  
 EIC3101 <= EIC2105 & EIC2200 & EIC2102 (115:81.0%, 0.913)  
 EIC3101 <= EIC2203 & EIC2102 (119:83.8%, 0.882)  
 EIC3101 <= EIC2105 & EIC2102 (119:83.8%, 0.882)  
 EIC3101 <= EIC2203 & EIC2200 & EIC2102 (116:81.7%, 0.905)  
 EIC3101 <= EIC2105 & EIC2203 & EIC2102 (117:82.4%, 0.897)  
 EIC3101 <= EIC2200 & EIC2102 (117:82.4%, 0.897)

**Figura III.51 - Regras de associação: disciplina EIC3101 (anual) após o filtro**

Melhor suporte: 131:92.3% (conf=80.2%)  
 Melhor confiança: 96.2% (sup=104:73.2%)

EIC3102	<=	EIC2104	&	EIC2200	&	EIC2100	&	EIC2204	(104:73.2%, 0.962)
EIC3102	<=	EIC2203	(131:92.3%, 0.802)						
EIC3102	<=	EIC2205	&	EIC2200	&	EIC2100	&	EIC2204	(105:73.9%, 0.952)
EIC3102	<=	EIC2203	&	EIC2200	&	EIC2100	&	EIC2204	(105:73.9%, 0.952)
EIC3102	<=	EIC2200	&	EIC2101	&	EIC2100	&	EIC2204	(105:73.9%, 0.952)
EIC3102	<=	EIC2200	&	EIC2100	&	EIC2204	(105:73.9%, 0.952)		
EIC3102	<=	EIC2105	&	EIC2200	&	EIC2100	&	EIC2204	(105:73.9%, 0.952)
EIC3102	<=	EIC2104	&	EIC2102	&	EIC2202	&	EIC2100	(105:73.9%, 0.952)
EIC3102	<=	EIC2205	&	EIC2203	(129:90.8%, 0.814)				
EIC3102	<=	EIC2104	&	EIC2200	&	EIC2101	&	EIC2100	(107:75.4%, 0.944)
EIC3102	<=	EIC2203	&	EIC2104	&	EIC2202	&	EIC2100	(107:75.4%, 0.944)
EIC3102	<=	EIC2205	&	EIC2104	&	EIC2102	&	EIC2100	(107:75.4%, 0.944)
EIC3102	<=	EIC2201	&	EIC2203	(127:89.4%, 0.819)				
EIC3102	<=	EIC2205	&	EIC2103	(127:89.4%, 0.819)				
EIC3102	<=	EIC2205	&	EIC2105	&	EIC2203	(127:89.4%, 0.819)		
EIC3102	<=	EIC2205	&	EIC2200	&	EIC2101	&	EIC2100	(108:76.1%, 0.935)
EIC3102	<=	EIC2205	&	EIC2104	&	EIC2202	&	EIC2100	(108:76.1%, 0.935)
EIC3102	<=	EIC2205	&	EIC2104	&	EIC2200	&	EIC2100	(108:76.1%, 0.935)
EIC3102	<=	EIC2205	&	EIC2104	&	EIC2100	&	EIC2204	(108:76.1%, 0.935)
EIC3102	<=	EIC2203	&	EIC2200	&	EIC2101	&	EIC2100	(108:76.1%, 0.935)
EIC3102	<=	EIC2203	&	EIC2104	&	EIC2200	&	EIC2100	(108:76.1%, 0.935)
EIC3102	<=	EIC2203	&	EIC2104	&	EIC2102	&	EIC2100	(108:76.1%, 0.935)
EIC3102	<=	EIC2203	&	EIC2104	&	EIC2100	&	EIC2204	(108:76.1%, 0.935)
EIC3102	<=	EIC2105	&	EIC2104	&	EIC2202	&	EIC2100	(108:76.1%, 0.935)
EIC3102	<=	EIC2105	&	EIC2104	&	EIC2102	&	EIC2100	(108:76.1%, 0.935)
EIC3102	<=	EIC2105	&	EIC2104	&	EIC2100	&	EIC2204	(108:76.1%, 0.935)
EIC3102	<=	EIC2104	&	EIC2202	&	EIC2100	(108:76.1%, 0.935)		
EIC3102	<=	EIC2104	&	EIC2102	&	EIC2100	(108:76.1%, 0.935)		
EIC3102	<=	EIC2104	&	EIC2101	&	EIC2100	&	EIC2204	(108:76.1%, 0.935)
EIC3102	<=	EIC2104	&	EIC2100	&	EIC2204	(108:76.1%, 0.935)		
EIC3102	<=	EIC2205	&	EIC2201	&	EIC2203	(125:88.0%, 0.832)		
EIC3102	<=	EIC2203	&	EIC2104	(125:88.0%, 0.832)				
EIC3102	<=	EIC2203	&	EIC2103	(125:88.0%, 0.832)				
EIC3102	<=	EIC2201	&	EIC2203	&	EIC2104	&	EIC2100	(111:78.2%, 0.928)
EIC3102	<=	EIC2205	&	EIC2201	&	EIC2104	&	EIC2100	(111:78.2%, 0.928)
EIC3102	<=	EIC2105	&	EIC2203	&	EIC2104	(124:87.3%, 0.839)		
EIC3102	<=	EIC2205	&	EIC2203	&	EIC2103	(124:87.3%, 0.839)		
EIC3102	<=	EIC2205	&	EIC2203	&	EIC2104	(124:87.3%, 0.839)		
EIC3102	<=	EIC2205	&	EIC2203	&	EIC2104	&	EIC2100	(113:79.6%, 0.92)
EIC3102	<=	EIC2205	&	EIC2105	&	EIC2203	&	EIC2104	(123:86.6%, 0.846)
EIC3102	<=	EIC2105	&	EIC2203	&	EIC2104	&	EIC2100	(114:80.3%, 0.912)
EIC3102	<=	EIC2203	&	EIC2104	&	EIC2100	(114:80.3%, 0.912)		
EIC3102	<=	EIC2205	&	EIC2104	&	EIC2100	(114:80.3%, 0.912)		
EIC3102	<=	EIC2205	&	EIC2105	&	EIC2104	&	EIC2100	(114:80.3%, 0.912)
EIC3102	<=	EIC2105	&	EIC2104	&	EIC2103	(121:85.2%, 0.851)		
EIC3102	<=	EIC2201	&	EIC2203	&	EIC2103	(121:85.2%, 0.851)		
EIC3102	<=	EIC2201	&	EIC2203	&	EIC2104	(121:85.2%, 0.851)		
EIC3102	<=	EIC2205	&	EIC2104	&	EIC2103	(121:85.2%, 0.851)		
EIC3102	<=	EIC2205	&	EIC2203	&	EIC2100	(115:81.0%, 0.904)		
EIC3102	<=	EIC2205	&	EIC2105	&	EIC2203	&	EIC2100	(115:81.0%, 0.904)
EIC3102	<=	EIC2105	&	EIC2104	&	EIC2100	(115:81.0%, 0.904)		
EIC3102	<=	EIC2104	&	EIC2100	(115:81.0%, 0.904)				
EIC3102	<=	EIC2205	&	EIC2201	&	EIC2203	&	EIC2104	(120:84.5%, 0.858)
EIC3102	<=	EIC2205	&	EIC2201	&	EIC2203	&	EIC2103	(120:84.5%, 0.858)
EIC3102	<=	EIC2205	&	EIC2105	&	EIC2104	&	EIC2103	(120:84.5%, 0.858)
EIC3102	<=	EIC2203	&	EIC2104	&	EIC2103	(120:84.5%, 0.858)		
EIC3102	<=	EIC2105	&	EIC2201	&	EIC2203	&	EIC2104	(120:84.5%, 0.858)
EIC3102	<=	EIC2105	&	EIC2203	&	EIC2100	(116:81.7%, 0.897)		
EIC3102	<=	EIC2203	&	EIC2100	(116:81.7%, 0.897)				
EIC3102	<=	EIC2205	&	EIC2100	(116:81.7%, 0.897)				
EIC3102	<=	EIC2205	&	EIC2105	&	EIC2100	(116:81.7%, 0.897)		
EIC3102	<=	EIC2105	&	EIC2203	&	EIC2104	&	EIC2103	(119:83.8%, 0.866)
EIC3102	<=	EIC2205	&	EIC2203	&	EIC2104	&	EIC2103	(119:83.8%, 0.866)
EIC3102	<=	EIC2105	&	EIC2100	(117:82.4%, 0.889)				
EIC3102	<=	EIC2100	(117:82.4%, 0.889)						

**Figura III.52 - Regras de associação: disciplina EIC3102 (anual) após o filtro**

Melhor suporte: 136:95.8% (conf=85.3%)  
 Melhor confiança: 98.1% (sup=108:76.1%)

EIC3103 <= EIC2104 & EIC2202 & EIC2100 (108:76.1%, 0.981)  
 EIC3103 <= EIC2105 & EIC2104 & EIC2202 & EIC2100 (108:76.1%, 0.981)  
 EIC3103 <= EIC2205 & EIC2104 & EIC2202 & EIC2100 (108:76.1%, 0.981)  
 EIC3103 <= EIC2205 (136:95.8%, 0.853)  
 EIC3103 <= EIC2201 & EIC2104 & EIC2100 (112:78.9%, 0.973)  
 EIC3103 <= EIC2105 & EIC2201 & EIC2104 & EIC2100 (112:78.9%, 0.973)  
 EIC3103 <= EIC2105 (134:94.4%, 0.866)  
 EIC3103 <= EIC2105 & EIC2201 & EIC2100 (114:80.3%, 0.965)  
 EIC3103 <= EIC2201 & EIC2100 (114:80.3%, 0.965)  
 EIC3103 <= EIC2205 & EIC2105 (132:93.0%, 0.871)  
 EIC3103 <= EIC2105 & EIC2104 & EIC2100 (115:81.0%, 0.957)  
 EIC3103 <= EIC2104 & EIC2100 (115:81.0%, 0.957)  
 EIC3103 <= EIC2203 (131:92.3%, 0.878)  
 EIC3103 <= EIC2100 (117:82.4%, 0.949)  
 EIC3103 <= EIC2105 & EIC2100 (117:82.4%, 0.949)  
 EIC3103 <= EIC2105 & EIC2200 & EIC2101 (117:82.4%, 0.949)  
 EIC3103 <= EIC2200 & EIC2101 (117:82.4%, 0.949)  
 EIC3103 <= EIC2105 & EIC2104 (129:90.8%, 0.884)  
 EIC3103 <= EIC2105 & EIC2201 (129:90.8%, 0.884)  
 EIC3103 <= EIC2205 & EIC2203 (129:90.8%, 0.884)  
 EIC3103 <= EIC2105 & EIC2201 & EIC2103 (120:84.5%, 0.933)  
 EIC3103 <= EIC2105 & EIC2203 (128:90.1%, 0.898)  
 EIC3103 <= EIC2105 & EIC2104 & EIC2103 (121:85.2%, 0.926)  
 EIC3103 <= EIC2105 & EIC2103 (125:88.0%, 0.912)  
 EIC3103 <= EIC2105 & EIC2203 & EIC2103 (123:86.6%, 0.919)

**Figura III.53 - Regras de associação: disciplina EIC3103 (anual) após o filtro**

Melhor suporte: 131:92.3% (conf=80.9%)  
 Melhor confiança: 96.1% (sup=103:72.5%)

EIC3104 <= EIC2103 & EIC2102 & EIC2100 & EIC2204 (103:72.5%, 0.961)  
 EIC3104 <= EIC2203 (131:92.3%, 0.809)  
 EIC3104 <= EIC2103 & EIC2200 & EIC2102 & EIC2204 (106:74.6%, 0.953)  
 EIC3104 <= EIC2205 & EIC2203 (129:90.8%, 0.814)  
 EIC3104 <= EIC2103 & EIC2102 & EIC2204 (108:76.1%, 0.944)  
 EIC3104 <= EIC2203 & EIC2103 & EIC2102 & EIC2204 (108:76.1%, 0.944)  
 EIC3104 <= EIC2205 & EIC2103 & EIC2102 & EIC2204 (108:76.1%, 0.944)  
 EIC3104 <= EIC2103 (128:90.1%, 0.82)  
 EIC3104 <= EIC2105 & EIC2203 (128:90.1%, 0.82)  
 EIC3104 <= EIC2205 & EIC2200 & EIC2102 & EIC2204 (109:76.8%, 0.936)  
 EIC3104 <= EIC2203 & EIC2200 & EIC2102 & EIC2204 (109:76.8%, 0.936)  
 EIC3104 <= EIC2200 & EIC2102 & EIC2204 (109:76.8%, 0.936)  
 EIC3104 <= EIC2201 & EIC2203 (127:89.4%, 0.827)  
 EIC3104 <= EIC2102 & EIC2204 (111:78.2%, 0.928)  
 EIC3104 <= EIC2203 & EIC2102 & EIC2204 (111:78.2%, 0.928)  
 EIC3104 <= EIC2205 & EIC2102 & EIC2204 (111:78.2%, 0.928)  
 EIC3104 <= EIC2205 & EIC2203 & EIC2102 & EIC2204 (111:78.2%, 0.928)  
 EIC3104 <= EIC2203 & EIC2103 (125:88.0%, 0.84)  
 EIC3104 <= EIC2205 & EIC2203 & EIC2103 & EIC2204 (112:78.9%, 0.92)  
 EIC3104 <= EIC2205 & EIC2103 & EIC2204 (112:78.9%, 0.92)  
 EIC3104 <= EIC2203 & EIC2103 & EIC2204 (112:78.9%, 0.92)  
 EIC3104 <= EIC2103 & EIC2204 (112:78.9%, 0.92)  
 EIC3104 <= EIC2203 & EIC2200 (123:86.6%, 0.846)  
 EIC3104 <= EIC2201 & EIC2103 (123:86.6%, 0.846)  
 EIC3104 <= EIC2105 & EIC2203 & EIC2103 (123:86.6%, 0.846)  
 EIC3104 <= EIC2103 & EIC2200 & EIC2102 (113:79.6%, 0.912)  
 EIC3104 <= EIC2203 & EIC2103 & EIC2200 & EIC2102 (113:79.6%, 0.912)  
 EIC3104 <= EIC2105 & EIC2203 & EIC2101 (122:85.9%, 0.852)  
 EIC3104 <= EIC2203 & EIC2101 (122:85.9%, 0.852)  
 EIC3104 <= EIC2205 & EIC2201 & EIC2204 (114:80.3%, 0.904)  
 EIC3104 <= EIC2205 & EIC2201 & EIC2203 & EIC2204 (114:80.3%, 0.904)  
 EIC3104 <= EIC2203 & EIC2101 & EIC2102 (114:80.3%, 0.904)  
 EIC3104 <= EIC2201 & EIC2203 & EIC2200 & EIC2102 (114:80.3%, 0.904)  
 EIC3104 <= EIC2201 & EIC2203 & EIC2103 & EIC2102 (114:80.3%, 0.904)  
 EIC3104 <= EIC2201 & EIC2103 & EIC2102 (114:80.3%, 0.904)  
 EIC3104 <= EIC2105 & EIC2203 & EIC2200 & EIC2102 (114:80.3%, 0.904)  
 EIC3104 <= EIC2105 & EIC2203 & EIC2104 & EIC2102 (114:80.3%, 0.904)  
 EIC3104 <= EIC2105 & EIC2203 & EIC2103 & EIC2102 (114:80.3%, 0.904)  
 EIC3104 <= EIC2105 & EIC2203 & EIC2101 & EIC2102 (114:80.3%, 0.904)  
 EIC3104 <= EIC2102 (121:85.2%, 0.868)  
 EIC3104 <= EIC2203 & EIC2103 & EIC2102 (116:81.7%, 0.897)  
 EIC3104 <= EIC2203 & EIC2200 & EIC2102 (116:81.7%, 0.897)  
 EIC3104 <= EIC2205 & EIC2203 & EIC2204 (116:81.7%, 0.897)  
 EIC3104 <= EIC2205 & EIC2204 (116:81.7%, 0.897)  
 EIC3104 <= EIC2203 & EIC2102 (119:83.8%, 0.882)  
 EIC3104 <= EIC2204 (117:82.4%, 0.889)  
 EIC3104 <= EIC2203 & EIC2204 (117:82.4%, 0.889)  
 EIC3104 <= EIC2201 & EIC2203 & EIC2102 (117:82.4%, 0.889)  
 EIC3104 <= EIC2200 & EIC2102 (117:82.4%, 0.889)  
 EIC3104 <= EIC2105 & EIC2203 & EIC2102 (117:82.4%, 0.889)  
 EIC3104 <= EIC2103 & EIC2102 (117:82.4%, 0.889)

**Figura III.54 - Regras de associação: disciplina EIC3104 (anual) após o filtro**

Melhor suporte: 136:95.8% (conf=88.2%)  
 Melhor confiança: 100.0% (sup=107:75.4%)

EIC3105 <= EIC2201 & EIC2103 & EIC2202 & EIC2100 (107:75.4%, 1.0)  
 EIC3105 <= EIC2205 (136:95.8%, 0.882)  
 EIC3105 <= EIC2105 & EIC2201 & EIC2103 & EIC2102 (112:78.9%, 0.991)  
 EIC3105 <= EIC2105 (134:94.4%, 0.903)  
 EIC3105 <= EIC2105 & EIC2201 & EIC2102 (116:81.7%, 0.983)  
 EIC3105 <= EIC2205 & EIC2105 (132:93.0%, 0.909)  
 EIC3105 <= EIC2105 & EIC2203 & EIC2102 (117:82.4%, 0.974)  
 EIC3105 <= EIC2105 & EIC2201 (129:90.8%, 0.922)  
 EIC3105 <= EIC2105 & EIC2102 (119:83.8%, 0.966)  
 EIC3105 <= EIC2105 & EIC2203 (128:90.1%, 0.93)  
 EIC3105 <= EIC2105 & EIC2203 & EIC2200 (120:84.5%, 0.958)  
 EIC3105 <= EIC2105 & EIC2201 & EIC2103 (120:84.5%, 0.958)  
 EIC3105 <= EIC2105 & EIC2200 (125:88.0%, 0.936)  
 EIC3105 <= EIC2105 & EIC2103 (125:88.0%, 0.936)  
 EIC3105 <= EIC2102 (121:85.2%, 0.95)  
 EIC3105 <= EIC2105 & EIC2201 & EIC2203 (124:87.3%, 0.944)

**Figura III.55 - Regras de associação: disciplina EIC3105 (anual) após o filtro**

Melhor suporte: 136:95.8% (conf=82.4%)  
 Melhor confiança: 98.1% (sup=104:73.2%)

EIC3200 <= EIC2201 & EIC2102 & EIC2100 & EIC2204 (104:73.2%, 0.981)  
 EIC3200 <= EIC2205 (136:95.8%, 0.824)  
 EIC3200 <= EIC2205 & EIC2201 & EIC2102 & EIC2100 (107:75.4%, 0.972)  
 EIC3200 <= EIC2201 (134:94.4%, 0.828)  
 EIC3200 <= EIC2105 (134:94.4%, 0.828)  
 EIC3200 <= EIC2105 & EIC2201 & EIC2102 & EIC2100 (108:76.1%, 0.963)  
 EIC3200 <= EIC2201 & EIC2102 & EIC2100 (108:76.1%, 0.963)  
 EIC3200 <= EIC2201 & EIC2203 & EIC2102 & EIC2100 (108:76.1%, 0.963)  
 EIC3200 <= EIC2201 & EIC2203 & EIC2202 & EIC2100 (108:76.1%, 0.963)  
 EIC3200 <= EIC2205 & EIC2105 (132:93.0%, 0.841)  
 EIC3200 <= EIC2205 & EIC2203 & EIC2202 & EIC2100 (109:76.8%, 0.954)  
 EIC3200 <= EIC2205 & EIC2203 & EIC2102 & EIC2100 (109:76.8%, 0.954)  
 EIC3200 <= EIC2205 & EIC2201 & EIC2202 & EIC2100 (109:76.8%, 0.954)  
 EIC3200 <= EIC2205 & EIC2105 & EIC2102 & EIC2100 (109:76.8%, 0.954)  
 EIC3200 <= EIC2205 & EIC2102 & EIC2100 (109:76.8%, 0.954)  
 EIC3200 <= EIC2203 & EIC2202 & EIC2100 (109:76.8%, 0.954)  
 EIC3200 <= EIC2201 & EIC2202 & EIC2100 (109:76.8%, 0.954)  
 EIC3200 <= EIC2105 & EIC2203 & EIC2202 & EIC2100 (109:76.8%, 0.954)  
 EIC3200 <= EIC2105 & EIC2201 & EIC2202 & EIC2100 (109:76.8%, 0.954)  
 EIC3200 <= EIC2205 & EIC2201 (130:91.5%, 0.854)  
 EIC3200 <= EIC2105 & EIC2201 & EIC2104 & EIC2102 (113:79.6%, 0.947)  
 EIC3200 <= EIC2205 & EIC2201 & EIC2104 & EIC2102 (113:79.6%, 0.947)  
 EIC3200 <= EIC2105 & EIC2201 (129:90.8%, 0.86)  
 EIC3200 <= EIC2205 & EIC2105 & EIC2201 & EIC2102 (115:81.0%, 0.939)  
 EIC3200 <= EIC2205 & EIC2105 & EIC2201 (127:89.4%, 0.874)  
 EIC3200 <= EIC2105 & EIC2201 & EIC2102 (116:81.7%, 0.931)  
 EIC3200 <= EIC2201 & EIC2104 (125:88.0%, 0.88)  
 EIC3200 <= EIC2205 & EIC2201 & EIC2102 (117:82.4%, 0.923)  
 EIC3200 <= EIC2205 & EIC2201 & EIC2104 (124:87.3%, 0.887)  
 EIC3200 <= EIC2105 & EIC2201 & EIC2104 (124:87.3%, 0.887)  
 EIC3200 <= EIC2201 & EIC2102 (118:83.1%, 0.915)  
 EIC3200 <= EIC2205 & EIC2105 & EIC2102 (118:83.1%, 0.915)  
 EIC3200 <= EIC2205 & EIC2105 & EIC2201 & EIC2104 (123:86.6%, 0.894)  
 EIC3200 <= EIC2105 & EIC2102 (119:83.8%, 0.908)  
 EIC3200 <= EIC2205 & EIC2201 & EIC2203 & EIC2104 (120:84.5%, 0.9)  
 EIC3200 <= EIC2205 & EIC2105 & EIC2201 & EIC2200 (120:84.5%, 0.9)  
 EIC3200 <= EIC2205 & EIC2102 (120:84.5%, 0.9)  
 EIC3200 <= EIC2105 & EIC2201 & EIC2203 & EIC2104 (120:84.5%, 0.9)

**Figura III.56 - Regras de associação: disciplina EIC3200 (anual) após o filtro**

Melhor suporte: 136:95.8% (conf=80.9%)  
 Melhor confiança: 99.0% (sup=104:73.2%)

EIC3201	<=	EIC2201	&	EIC2102	&	EIC2100	&	EIC2204	(104:73.2%, 0.99)
EIC3201	<=	EIC2205	(136:95.8%, 0.809)						
EIC3201	<=	EIC2205	&	EIC2102	&	EIC2100	&	EIC2204	(105:73.9%, 0.981)
EIC3201	<=	EIC2203	&	EIC2102	&	EIC2100	&	EIC2204	(105:73.9%, 0.981)
EIC3201	<=	EIC2105	&	EIC2102	&	EIC2100	&	EIC2204	(105:73.9%, 0.981)
EIC3201	<=	EIC2102	&	EIC2100	&	EIC2204	(105:73.9%, 0.981)		
EIC3201	<=	EIC2101	&	EIC2102	&	EIC2100	&	EIC2204	(105:73.9%, 0.981)
EIC3201	<=	EIC2201	(134:94.4%, 0.813)						
EIC3201	<=	EIC2105	(134:94.4%, 0.813)						
EIC3201	<=	EIC2105	&	EIC2201	&	EIC2100	&	EIC2204	(107:75.4%, 0.972)
EIC3201	<=	EIC2201	&	EIC2100	&	EIC2204	(107:75.4%, 0.972)		
EIC3201	<=	EIC2201	&	EIC2101	&	EIC2100	&	EIC2204	(107:75.4%, 0.972)
EIC3201	<=	EIC2201	&	EIC2101	&	EIC2102	&	EIC2100	(107:75.4%, 0.972)
EIC3201	<=	EIC2201	&	EIC2104	&	EIC2102	&	EIC2204	(107:75.4%, 0.972)
EIC3201	<=	EIC2201	&	EIC2203	&	EIC2100	&	EIC2204	(107:75.4%, 0.972)
EIC3201	<=	EIC2205	&	EIC2101	&	EIC2102	&	EIC2100	(107:75.4%, 0.972)
EIC3201	<=	EIC2205	&	EIC2201	&	EIC2100	&	EIC2204	(107:75.4%, 0.972)
EIC3201	<=	EIC2205	&	EIC2201	&	EIC2102	&	EIC2100	(107:75.4%, 0.972)
EIC3201	<=	EIC2205	&	EIC2105	(132:93.0%, 0.826)				
EIC3201	<=	EIC2201	&	EIC2104	&	EIC2101	&	EIC2102	(110:77.5%, 0.964)
EIC3201	<=	EIC2205	&	EIC2201	(130:91.5%, 0.838)				
EIC3201	<=	EIC2205	&	EIC2201	&	EIC2101	&	EIC2102	(112:78.9%, 0.955)
EIC3201	<=	EIC2105	&	EIC2201	(129:90.8%, 0.845)				
EIC3201	<=	EIC2205	&	EIC2203	&	EIC2101	&	EIC2102	(113:79.6%, 0.947)
EIC3201	<=	EIC2205	&	EIC2201	&	EIC2104	&	EIC2102	(113:79.6%, 0.947)
EIC3201	<=	EIC2201	&	EIC2203	&	EIC2101	&	EIC2102	(113:79.6%, 0.947)
EIC3201	<=	EIC2201	&	EIC2101	&	EIC2102	(113:79.6%, 0.947)		
EIC3201	<=	EIC2105	&	EIC2201	&	EIC2200	&	EIC2102	(113:79.6%, 0.947)
EIC3201	<=	EIC2105	&	EIC2201	&	EIC2104	&	EIC2102	(113:79.6%, 0.947)
EIC3201	<=	EIC2105	&	EIC2201	&	EIC2101	&	EIC2102	(113:79.6%, 0.947)
EIC3201	<=	EIC2205	&	EIC2105	&	EIC2201	(127:89.4%, 0.858)		
EIC3201	<=	EIC2205	&	EIC2105	&	EIC2201	&	EIC2102	(115:81.0%, 0.939)
EIC3201	<=	EIC2101	(125:88.0%, 0.864)						
EIC3201	<=	EIC2201	&	EIC2104	(125:88.0%, 0.864)				
EIC3201	<=	EIC2205	&	EIC2201	&	EIC2203	(125:88.0%, 0.864)		
EIC3201	<=	EIC2105	&	EIC2201	&	EIC2102	(116:81.7%, 0.931)		
EIC3201	<=	EIC2205	&	EIC2201	&	EIC2104	(124:87.3%, 0.871)		
EIC3201	<=	EIC2205	&	EIC2101	(124:87.3%, 0.871)				
EIC3201	<=	EIC2105	&	EIC2201	&	EIC2203	(124:87.3%, 0.871)		
EIC3201	<=	EIC2105	&	EIC2201	&	EIC2104	(124:87.3%, 0.871)		
EIC3201	<=	EIC2105	&	EIC2101	(124:87.3%, 0.871)				
EIC3201	<=	EIC2205	&	EIC2201	&	EIC2102	(117:82.4%, 0.923)		
EIC3201	<=	EIC2205	&	EIC2105	&	EIC2101	(123:86.6%, 0.878)		
EIC3201	<=	EIC2205	&	EIC2105	&	EIC2201	&	EIC2104	(123:86.6%, 0.878)
EIC3201	<=	EIC2205	&	EIC2201	&	EIC2203	&	EIC2101	(118:83.1%, 0.915)
EIC3201	<=	EIC2205	&	EIC2105	&	EIC2102	(118:83.1%, 0.915)		
EIC3201	<=	EIC2201	&	EIC2102	(118:83.1%, 0.915)				
EIC3201	<=	EIC2203	&	EIC2101	(122:85.9%, 0.885)				
EIC3201	<=	EIC2105	&	EIC2203	&	EIC2101	(122:85.9%, 0.885)		
EIC3201	<=	EIC2105	&	EIC2102	(119:83.8%, 0.908)				
EIC3201	<=	EIC2105	&	EIC2201	&	EIC2203	&	EIC2101	(119:83.8%, 0.908)
EIC3201	<=	EIC2201	&	EIC2203	&	EIC2101	(119:83.8%, 0.908)		
EIC3201	<=	EIC2205	&	EIC2105	&	EIC2201	&	EIC2101	(119:83.8%, 0.908)
EIC3201	<=	EIC2102	(121:85.2%, 0.893)						
EIC3201	<=	EIC2201	&	EIC2101	(121:85.2%, 0.893)				
EIC3201	<=	EIC2205	&	EIC2105	&	EIC2203	&	EIC2101	(121:85.2%, 0.893)
EIC3201	<=	EIC2205	&	EIC2203	&	EIC2101	(121:85.2%, 0.893)		
EIC3201	<=	EIC2205	&	EIC2201	&	EIC2101	(120:84.5%, 0.9)		
EIC3201	<=	EIC2205	&	EIC2102	(120:84.5%, 0.9)				
EIC3201	<=	EIC2105	&	EIC2201	&	EIC2101	(120:84.5%, 0.9)		

Figura III.57 - Regras de associação: disciplina EIC3201 (anual) após o filtro

Melhor suporte: 132:93.0% (conf=80.3%)  
Melhor confiança: 95.2% (sup=104:73.2%)

EIC3202	<=	EIC2201	&	EIC2102	&	EIC2100	&	EIC2204	(104:73.2%, 0.952)
EIC3202	<=	EIC2201	&	EIC2200	&	EIC2100	&	EIC2204	(104:73.2%, 0.952)
EIC3202	<=	EIC2205	&	EIC2105	(132:93.0%, 0.803)				
EIC3202	<=	EIC2205	&	EIC2201	&	EIC2102	&	EIC2100	(107:75.4%, 0.944)
EIC3202	<=	EIC2205	&	EIC2201	&	EIC2100	&	EIC2204	(107:75.4%, 0.944)
EIC3202	<=	EIC2201	&	EIC2203	&	EIC2100	&	EIC2204	(107:75.4%, 0.944)
EIC3202	<=	EIC2201	&	EIC2101	&	EIC2100	&	EIC2204	(107:75.4%, 0.944)
EIC3202	<=	EIC2201	&	EIC2100	&	EIC2204	(107:75.4%, 0.944)		
EIC3202	<=	EIC2105	&	EIC2201	&	EIC2100	&	EIC2204	(107:75.4%, 0.944)
EIC3202	<=	EIC2203	(131:92.3%, 0.809)						
EIC3202	<=	EIC2105	&	EIC2201	&	EIC2102	&	EIC2100	(108:76.1%, 0.935)
EIC3202	<=	EIC2105	&	EIC2201	&	EIC2200	&	EIC2204	(108:76.1%, 0.935)
EIC3202	<=	EIC2201	&	EIC2102	&	EIC2100	(108:76.1%, 0.935)		
EIC3202	<=	EIC2201	&	EIC2200	&	EIC2101	&	EIC2204	(108:76.1%, 0.935)
EIC3202	<=	EIC2201	&	EIC2203	&	EIC2102	&	EIC2100	(108:76.1%, 0.935)
EIC3202	<=	EIC2201	&	EIC2203	&	EIC2202	&	EIC2100	(108:76.1%, 0.935)
EIC3202	<=	EIC2205	&	EIC2201	(130:91.5%, 0.815)				
EIC3202	<=	EIC2205	&	EIC2201	&	EIC2203	&	EIC2100	(112:78.9%, 0.929)
EIC3202	<=	EIC2205	&	EIC2203	(129:90.8%, 0.822)				
EIC3202	<=	EIC2105	&	EIC2201	(129:90.8%, 0.822)				
EIC3202	<=	EIC2105	&	EIC2201	&	EIC2203	&	EIC2100	(113:79.6%, 0.92)
EIC3202	<=	EIC2201	&	EIC2203	&	EIC2100	(113:79.6%, 0.92)		
EIC3202	<=	EIC2205	&	EIC2105	&	EIC2201	&	EIC2100	(113:79.6%, 0.92)
EIC3202	<=	EIC2205	&	EIC2201	&	EIC2100	(113:79.6%, 0.92)		
EIC3202	<=	EIC2105	&	EIC2203	(128:90.1%, 0.828)				
EIC3202	<=	EIC2201	&	EIC2100	(114:80.3%, 0.912)				
EIC3202	<=	EIC2105	&	EIC2201	&	EIC2100	(114:80.3%, 0.912)		
EIC3202	<=	EIC2205	&	EIC2105	&	EIC2203	(127:89.4%, 0.835)		
EIC3202	<=	EIC2205	&	EIC2105	&	EIC2201	(127:89.4%, 0.835)		
EIC3202	<=	EIC2201	&	EIC2203	(127:89.4%, 0.835)				
EIC3202	<=	EIC2205	&	EIC2105	&	EIC2203	&	EIC2100	(115:81.0%, 0.904)
EIC3202	<=	EIC2205	&	EIC2203	&	EIC2100	(115:81.0%, 0.904)		
EIC3202	<=	EIC2205	&	EIC2201	&	EIC2203	(125:88.0%, 0.848)		
EIC3202	<=	EIC2205	&	EIC2105	&	EIC2100	(116:81.7%, 0.897)		
EIC3202	<=	EIC2205	&	EIC2100	(116:81.7%, 0.897)				
EIC3202	<=	EIC2203	&	EIC2100	(116:81.7%, 0.897)				
EIC3202	<=	EIC2105	&	EIC2203	&	EIC2100	(116:81.7%, 0.897)		
EIC3202	<=	EIC2105	&	EIC2201	&	EIC2203	(124:87.3%, 0.855)		
EIC3202	<=	EIC2100	(117:82.4%, 0.889)						
EIC3202	<=	EIC2105	&	EIC2100	(117:82.4%, 0.889)				
EIC3202	<=	EIC2205	&	EIC2105	&	EIC2201	&	EIC2203	(123:86.6%, 0.862)
EIC3202	<=	EIC2205	&	EIC2201	&	EIC2203	&	EIC2101	(118:83.1%, 0.881)
EIC3202	<=	EIC2201	&	EIC2203	&	EIC2104	(121:85.2%, 0.868)		
EIC3202	<=	EIC2105	&	EIC2201	&	EIC2203	&	EIC2104	(120:84.5%, 0.875)
EIC3202	<=	EIC2205	&	EIC2201	&	EIC2203	&	EIC2104	(120:84.5%, 0.875)

**Figura III.58 - Regras de associação: disciplina EIC3202 (anual) após o filtro**

Melhor suporte: 130:91.5% (conf=80.0%)  
Melhor confiança: 97.1% (sup=103:72.5%)

EIC3203	<=	EIC2200	&	EIC2202	&	EIC2100	&	EIC2204	(103:72.5%, 0.971)
EIC3203	<=	EIC2201	&	EIC2202	&	EIC2100	&	EIC2204	(103:72.5%, 0.971)
EIC3203	<=	EIC2205	&	EIC2201	(130:91.5%, 0.8)				
EIC3203	<=	EIC2201	&	EIC2200	&	EIC2202	&	EIC2204	(106:74.6%, 0.962)
EIC3203	<=	EIC2201	&	EIC2101	&	EIC2202	&	EIC2204	(106:74.6%, 0.962)
EIC3203	<=	EIC2200	&	EIC2101	&	EIC2202	&	EIC2204	(106:74.6%, 0.962)
EIC3203	<=	EIC2105	&	EIC2201	&	EIC2202	&	EIC2204	(106:74.6%, 0.962)
EIC3203	<=	EIC2105	&	EIC2200	&	EIC2202	&	EIC2204	(106:74.6%, 0.962)
EIC3203	<=	EIC2205	&	EIC2203	(129:90.8%, 0.806)				
EIC3203	<=	EIC2105	&	EIC2201	(129:90.8%, 0.806)				
EIC3203	<=	EIC2105	&	EIC2200	&	EIC2102	&	EIC2204	(108:76.1%, 0.954)
EIC3203	<=	EIC2105	&	EIC2201	&	EIC2200	&	EIC2204	(108:76.1%, 0.954)
EIC3203	<=	EIC2200	&	EIC2101	&	EIC2102	&	EIC2204	(108:76.1%, 0.954)
EIC3203	<=	EIC2201	&	EIC2200	&	EIC2101	&	EIC2204	(108:76.1%, 0.954)
EIC3203	<=	EIC2201	&	EIC2200	&	EIC2102	&	EIC2204	(108:76.1%, 0.954)
EIC3203	<=	EIC2105	&	EIC2203	(128:90.1%, 0.812)				
EIC3203	<=	EIC2205	&	EIC2201	&	EIC2200	&	EIC2204	(109:76.8%, 0.945)
EIC3203	<=	EIC2205	&	EIC2200	&	EIC2102	&	EIC2204	(109:76.8%, 0.945)
EIC3203	<=	EIC2205	&	EIC2200	&	EIC2101	&	EIC2204	(109:76.8%, 0.945)
EIC3203	<=	EIC2205	&	EIC2105	&	EIC2200	&	EIC2204	(109:76.8%, 0.945)
EIC3203	<=	EIC2203	&	EIC2200	&	EIC2102	&	EIC2204	(109:76.8%, 0.945)
EIC3203	<=	EIC2203	&	EIC2200	&	EIC2101	&	EIC2204	(109:76.8%, 0.945)
EIC3203	<=	EIC2201	&	EIC2101	&	EIC2102	&	EIC2204	(109:76.8%, 0.945)
EIC3203	<=	EIC2200	&	EIC2102	&	EIC2204	(109:76.8%, 0.945)		
EIC3203	<=	EIC2200	&	EIC2101	&	EIC2204	(109:76.8%, 0.945)		
EIC3203	<=	EIC2105	&	EIC2203	&	EIC2200	&	EIC2204	(109:76.8%, 0.945)
EIC3203	<=	EIC2105	&	EIC2201	&	EIC2102	&	EIC2204	(109:76.8%, 0.945)
EIC3203	<=	EIC2105	&	EIC2200	&	EIC2204	(109:76.8%, 0.945)		
EIC3203	<=	EIC2105	&	EIC2200	&	EIC2101	&	EIC2204	(109:76.8%, 0.945)

**Figura III.59 - Regras de associação: disciplina EIC3203 (anual) após o filtro**



EIC3203	<=	EIC2205	&	EIC2105	&	EIC2203	(127:89.4%, 0.819)		
EIC3203	<=	EIC2205	&	EIC2105	&	EIC2201	(127:89.4%, 0.819)		
EIC3203	<=	EIC2201	&	EIC2203			(127:89.4%, 0.819)		
EIC3203	<=	EIC2101	&	EIC2102	&	EIC2204	(110:77.5%, 0.936)		
EIC3203	<=	EIC2105	&	EIC2101	&	EIC2102	&	EIC2204	(110:77.5%, 0.936)
EIC3203	<=	EIC2105	&	EIC2102	&	EIC2204	(110:77.5%, 0.936)		
EIC3203	<=	EIC2105	&	EIC2203	&	EIC2102	&	EIC2204	(110:77.5%, 0.936)
EIC3203	<=	EIC2201	&	EIC2102	&	EIC2204	(110:77.5%, 0.936)		
EIC3203	<=	EIC2201	&	EIC2200	&	EIC2204	(110:77.5%, 0.936)		
EIC3203	<=	EIC2201	&	EIC2203	&	EIC2102	&	EIC2204	(110:77.5%, 0.936)
EIC3203	<=	EIC2201	&	EIC2203	&	EIC2200	&	EIC2204	(110:77.5%, 0.936)
EIC3203	<=	EIC2203	&	EIC2101	&	EIC2102	&	EIC2204	(110:77.5%, 0.936)
EIC3203	<=	EIC2205	&	EIC2101	&	EIC2102	&	EIC2204	(110:77.5%, 0.936)
EIC3203	<=	EIC2205	&	EIC2105	&	EIC2102	&	EIC2204	(110:77.5%, 0.936)
EIC3203	<=	EIC2205	&	EIC2200	&	EIC2204	(110:77.5%, 0.936)		
EIC3203	<=	EIC2205	&	EIC2201	&	EIC2102	&	EIC2204	(110:77.5%, 0.936)
EIC3203	<=	EIC2205	&	EIC2203	&	EIC2200	&	EIC2204	(110:77.5%, 0.936)
EIC3203	<=	EIC2205	&	EIC2201	&	EIC2203	(125:88.0%, 0.832)		
EIC3203	<=	EIC2205	&	EIC2203	&	EIC2102	&	EIC2204	(111:78.2%, 0.928)
EIC3203	<=	EIC2205	&	EIC2200	&	EIC2101	&	EIC2102	(111:78.2%, 0.928)
EIC3203	<=	EIC2205	&	EIC2102	&	EIC2204	(111:78.2%, 0.928)		
EIC3203	<=	EIC2203	&	EIC2102	&	EIC2204	(111:78.2%, 0.928)		
EIC3203	<=	EIC2201	&	EIC2200	&	EIC2101	&	EIC2102	(111:78.2%, 0.928)
EIC3203	<=	EIC2200	&	EIC2204	(111:78.2%, 0.928)				
EIC3203	<=	EIC2102	&	EIC2204	(111:78.2%, 0.928)				
EIC3203	<=	EIC2105	&	EIC2201	&	EIC2203	(124:87.3%, 0.839)		
EIC3203	<=	EIC2105	&	EIC2200	&	EIC2101	&	EIC2102	(112:78.9%, 0.92)
EIC3203	<=	EIC2105	&	EIC2201	&	EIC2102	&	EIC2202	(112:78.9%, 0.92)
EIC3203	<=	EIC2105	&	EIC2203	&	EIC2102	&	EIC2202	(112:78.9%, 0.92)
EIC3203	<=	EIC2200	&	EIC2101	&	EIC2102	(112:78.9%, 0.92)		
EIC3203	<=	EIC2203	&	EIC2200	&	EIC2101	&	EIC2102	(112:78.9%, 0.92)
EIC3203	<=	EIC2205	&	EIC2201	&	EIC2101	&	EIC2102	(112:78.9%, 0.92)
EIC3203	<=	EIC2205	&	EIC2105	&	EIC2201	&	EIC2203	(123:86.6%, 0.846)
EIC3203	<=	EIC2205	&	EIC2203	&	EIC2101	&	EIC2102	(113:79.6%, 0.912)
EIC3203	<=	EIC2205	&	EIC2201	&	EIC2104	&	EIC2102	(113:79.6%, 0.912)
EIC3203	<=	EIC2205	&	EIC2201	&	EIC2101	&	EIC2204	(113:79.6%, 0.912)
EIC3203	<=	EIC2205	&	EIC2105	&	EIC2201	&	EIC2204	(113:79.6%, 0.912)
EIC3203	<=	EIC2205	&	EIC2105	&	EIC2102	&	EIC2202	(113:79.6%, 0.912)
EIC3203	<=	EIC2201	&	EIC2203	&	EIC2104	&	EIC2102	(113:79.6%, 0.912)
EIC3203	<=	EIC2201	&	EIC2203	&	EIC2102	&	EIC2202	(113:79.6%, 0.912)
EIC3203	<=	EIC2201	&	EIC2203	&	EIC2101	&	EIC2204	(113:79.6%, 0.912)
EIC3203	<=	EIC2201	&	EIC2203	&	EIC2101	&	EIC2102	(113:79.6%, 0.912)
EIC3203	<=	EIC2201	&	EIC2101	&	EIC2204	(113:79.6%, 0.912)		
EIC3203	<=	EIC2201	&	EIC2101	&	EIC2102	(113:79.6%, 0.912)		
EIC3203	<=	EIC2105	&	EIC2201	&	EIC2204	(113:79.6%, 0.912)		
EIC3203	<=	EIC2105	&	EIC2201	&	EIC2203	&	EIC2204	(113:79.6%, 0.912)
EIC3203	<=	EIC2105	&	EIC2201	&	EIC2203	&	EIC2202	(113:79.6%, 0.912)
EIC3203	<=	EIC2105	&	EIC2201	&	EIC2200	&	EIC2102	(113:79.6%, 0.912)
EIC3203	<=	EIC2105	&	EIC2201	&	EIC2104	&	EIC2102	(113:79.6%, 0.912)
EIC3203	<=	EIC2105	&	EIC2201	&	EIC2101	&	EIC2204	(113:79.6%, 0.912)
EIC3203	<=	EIC2105	&	EIC2201	&	EIC2101	&	EIC2102	(113:79.6%, 0.912)
EIC3203	<=	EIC2105	&	EIC2102	&	EIC2202	(113:79.6%, 0.912)		
EIC3203	<=	EIC2102	(121:85.2%, 0.86)						
EIC3203	<=	EIC2105	&	EIC2201	&	EIC2203	&	EIC2102	(115:81.0%, 0.904)
EIC3203	<=	EIC2205	&	EIC2105	&	EIC2201	&	EIC2102	(115:81.0%, 0.904)
EIC3203	<=	EIC2205	&	EIC2102	(120:84.5%, 0.867)				
EIC3203	<=	EIC2205	&	EIC2201	&	EIC2203	&	EIC2102	(116:81.7%, 0.897)
EIC3203	<=	EIC2205	&	EIC2105	&	EIC2203	&	EIC2102	(116:81.7%, 0.897)
EIC3203	<=	EIC2105	&	EIC2201	&	EIC2102	(116:81.7%, 0.897)		
EIC3203	<=	EIC2203	&	EIC2102	(119:83.8%, 0.874)				
EIC3203	<=	EIC2105	&	EIC2102	(119:83.8%, 0.874)				
EIC3203	<=	EIC2105	&	EIC2203	&	EIC2102	(117:82.4%, 0.889)		
EIC3203	<=	EIC2201	&	EIC2203	&	EIC2102	(117:82.4%, 0.889)		
EIC3203	<=	EIC2205	&	EIC2201	&	EIC2102	(117:82.4%, 0.889)		
EIC3203	<=	EIC2201	&	EIC2102	(118:83.1%, 0.881)				
EIC3203	<=	EIC2205	&	EIC2105	&	EIC2102	(118:83.1%, 0.881)		
EIC3203	<=	EIC2205	&	EIC2203	&	EIC2102	(118:83.1%, 0.881)		

Figura III.60 - Regras de associação: disciplina EIC3203 (anual) após o filtro (cont.)

Melhor suporte: 136:95.8% (conf=86.8%)  
 Melhor confiança: 99.1% (sup=107:75.4%)

EIC3204 <= EIC2201 & EIC2104 & EIC2102 & EIC2204 (107:75.4%, 0.991)  
 EIC3204 <= EIC2205 & EIC2201 & EIC2102 & EIC2100 (107:75.4%, 0.991)  
 EIC3204 <= EIC2205 (136:95.8%, 0.868)  
 EIC3204 <= EIC2205 & EIC2201 & EIC2101 & EIC2202 (110:77.5%, 0.982)  
 EIC3204 <= EIC2201 & EIC2101 & EIC2202 (110:77.5%, 0.982)  
 EIC3204 <= EIC2105 & EIC2201 & EIC2101 & EIC2202 (110:77.5%, 0.982)  
 EIC3204 <= EIC2201 (134:94.4%, 0.881)  
 EIC3204 <= EIC2105 & EIC2201 & EIC2102 & EIC2202 (112:78.9%, 0.973)  
 EIC3204 <= EIC2105 & EIC2201 & EIC2200 & EIC2202 (112:78.9%, 0.973)  
 EIC3204 <= EIC2205 & EIC2201 & EIC2101 & EIC2102 (112:78.9%, 0.973)  
 EIC3204 <= EIC2205 & EIC2105 (132:93.0%, 0.886)  
 EIC3204 <= EIC2205 & EIC2105 & EIC2201 & EIC2202 (115:81.0%, 0.965)  
 EIC3204 <= EIC2205 & EIC2105 & EIC2201 & EIC2102 (115:81.0%, 0.965)  
 EIC3204 <= EIC2105 & EIC2201 & EIC2202 (115:81.0%, 0.965)  
 EIC3204 <= EIC2205 & EIC2201 (130:91.5%, 0.9)  
 EIC3204 <= EIC2105 & EIC2201 & EIC2102 (116:81.7%, 0.957)  
 EIC3204 <= EIC2105 & EIC2202 (116:81.7%, 0.957)  
 EIC3204 <= EIC2205 & EIC2105 & EIC2202 (116:81.7%, 0.957)  
 EIC3204 <= EIC2205 & EIC2105 & EIC2201 (127:89.4%, 0.913)  
 EIC3204 <= EIC2205 & EIC2201 & EIC2202 (117:82.4%, 0.949)  
 EIC3204 <= EIC2205 & EIC2201 & EIC2102 (117:82.4%, 0.949)  
 EIC3204 <= EIC2204 (117:82.4%, 0.949)  
 EIC3204 <= EIC2203 & EIC2204 (117:82.4%, 0.949)  
 EIC3204 <= EIC2201 & EIC2202 (117:82.4%, 0.949)  
 EIC3204 <= EIC2205 & EIC2201 & EIC2104 (124:87.3%, 0.919)  
 EIC3204 <= EIC2205 & EIC2101 (124:87.3%, 0.919)  
 EIC3204 <= EIC2105 & EIC2201 & EIC2203 (124:87.3%, 0.919)  
 EIC3204 <= EIC2105 & EIC2201 & EIC2104 (124:87.3%, 0.919)  
 EIC3204 <= EIC2205 & EIC2201 & EIC2101 (120:84.5%, 0.942)  
 EIC3204 <= EIC2205 & EIC2105 & EIC2201 & EIC2104 (123:86.6%, 0.927)  
 EIC3204 <= EIC2205 & EIC2105 & EIC2201 & EIC2203 (123:86.6%, 0.927)  
 EIC3204 <= EIC2201 & EIC2101 (121:85.2%, 0.934)

**Figura III.61 - Regras de associação: disciplina EIC3204 (anual) após o filtro**

Melhor suporte: 136:95.8% (conf=82.4%)  
 Melhor confiança: 100.0% (sup=104:73.2%)

EIC3205 <= EIC2201 & EIC2102 & EIC2100 & EIC2204 (104:73.2%, 1.0)  
 EIC3205 <= EIC2205 (136:95.8%, 0.824)  
 EIC3205 <= EIC2205 & EIC2201 & EIC2102 & EIC2100 (107:75.4%, 0.991)  
 EIC3205 <= EIC2201 (134:94.4%, 0.843)  
 EIC3205 <= EIC2105 & EIC2201 & EIC2102 & EIC2204 (109:76.8%, 0.982)  
 EIC3205 <= EIC2201 & EIC2101 & EIC2102 & EIC2204 (109:76.8%, 0.982)  
 EIC3205 <= EIC2205 & EIC2105 (132:93.0%, 0.848)  
 EIC3205 <= EIC2205 & EIC2201 & EIC2101 & EIC2102 (112:78.9%, 0.973)  
 EIC3205 <= EIC2203 (131:92.3%, 0.863)  
 EIC3205 <= EIC2105 & EIC2201 & EIC2203 & EIC2102 (115:81.0%, 0.965)  
 EIC3205 <= EIC2205 & EIC2105 & EIC2201 & EIC2102 (115:81.0%, 0.965)  
 EIC3205 <= EIC2105 & EIC2201 (129:90.8%, 0.868)  
 EIC3205 <= EIC2205 & EIC2203 (129:90.8%, 0.868)  
 EIC3205 <= EIC2205 & EIC2201 & EIC2203 & EIC2102 (116:81.7%, 0.957)  
 EIC3205 <= EIC2205 & EIC2105 & EIC2203 & EIC2102 (116:81.7%, 0.957)  
 EIC3205 <= EIC2105 & EIC2201 & EIC2102 (116:81.7%, 0.957)  
 EIC3205 <= EIC2105 & EIC2203 (128:90.1%, 0.875)  
 EIC3205 <= EIC2105 & EIC2203 & EIC2102 (117:82.4%, 0.949)  
 EIC3205 <= EIC2201 & EIC2203 & EIC2102 (117:82.4%, 0.949)  
 EIC3205 <= EIC2205 & EIC2201 & EIC2102 (117:82.4%, 0.949)  
 EIC3205 <= EIC2201 & EIC2203 (127:89.4%, 0.89)  
 EIC3205 <= EIC2205 & EIC2203 & EIC2102 (118:83.1%, 0.941)  
 EIC3205 <= EIC2205 & EIC2105 & EIC2102 (118:83.1%, 0.941)  
 EIC3205 <= EIC2201 & EIC2102 (118:83.1%, 0.941)  
 EIC3205 <= EIC2205 & EIC2201 & EIC2203 (125:88.0%, 0.896)  
 EIC3205 <= EIC2105 & EIC2102 (119:83.8%, 0.933)  
 EIC3205 <= EIC2203 & EIC2102 (119:83.8%, 0.933)  
 EIC3205 <= EIC2105 & EIC2201 & EIC2203 (124:87.3%, 0.903)  
 EIC3205 <= EIC2205 & EIC2102 (120:84.5%, 0.925)  
 EIC3205 <= EIC2205 & EIC2105 & EIC2201 & EIC2203 (123:86.6%, 0.911)  
 EIC3205 <= EIC2102 (121:85.2%, 0.917)

**Figura III.62 - Regras de associação: disciplina EIC3205 (anual) após o filtro**

EIC4200	←=	EIC4100	&	EIC4101	&	EIC4104	(51:68.9%, 0.882)		
EIC4200	←=	EIC4100	&	EIC4101	(61:82.4%, 0.869)				
EIC4200	←=	EIC4100	&	EIC4102	&	EIC4101	&	EIC4104	(51:68.9%, 0.882)
EIC4200	←=	EIC4100	&	EIC4102	&	EIC4101	(61:82.4%, 0.869)		
EIC4200	←=	EIC4100	&	EIC4102	&	EIC4103	&	EIC4101	(58:78.4%, 0.879)
EIC4200	←=	EIC4100	&	EIC4102	&	EIC4103	&	EIC4104	(49:66.2%, 0.898)
EIC4200	←=	EIC4100	&	EIC4102	&	EIC4103	(60:81.1%, 0.867)		
EIC4200	←=	EIC4100	&	EIC4102	&	EIC4104	(53:71.6%, 0.868)		
EIC4200	←=	EIC4100	&	EIC4102	(64:86.5%, 0.844)				
EIC4200	←=	EIC4100	&	EIC4103	&	EIC4101	&	EIC4104	(48:64.9%, 0.896)
EIC4200	←=	EIC4100	&	EIC4103	&	EIC4101	(58:78.4%, 0.879)		
EIC4200	←=	EIC4100	&	EIC4103	&	EIC4104	(50:67.6%, 0.88)		
EIC4200	←=	EIC4100	&	EIC4103	(61:82.4%, 0.852)				
EIC4200	←=	EIC4100	&	EIC4104	(54:73.0%, 0.852)				
EIC4200	←=	EIC4100	(65:87.8%, 0.831)						
EIC4200	←=	EIC4101	&	EIC4104	(51:68.9%, 0.882)				
EIC4200	←=	EIC4101	(61:82.4%, 0.869)						
EIC4200	←=	EIC4102	&	EIC4101	&	EIC4104	(51:68.9%, 0.882)		
EIC4200	←=	EIC4102	&	EIC4101	(61:82.4%, 0.869)				
EIC4200	←=	EIC4102	&	EIC4103	&	EIC4101	&	EIC4104	(48:64.9%, 0.896)
EIC4200	←=	EIC4102	&	EIC4103	&	EIC4101	(58:78.4%, 0.879)		
EIC4200	←=	EIC4102	&	EIC4103	&	EIC4104	(49:66.2%, 0.898)		
EIC4200	←=	EIC4102	&	EIC4103	(61:82.4%, 0.852)				
EIC4200	←=	EIC4102	&	EIC4104	(53:71.6%, 0.868)				
EIC4200	←=	EIC4102	(65:87.8%, 0.831)						
EIC4200	←=	EIC4103	&	EIC4101	&	EIC4104	(48:64.9%, 0.896)		
EIC4200	←=	EIC4103	&	EIC4101	(58:78.4%, 0.879)				
EIC4200	←=	EIC4103	&	EIC4104	(50:67.6%, 0.88)				
EIC4200	←=	EIC4103	(62:83.8%, 0.839)						
EIC4200	←=	EIC4104	(56:75.7%, 0.839)						
EIC4200	←=	EIC4105	&	EIC4100	&	EIC4101	&	EIC4104	(51:68.9%, 0.882)
EIC4200	←=	EIC4105	&	EIC4100	&	EIC4101	(61:82.4%, 0.869)		
EIC4200	←=	EIC4105	&	EIC4100	&	EIC4102	&	EIC4101	(61:82.4%, 0.869)
EIC4200	←=	EIC4105	&	EIC4100	&	EIC4102	&	EIC4103	(60:81.1%, 0.867)
EIC4200	←=	EIC4105	&	EIC4100	&	EIC4102	&	EIC4104	(52:70.3%, 0.885)
EIC4200	←=	EIC4105	&	EIC4100	&	EIC4102	(63:85.1%, 0.857)		
EIC4200	←=	EIC4105	&	EIC4100	&	EIC4103	&	EIC4101	(58:78.4%, 0.879)
EIC4200	←=	EIC4105	&	EIC4100	&	EIC4103	&	EIC4104	(50:67.6%, 0.88)
EIC4200	←=	EIC4105	&	EIC4100	&	EIC4103	(61:82.4%, 0.852)		
EIC4200	←=	EIC4105	&	EIC4100	&	EIC4104	(53:71.6%, 0.868)		
EIC4200	←=	EIC4105	&	EIC4100	(64:86.5%, 0.844)				
EIC4200	←=	EIC4105	&	EIC4101	&	EIC4104	(51:68.9%, 0.882)		
EIC4200	←=	EIC4105	&	EIC4101	(61:82.4%, 0.869)				
EIC4200	←=	EIC4105	&	EIC4102	&	EIC4101	&	EIC4104	(51:68.9%, 0.882)
EIC4200	←=	EIC4105	&	EIC4102	&	EIC4101	(61:82.4%, 0.869)		
EIC4200	←=	EIC4105	&	EIC4102	&	EIC4103	&	EIC4101	(58:78.4%, 0.879)
EIC4200	←=	EIC4105	&	EIC4102	&	EIC4103	&	EIC4104	(49:66.2%, 0.898)
EIC4200	←=	EIC4105	&	EIC4102	&	EIC4103	(61:82.4%, 0.852)		
EIC4200	←=	EIC4105	&	EIC4102	&	EIC4104	(52:70.3%, 0.885)		
EIC4200	←=	EIC4105	&	EIC4102	(64:86.5%, 0.844)				
EIC4200	←=	EIC4105	&	EIC4103	&	EIC4101	&	EIC4104	(48:64.9%, 0.896)
EIC4200	←=	EIC4105	&	EIC4103	&	EIC4101	(58:78.4%, 0.879)		
EIC4200	←=	EIC4105	&	EIC4103	&	EIC4104	(50:67.6%, 0.88)		
EIC4200	←=	EIC4105	&	EIC4103	(62:83.8%, 0.839)				
EIC4200	←=	EIC4105	&	EIC4104	(55:74.3%, 0.855)				

Figura III.63 - Regras de associação: disciplina EIC4200 (sem.)

Melhor suporte: 65:87.8% (conf=83.1%)									
Melhor confiança: 89.8% (sup=49:66.2%)									
EIC4200	←=	EIC4100	&	EIC4102	&	EIC4103	&	EIC4104	(49:66.2%, 0.898)
EIC4200	←=	EIC4102	&	EIC4103	&	EIC4104	(49:66.2%, 0.898)		
EIC4200	←=	EIC4105	&	EIC4102	&	EIC4103	&	EIC4104	(49:66.2%, 0.898)
EIC4200	←=	EIC4100	(65:87.8%, 0.831)						
EIC4200	←=	EIC4102	(65:87.8%, 0.831)						
EIC4200	←=	EIC4105	&	EIC4102	&	EIC4104	(52:70.3%, 0.885)		
EIC4200	←=	EIC4105	&	EIC4100	&	EIC4102	&	EIC4104	(52:70.3%, 0.885)
EIC4200	←=	EIC4105	&	EIC4102	(64:86.5%, 0.844)				
EIC4200	←=	EIC4105	&	EIC4100	(64:86.5%, 0.844)				
EIC4200	←=	EIC4100	&	EIC4102	(64:86.5%, 0.844)				
EIC4200	←=	EIC4100	&	EIC4102	&	EIC4103	&	EIC4101	(58:78.4%, 0.879)
EIC4200	←=	EIC4100	&	EIC4103	&	EIC4101	(58:78.4%, 0.879)		
EIC4200	←=	EIC4102	&	EIC4103	&	EIC4101	(58:78.4%, 0.879)		
EIC4200	←=	EIC4103	&	EIC4101	(58:78.4%, 0.879)				
EIC4200	←=	EIC4105	&	EIC4100	&	EIC4103	&	EIC4101	(58:78.4%, 0.879)
EIC4200	←=	EIC4105	&	EIC4102	&	EIC4103	&	EIC4101	(58:78.4%, 0.879)
EIC4200	←=	EIC4105	&	EIC4103	&	EIC4101	(58:78.4%, 0.879)		
EIC4200	←=	EIC4105	&	EIC4100	&	EIC4102	(63:85.1%, 0.857)		
EIC4200	←=	EIC4105	&	EIC4102	&	EIC4101	(61:82.4%, 0.869)		
EIC4200	←=	EIC4105	&	EIC4101	(61:82.4%, 0.869)				
EIC4200	←=	EIC4105	&	EIC4100	&	EIC4102	&	EIC4101	(61:82.4%, 0.869)
EIC4200	←=	EIC4105	&	EIC4100	&	EIC4101	(61:82.4%, 0.869)		
EIC4200	←=	EIC4102	&	EIC4101	(61:82.4%, 0.869)				
EIC4200	←=	EIC4101	(61:82.4%, 0.869)						
EIC4200	←=	EIC4100	&	EIC4102	&	EIC4101	(61:82.4%, 0.869)		
EIC4200	←=	EIC4100	&	EIC4101	(61:82.4%, 0.869)				

Figura III.64 - Regras de associação: disciplina EIC4200 (sem.) após o filtro

EIC4201	<=	EIC4100	&	EIC4101	&	EIC4104	(51:68.9%, 0.961)		
EIC4201	<=	EIC4100	&	EIC4101	(61:82.4%, 0.967)				
EIC4201	<=	EIC4100	&	EIC4102	&	EIC4101	&	EIC4104	(51:68.9%, 0.961)
EIC4201	<=	EIC4100	&	EIC4102	&	EIC4101	(61:82.4%, 0.967)		
EIC4201	<=	EIC4100	&	EIC4102	&	EIC4103	&	EIC4101	(58:78.4%, 0.966)
EIC4201	<=	EIC4100	&	EIC4102	&	EIC4103	&	EIC4104	(49:66.2%, 0.959)
EIC4201	<=	EIC4100	&	EIC4102	&	EIC4103	(60:81.1%, 0.967)		
EIC4201	<=	EIC4100	&	EIC4102	&	EIC4104	(53:71.6%, 0.962)		
EIC4201	<=	EIC4100	&	EIC4102	(64:86.5%, 0.969)				
EIC4201	<=	EIC4100	&	EIC4103	&	EIC4101	&	EIC4104	(48:64.9%, 0.958)
EIC4201	<=	EIC4100	&	EIC4103	&	EIC4101	(58:78.4%, 0.966)		
EIC4201	<=	EIC4100	&	EIC4103	&	EIC4104	(50:67.6%, 0.96)		
EIC4201	<=	EIC4100	&	EIC4103	(61:82.4%, 0.967)				
EIC4201	<=	EIC4100	&	EIC4104	(54:73.0%, 0.963)				
EIC4201	<=	EIC4100	(65:87.8%, 0.969)						
EIC4201	<=	EIC4101	&	EIC4104	(51:68.9%, 0.961)				
EIC4201	<=	EIC4101	(61:82.4%, 0.967)						
EIC4201	<=	EIC4102	&	EIC4101	&	EIC4104	(51:68.9%, 0.961)		
EIC4201	<=	EIC4102	&	EIC4101	(61:82.4%, 0.967)				
EIC4201	<=	EIC4102	&	EIC4103	&	EIC4101	&	EIC4104	(48:64.9%, 0.958)
EIC4201	<=	EIC4102	&	EIC4103	&	EIC4101	(58:78.4%, 0.966)		
EIC4201	<=	EIC4102	&	EIC4103	&	EIC4104	(49:66.2%, 0.959)		
EIC4201	<=	EIC4102	&	EIC4103	(61:82.4%, 0.967)				
EIC4201	<=	EIC4102	&	EIC4104	(53:71.6%, 0.962)				
EIC4201	<=	EIC4102	(65:87.8%, 0.969)						
EIC4201	<=	EIC4103	&	EIC4101	&	EIC4104	(48:64.9%, 0.958)		
EIC4201	<=	EIC4103	&	EIC4101	(58:78.4%, 0.966)				
EIC4201	<=	EIC4103	&	EIC4104	(50:67.6%, 0.96)				
EIC4201	<=	EIC4103	(62:83.8%, 0.968)						
EIC4201	<=	EIC4104	(56:75.7%, 0.946)						
EIC4201	<=	EIC4105	&	EIC4100	&	EIC4101	&	EIC4104	(51:68.9%, 0.961)
EIC4201	<=	EIC4105	&	EIC4100	&	EIC4101	(61:82.4%, 0.967)		
EIC4201	<=	EIC4105	&	EIC4100	&	EIC4102	&	EIC4101	(61:82.4%, 0.967)
EIC4201	<=	EIC4105	&	EIC4100	&	EIC4102	&	EIC4103	(60:81.1%, 0.967)
EIC4201	<=	EIC4105	&	EIC4100	&	EIC4102	&	EIC4104	(52:70.3%, 0.962)
EIC4201	<=	EIC4105	&	EIC4100	&	EIC4102	(63:85.1%, 0.968)		
EIC4201	<=	EIC4105	&	EIC4100	&	EIC4103	&	EIC4101	(58:78.4%, 0.966)
EIC4201	<=	EIC4105	&	EIC4100	&	EIC4103	&	EIC4104	(50:67.6%, 0.96)
EIC4201	<=	EIC4105	&	EIC4100	&	EIC4103	(61:82.4%, 0.967)		
EIC4201	<=	EIC4105	&	EIC4100	&	EIC4104	(53:71.6%, 0.962)		
EIC4201	<=	EIC4105	&	EIC4100	(64:86.5%, 0.969)				
EIC4201	<=	EIC4105	&	EIC4101	&	EIC4104	(51:68.9%, 0.961)		
EIC4201	<=	EIC4105	&	EIC4101	(61:82.4%, 0.967)				
EIC4201	<=	EIC4105	&	EIC4102	&	EIC4101	&	EIC4104	(51:68.9%, 0.961)
EIC4201	<=	EIC4105	&	EIC4102	&	EIC4101	(61:82.4%, 0.967)		
EIC4201	<=	EIC4105	&	EIC4102	&	EIC4103	&	EIC4101	(58:78.4%, 0.966)
EIC4201	<=	EIC4105	&	EIC4102	&	EIC4103	&	EIC4104	(49:66.2%, 0.959)
EIC4201	<=	EIC4105	&	EIC4102	&	EIC4103	(61:82.4%, 0.967)		
EIC4201	<=	EIC4105	&	EIC4102	&	EIC4104	(52:70.3%, 0.962)		
EIC4201	<=	EIC4105	&	EIC4102	(64:86.5%, 0.969)				
EIC4201	<=	EIC4105	&	EIC4103	&	EIC4101	&	EIC4104	(48:64.9%, 0.958)
EIC4201	<=	EIC4105	&	EIC4103	&	EIC4101	(58:78.4%, 0.966)		
EIC4201	<=	EIC4105	&	EIC4103	&	EIC4104	(50:67.6%, 0.96)		
EIC4201	<=	EIC4105	&	EIC4103	(62:83.8%, 0.968)				
EIC4201	<=	EIC4105	&	EIC4104	(55:74.3%, 0.945)				
EIC4201	<=	EIC4105	(69:93.2%, 0.942)						

**Figura III.65 - Regras de associação: disciplina EIC4201 (sem.)**

Melhor suporte: 69:93.2% (conf=94.2%)  
 Melhor confiança: 96.9% (sup=65:87.8%)

EIC4201 <= EIC4100 (65:87.8%, 0.969)  
 EIC4201 <= EIC4102 (65:87.8%, 0.969)  
 EIC4201 <= EIC4105 (69:93.2%, 0.942)

**Figura III.66 - Regras de associação: disciplina EIC4201 (sem.) após o filtro**

EIC4202	<=	EIC4100	&	EIC4101	&	EIC4104	(51:68.9%, 0.98)		
EIC4202	<=	EIC4100	&	EIC4101	(61:82.4%, 0.951)				
EIC4202	<=	EIC4100	&	EIC4102	&	EIC4104	(51:68.9%, 0.98)		
EIC4202	<=	EIC4100	&	EIC4102	&	EIC4101	(61:82.4%, 0.951)		
EIC4202	<=	EIC4100	&	EIC4102	&	EIC4103	&	EIC4101	(58:78.4%, 0.948)
EIC4202	<=	EIC4100	&	EIC4102	&	EIC4103	&	EIC4104	(49:66.2%, 0.98)
EIC4202	<=	EIC4100	&	EIC4102	&	EIC4103	(60:81.1%, 0.95)		
EIC4202	<=	EIC4100	&	EIC4102	&	EIC4104	(53:71.6%, 0.981)		
EIC4202	<=	EIC4100	&	EIC4102	(64:86.5%, 0.953)				
EIC4202	<=	EIC4100	&	EIC4103	&	EIC4101	&	EIC4104	(48:64.9%, 0.979)
EIC4202	<=	EIC4100	&	EIC4103	&	EIC4101	(58:78.4%, 0.948)		
EIC4202	<=	EIC4100	&	EIC4103	&	EIC4104	(50:67.6%, 0.96)		
EIC4202	<=	EIC4100	&	EIC4103	(61:82.4%, 0.934)				
EIC4202	<=	EIC4100	&	EIC4104	(54:73.0%, 0.963)				
EIC4202	<=	EIC4100	(65:87.8%, 0.938)						
EIC4202	<=	EIC4101	&	EIC4104	(51:68.9%, 0.98)				
EIC4202	<=	EIC4101	(61:82.4%, 0.951)						
EIC4202	<=	EIC4102	&	EIC4101	&	EIC4104	(51:68.9%, 0.98)		
EIC4202	<=	EIC4102	&	EIC4101	(61:82.4%, 0.951)				
EIC4202	<=	EIC4102	&	EIC4103	&	EIC4101	&	EIC4104	(48:64.9%, 0.979)
EIC4202	<=	EIC4102	&	EIC4103	&	EIC4101	(58:78.4%, 0.948)		
EIC4202	<=	EIC4102	&	EIC4103	&	EIC4104	(49:66.2%, 0.98)		
EIC4202	<=	EIC4102	&	EIC4103	(61:82.4%, 0.951)				
EIC4202	<=	EIC4102	&	EIC4104	(53:71.6%, 0.981)				
EIC4202	<=	EIC4102	(65:87.8%, 0.954)						
EIC4202	<=	EIC4103	&	EIC4101	&	EIC4104	(48:64.9%, 0.979)		
EIC4202	<=	EIC4103	&	EIC4101	(58:78.4%, 0.948)				
EIC4202	<=	EIC4103	&	EIC4104	(50:67.6%, 0.96)				
EIC4202	<=	EIC4103	(62:83.8%, 0.935)						
EIC4202	<=	EIC4104	(56:75.7%, 0.964)						
EIC4202	<=	EIC4105	&	EIC4100	&	EIC4101	&	EIC4104	(51:68.9%, 0.98)
EIC4202	<=	EIC4105	&	EIC4100	&	EIC4101	(61:82.4%, 0.951)		
EIC4202	<=	EIC4105	&	EIC4100	&	EIC4102	&	EIC4101	(61:82.4%, 0.951)
EIC4202	<=	EIC4105	&	EIC4100	&	EIC4102	&	EIC4103	(60:81.1%, 0.95)
EIC4202	<=	EIC4105	&	EIC4100	&	EIC4102	&	EIC4104	(52:70.3%, 0.981)
EIC4202	<=	EIC4105	&	EIC4100	&	EIC4102	(63:85.1%, 0.952)		
EIC4202	<=	EIC4105	&	EIC4100	&	EIC4103	&	EIC4101	(58:78.4%, 0.948)
EIC4202	<=	EIC4105	&	EIC4100	&	EIC4103	&	EIC4104	(50:67.6%, 0.96)
EIC4202	<=	EIC4105	&	EIC4100	&	EIC4103	(61:82.4%, 0.934)		
EIC4202	<=	EIC4105	&	EIC4100	&	EIC4104	(53:71.6%, 0.962)		
EIC4202	<=	EIC4105	&	EIC4100	(64:86.5%, 0.938)				
EIC4202	<=	EIC4105	&	EIC4101	&	EIC4104	(51:68.9%, 0.98)		
EIC4202	<=	EIC4105	&	EIC4101	(61:82.4%, 0.951)				
EIC4202	<=	EIC4105	&	EIC4102	&	EIC4101	&	EIC4104	(51:68.9%, 0.98)
EIC4202	<=	EIC4105	&	EIC4102	&	EIC4101	(61:82.4%, 0.951)		
EIC4202	<=	EIC4105	&	EIC4102	&	EIC4103	&	EIC4101	(58:78.4%, 0.948)
EIC4202	<=	EIC4105	&	EIC4102	&	EIC4103	&	EIC4104	(49:66.2%, 0.98)
EIC4202	<=	EIC4105	&	EIC4102	&	EIC4103	(61:82.4%, 0.951)		
EIC4202	<=	EIC4105	&	EIC4102	&	EIC4104	(52:70.3%, 0.981)		
EIC4202	<=	EIC4105	&	EIC4102	(64:86.5%, 0.953)				
EIC4202	<=	EIC4105	&	EIC4103	&	EIC4101	&	EIC4104	(48:64.9%, 0.979)
EIC4202	<=	EIC4105	&	EIC4103	&	EIC4101	(58:78.4%, 0.948)		
EIC4202	<=	EIC4105	&	EIC4103	&	EIC4104	(50:67.6%, 0.96)		
EIC4202	<=	EIC4105	&	EIC4103	(62:83.8%, 0.935)				
EIC4202	<=	EIC4105	&	EIC4104	(55:74.3%, 0.964)				
EIC4202	<=	EIC4105	(69:93.2%, 0.928)						

Figura III.67 - Regras de associação: disciplina EIC4202 (sem.)

Melhor suporte:	69:93.2%	{conf=92.8%}					
Melhor confiança:	98.1%	{sup=53:71.6%}					
EIC4202	<=	EIC4100	&	EIC4102	&	EIC4104	(53:71.6%, 0.981)
EIC4202	<=	EIC4102	&	EIC4104	(53:71.6%, 0.981)		
EIC4202	<=	EIC4105	(69:93.2%, 0.928)				
EIC4202	<=	EIC4104	(56:75.7%, 0.964)				
EIC4202	<=	EIC4102	(65:87.8%, 0.954)				

Figura III.68 - Regras de associação: disciplina EIC4202 (sem.) após o filtro

EIC4203	<=	EIC4100	&	EIC4101	&	EIC4104	(51:68.9%, 0.922)		
EIC4203	<=	EIC4100	&	EIC4101	(61:82.4%, 0.918)				
EIC4203	<=	EIC4100	&	EIC4102	&	EIC4101	&	EIC4104	(51:68.9%, 0.922)
EIC4203	<=	EIC4100	&	EIC4102	&	EIC4101	(61:82.4%, 0.918)		
EIC4203	<=	EIC4100	&	EIC4102	&	EIC4103	&	EIC4101	(58:78.4%, 0.948)
EIC4203	<=	EIC4100	&	EIC4102	&	EIC4103	&	EIC4104	(49:66.2%, 0.959)
EIC4203	<=	EIC4100	&	EIC4102	&	EIC4103	(60:81.1%, 0.933)		
EIC4203	<=	EIC4100	&	EIC4102	&	EIC4104	(53:71.6%, 0.925)		
EIC4203	<=	EIC4100	&	EIC4102	(64:86.5%, 0.906)				
EIC4203	<=	EIC4100	&	EIC4103	&	EIC4101	&	EIC4104	(48:64.9%, 0.958)
EIC4203	<=	EIC4100	&	EIC4103	&	EIC4101	(58:78.4%, 0.948)		
EIC4203	<=	EIC4100	&	EIC4103	&	EIC4104	(50:67.6%, 0.96)		
EIC4203	<=	EIC4100	&	EIC4103	(61:82.4%, 0.934)				
EIC4203	<=	EIC4100	&	EIC4104	(54:73.0%, 0.926)				
EIC4203	<=	EIC4100	(65:87.8%, 0.908)						
EIC4203	<=	EIC4101	&	EIC4104	(51:68.9%, 0.922)				
EIC4203	<=	EIC4101	(61:82.4%, 0.918)						
EIC4203	<=	EIC4102	&	EIC4101	&	EIC4104	(51:68.9%, 0.922)		
EIC4203	<=	EIC4102	&	EIC4101	(61:82.4%, 0.918)				
EIC4203	<=	EIC4102	&	EIC4103	&	EIC4101	&	EIC4104	(48:64.9%, 0.958)
EIC4203	<=	EIC4102	&	EIC4103	&	EIC4101	(58:78.4%, 0.948)		
EIC4203	<=	EIC4102	&	EIC4103	&	EIC4104	(49:66.2%, 0.959)		
EIC4203	<=	EIC4102	&	EIC4103	(61:82.4%, 0.934)				
EIC4203	<=	EIC4102	&	EIC4104	(53:71.6%, 0.925)				
EIC4203	<=	EIC4102	(65:87.8%, 0.908)						
EIC4203	<=	EIC4103	&	EIC4101	&	EIC4104	(48:64.9%, 0.958)		
EIC4203	<=	EIC4103	&	EIC4101	(58:78.4%, 0.948)				
EIC4203	<=	EIC4103	&	EIC4104	(50:67.6%, 0.96)				
EIC4203	<=	EIC4103	(62:83.8%, 0.935)						
EIC4203	<=	EIC4104	(56:75.7%, 0.893)						
EIC4203	<=	EIC4105	&	EIC4100	&	EIC4101	&	EIC4104	(51:68.9%, 0.922)
EIC4203	<=	EIC4105	&	EIC4100	&	EIC4101	(61:82.4%, 0.918)		
EIC4203	<=	EIC4105	&	EIC4100	&	EIC4102	&	EIC4101	(61:82.4%, 0.918)
EIC4203	<=	EIC4105	&	EIC4100	&	EIC4102	&	EIC4103	(60:81.1%, 0.933)
EIC4203	<=	EIC4105	&	EIC4100	&	EIC4102	&	EIC4104	(52:70.3%, 0.923)
EIC4203	<=	EIC4105	&	EIC4100	&	EIC4102	(63:85.1%, 0.905)		
EIC4203	<=	EIC4105	&	EIC4100	&	EIC4103	&	EIC4101	(58:78.4%, 0.948)
EIC4203	<=	EIC4105	&	EIC4100	&	EIC4103	&	EIC4104	(50:67.6%, 0.96)
EIC4203	<=	EIC4105	&	EIC4100	&	EIC4103	(61:82.4%, 0.934)		
EIC4203	<=	EIC4105	&	EIC4100	&	EIC4104	(53:71.6%, 0.925)		
EIC4203	<=	EIC4105	&	EIC4100	(64:86.5%, 0.906)				
EIC4203	<=	EIC4105	&	EIC4101	&	EIC4104	(51:68.9%, 0.922)		
EIC4203	<=	EIC4105	&	EIC4101	(61:82.4%, 0.918)				
EIC4203	<=	EIC4105	&	EIC4102	&	EIC4101	&	EIC4104	(51:68.9%, 0.922)
EIC4203	<=	EIC4105	&	EIC4102	&	EIC4101	(61:82.4%, 0.918)		
EIC4203	<=	EIC4105	&	EIC4102	&	EIC4103	&	EIC4101	(58:78.4%, 0.948)
EIC4203	<=	EIC4105	&	EIC4102	&	EIC4103	&	EIC4104	(49:66.2%, 0.959)
EIC4203	<=	EIC4105	&	EIC4102	&	EIC4103	(61:82.4%, 0.934)		
EIC4203	<=	EIC4105	&	EIC4102	&	EIC4104	(52:70.3%, 0.923)		
EIC4203	<=	EIC4105	&	EIC4102	(64:86.5%, 0.906)				
EIC4203	<=	EIC4105	&	EIC4103	&	EIC4101	&	EIC4104	(48:64.9%, 0.958)
EIC4203	<=	EIC4105	&	EIC4103	&	EIC4101	(58:78.4%, 0.948)		
EIC4203	<=	EIC4105	&	EIC4103	&	EIC4104	(50:67.6%, 0.96)		
EIC4203	<=	EIC4105	&	EIC4103	(62:83.8%, 0.935)				
EIC4203	<=	EIC4105	&	EIC4104	(55:74.3%, 0.891)				
EIC4203	<=	EIC4105	(69:93.2%, 0.855)						

Figura III.69 - Regras de associação: disciplina EIC4203 (sem.)

Melhor suporte:	69:93.2%	(conf=85.5%)							
Melhor confiança:	96.0%	(sup=50:67.6%)							
EIC4203	<=	EIC4100	&	EIC4103	&	EIC4104	(50:67.6%, 0.96)		
EIC4203	<=	EIC4103	&	EIC4104	(50:67.6%, 0.96)				
EIC4203	<=	EIC4105	&	EIC4100	&	EIC4103	&	EIC4104	(50:67.6%, 0.96)
EIC4203	<=	EIC4105	&	EIC4103	&	EIC4104	(50:67.6%, 0.96)		
EIC4203	<=	EIC4105	(69:93.2%, 0.855)						
EIC4203	<=	EIC4105	&	EIC4103	&	EIC4101	(58:78.4%, 0.948)		
EIC4203	<=	EIC4105	&	EIC4102	&	EIC4103	&	EIC4101	(58:78.4%, 0.948)
EIC4203	<=	EIC4105	&	EIC4100	&	EIC4103	&	EIC4101	(58:78.4%, 0.948)
EIC4203	<=	EIC4103	&	EIC4101	(58:78.4%, 0.948)				
EIC4203	<=	EIC4102	&	EIC4103	&	EIC4101	(58:78.4%, 0.948)		
EIC4203	<=	EIC4100	&	EIC4103	&	EIC4101	(58:78.4%, 0.948)		
EIC4203	<=	EIC4100	&	EIC4102	&	EIC4103	&	EIC4101	(58:78.4%, 0.948)
EIC4203	<=	EIC4102	(65:87.8%, 0.908)						
EIC4203	<=	EIC4100	(65:87.8%, 0.908)						
EIC4203	<=	EIC4103	(62:83.8%, 0.935)						
EIC4203	<=	EIC4105	&	EIC4103	(62:83.8%, 0.935)				

Figura III.70 - Regras de associação: disciplina EIC4203 (sem.) após o filtro

EIC4205	<=	EIC4100	&	EIC4101	&	EIC4104	(51:68.9%, 0.961)		
EIC4205	<=	EIC4100	&	EIC4101	(61:82.4%, 0.951)				
EIC4205	<=	EIC4100	&	EIC4102	&	EIC4101	&	EIC4104	(51:68.9%, 0.961)
EIC4205	<=	EIC4100	&	EIC4102	&	EIC4101	(61:82.4%, 0.951)		
EIC4205	<=	EIC4100	&	EIC4102	&	EIC4103	&	EIC4101	(58:78.4%, 0.948)
EIC4205	<=	EIC4100	&	EIC4102	&	EIC4103	&	EIC4104	(49:66.2%, 0.959)
EIC4205	<=	EIC4100	&	EIC4102	&	EIC4103	(60:81.1%, 0.933)		
EIC4205	<=	EIC4100	&	EIC4102	&	EIC4104	(53:71.6%, 0.943)		
EIC4205	<=	EIC4100	&	EIC4102	(64:86.5%, 0.922)				
EIC4205	<=	EIC4100	&	EIC4103	&	EIC4101	&	EIC4104	(48:64.9%, 0.958)
EIC4205	<=	EIC4100	&	EIC4103	&	EIC4101	(58:78.4%, 0.948)		
EIC4205	<=	EIC4100	&	EIC4103	&	EIC4104	(50:67.6%, 0.96)		
EIC4205	<=	EIC4100	&	EIC4103	(61:82.4%, 0.934)				
EIC4205	<=	EIC4100	&	EIC4104	(54:73.0%, 0.944)				
EIC4205	<=	EIC4100	(65:87.8%, 0.923)						
EIC4205	<=	EIC4101	&	EIC4104	(51:68.9%, 0.961)				
EIC4205	<=	EIC4101	(61:82.4%, 0.951)						
EIC4205	<=	EIC4102	&	EIC4101	&	EIC4104	(51:68.9%, 0.961)		
EIC4205	<=	EIC4102	&	EIC4101	(61:82.4%, 0.951)				
EIC4205	<=	EIC4102	&	EIC4103	&	EIC4101	&	EIC4104	(48:64.9%, 0.958)
EIC4205	<=	EIC4102	&	EIC4103	&	EIC4101	(58:78.4%, 0.948)		
EIC4205	<=	EIC4102	&	EIC4103	&	EIC4104	(49:66.2%, 0.959)		
EIC4205	<=	EIC4102	&	EIC4103	(61:82.4%, 0.934)				
EIC4205	<=	EIC4102	&	EIC4104	(53:71.6%, 0.943)				
EIC4205	<=	EIC4102	(65:87.8%, 0.923)						
EIC4205	<=	EIC4103	&	EIC4101	&	EIC4104	(48:64.9%, 0.958)		
EIC4205	<=	EIC4103	&	EIC4101	(58:78.4%, 0.948)				
EIC4205	<=	EIC4103	&	EIC4104	(50:67.6%, 0.96)				
EIC4205	<=	EIC4103	(62:83.8%, 0.935)						
EIC4205	<=	EIC4104	(56:75.7%, 0.911)						
EIC4205	<=	EIC4105	&	EIC4100	&	EIC4101	&	EIC4104	(51:68.9%, 0.961)
EIC4205	<=	EIC4105	&	EIC4100	&	EIC4101	(61:82.4%, 0.951)		
EIC4205	<=	EIC4105	&	EIC4100	&	EIC4102	&	EIC4101	(61:82.4%, 0.951)
EIC4205	<=	EIC4105	&	EIC4100	&	EIC4102	&	EIC4103	(60:81.1%, 0.933)
EIC4205	<=	EIC4105	&	EIC4100	&	EIC4102	&	EIC4104	(52:70.3%, 0.962)
EIC4205	<=	EIC4105	&	EIC4100	&	EIC4102	(63:85.1%, 0.937)		
EIC4205	<=	EIC4105	&	EIC4100	&	EIC4103	&	EIC4101	(58:78.4%, 0.948)
EIC4205	<=	EIC4105	&	EIC4100	&	EIC4103	&	EIC4104	(50:67.6%, 0.96)
EIC4205	<=	EIC4105	&	EIC4100	&	EIC4103	(61:82.4%, 0.934)		
EIC4205	<=	EIC4105	&	EIC4100	&	EIC4104	(53:71.6%, 0.962)		
EIC4205	<=	EIC4105	&	EIC4100	(64:86.5%, 0.938)				
EIC4205	<=	EIC4105	&	EIC4101	&	EIC4104	(51:68.9%, 0.961)		
EIC4205	<=	EIC4105	&	EIC4101	(61:82.4%, 0.951)				
EIC4205	<=	EIC4105	&	EIC4102	&	EIC4101	&	EIC4104	(51:68.9%, 0.961)
EIC4205	<=	EIC4105	&	EIC4102	&	EIC4101	(61:82.4%, 0.951)		
EIC4205	<=	EIC4105	&	EIC4102	&	EIC4103	&	EIC4101	(58:78.4%, 0.948)
EIC4205	<=	EIC4105	&	EIC4102	&	EIC4103	&	EIC4104	(49:66.2%, 0.959)
EIC4205	<=	EIC4105	&	EIC4102	&	EIC4103	(61:82.4%, 0.934)		
EIC4205	<=	EIC4105	&	EIC4102	&	EIC4104	(52:70.3%, 0.962)		
EIC4205	<=	EIC4105	&	EIC4102	(64:86.5%, 0.938)				
EIC4205	<=	EIC4105	&	EIC4103	&	EIC4101	&	EIC4104	(48:64.9%, 0.958)
EIC4205	<=	EIC4105	&	EIC4103	&	EIC4101	(58:78.4%, 0.948)		
EIC4205	<=	EIC4105	&	EIC4103	&	EIC4104	(50:67.6%, 0.96)		
EIC4205	<=	EIC4105	&	EIC4103	(62:83.8%, 0.935)				
EIC4205	<=	EIC4105	&	EIC4104	(55:74.3%, 0.927)				
EIC4205	<=	EIC4105	(69:93.2%, 0.884)						

Figura III.71 - Regras de associação: disciplina EIC4205 (sem.)

Melhor suporte:	69:93.2%	(conf=88.4%)							
Melhor confiança:	96.2%	(sup=53:71.6%)							
EIC4205	<=	EIC4105	&	EIC4100	&	EIC4104	(53:71.6%, 0.962)		
EIC4205	<=	EIC4105	(69:93.2%, 0.884)						
EIC4205	<=	EIC4105	&	EIC4102	&	EIC4101	(61:82.4%, 0.951)		
EIC4205	<=	EIC4105	&	EIC4101	(61:82.4%, 0.951)				
EIC4205	<=	EIC4105	&	EIC4100	&	EIC4102	&	EIC4101	(61:82.4%, 0.951)
EIC4205	<=	EIC4105	&	EIC4100	&	EIC4101	(61:82.4%, 0.951)		
EIC4205	<=	EIC4102	&	EIC4101	(61:82.4%, 0.951)				
EIC4205	<=	EIC4101	(61:82.4%, 0.951)						
EIC4205	<=	EIC4100	&	EIC4102	&	EIC4101	(61:82.4%, 0.951)		
EIC4205	<=	EIC4100	&	EIC4101	(61:82.4%, 0.951)				
EIC4205	<=	EIC4102	(65:87.8%, 0.923)						
EIC4205	<=	EIC4100	(65:87.8%, 0.923)						
EIC4205	<=	EIC4105	&	EIC4100	(64:86.5%, 0.938)				
EIC4205	<=	EIC4105	&	EIC4102	(64:86.5%, 0.938)				

Figura III.72 - Regras de associação: disciplina EIC4205 (sem.) após o filtro













Melhor suporte: 72:97.3% (conf=90.3%)  
 Melhor confiança: 100.0% (sup=64:86.5%)

EIC4102	<=	EIC3105 & EIC3202 & EIC3201 & EIC3102	(64:86.5%, 1.0)
EIC4102	<=	EIC3202 & EIC3200 & EIC3201 & EIC3102	(64:86.5%, 1.0)
EIC4102	<=	EIC3205 & EIC3202 & EIC3201 & EIC3102	(64:86.5%, 1.0)
EIC4102	<=	EIC3105	(72:97.3%, 0.903)
EIC4102	<=	EIC3205 & EIC3202 & EIC3200 & EIC3102	(66:89.2%, 0.985)
EIC4102	<=	EIC3205 & EIC3202 & EIC3102	(66:89.2%, 0.985)
EIC4102	<=	EIC3205 & EIC3103 & EIC3202 & EIC3102	(66:89.2%, 0.985)
EIC4102	<=	EIC3202 & EIC3200 & EIC3102	(66:89.2%, 0.985)
EIC4102	<=	EIC3105 & EIC3205 & EIC3202 & EIC3102	(66:89.2%, 0.985)
EIC4102	<=	EIC3105 & EIC3202 & EIC3200 & EIC3102	(66:89.2%, 0.985)
EIC4102	<=	EIC3105 & EIC3202 & EIC3102	(66:89.2%, 0.985)
EIC4102	<=	EIC3105 & EIC3103 & EIC3202 & EIC3102	(66:89.2%, 0.985)
EIC4102	<=	EIC3103 & EIC3202 & EIC3200 & EIC3102	(66:89.2%, 0.985)
EIC4102	<=	EIC3205	(71:95.9%, 0.915)
EIC4102	<=	EIC3105 & EIC3205	(71:95.9%, 0.915)
EIC4102	<=	EIC3103	(71:95.9%, 0.915)
EIC4102	<=	EIC3103 & EIC3200 & EIC3102	(67:90.5%, 0.97)
EIC4102	<=	EIC3103 & EIC3202 & EIC3200	(67:90.5%, 0.97)
EIC4102	<=	EIC3105 & EIC3102	(67:90.5%, 0.97)
EIC4102	<=	EIC3105 & EIC3103 & EIC3102	(67:90.5%, 0.97)
EIC4102	<=	EIC3105 & EIC3103 & EIC3200 & EIC3102	(67:90.5%, 0.97)
EIC4102	<=	EIC3105 & EIC3103 & EIC3202	(67:90.5%, 0.97)
EIC4102	<=	EIC3105 & EIC3200 & EIC3102	(67:90.5%, 0.97)
EIC4102	<=	EIC3105 & EIC3205 & EIC3102	(67:90.5%, 0.97)
EIC4102	<=	EIC3105 & EIC3205 & EIC3103 & EIC3102	(67:90.5%, 0.97)
EIC4102	<=	EIC3105 & EIC3205 & EIC3103 & EIC3202	(67:90.5%, 0.97)
EIC4102	<=	EIC3105 & EIC3205 & EIC3200 & EIC3102	(67:90.5%, 0.97)
EIC4102	<=	EIC3200 & EIC3102	(67:90.5%, 0.97)
EIC4102	<=	EIC3205 & EIC3102	(67:90.5%, 0.97)
EIC4102	<=	EIC3205 & EIC3103 & EIC3102	(67:90.5%, 0.97)
EIC4102	<=	EIC3205 & EIC3103 & EIC3200 & EIC3102	(67:90.5%, 0.97)
EIC4102	<=	EIC3205 & EIC3103 & EIC3202	(67:90.5%, 0.97)
EIC4102	<=	EIC3205 & EIC3103 & EIC3202	(67:90.5%, 0.97)
EIC4102	<=	EIC3205 & EIC3200 & EIC3102	(67:90.5%, 0.97)
EIC4102	<=	EIC3105 & EIC3200	(70:94.6%, 0.929)
EIC4102	<=	EIC3105 & EIC3205 & EIC3200	(70:94.6%, 0.929)
EIC4102	<=	EIC3200	(70:94.6%, 0.929)
EIC4102	<=	EIC3202	(70:94.6%, 0.929)
EIC4102	<=	EIC3205 & EIC3200	(70:94.6%, 0.929)
EIC4102	<=	EIC3205 & EIC3202	(68:91.9%, 0.956)
EIC4102	<=	EIC3205 & EIC3202 & EIC3200	(68:91.9%, 0.956)
EIC4102	<=	EIC3202 & EIC3200	(68:91.9%, 0.956)
EIC4102	<=	EIC3202 & EIC3102	(68:91.9%, 0.956)
EIC4102	<=	EIC3105 & EIC3205 & EIC3202	(68:91.9%, 0.956)
EIC4102	<=	EIC3105 & EIC3202	(68:91.9%, 0.956)
EIC4102	<=	EIC3105 & EIC3202 & EIC3200	(68:91.9%, 0.956)
EIC4102	<=	EIC3103 & EIC3202 & EIC3102	(68:91.9%, 0.956)
EIC4102	<=	EIC3205 & EIC3103	(69:93.2%, 0.942)
EIC4102	<=	EIC3205 & EIC3103 & EIC3200	(69:93.2%, 0.942)
EIC4102	<=	EIC3105 & EIC3205 & EIC3103	(69:93.2%, 0.942)
EIC4102	<=	EIC3105 & EIC3205 & EIC3103 & EIC3200	(69:93.2%, 0.942)
EIC4102	<=	EIC3105 & EIC3103	(69:93.2%, 0.942)
EIC4102	<=	EIC3105 & EIC3103 & EIC3200	(69:93.2%, 0.942)
EIC4102	<=	EIC3103 & EIC3202	(69:93.2%, 0.942)
EIC4102	<=	EIC3103 & EIC3200	(69:93.2%, 0.942)
EIC4102	<=	EIC3103 & EIC3102	(69:93.2%, 0.942)
EIC4102	<=	EIC3102	(69:93.2%, 0.942)

**Figura III.78 - Regras de associação: disciplina EIC4102 (anual) após o filtro**

Melhor suporte: 72:97.3% (conf=86.1%)	
Melhor confiança: 93.8% (sup=65:87.8%)	
EIC4103	<= EIC3103 & EIC3200 & EIC3201 & EIC3102 (65:87.8%, 0.938)
EIC4103	<= EIC3103 & EIC3204 & EIC3200 & EIC3201 (65:87.8%, 0.938)
EIC4103	<= EIC3105 & EIC3100 & EIC3201 & EIC3102 (65:87.8%, 0.938)
EIC4103	<= EIC3105 & EIC3103 & EIC3201 & EIC3102 (65:87.8%, 0.938)
EIC4103	<= EIC3105 & EIC3103 & EIC3204 & EIC3201 (65:87.8%, 0.938)
EIC4103	<= EIC3105 & EIC3200 & EIC3201 & EIC3102 (65:87.8%, 0.938)
EIC4103	<= EIC3105 & EIC3201 & EIC3102 (65:87.8%, 0.938)
EIC4103	<= EIC3105 & EIC3204 & EIC3201 & EIC3102 (65:87.8%, 0.938)
EIC4103	<= EIC3105 & EIC3205 & EIC3201 & EIC3102 (65:87.8%, 0.938)
EIC4103	<= EIC3200 & EIC3100 & EIC3201 & EIC3102 (65:87.8%, 0.938)
EIC4103	<= EIC3200 & EIC3201 & EIC3102 (65:87.8%, 0.938)
EIC4103	<= EIC3204 & EIC3200 & EIC3201 & EIC3102 (65:87.8%, 0.938)
EIC4103	<= EIC3205 & EIC3100 & EIC3201 & EIC3102 (65:87.8%, 0.938)
EIC4103	<= EIC3205 & EIC3103 & EIC3201 & EIC3102 (65:87.8%, 0.938)
EIC4103	<= EIC3205 & EIC3103 & EIC3204 & EIC3201 (65:87.8%, 0.938)
EIC4103	<= EIC3205 & EIC3200 & EIC3201 & EIC3102 (65:87.8%, 0.938)
EIC4103	<= EIC3205 & EIC3201 & EIC3102 (65:87.8%, 0.938)
EIC4103	<= EIC3205 & EIC3204 & EIC3201 & EIC3102 (65:87.8%, 0.938)
EIC4103	<= EIC3105 (72:97.3%, 0.861)
EIC4103	<= EIC3205 & EIC3200 & EIC3102 (67:90.5%, 0.925)
EIC4103	<= EIC3205 & EIC3103 & EIC3200 & EIC3102 (67:90.5%, 0.925)
EIC4103	<= EIC3205 & EIC3103 & EIC3102 (67:90.5%, 0.925)
EIC4103	<= EIC3205 & EIC3102 (67:90.5%, 0.925)
EIC4103	<= EIC3200 & EIC3102 (67:90.5%, 0.925)
EIC4103	<= EIC3105 & EIC3205 & EIC3200 & EIC3102 (67:90.5%, 0.925)
EIC4103	<= EIC3105 & EIC3205 & EIC3103 & EIC3102 (67:90.5%, 0.925)
EIC4103	<= EIC3105 & EIC3205 & EIC3102 (67:90.5%, 0.925)
EIC4103	<= EIC3105 & EIC3200 & EIC3102 (67:90.5%, 0.925)
EIC4103	<= EIC3105 & EIC3103 & EIC3200 & EIC3102 (67:90.5%, 0.925)
EIC4103	<= EIC3105 & EIC3103 & EIC3102 (67:90.5%, 0.925)
EIC4103	<= EIC3105 & EIC3102 (67:90.5%, 0.925)
EIC4103	<= EIC3103 & EIC3200 & EIC3102 (67:90.5%, 0.925)
EIC4103	<= EIC3205 (71:95.9%, 0.873)
EIC4103	<= EIC3105 & EIC3205 (71:95.9%, 0.873)
EIC4103	<= EIC3103 (71:95.9%, 0.873)
EIC4103	<= EIC3102 (69:93.2%, 0.899)
EIC4103	<= EIC3103 & EIC3102 (69:93.2%, 0.899)
EIC4103	<= EIC3103 & EIC3200 (69:93.2%, 0.899)
EIC4103	<= EIC3105 & EIC3103 & EIC3200 (69:93.2%, 0.899)
EIC4103	<= EIC3105 & EIC3103 (69:93.2%, 0.899)
EIC4103	<= EIC3105 & EIC3205 & EIC3103 & EIC3200 (69:93.2%, 0.899)
EIC4103	<= EIC3105 & EIC3205 & EIC3103 (69:93.2%, 0.899)
EIC4103	<= EIC3205 & EIC3103 & EIC3200 (69:93.2%, 0.899)
EIC4103	<= EIC3205 & EIC3103 (69:93.2%, 0.899)
EIC4103	<= EIC3105 & EIC3200 (70:94.6%, 0.886)
EIC4103	<= EIC3105 & EIC3205 & EIC3200 (70:94.6%, 0.886)
EIC4103	<= EIC3200 (70:94.6%, 0.886)
EIC4103	<= EIC3205 & EIC3200 (70:94.6%, 0.886)

**Figura III.79 - Regras de associação: disciplina EIC4103 (anual) após o filtro**









EIC4200	<=	EIC3205	&	EIC3202	&	EIC3200	&	EIC3100	(67:90.5%, 0.836)
EIC4200	<=	EIC3205	&	EIC3204	&	EIC3100	(67:90.5%, 0.836)		
EIC4200	<=	EIC3205	&	EIC3204	&	EIC3200	&	EIC3100	(67:90.5%, 0.836)
EIC4200	<=	EIC3205	&	EIC3204	&	EIC3200	(67:90.5%, 0.836)		
EIC4200	<=	EIC3105	&	EIC3202	&	EIC3200	(68:91.9%, 0.824)		
EIC4200	<=	EIC3105	&	EIC3202	(68:91.9%, 0.824)				
EIC4200	<=	EIC3105	&	EIC3205	&	EIC3202	&	EIC3200	(68:91.9%, 0.824)
EIC4200	<=	EIC3105	&	EIC3205	&	EIC3202	(68:91.9%, 0.824)		
EIC4200	<=	EIC3105	&	EIC3205	&	EIC3204	(68:91.9%, 0.824)		
EIC4200	<=	EIC3202	&	EIC3100	(68:91.9%, 0.824)				
EIC4200	<=	EIC3202	&	EIC3200	(68:91.9%, 0.824)				
EIC4200	<=	EIC3204	&	EIC3100	(68:91.9%, 0.824)				
EIC4200	<=	EIC3205	&	EIC3202	&	EIC3200	(68:91.9%, 0.824)		
EIC4200	<=	EIC3205	&	EIC3202	(68:91.9%, 0.824)				
EIC4200	<=	EIC3205	&	EIC3204	(68:91.9%, 0.824)				

**Figura III.84 - Regras de associação: disciplina EIC4200 (anual) após o filtro (cont.)**

Melhor suporte: 72:97.3% (conf=93.1%)									
Melhor confiança: 97.1% (sup=68:91.9%)									
EIC4201	<=	EIC3105	&	EIC3202	&	EIC3200	(68:91.9%, 0.971)		
EIC4201	<=	EIC3105	&	EIC3202	(68:91.9%, 0.971)				
EIC4201	<=	EIC3105	&	EIC3205	&	EIC3202	&	EIC3200	(68:91.9%, 0.971)
EIC4201	<=	EIC3105	&	EIC3205	&	EIC3202	(68:91.9%, 0.971)		
EIC4201	<=	EIC3202	&	EIC3200	(68:91.9%, 0.971)				
EIC4201	<=	EIC3205	&	EIC3202	&	EIC3200	(68:91.9%, 0.971)		
EIC4201	<=	EIC3205	&	EIC3202	(68:91.9%, 0.971)				
EIC4201	<=	EIC3105	(72:97.3%, 0.931)						
EIC4201	<=	EIC3205	&	EIC3200	(70:94.6%, 0.957)				
EIC4201	<=	EIC3200	(70:94.6%, 0.957)						
EIC4201	<=	EIC3105	&	EIC3205	&	EIC3200	(70:94.6%, 0.957)		
EIC4201	<=	EIC3105	&	EIC3200	(70:94.6%, 0.957)				
EIC4201	<=	EIC3205	(71:95.9%, 0.944)						
EIC4201	<=	EIC3105	&	EIC3205	(71:95.9%, 0.944)				

**Figura III.85 - Regras de associação: disciplina EIC4201 (anual) após o filtro**









Melhor suporte: 72:97.3% (conf=84.7%)	
Melhor confiança: 92.3% (sup=65:87.8%)	
EIC4205	<= EIC3103 & EIC3200 & EIC3201 & EIC3102 (65:87.8%, 0.923)
EIC4205	<= EIC3103 & EIC3204 & EIC3200 & EIC3201 (65:87.8%, 0.923)
EIC4205	<= EIC3105 & EIC3100 & EIC3201 & EIC3102 (65:87.8%, 0.923)
EIC4205	<= EIC3105 & EIC3103 & EIC3201 & EIC3102 (65:87.8%, 0.923)
EIC4205	<= EIC3105 & EIC3103 & EIC3204 & EIC3201 (65:87.8%, 0.923)
EIC4205	<= EIC3105 & EIC3200 & EIC3201 & EIC3102 (65:87.8%, 0.923)
EIC4205	<= EIC3105 & EIC3201 & EIC3102 (65:87.8%, 0.923)
EIC4205	<= EIC3105 & EIC3204 & EIC3201 & EIC3102 (65:87.8%, 0.923)
EIC4205	<= EIC3105 & EIC3205 & EIC3201 & EIC3102 (65:87.8%, 0.923)
EIC4205	<= EIC3200 & EIC3100 & EIC3201 & EIC3102 (65:87.8%, 0.923)
EIC4205	<= EIC3200 & EIC3201 & EIC3102 (65:87.8%, 0.923)
EIC4205	<= EIC3204 & EIC3200 & EIC3201 & EIC3102 (65:87.8%, 0.923)
EIC4205	<= EIC3205 & EIC3100 & EIC3201 & EIC3102 (65:87.8%, 0.923)
EIC4205	<= EIC3205 & EIC3103 & EIC3201 & EIC3102 (65:87.8%, 0.923)
EIC4205	<= EIC3205 & EIC3103 & EIC3204 & EIC3201 (65:87.8%, 0.923)
EIC4205	<= EIC3205 & EIC3200 & EIC3201 & EIC3102 (65:87.8%, 0.923)
EIC4205	<= EIC3205 & EIC3201 & EIC3102 (65:87.8%, 0.923)
EIC4205	<= EIC3205 & EIC3204 & EIC3201 & EIC3102 (65:87.8%, 0.923)
EIC4205	<= EIC3105 (72:97.3%, 0.847)
EIC4205	<= EIC3205 & EIC3200 & EIC3102 (67:90.5%, 0.91)
EIC4205	<= EIC3205 & EIC3103 & EIC3200 & EIC3102 (67:90.5%, 0.91)
EIC4205	<= EIC3205 & EIC3103 & EIC3102 (67:90.5%, 0.91)
EIC4205	<= EIC3205 & EIC3102 (67:90.5%, 0.91)
EIC4205	<= EIC3200 & EIC3102 (67:90.5%, 0.91)
EIC4205	<= EIC3105 & EIC3205 & EIC3200 & EIC3102 (67:90.5%, 0.91)
EIC4205	<= EIC3105 & EIC3205 & EIC3103 & EIC3102 (67:90.5%, 0.91)
EIC4205	<= EIC3105 & EIC3205 & EIC3102 (67:90.5%, 0.91)
EIC4205	<= EIC3105 & EIC3200 & EIC3102 (67:90.5%, 0.91)
EIC4205	<= EIC3105 & EIC3103 & EIC3200 & EIC3102 (67:90.5%, 0.91)
EIC4205	<= EIC3105 & EIC3103 & EIC3102 (67:90.5%, 0.91)
EIC4205	<= EIC3105 & EIC3102 (67:90.5%, 0.91)
EIC4205	<= EIC3103 & EIC3200 & EIC3102 (67:90.5%, 0.91)
EIC4205	<= EIC3205 (71:95.9%, 0.859)
EIC4205	<= EIC3105 & EIC3205 (71:95.9%, 0.859)
EIC4205	<= EIC3103 (71:95.9%, 0.859)
EIC4205	<= EIC3102 (69:93.2%, 0.884)
EIC4205	<= EIC3103 & EIC3102 (69:93.2%, 0.884)
EIC4205	<= EIC3103 & EIC3200 (69:93.2%, 0.884)
EIC4205	<= EIC3105 & EIC3103 & EIC3200 (69:93.2%, 0.884)
EIC4205	<= EIC3105 & EIC3103 (69:93.2%, 0.884)
EIC4205	<= EIC3105 & EIC3205 & EIC3103 & EIC3200 (69:93.2%, 0.884)
EIC4205	<= EIC3105 & EIC3205 & EIC3103 (69:93.2%, 0.884)
EIC4205	<= EIC3205 & EIC3103 & EIC3200 (69:93.2%, 0.884)
EIC4205	<= EIC3205 & EIC3103 (69:93.2%, 0.884)
EIC4205	<= EIC3105 & EIC3200 (70:94.6%, 0.871)
EIC4205	<= EIC3105 & EIC3205 & EIC3200 (70:94.6%, 0.871)
EIC4205	<= EIC3200 (70:94.6%, 0.871)
EIC4205	<= EIC3205 & EIC3200 (70:94.6%, 0.871)

**Figura III.90 - Regras de associação: disciplina EIC4205 (anual) após o filtro**







```

EIC5101 <= EIC4102 & EIC4103 & EIC4101 & EIC4104 (29:87.9%, 0.931)
EIC5101 <= EIC4102 & EIC4101 & EIC4104 (29:87.9%, 0.931)
EIC5101 <= EIC4102 & EIC4100 & EIC4101 & EIC4104 (29:87.9%, 0.931)
EIC5101 <= EIC4101 & EIC4104 (29:87.9%, 0.931)
EIC5101 <= EIC4100 & EIC4101 & EIC4104 (29:87.9%, 0.931)
EIC5101 <= EIC4105 & EIC4103 & EIC4101 (32:97.0%, 0.875)
EIC5101 <= EIC4105 & EIC4103 & EIC4100 & EIC4101 (32:97.0%, 0.875)
EIC5101 <= EIC4105 & EIC4101 (32:97.0%, 0.875)
EIC5101 <= EIC4105 & EIC4100 & EIC4101 (32:97.0%, 0.875)
EIC5101 <= EIC4103 & EIC4101 (32:97.0%, 0.875)
EIC5101 <= EIC4103 & EIC4100 & EIC4101 (32:97.0%, 0.875)
EIC5101 <= EIC4102 & EIC4105 & EIC4103 & EIC4101 (32:97.0%, 0.875)
EIC5101 <= EIC4102 & EIC4105 & EIC4101 (32:97.0%, 0.875)
EIC5101 <= EIC4102 & EIC4105 & EIC4100 & EIC4101 (32:97.0%, 0.875)
EIC5101 <= EIC4102 & EIC4103 & EIC4101 (32:97.0%, 0.875)
EIC5101 <= EIC4102 & EIC4103 & EIC4100 & EIC4101 (32:97.0%, 0.875)
EIC5101 <= EIC4102 & EIC4101 (32:97.0%, 0.875)
EIC5101 <= EIC4102 & EIC4100 & EIC4101 (32:97.0%, 0.875)
EIC5101 <= EIC4101 (32:97.0%, 0.875)
EIC5101 <= EIC4100 & EIC4101 (32:97.0%, 0.875)
EIC5101 <= EIC4100 & EIC4104 (30:90.9%, 0.9)
EIC5101 <= EIC4102 & EIC4100 & EIC4104 (30:90.9%, 0.9)
EIC5101 <= EIC4102 & EIC4103 & EIC4100 & EIC4104 (30:90.9%, 0.9)
EIC5101 <= EIC4102 & EIC4103 & EIC4104 (30:90.9%, 0.9)
EIC5101 <= EIC4102 & EIC4104 (30:90.9%, 0.9)
EIC5101 <= EIC4102 & EIC4105 & EIC4100 & EIC4104 (30:90.9%, 0.9)
EIC5101 <= EIC4102 & EIC4105 & EIC4103 & EIC4104 (30:90.9%, 0.9)
EIC5101 <= EIC4102 & EIC4105 & EIC4104 (30:90.9%, 0.9)
EIC5101 <= EIC4103 & EIC4100 & EIC4104 (30:90.9%, 0.9)
EIC5101 <= EIC4103 & EIC4104 (30:90.9%, 0.9)
EIC5101 <= EIC4104 (30:90.9%, 0.9)
EIC5101 <= EIC4105 & EIC4100 & EIC4104 (30:90.9%, 0.9)
EIC5101 <= EIC4105 & EIC4103 & EIC4100 & EIC4104 (30:90.9%, 0.9)
EIC5101 <= EIC4105 & EIC4103 & EIC4104 (30:90.9%, 0.9)
EIC5101 <= EIC4105 & EIC4104 (30:90.9%, 0.9)

```

**Figura III.93 - Regras de associação: disciplina EIC5101 (anual) após o filtro (cont.)**





```

Melhor suporte: 16:48.5% (conf=81.2%)
Melhor confiança: 86.7% (sup=15:45.5%)

EIC5104 <= EIC4101 & EIC4104 & EIC4200 & EIC4204 (15:45.5%, 0.867)
EIC5104 <= EIC4100 & EIC4101 & EIC4104 & EIC4204 (16:48.5%, 0.812)
EIC5104 <= EIC4100 & EIC4104 & EIC4200 & EIC4204 (16:48.5%, 0.812)
EIC5104 <= EIC4101 & EIC4104 & EIC4204 (16:48.5%, 0.812)
EIC5104 <= EIC4101 & EIC4201 & EIC4104 & EIC4204 (16:48.5%, 0.812)
EIC5104 <= EIC4101 & EIC4203 & EIC4104 & EIC4204 (16:48.5%, 0.812)
EIC5104 <= EIC4101 & EIC4205 & EIC4104 & EIC4204 (16:48.5%, 0.812)
EIC5104 <= EIC4101 & EIC4205 & EIC4200 & EIC4204 (16:48.5%, 0.812)
EIC5104 <= EIC4102 & EIC4101 & EIC4104 & EIC4204 (16:48.5%, 0.812)
EIC5104 <= EIC4102 & EIC4104 & EIC4200 & EIC4204 (16:48.5%, 0.812)
EIC5104 <= EIC4103 & EIC4101 & EIC4104 & EIC4204 (16:48.5%, 0.812)
EIC5104 <= EIC4103 & EIC4104 & EIC4200 & EIC4204 (16:48.5%, 0.812)
EIC5104 <= EIC4104 & EIC4200 & EIC4204 (16:48.5%, 0.812)
EIC5104 <= EIC4105 & EIC4101 & EIC4104 & EIC4204 (16:48.5%, 0.812)
EIC5104 <= EIC4105 & EIC4104 & EIC4200 & EIC4204 (16:48.5%, 0.812)
EIC5104 <= EIC4201 & EIC4104 & EIC4200 & EIC4204 (16:48.5%, 0.812)
EIC5104 <= EIC4202 & EIC4101 & EIC4104 & EIC4204 (16:48.5%, 0.812)
EIC5104 <= EIC4202 & EIC4104 & EIC4200 & EIC4204 (16:48.5%, 0.812)
EIC5104 <= EIC4203 & EIC4104 & EIC4200 & EIC4204 (16:48.5%, 0.812)
EIC5104 <= EIC4205 & EIC4104 & EIC4200 & EIC4204 (16:48.5%, 0.812)

```

**Figura III.96 - Regras de associação: disciplina EIC5104 (anual) após o filtro**





EIC5201	⇐	EIC4103	&	EIC4104	(30:90.9%, 0.967)				
EIC5201	⇐	EIC4103	&	EIC4100	&	EIC4205	(30:90.9%, 0.967)		
EIC5201	⇐	EIC4103	&	EIC4100	&	EIC4203	&	EIC4205	(30:90.9%, 0.967)
EIC5201	⇐	EIC4103	&	EIC4100	&	EIC4202	&	EIC4205	(30:90.9%, 0.967)
EIC5201	⇐	EIC4103	&	EIC4100	&	EIC4201	&	EIC4205	(30:90.9%, 0.967)
EIC5201	⇐	EIC4103	&	EIC4100	&	EIC4104	(30:90.9%, 0.967)		
EIC5201	⇐	EIC4102	&	EIC4205	(30:90.9%, 0.967)				
EIC5201	⇐	EIC4102	&	EIC4203	&	EIC4205	(30:90.9%, 0.967)		
EIC5201	⇐	EIC4102	&	EIC4202	&	EIC4205	(30:90.9%, 0.967)		
EIC5201	⇐	EIC4102	&	EIC4202	&	EIC4203	&	EIC4205	(30:90.9%, 0.967)
EIC5201	⇐	EIC4102	&	EIC4202	&	EIC4201	&	EIC4205	(30:90.9%, 0.967)
EIC5201	⇐	EIC4102	&	EIC4201	&	EIC4205	(30:90.9%, 0.967)		
EIC5201	⇐	EIC4102	&	EIC4201	&	EIC4203	&	EIC4205	(30:90.9%, 0.967)
EIC5201	⇐	EIC4102	&	EIC4105	&	EIC4205	(30:90.9%, 0.967)		
EIC5201	⇐	EIC4102	&	EIC4105	&	EIC4203	&	EIC4205	(30:90.9%, 0.967)
EIC5201	⇐	EIC4102	&	EIC4105	&	EIC4202	&	EIC4205	(30:90.9%, 0.967)
EIC5201	⇐	EIC4102	&	EIC4105	&	EIC4201	&	EIC4205	(30:90.9%, 0.967)
EIC5201	⇐	EIC4102	&	EIC4105	&	EIC4104	(30:90.9%, 0.967)		
EIC5201	⇐	EIC4102	&	EIC4105	&	EIC4103	&	EIC4205	(30:90.9%, 0.967)
EIC5201	⇐	EIC4102	&	EIC4105	&	EIC4103	&	EIC4104	(30:90.9%, 0.967)
EIC5201	⇐	EIC4102	&	EIC4105	&	EIC4100	&	EIC4205	(30:90.9%, 0.967)
EIC5201	⇐	EIC4102	&	EIC4105	&	EIC4100	&	EIC4104	(30:90.9%, 0.967)
EIC5201	⇐	EIC4102	&	EIC4104	(30:90.9%, 0.967)				
EIC5201	⇐	EIC4102	&	EIC4103	&	EIC4205	(30:90.9%, 0.967)		
EIC5201	⇐	EIC4102	&	EIC4103	&	EIC4203	&	EIC4205	(30:90.9%, 0.967)
EIC5201	⇐	EIC4102	&	EIC4103	&	EIC4202	&	EIC4205	(30:90.9%, 0.967)
EIC5201	⇐	EIC4102	&	EIC4103	&	EIC4201	&	EIC4205	(30:90.9%, 0.967)
EIC5201	⇐	EIC4102	&	EIC4103	&	EIC4104	(30:90.9%, 0.967)		
EIC5201	⇐	EIC4102	&	EIC4103	&	EIC4100	&	EIC4205	(30:90.9%, 0.967)
EIC5201	⇐	EIC4102	&	EIC4103	&	EIC4100	&	EIC4104	(30:90.9%, 0.967)
EIC5201	⇐	EIC4102	&	EIC4100	&	EIC4205	(30:90.9%, 0.967)		
EIC5201	⇐	EIC4102	&	EIC4100	&	EIC4203	&	EIC4205	(30:90.9%, 0.967)
EIC5201	⇐	EIC4102	&	EIC4100	&	EIC4202	&	EIC4205	(30:90.9%, 0.967)
EIC5201	⇐	EIC4102	&	EIC4100	&	EIC4201	&	EIC4205	(30:90.9%, 0.967)
EIC5201	⇐	EIC4102	&	EIC4100	&	EIC4104	(30:90.9%, 0.967)		
EIC5201	⇐	EIC4100	&	EIC4205	(30:90.9%, 0.967)				
EIC5201	⇐	EIC4100	&	EIC4203	&	EIC4205	(30:90.9%, 0.967)		
EIC5201	⇐	EIC4100	&	EIC4202	&	EIC4205	(30:90.9%, 0.967)		
EIC5201	⇐	EIC4100	&	EIC4202	&	EIC4203	&	EIC4205	(30:90.9%, 0.967)
EIC5201	⇐	EIC4100	&	EIC4202	&	EIC4201	&	EIC4205	(30:90.9%, 0.967)
EIC5201	⇐	EIC4100	&	EIC4201	&	EIC4205	(30:90.9%, 0.967)		
EIC5201	⇐	EIC4100	&	EIC4201	&	EIC4203	&	EIC4205	(30:90.9%, 0.967)
EIC5201	⇐	EIC4100	&	EIC4104	(30:90.9%, 0.967)				

Figura III.99 - Regras de associação: disciplina EIC5201 (anual) após o filtro (cont.)

NOTAS	
Nº	1824
Pa de aquisição	OFERTA
	Do AUTOR
Proceder	



