

Auditoria na Banca Utilizando Redes Bayesianas

Por:

Maria Margarida Oliveira Camarinha

Orientada por:

Professor Doutor João Manuel Portela da Gama

Tese de Mestrado em
Análise de Dados e Sistemas de Apoio à Decisão

Dezembro, 2009

NOTA BIOGRÁFICA

MARGARIDA CAMARINHA

Licenciatura em Economia na Faculdade de Economia do Porto.
Experiência de 10 anos no sector da banca, 3 dos quais em Auditoria Interna.
Interesse académico e profissional nas áreas de Inteligência Artificial e Auditoria.

AGRADECIMENTOS

Ao meu orientador, Professor Doutor João Manuel Portela da Gama, pela
motivação, pelo “*anda prá frente*”, por tudo.
À minha irmã Isabel, e aos meus amigos, Sandra, João, Manela, Camila, Ramiro,
João, Hugo, Manel e Paulo porque, em tão diferentes moldes, me empurraram
para a conclusão desta tese.
À minha mãe e ao meu pai.
Aos meus colegas de trabalho.

RESUMO

O propósito deste trabalho foi analisar a aplicabilidade de Redes Bayesianas Normais e Dinâmicas no contexto de Auditoria Interna de uma Instituição Bancária, como método para a obtenção de conhecimento, a partir de bases de dados, sobre o impacto produzido pelas recomendações efectuadas no decurso dos trabalhos efectuados. Inicialmente apresentaram-se os conceitos teóricos sobre o tema, seguindo-se a análise de casos práticos. Os resultados experimentais mostraram que a estratégia levada a cabo pela Auditoria parece estar a produzir resultados satisfatórios considerando as variáveis em análise no estudo de casos.

ABSTRACT

The purpose of this study was to analyze the applicability of Normal and Dynamic Bayesian Networks in the context of Internal Audit in a Banking Institution, as a method for getting knowledge from a dataset, on the impact of the recommendations made in the course of the works done. Initially it was presented the theoretical concepts on the subject, followed by the analysis of practical cases. The experimental results showed that the strategy undertaken by the Audit seems to produce satisfactory results considering the variables under account in the case studies.

ÍNDICE

ÍNDICE DE TABELAS.....	7
ÍNDICE DE FIGURAS.....	8
ÍNDICE DE ALGORITMOS.....	9
1. ENQUADRAMENTO.....	1
2. REDES BAYESIANAS.....	3
2.1. PROBABILIDADE BAYESIANA.....	3
2.2. REDES BAYESIANAS.....	6
2.2.1. Definição.....	6
2.2.2. Construção.....	8
2.3. INFERÊNCIA EM REDES BAYESIANAS.....	10
2.3.1. Definição.....	11
2.3.2. Algoritmos.....	11
2.3.3. Aprendizagem.....	18
3. MODELOS DINÂMICOS.....	22
3.1. RACIOCÍNIO PROBABILÍSTICO TEMPORAL.....	22
3.2. CLASSIFICAÇÃO DOS MODELOS DINÂMICOS.....	23
3.2.1. Premissa de Markov.....	23
3.2.2. HMMs.....	26
3.3. REDES BAYESIANAS DINÂMICAS.....	26
3.3.1. Definição.....	27
3.3.2. Construção.....	28
3.3.3. Premissas adoptadas.....	29
3.3.4. Inferência exacta em Redes Bayesianas Dinâmicas.....	30
3.3.4.1. Filtering.....	30
3.3.4.2. Previsão.....	32
3.3.4.3. Smoothing.....	32
3.3.4.4. Sequência mais provável.....	34
3.3.5. Inferência aproximada nas Redes Bayesianas Dinâmicas.....	34
4. TRABALHOS RELACIONADOS.....	36
5. AUDITORIA NA BANCA – APLICAÇÕES.....	40
5.1. O SOFTWARE – GENIE 2.0.....	41
5.1.1. Vantagens e Inconvenientes do GeNIe 2.0.....	43
5.2. AUDITORIA NA BANCA USANDO REDES BAYESIANAS NORMAIS.....	44
5.2.1. Estudo de um caso teórico – Falhas de Crédito em contas com indicadores de risco ...	44
5.2.2. Estudo de um caso prático – Impacto da Auditoria em Operações de Crédito.....	50
5.2.2.1. Tratamento dos dados.....	51

5.2.2.2. Caracterização das variáveis	51
5.2.2.3. Métodos e algoritmos de aprendizagem utilizados.....	52
5.2.2.4. Interface gráfico do modelo	52
5.2.2.5. Descrição da topologia / arquitectura da rede automática	53
5.2.2.6. Topologia da rede com intervenção manual	55
5.2.2.7. Simulação, propagação e discussão dos resultados	56
5.3. <i>AUDITORIA NA BANCA USANDO REDES BAYESIANAS DINÂMICAS</i>	57
5.3.1. Estudo de um caso prático – Impacto da Auditoria em Operações de Crédito.....	58
5.3.1.1. Tratamento dos dados.....	59
5.3.1.2. Caracterização das variáveis	60
5.3.1.3. Metodologia utilizada.....	60
5.3.1.4. Interface gráfico do modelo	61
5.3.1.5. Simulação, propagação e discussão dos resultados	68
6. <i>CONSIDERAÇÕES FINAIS E TRABALHO FUTURO</i>	74

ÍNDICE DE TABELAS

Tabela 1 – $P(F)$, $P(I)$ e $P(S)$	10
Tabela 2 – $P(G F)$	10
Tabela 3 - $P(J F,I,S)$	10
Tabela 4 – Classificação dos atributos	24
Tabela 5 – Modelo de matriz de transição de estados a ser construída	25
Tabela 6 – Variáveis do modelo	45
Tabela 7 – Caracterização das variáveis do caso em estudo	52
Tabela 8 – Caracterização das variáveis em estudo	60

ÍNDICE DE FIGURAS

Figura 1 – Três exemplos de cálculo de probabilidades em diferentes estruturas de redes Bayesianas	7
Figura 2 – Tipos de estruturas de grafos	8
Figura 3 - Rede Bayesianiana para detecção de problemas de fraude	9
Figura 4 – Tipos de Inferência com Redes Bayesianas	11
Figura 5 – Classificação dos Algoritmos	12
Figura 6 – Exemplo da estrutura de uma <i>polytree</i>	13
Figura 7 - Exemplo de estrutura sujeita à utilização de um algoritmo de <i>clustering</i>	14
Figura 8 – Rede Bayesianiana	24
Figura 9 – Exemplo de uma Rede Bayesianiana Dinâmica	28
Figura 10 – Rede Bayesianiana	28
Figura 11 – Diferentes Topologias de Redes Bayesianas Dinâmicas	29
Figura 12 – Interface gráfico do GeNIe 2.0	41
Figura 13 - Exemplo de uma Rede Bayesianiana no GeNIe	42
Figura 14 – Propriedades dos nós e da rede no GeNIe	42
Figura 15 – Tabelas de Probabilidades no GeNIe	43
Figura 16 – Importação dos dados	45
Figura 17 – Aprendizagem da estrutura da rede no GeNIe	46
Figura 18 – Aplicação de vários métodos de aprendizagem no GeNIe	46
Figura 19 – Algoritmos de propagação dos dados no GeNIe	47
Figura 20 – Alertas no GeNIe	48
Figura 21 – Gráficos de Barras no GeNIe	48
Figura 22 – Simulações no GeNIe	49
Figura 23 – Rede Bayesianiana gerada pelo GeNIe	53
Figura 24 – Tabela de Probabilidades da variável <i>Situacao_Actuall</i>	53
Figura 25 – Nova rede com intervenção do perito	55
Figura 26 – Tabela de Probabilidades da variável <i>Situacao_Actual</i> após a intervenção do perito	55
Figura 27 – Actualização Bayesianiana para novas evidências da variável <i>Analizado</i>	56
Figura 28 - Actualização Bayesianiana para novas evidências da variável <i>Recomend</i>	57
Figura 29 – Importação de dados pelo GeNIe	61
Figura 30- Janela <i>Learn New Network</i>	62
Figura 31 – Janela <i>Knowledge Editor</i>	62
Figura 32 – Tabela de probabilidades da variável <i>rec</i>	63
Figura 33 – Tabela de probabilidades da variável <i>rec_1</i>	63
Figura 34 – Janela <i>Learn New Network</i>	64
Figura 35 – Menu <i>Enable Temporal Plate</i>	65
Figura 36 – Desenho do arco temporal	65
Figura 37 – Definição do número de <i>fatias de tempo</i>	66
Figura 38 – Rede Bayesianiana Dinâmica	66
Figura 39 – Tabela de probabilidades para $t=0$	67
Figura 40 – Tabela de probabilidades para $t=1$	67
Figura 41 – Nova tabela de probabilidades para $t=1$	67
Figura 42 – Probabilidades temporais	67
Figura 43 – Simulação de evidências na Rede Bayesianiana Dinâmica	68
Figura 44 – Análise de situações extremas	69
Figura 45 – <i>Unroll</i> da rede	69
Figura 46 – Outra forma de representar a Rede Bayesianiana Dinâmica	70
Figura 47 – Simulação de cenários utilizando o <i>Clustering</i>	71
Figura 48 – Simulação de cenários utilizando o <i>Likelihood Sampling</i>	72

ÍNDICE DE ALGORITMOS

Algoritmo 1 - <i>Polytrees</i>	13
Algoritmo 2 – Árvore de Junção.....	14
Algoritmo 3 – <i>Logic Sampling</i>	15
Algoritmo 4 – <i>Likelihood Weighting</i>	16
Algoritmo 5 – <i>Self Importance Sampling</i>	17
Algoritmo 6 – AIS Sampling	18
Algoritmo 7 – EPIS - <i>Sampling</i>	18

1. ENQUADRAMENTO

A auditoria interna realizada a uma Instituição Bancária, no âmbito do presente estudo, centra-se na análise e avaliação dos procedimentos implementados, desde a recolha de informação de base até ao apuramento final dos resultados e respectiva divulgação. Deve, do mesmo modo, conferir a metodologia e o processo de monitorização utilizado pelo banco para quantificar o grau de exposição a riscos.

No processo de identificação e avaliação do risco, os auditores internos desenvolvem um discernimento importante sobre as actividades, o que pode resultar numa oportunidade de aperfeiçoamento operacional extremamente benéfica para a empresa, e o que pode concretizar-se na forma de diagramas que organizem o conhecimento através de um mapa de causas e efeitos.

Este trabalho propõe-se a aumentar a percepção dos benefícios decorrentes dos trabalhos de auditoria interna quanto ao acompanhamento das actividades realizadas pelas agências bancárias, de forma a auxiliar e a alcançar um dos objectivos da sua visão estratégica, que não é mais do que o passar uma imagem positiva e de actuação preventiva.

Para que esse objectivo seja alcançado, este estudo serve-se da aplicação de uma Rede Bayesiana, como forma de representação de conhecimento, capaz de gerar automaticamente resultados mesmo em situações de inexistência de algumas peças de informação.

Salientar a importância da auditoria bancária como instrumento de credibilidade, bem como perceber o nível de eficácia com que actua, constituíram a base motivacional desta dissertação. O reconhecimento de que este é um tema abrangente, e de que uma abordagem generalista reduziria a utilidade do modelo, fez concentrar a análise num conjunto específico de operações relacionadas com a área de crédito. Em que medida um modelo pode capturar a assertividade de um departamento é matéria que nem sempre é trivial, principalmente quando existe um conjunto de factores de incerteza alheios à actuação do próprio departamento. Este contexto é pois propício à utilização

de Redes Bayesianas já que estas representam domínios de conhecimento com determinado grau de incerteza e para além disso são capazes de modelar o conhecimento de um perito nesse domínio de uma forma intuitiva. O reconhecimento da dificuldade do tema e a ausência de publicações que o relacionem com Redes Bayesianas, constituíram também uma motivação deste trabalho.

Este estudo introduziu as Redes Bayesianas como um dos principais métodos para modelar a incerteza e com grande utilidade para representar processos probabilísticos, através de exposições teóricas e análise de casos. A introdução do raciocínio probabilístico temporal tornou possível ampliar o poder de interpretação das Redes Bayesianas – agora Redes Bayesianas Dinâmicas – adequando-as mais ainda a aplicações do mundo real.

Numa primeira abordagem, a investigação incidiu em técnicas de aprendizagem ou actualização de variáveis de uma forma automática, baseada em dados que evidenciem experiências passadas – montagem da rede a partir dos dados – seguiu-se uma outra abordagem que incorpora a intervenção de um perito no desenho da estrutura da rede. Esta interacção entre automatismo e experiência de negócio intenta quanto ao aproveitamento dos recursos já existentes na organização bem como à obtenção de ganhos de desempenho através do estudo da utilização de uma ferramenta capaz de evidenciar e até melhorar o processo de auditoria.

Finalmente, e porque as eventuais repercussões de uma auditoria a qualquer área não assumem carácter imediato, foi introduzido o âmbito temporal na análise, desta vez utilizando um indicador mais genérico e compensando o mesmo com uma demonstração mais simplificada de uma Rede Bayesiana Dinâmica.

2. REDES BAYESIANAS

A linha de fundamentos desenvolvida no contexto deste trabalho compreende as abordagens teóricas no que respeita a Redes Bayesianas, os conceitos e temáticas que a estas e ao seu funcionamento se encontram associados, a sua extensão para a definição de Redes Bayesianas Dinâmicas, bem como o seu âmbito de actuação e aplicabilidade empírica.

Primeiro rememora-se o conceito de Probabilidade, fundamental para uma compreensão basilar, segue-se a exploração do conceito de Redes Bayesianas e estende-se, no próximo capítulo, a abordagem às Redes Bayesianas Dinâmicas.

2.1. PROBABILIDADE BAYESIANA

Thomas Bayes (1702-1761) foi um filósofo, matemático e pastor presbiteriano inglês considerado um dos pais do cálculo da probabilidade. Publicou um trabalho onde defendeu o cálculo diferencial de Newton contra as críticas de Berkeley. O resultado do seu trabalho passou a ser conhecido como o Teorema de Bayes (ou Regra de Bayes ou Lei de Bayes), uma fórmula que se transformou numa lei fundamental da Matemática. Este Teorema faz parte do ensaio *Essay Towards Solving a Problem in the Doctrine of Chances* publicado a título póstumo em 1764 na *Philosophical Transactions of the Royal Society of London* (Wikipedia).

As incertezas existentes no mundo real são tratadas pela mente humana de maneira que seja possível alguma tomada de decisão baseada em factos que não representem uma verdade absoluta (Cowell, *et al.*, 1999). Segundo (Charniak, 1991) a principal vantagem de raciocínio probabilístico sobre raciocínio lógico é o facto de que os agentes podem tomar decisões racionais mesmo quando não existe informação suficiente para se provar que uma acção funcionará. Na interpretação Bayesiana, a transformação de incerteza, como forma de avaliação crítica sobre a ocorrência de um evento incerto, em números, passa por perceber e respeitar o cálculo de probabilidades.

Em termos formais, sendo Ω o espaço amostral, conjunto de todos os resultados possíveis, e A um evento, subconjunto desse espaço, o que se deseja é atribuir a esse

evento um número não negativo chamado probabilidade. A notação adoptada é $P(A)$, para “a probabilidade de ocorrência do evento A ” e os axiomas importantes associados ao cálculo de probabilidades são:

- (i) $P(\Omega) = 1 \Leftrightarrow$ isto é, a probabilidade de ocorrência de um evento definido como sendo o espaço amostral Ω é sempre igual a 1.
- (ii) $0 \leq P(A) \leq 1 \Leftrightarrow$ sendo A um evento qualquer, os seus valores probabilísticos não podem ser menores que zero nem maiores que um, sendo igual a 1 a probabilidade do evento certo e zero o valor atribuído ao evento impossível. Os valores intermédios correspondem a eventos possíveis ou incertos.
- (iii) $P(A \text{ ou } B) = P(A) + P(B) \Leftrightarrow$ a probabilidade de qualquer evento ocorrer de entre um número n de eventos alternativos mutuamente exclusivos, A e B neste caso, é igual à soma das probabilidades dos eventos individuais.
- (iv) $P(A \text{ e } B) = P(A|B) \times P(B) \Leftrightarrow$ a probabilidade de dois eventos A e B ocorrerem é igual à multiplicação da probabilidade condicional $P(A|B)$ – lê-se “probabilidade de A dado B ” – pela probabilidade de B .
- (v) $P(A) = P(A|B) \times P(B) + P(A|\sim B) \times P(\sim B) \Leftrightarrow$ probabilidade total de um evento A é igual à probabilidade de ocorrência desse evento dada a ocorrência de dois eventos disjuntos, isto é, neste caso dada a ocorrência de B ou não B multiplicados pelas respectivas probabilidades *a priori*.

Generalizando, seja B_n uma sequência de eventos disjuntos tais que

$$\sum_n P(B_n) = 1, \text{ então, para qualquer evento } A, P(A) = \sum_n P(A|B_n)P(B_n)$$

Pela regra do produto e assumindo, por comutação, que $P(A \text{ e } B) = P(B \text{ e } A)$ então:

$$P(A \text{ e } B) = P(B|A) \times P(A) \quad \text{e} \quad P(A \text{ e } B) = P(A|B) \times P(B)$$

Igualando e dividindo estas duas últimas equações por $P(B)$, verifica-se que a probabilidade condicional pode ser definida em termos da probabilidade *a priori* e denotada pela equação que evidencia o Teorema de Bayes:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B|A)P(A)}{P(B)}$$

O modelo Bayesiano interpreta pois a probabilidade condicional definida em termos de probabilidade *a priori*, onde o grau de crença de um agente causa um efeito noutro agente. Nesta fórmula, a probabilidade do evento A ocorrer SE ocorrer o evento B – $P(A|B)$ – é dada por três factores: a probabilidade condicional, de B dado A ; a probabilidade *a priori* de B ; e a probabilidade *a priori* de A .

O exemplo a seguir exposto foi inspirado em (Johnstone, 1994) e ilustra, de forma simplificada, o uso do Teorema de Bayes para o cálculo da probabilidade de um evento.

Considerando:

B: A realização de uma operação não conforme

A1: O auditor detecta o erro sendo a $P(A_1) = 0.8$

A2: O auditor não detecta o erro sendo a $P(A_2) = 0.2$

Sabe-se ainda que a $P(B|A_1) = 0.1$ e $P(B|A_2) = 0.2$

Assim, de acordo com (v), é possível obter a probabilidade de ocorrência de uma inconformidade (o mesmo que operação não conforme), ou seja:

$$P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2) = 0.1 \times 0.8 + 0.2 \times 0.2 = 0.12$$

Aplicando o Teorema de Bayes, obtém-se:

$$P(A_1|B) = \frac{P(B|A_1)P(A_1)}{P(B)} = \frac{0.1 \times 0.8}{0.12} = \frac{2}{3}$$

$$P(A_2|B) = \frac{P(B|A_2)P(A_2)}{P(B)} = \frac{0.2 \times 0.2}{0.12} = \frac{1}{3} < P(A_1|B)$$

De onde se conclui que provavelmente o auditor irá detectar o erro de uma operação não conforme.

Introduziram-se desta forma os conceitos da teoria de probabilidade considerados relevantes para a percepção de base das Redes Bayesianas, que reflectem nos valores de probabilidade o que se poderá chamar a crença do especialista sobre o que ele espera que ocorra em determinadas situações.

2.2. REDES BAYESIANAS

As Redes Bayesianas oferecem uma estrutura intuitiva de representar o conhecimento, permitindo o raciocínio e a inferência quer em termos de diagnóstico, quer em termos de previsão. O conhecimento representado pode ser aprendido a partir de exemplos observados e/ou através de mecanismos de raciocínio de agentes vocacionados ou especialistas na matéria em causa.

A vantagem da sua utilização concentra-se no facto de permitir a representação e manipulação da incerteza com base na interpretação de conhecimento sem desprezar princípios matemáticos fundamentados. Constituem portanto um formalismo gráfico que explora o relacionamento de causalidade e independência condicional entre variáveis, com base na visão de Bayes da teoria de probabilidades.

2.2.1. Definição

Redes Bayesianas, também conhecidas por redes probabilísticas ou redes causais, são tipos específicos de redes de conhecimento, modelos que codificam os relacionamentos probabilísticos entre as variáveis que definem um determinado domínio (Russell e Norvig, 2003) e que são utilizadas para representar processos probabilísticos e causais. A sua representação é feita através de um grafo direccionado acíclico no qual os nós representam variáveis de um domínio e os arcos representam a dependência condicional ou informativa entre as variáveis. Para representação da intensidade da dependência, são utilizadas probabilidades, associadas a cada grupo de nós pais-filhos na rede (Pearl, 1988). A grande vantagem no uso deste tipo de estrutura está em conseguir representar incerteza de forma gráfica através de nós e grafos (Charniak, 1991).

Formalmente, as Redes Bayesianas condizem com a seguinte especificação:

- (i) Um conjunto de variáveis discretas $U = \{X_1, X_2, \dots, X_n\}$. Note-se que a natureza das variáveis pode ser discreta (categórica), por exemplo, com dois estados possíveis, verdadeiro e falso, sendo suas probabilidades condicionais facilmente representadas em tabelas de probabilidades. Ou podem ser de natureza contínua, neste caso representadas por funções de densidade ou então discretizadas;

- (ii) Um conjunto de arcos que ligam as variáveis;
- (iii) Não pode haver ciclos direccionados, as variáveis e os arcos formam um grafo acíclico direccionado (*DAG – Directed Acyclic Graph*);
- (iv) Cada variável tem um conjunto finito de estados mutuamente exclusivos ;
- (v) Para cada variável X com pais $pa(X) = \{Y_1, Y_2, \dots, Y_m\}$ está associada uma tabela $P(X|pa(X))$. Note-se que um arco de Y para X significa, nesta notação, que Y é pai de X . Se X não tem pais, a tabela será uma tabela de probabilidade *a priori* $P(X)$.
- (vi) A probabilidade conjunta das variáveis é o produto de todas as tabelas de probabilidade condicional especificadas na rede sendo $pa(X_i)$ o conjunto de pais de X_i :

$$P(U) = \prod_{i=1}^n P(X_i|pa(X_i))$$

Uma das características mais importantes das Redes Bayesianas é o facto de estas proporcionarem uma estrutura matemática sóbria capaz de modelar relações complexas entre variáveis ao mesmo tempo que permitem a visualização da representação gráfica dessas relações. A Figura 1 ilustra três exemplos simples de probabilidades associadas a diferentes relações existentes entre variáveis.

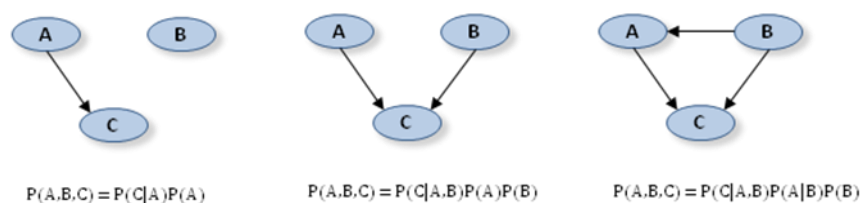


Figura 1 – Três exemplos de cálculo de probabilidades em diferentes estruturas de redes Bayesianas

A Figura 2 ilustra as diferentes estruturas de grafos, a topologia de uma Rede Bayesiana não se enquadra nos exemplos assinalados a cinzento:

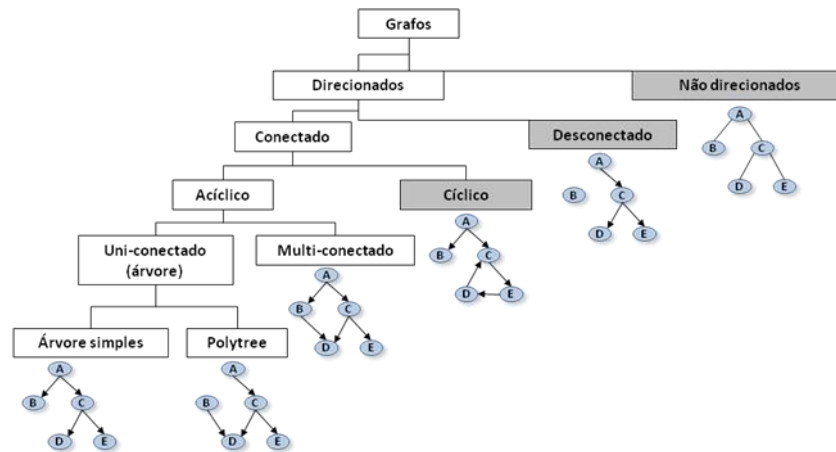


Figura 2 – Tipos de estruturas de grafos

Em resumo, a topologia de uma Rede Bayesiana representa um modelo probabilístico completo de um domínio, com a descrição das informações qualitativas (relações de dependências) e quantitativas (distribuições de probabilidades) e uma estrutura de controle para a inferência.

2.2.2. Construção

Segundo (Pearl, 1988), ao construir uma Rede Bayesiana nem sempre é possível identificar de forma numérica as dependências entre as variáveis, sendo na maioria das vezes essa tarefa realizada por meio do julgamento humano. Assim, para designar que uma determinada variável é descendente directa de outra, deve-se constatar se ela é causa directa ou influencia a outra variável directamente.

As Redes Bayesianas, no que respeita ao desenho da estrutura e ao cálculo das probabilidades, podem ser construídas:

- Manualmente, por peritos que intervenham na definição das probabilidades bem como das relações de dependência estabelecidas;
- Também pode suceder que haja intervenção manual de especialistas no que respeita à sua construção, mas as probabilidades serem obtidas a partir dos dados e vice-versa, ou um misto de ambos;
- Casos existem ainda em que estrutura e probabilidades são obtidas a partir dos dados através de um processo de aprendizagem (ver 2.3).

Dada uma distribuição P e um dado ordenamento de variáveis de U , existe um procedimento recursivo simples para construir (manualmente) uma Rede Bayesiana. E dada uma Rede Bayesiana, existe um procedimento simples para reconstruir $P(X_1, X_2, \dots, X_n)$. Na prática, o problema é definir as variáveis e suas interdependências (quase nunca se dispõe de P). Entretanto, é preciso identificar as relações directas. Assim, para a construção de uma rede importa seguir os seguintes passos:

- (i) Escolha das variáveis relevantes X_i que descrevem o domínio;
- (ii) Escolha de uma ordenação para as variáveis:
- (iii) Enquanto existirem variáveis (para $i = 1$ a n):
 - a. Adicionar um nó à rede para a variável seleccionada X_i ;
 - b. Definir os $pa(X_i)$ como o conjunto mínimo de nós já existentes na rede para os quais a propriedade de independência condicionada se verifique, isto é, seleccionar pais para $X_1, \dots, X_{i-1}$ tais que $P(X_i | pa(X_i)) = P(X_i | X_1, \dots, X_{i-1})$;
 - c. Definir a tabela de probabilidade condicionada para X_i .

A Figura 3 ilustra os conceitos acabados de teorizar. Neste exemplo (Heckerman, 1995), pretende-se estabelecer a influência da variável Cartão Fraudado (F), Idade (I) e Sexo (S) sobre as compras de Gasolina (G) e Jóias (J). O conjunto destas variáveis constitui as variáveis do modelo, sendo que os círculos representam os nós das variáveis da rede e os arcos ou setas representam o relacionamento causal entre as variáveis. A representação é intuitiva e evidencia que os cartões fraudados influenciam a compra de gasolina e que os cartões fraudados, a idade e o sexo influenciam a compra de jóias.

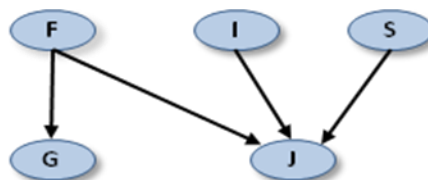


Figura 3 - Rede Bayesiana para detecção de problemas de fraude

As Tabelas 1, 2 e 3 descrevem a estrutura quantitativa da rede, ou seja, os valores representados pelas distribuições marginais ou condicionais do modelo, avaliando, em termos probabilísticos, as dependências que se enquadram na sua estrutura qualitativa.

Tabela 1 – $P(F)$, $P(I)$ e $P(S)$

$P(F=\text{Falso})$	$P(F=\text{Verdadeiro})$	$P(I \leq 30)$	$P(50 > I > 30)$	$P(I \geq 50)$	$P(S=\text{Masculino})$	$P(S=\text{Feminino})$
0.99999	0.00001	0.25	0.4	0.35	0.5	0.5

Tabela 2 – $P(G|F)$

F	$P(G=\text{Falso})$	$P(G=\text{Verdadeiro})$
Falso	0.99	0.01
Verdadeiro	0.8	0.2

Tabela 3 – $P(J|F,I,S)$

F	I	S	$P(J=\text{Falso})$	$P(J=\text{Verdadeiro})$
Falso	<30	Feminino	0.9995	0.0005
Verdadeiro	<30	Feminino	0.95	0.05
Falso	<30	Masculino	0.9999	0.0001
Verdadeiro	<30	Masculino	0.95	0.05
Falso	30-50	Feminino	0.998	0.002
Verdadeiro	30-50	Feminino	0.95	0.05
Falso	30-50	Masculino	0.9996	0.0004
Verdadeiro	30-50	Masculino	0.95	0.05
Falso	>50	Feminino	0.999	0.001
Verdadeiro	>50	Feminino	0.95	0.05
Falso	>50	Masculino	0.9998	0.0002
Verdadeiro	>50	Masculino	0.95	0.05

2.3. INFERÊNCIA EM REDES BAYESIANAS

Uma vez construída uma representação probabilística através do modelo de Redes Bayesianas, para a incerteza presente no relacionamento entre variáveis de um domínio de dados, uma das tarefas mais importantes consiste em obter estimativas de probabilidades dos eventos relacionados, à medida que novas informações ou evidências sejam conhecidas.

Inferências podem ser utilizadas sobre Redes Bayesianas para:

- Diagnóstico: dos efeitos para as causas, o que vai no sentido das relações definidas, por exemplo, qual a influência de cartões fraudados na compra de Gasolina (20%) e na compra de Jóias (5%);
- Prognóstico ou Inferência causal: das causas para os efeitos, por exemplo, dado que se compra gasolina e jóias, qual a probabilidade de isso acontecer com cartões fraudados (1%);

A Figura 4 ilustra estes dois tipos de inferência com base no exemplo do ponto anterior.

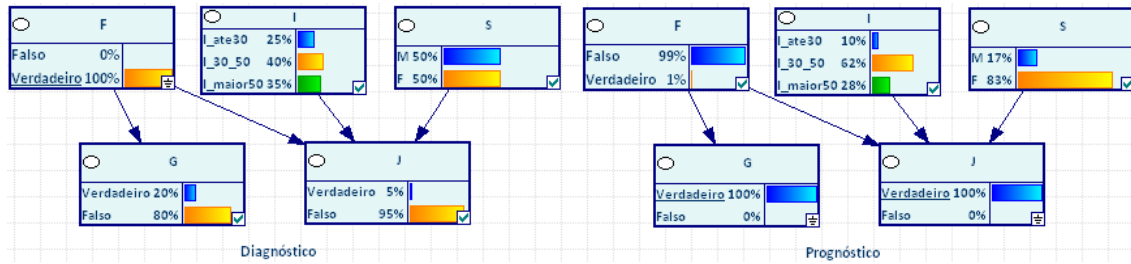


Figura 4 – Tipos de Inferência com Redes Bayesianas

2.3.1. Definição

A Inferência é uma das operações mais importantes no que respeita ao trabalho investido no desenvolvimento de Redes Bayesianas e certamente uma das mais difíceis. Inferência Bayesiana é um tipo de inferência estatística que descreve as incertezas sobre quantidades invisíveis de forma probabilística. Incertezas são modificadas periodicamente após observações de novos dados ou resultados. A operação que calibra a medida das incertezas é conhecida como operação Bayesiana e é baseada na fórmula de Bayes (Wikipedia).

A tarefa básica de uma inferência probabilística é chegar a uma distribuição de probabilidades *a posteriori* para um conjunto de variáveis de consulta (Variável_Consulta), dado um conjunto de variáveis de evidência, ou seja, o sistema computa $P(\text{Variável_Consulta} | \text{Variáveis_Evidência})$, denominação atribuída por (Russell e Norvig, 2003). Trata-se de um procedimento geral, realizado através de cálculos probabilísticos, que utiliza as informações existentes na Rede Bayesiana para a construção uma distribuição de probabilidades *a posteriori* dadas as probabilidades *a priori*.

Existem vários algoritmos de Inferência que merecem referência e que se passa a referenciar de seguida.

2.3.2. Algoritmos

O processo de inferência é a base da utilidade das Redes Bayesianas. Existem algoritmos eficientes em tempo e espaço, que não exigem o cálculo de toda a tabela

conjunta (Jensen, 2001). Os algoritmos de propagação são classificados, segundo (Castillo, *et al.*, 1997) como podendo ser exactos, aproximados (normalmente utilizados em situações em que os métodos exactos não sejam os mais apropriados) e ainda simbólicos.

A propagação simbólica é útil nos casos em que não existe informação sobre a especificação numérica dos parâmetros do modelo probabilístico e/ou quando os especialistas por si só não conseguem especificar intervalos para esses parâmetros. Ainda que mereçam referência, não serão abordados neste trabalho e portanto não será efectuada uma análise detalhada sobre os mesmos.

Os métodos exactos e aproximados podem ser vários, pelo que se faz mencionar de seguida os que se constam da ferramenta GeNIe (GeNIe) que irá ser utilizada e abordada mais à frente.

O leque de algoritmos é vasto e a estrutura na qual se inserem nem sempre é intuitiva pelo que se compila na Figura 5 um resumo dos algoritmos abordados no âmbito deste trabalho.

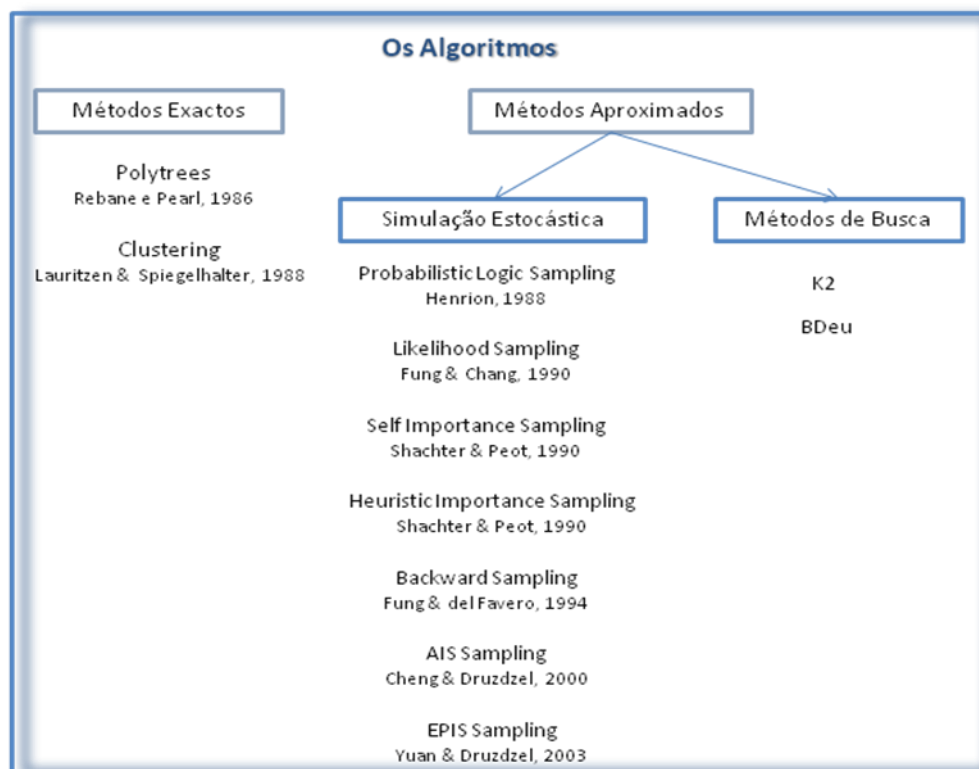


Figura 5 – Classificação dos Algoritmos

ALGORITMOS EXACTOS

Diz-se que um algoritmo de inferência é exacto se os cálculos de probabilidades realizados *a posteriori* através de somatórios e combinação de valores são produzidos contendo apenas erro de arredondamento. São caracterizados por possuírem uma complexidade computacional relativamente baixa, e pela consequente rapidez, no entanto tornam-se ineficientes para cálculos com dezenas ou centenas de variáveis como é o caso de situações reais (Castillo, *et al.*, 1997). Exemplos dos métodos exactos são as *Polytrees* e os Algoritmos de *Clustering*:

✦ *Polytrees*

Uma *polytree*, que é um grafo simplesmente conectado, não possui ciclos; portanto, existe no máximo um caminho entre quaisquer dois nós da estrutura. (Rebane e Pearl, 1987) conseguiram provar que quando a distribuição de probabilidade tem um mapa perfeito e este mapa é uma *polytree*, então o algoritmo desenvolvido sempre o encontra. Utiliza a programação dinâmica, para evitar calcular várias vezes um mesmo factor. Como algoritmo recursivo, utiliza o Teorema de Bayes a cada passo. A Figura 6 ilustra um exemplo deste tipo de estrutura:

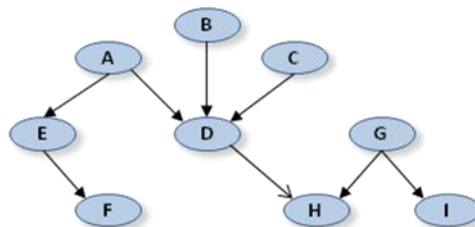


Figura 6 – Exemplo da estrutura de uma *polytree*

O algoritmo que gera *polytrees* segue os seguintes passos:

Algoritmo *Polytrees*

0. Iniciar parâmetros: n (número de nós), $\max D$ (número máximo de arcos por nó), N (número de iterações).
 1. Iniciar grafo com uma árvore simples ordenada.
 2. Repetir N vezes:
 - a. Gerar um par distinto de nós, i e j , distribuído uniformemente;
 - b. Se o arco (i,j) existir no grafo actual ou se o grafo resultante não satisfizer a condição $\max D$:
 - i. Então manter o mesmo estado;
 - ii. Senão:
 1. Inverter o arco (i,j) com probabilidade de $\frac{1}{2}$;
 2. Encontrar o predecessor, nó k , no caminho entre os nós i e j , remover o arco entre k e j , e adicionar o arco (i,j) ou (j,i) , dependendo do sorteio do ponto anterior;
 3. Voltar ao grafo actual após as N iterações
-

Algoritmo 1 - *Polytrees*

+ *Algoritmos de clustering (algoritmos de agrupamento, também conhecidos como árvores de junção - junction trees)*

Tal como o nome indica, estes algoritmos dividem dados em grupos ou clusters de itens que têm propriedades semelhantes, obtendo uma sequência óptima de eliminação e criando uma estrutura para propagar as multiplicações e marginalizações das tabelas. A Figura 7 ilustra um exemplo deste tipo de estrutura:

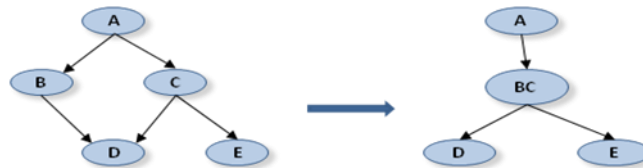


Figura 7 - Exemplo de estrutura sujeita à utilização de um algoritmo de *clustering*

Para se obter uma representação de uma árvore de junção, para uma Rede Bayesiana é necessário construir um grafo não dirigido, chamado grafo moralizado, e triangular (adicionando arcos ao grafo moralizado para formar um grafo triangular), seleccionar subconjuntos de nós do grafo triangular, os chamados cliques, montar a árvore de junção (Algoritmo Árvore de Junção) e atribuir distribuições de probabilidades ao cliques. Estes procedimentos devem ser aplicados sequencialmente e na ordem em que são citados.

Algoritmo Árvore de Junção: Algoritmo de Frank Jensen, 1994

0. Definir numeração $\alpha: U \rightarrow \{1, 2, \dots, |U|\}$ em função inversa da ordem da eliminação, tal que para $u, v \in U$, se u for eliminado após v :
 - a. Então $\alpha(u) < \alpha(v)$;
 1. Seja v , a variável de C de maior numeração, tal que as w , com numeração menor do que v , isto é, $u \in C$ e $\alpha(w) < \alpha(v)$. Se v existir:
 - a. Então o índice de C é $\alpha(v)$
 - b. Senão é 1.
 Após obter todos os índices, numerar os cliques por ordem crescente de índice, iniciando com 1;
 2. Considerar o clique de número 1, isto é, C_1 , como a raiz da árvore;
 3. Ligar C_k a algum clique C_j , $j < k > 1$, já na árvore, que contenha $S_k = C_k \cap U^{k-1}_{i=1} C_i$. Se existir mais de um, ligar ao de menor índice.
-

Algoritmo 2 - Árvore de Junção

ALGORITMOS APROXIMADOS (OU DE SIMULAÇÃO ESTOCÁTICA)

Os algoritmos de inferência aproximados são aplicáveis a qualquer tipo de estrutura de rede e baseiam-se na simulação para calcular de forma aproximada as probabilidades condicionais. Relativamente à inferência exacta, permitem obter soluções mais eficientes, nomeadamente nos casos em que as redes são grandes e muito conectadas,

porém trata-se de soluções aproximadas. Basicamente, esses métodos geram realizações até um determinado tamanho N através das funções de probabilidade conjunta das variáveis e ao final das N realizações aproximam as probabilidades através do quociente entre a frequência de aparição das realizações e o número total de amostras geradas. Segue-se a referência de alguns exemplos que constam no GeNIe:

✦ *Forward Sampling (Logic Sampling)*

(Henrion, 1988) descreve este algoritmo. É o algoritmo de inferência aproximado mais simples. No entanto, é ineficiente por ser aproximado e pelo número excessivo de simulações necessárias para obter resultados razoáveis em certos casos. Sorteia as variáveis sem pais, segundo suas probabilidades *a priori*. Em seguida, sorteia os seus filhos, segundo as probabilidades condicionadas aos pais, e assim sucessivamente. Essa geração de configurações para as variáveis é repetida várias vezes. Com a introdução de evidências, o algoritmo simplesmente descarta as configurações inteiras que não se adequam à evidência. O *Forward Sampling* é um algoritmo que produz amostras de uma distribuição difícil de amostrar a partir de uma distribuição fácil de amostrar.

Refira-se no entanto que existem problemas associados a esta lógica. Isto porque se uma evidência é muito rara, a maioria das configurações geradas serão rejeitadas e serão necessárias muitas simulações para se gerar um número razoável de configurações compatíveis. Por outro lado, a fracção de amostras consistentes com a evidência cai exponencialmente conforme o número de variáveis de evidência cresce, assim, esse algoritmo é simplesmente inútil para problemas complexos. Segue-se a descrição do pseudo-código deste algoritmo:

Algoritmo Logic Sampling

0. Iniciar parâmetros: N (número de amostras)
 1. Estabelecer uma ordenação O das variáveis da rede tal que nenhuma variável apareça antes do seu predecessor:
 2. Repetir N vezes:
 - a. De $i=1$ até $|O|$:
 - i. $X_i = O(i)$
 - ii. Obter um valor para X_i tal que $P^{(x_i)} = P(X_i | c)$, onde c é a configuração dos valores já obtidos para $pa(X_i)$
 3. Voltar ao grafo após as N repetições
-

Algoritmo 3 – Logic Sampling

+ *Likelihood Sampling (Likelihood weighting)*

Bastante usado, funciona bem e soluciona o problema anterior. As variáveis de evidência não são sorteadas, mas recebem o valor da evidência. No entanto, utiliza-se como *peso* as probabilidades condicionais (sobre as evidências) para saber o quão provável é a configuração gerada. Desta forma é possível afirmar que este algoritmo resolve o problema de rejeições do *Forward Sampling*, gerando apenas amostras consistentes, onde cada evento gerado é ponderado pela probabilidade de que o evento concorde com a evidência, medida pela função de ponderação. O peso das probabilidades condicionais (sobre as evidências) é utilizado para saber o quão provável é a configuração gerada. (Kanazawa, *et al.*, 1995) descrevem este algoritmo da seguinte forma:

Algoritmo *Likelihood weighting*

```

Para i = 1 até N
     $w_{x_i} \leftarrow 1.0$ 
Termina ciclo
Para t = 0 até T
    Definir  $E_t$ 
    Para i = 1 até N
        Adicionar amostra de  $X_t$  a  $S_t$ 
         $w_{x_i} \leftarrow w_{x_i} \times \text{Probabilidade}(E_t | s_i)$ 
        Acrescentar  $w_{x_i}$  à pontuação dos valores das amostras  $X_t$ 
    Termina ciclo
Termina ciclo

```

Algoritmo 4 – Likelihood Weighting

+ *Self Importance Sampling / Heuristic Importance Sampling*

(Shachter e Peot, 1990) definem duas variantes de amostragem: *Self Importance* e *Heuristic Importance*. A função importância utilizada - $Pr^o(X|E)$ – é actualizada no passo 3.a do algoritmo que se segue. Este algoritmo procura rever periodicamente as tabelas de probabilidade condicional de forma a aproximar a distribuição da amostra à distribuição obtida *a posteriori*.

Este algoritmo controla portanto a geração de amostras procurando obter uma redução de várias ordens no número de amostras necessárias por simulação para se obter uma estimativa confiável, reduzindo a variância e erros nos resultados com a ajuda de uma função de importância. Esta é uma função densidade probabilidade de um sistema num dado domínio. A distribuição de probabilidade das variáveis de um modelo de Redes Bayesianas é dada pelo produto da probabilidade de cada nó

dados os seus pais, ou seja: $\Pr(X) = \prod_{i=1 \text{ to } n} \Pr(X_i | Pa(X_i))$, onde $Pa(X_i)$ representa os pais do nó X_i .

A distribuição de probabilidade para os nós com variáveis de consulta (ao invés de variáveis evidência) é dada por $\Pr(X|E, E = e) = \prod_{i=X|E} \Pr(X_i | Pa(X_i))$.

O *Heuristic Importance* primeiro transforma a rede numa *polytree* e utiliza uma versão deste algoritmo para calcular as funções de probabilidade para cada nó não observado. $\Pr^o(X|E)$ é uma combinação dessas funções de probabilidade com $\Pr(X|E, e)$. No passo 3.a do algoritmo, a função importância não é actualizada. Os algoritmos baseados na busca heurística têm como importante característica a dependência do processo de aprendizagem em relação à ordenação das variáveis.

Algoritmo Self Importance Sampling

0. Ordenar os nós de acordo com a topologia da rede.
 1. Iniciar parâmetros: m (número de amostras, l (intervalo de actualização), vector de pontuação dos nós e função de importância. $\Pr^o(X|E)$
 2. $k \leftarrow 0, T \leftarrow \emptyset$
 3. Para $i \leftarrow 1$ até m :
 - a. Se $(i \bmod l = 0)$ então
 - $k \leftarrow k + 1$
 - Actualizar a função importância $\Pr^k(X|E)$ baseado em T
 - b. Termina ciclo
 - $S_i \leftarrow$ Gera uma amostra de acordo com $\Pr^k(X|E)$
 - $T \leftarrow T \cup \{S_i\}$
 - Calcular as pontuações $(s_i, \Pr(X|E, e), \Pr^k(X|E))$ e acrescentar ao registo correspondente do vector de pontuações dos nós
 4. Termina ciclo
 5. Normalizar os vectores de pontuação para cada nó.
-

Algoritmo 5 – Self Importance Sampling

✦ Backward Sampling

Considerando ainda o algoritmo do ponto anterior, considere-se pois, genericamente, que, no passo 0, e com *Backward Sampling*, a ordenação é efectuada no sentido oposto. No passo 1 utiliza a probabilidade de evidência observada e de nós declarados para calcular $\Pr^o(X|E)$. Embora (Fung e Favero, 1994) mencionassem a possibilidade de uma ordenação dinâmica dos nós, não sugeriram no entanto qualquer forma de actualização da função importância, o que poderá conduzir a uma convergência fraca entre esta e a função óptima.

+ AIS Sampling

Este algoritmo – *Adaptive Importance Sampling for Bayesian Networks* (AIS-BN) – é descrito em (Cheng e Druzdzel, 2000):

Algoritmo AIS_BN

0. Ordenar os nós de acordo com a topologia da rede
 1. Iniciar parâmetros: função importância $Pr^k(X|E)$ utilizando métodos heurísticos, w^k (peso), m (número desejado de amostras), l (intervalo de actualização), vector de pontuação dos nós
 2. $k \leftarrow 0, T \leftarrow 0, w_{score} \leftarrow 0, w_{sum} \leftarrow 0$
 3. Para $i \leftarrow 1$ até m :
 - a. Se $(i \bmod l = 0)$ então
 - $k \leftarrow k + 1$
 - Actualizar a função importância $Pr^k(X|E)$ e w^k baseado em T
 - b. Termina ciclo
 - $S_i \leftarrow$ Gera uma amostra de acordo com $Pr^k(X|E)$
 - $T \leftarrow T \cup \{S_i\}$
 - $w_{score} \leftarrow$ Pontuações $(S_i, Pr(X|E, s), Pr^k(X|E), w^k)$
 - $w_{sum} \leftarrow w_{score} + w_{score}$
 4. Termina ciclo
 5. Obter estimativa de $Pr(E)$ como w_{score}/w_{sum} e normalizar os vectores de pontuação para cada nó (opcional).
-

Algoritmo 6 – AIS Sampling

+ EPIS Sampling

(Yuan e Druzdzel, 2004) reproduzem este algoritmo – *Estimated Posterior Importance Sampling algorithm for Bayesian Networks* (EPIS-BN) – nos seguintes moldes:

Algoritmo EPIS_BN

0. Ordenar os nós de acordo com a topologia da rede
 1. Iniciar parâmetros: m (número de amostras), ϵ e d (numero de iterações)
 2. Os nós de evidencia são vectores que assumem o valor 1 no estado observado e 0 nos restantes estados; os outros nós são vectores de 1's.
 3. Para $i \leftarrow 1$ até d :
 - Para todos os nós, recalculer mensagens enviadas com base em todas as mensagens recebidas na última iteração
 4. Terminar ciclo
 5. Calcular a função importância com base nas mensagens finais
 6. Reproduzir a função importância pelo método heurístico
 7. Para $i \leftarrow 1$ até m :
 - a. $S_i \leftarrow$ gerar amostra de acordo com $P(X|E)$
 - b. Computar w_{score} a partir de S_i
 - c. Adicionar w_{score} ao correspondente registo de cada tabela de pontuação
 8. Terminar ciclo
 9. Normalizar cada tabela de pontuação e obter estimativas para cada nó.
-

Algoritmo 7 – EPIS - Sampling

2.3.3. Aprendizagem

A construção manual de uma Rede Bayesiana pode ser um processo bastante trabalhoso e caro para grandes aplicações (Friedman e Goldszmidt, 1996). Em áreas mais complexas, a sua especificação pode consumir bastante tempo e sujeitar-se à ocorrência

de erros. E é por esta razão que os esforços dirigidos para o desenvolvimento de métodos que possam construir directamente Redes Bayesianas a partir de uma base de dados, em detrimento do discernimento de especialistas humanos, têm vindo a crescer continuamente.

Dada uma base de dados, é possível aprender a estrutura e os parâmetros de uma Rede Bayesiana. A aprendizagem da estrutura é muito mais difícil e imprecisa. Existem vários algoritmos de busca no GeNIe e não há consenso sobre o melhor: K2 e Bdeu (nos métodos *Greedy Thick Thinning* (Chickering, 2002) e Naive), PC (Kalisch, 2006) Essencial *Graph Search*.

No contexto de Redes Bayesianas, aprendizagem é um processo que tem como entrada um conjunto de dados (informação *a priori*) e como saída uma Rede Bayesiana. O problema de aprendizagem Bayesiana pode então ser enunciado como: dado um conjunto de treino $D = \{u_1, u_2, \dots, u_n\}$, de U , onde U é o conjunto das variáveis do domínio e cada u_i corresponde aos valores das ocorrências de cada uma das variáveis de U , encontrar a rede B que melhor se adequa ao conjunto D (Friedman e Goldszmidt, 1996). A formalização do conceito de adequação e a maneira utilizada para encontrar B é que vão diferenciar os métodos de aprendizagem Bayesiana.

De maneira geral a aprendizagem do pode ocorrer com ou sem supervisão. Quando a aquisição dos conceitos e do conhecimento estruturado é orientada por um supervisor ou uma função de supervisão que classifica os resultados apresentados em certos ou errados, a aprendizagem é supervisionada. Quando a aprendizagem ocorre sem esse tipo de orientação então diz-se não supervisionada.

Aprender o quê?

- Aprender parâmetros – consiste em calcular automaticamente as tabelas de probabilidade condicional de uma Rede Bayesiana, dada a sua estrutura e uma base de dados.
- Aprender a estrutura – consiste em desenhar automaticamente a estrutura gráfica de uma Rede Bayesiana dada uma base de dados. Para aprender a estrutura, existem duas abordagens principais:

- Métodos de busca e pontuação, que procuram no espaço de estruturas, a “melhor” estrutura e que definem uma pontuação prosseguindo enquanto a pontuação de uma rede for significativamente melhor que a anterior. Estes métodos utilizam os algoritmos K2 e BDeu.
 - O K2 procura entre as $2^{n(n-1)/2}$ configurações possíveis a que maximiza a função de pontuação, sendo n o número de variáveis. Necessita pois de um limite superior para o número de pais das variáveis e de uma ordenação prévia das mesmas para garantir a ausência de ciclos na estrutura da rede.
 - O BDeu: O método *Bayesian Dirichlet Equivalente com Uniforme a priori* (BDeu) assume um caso especial da distribuição Dirichlet, na qual uma constante matemática é usada para criar uma distribuição fixa ou uniforme de estados anteriores. A pontuação *Bayesian Dirichlet Equivalent* (BDE) também usa análise de Bayesiana para avaliar uma rede dado um conjunto de dados. O método de pontuação do BDE foi desenvolvido por Heckerman e baseia-se na métrica de BD desenvolvida por Cooper e Herskovits. A distribuição Dirichlet é uma distribuição multinomial que descreve a probabilidade condicional de cada variável da rede. A pontuação do BDE também assume equivalência de probabilidade, o que significa que não se pode esperar que os dados separem estruturas equivalentes. Por outras palavras, se a pontuação de *If A Then B* for igual à pontuação de *If B Then A*, não será possível distinguir as estruturas com base nos dados nem deduzir a causa.
- Métodos baseados em análise de dependência que procuram uma rede que represente da melhor maneira possível a distribuição conjunta que surge da amostra aleatória. É fundamental que esta rede represente todas as relações de independência e dependência da distribuição conjunta induzida pela amostra. Estes métodos utilizam os algoritmos PC.

Existem diversas variantes da tarefa de aprendizagem porque a estrutura de uma rede pode ser conhecida ou desconhecida e as variáveis podem ser todas observáveis ou não.

O quadro abaixo representado enumera alguns problemas associados aos vários tipos de aprendizagem:

		Estrutura	
		Conhecida	Desconhecida
Variáveis	Observáveis	As tabelas de probabilidade condicional podem ser estimadas usando o conjunto de exemplos com classificador de Bayes; as probabilidades condicionadas são estimadas a partir de frequências relativas observadas nos dados.	O problema é construir a topologia da rede. Busca no espaço de estruturas. Os algoritmos K2 e BDeu são exemplos de métodos que procuram aprender a estrutura da rede.
	Não Observáveis	Caso parecido com aprendizagem em Redes Neurais	Não se conhece algoritmos para este tipo de problema

Os algoritmos de aprendizagem fornecem pois uma ajuda valiosa para a modelação de um problema que utilize Redes Bayesianas, porém, nunca a ajuda de um especialista para validar ou re-calibrar os dados aprendidos, deverá ser descartada.

3. MODELOS DINÂMICOS

O estado de mudança dos acontecimentos é tratado com recurso ao uso de um conjunto de variáveis aleatórias que representam um estado em cada instante de tempo. Os agentes que actuam em ambientes incertos têm que se manter actualizados sobre o estado do ambiente. A tarefa complica-se consideradas percepções parciais e/ou ruidosas e a incerteza quanto às alterações ambientais que vão ocorrendo ao longo do tempo.

Neste capítulo introduz-se o factor tempo no raciocínio probabilístico e são apresentados alguns modelos úteis que se manifestam mais apropriados quando se pretende considerar dependências entre as variáveis incorporando simultaneamente um raciocínio probabilístico temporal. São eles os Models Ocultos de Markov (*Hidden Markov Models* - HMM) e as Redes Bayesianas Dinâmicas. Estas últimas, como caso geral dos anteriores.

3.1. RACIOCÍNIO PROBABILISTICO TEMPORAL

O processo de mudança do ambiente pode ser visto como uma série de intervalos (fatias) de tempo. Cada instante contém variáveis aleatórias, algumas observáveis, outras não.

O modelo de dados de uma série temporal pode ser representado por uma estrutura formada por uma escala de tempo composta por determinado número de casos k , onde $k = 1, 2, \dots, t$, um número j de atributos, com $j = 1, 2, \dots, p$ normalmente divididos em i objectos discretizáveis ou então intervalos de tempo que se repetem no decorrer do período de tempo em estudo.

As séries temporais são regularmente utilizadas para efeitos de previsão. Nos modelos de séries temporais a reacção da variável dependente (Y_t), devida a uma variação da variável (ou variáveis) explicativa (X_t), raramente é instantânea. Normalmente, a variável dependente (Y_t) reage às alterações da variável (ou variáveis) explicativa (X_t) com um lapso de tempo. As razões que explicam este lapso de tempo podem ser de

ordem psicológica, ou seja, relacionadas por exemplo com o comportamento humano, gostos diferentes, expectativas ou incerteza.

3.2.CLASSIFICAÇÃO DOS MODELOS DINÂMICOS

Dependendo da forma dos modelos de transição e observação e dos tipos das variáveis aleatórias utilizadas, existem modelos mais específicos que podem ser usados para os processos de inferência temporal. Estes modelos podem ser classificados da seguinte forma:

- ✦ Modelo Temporal Repetitivo: se as fatias de tempo e os *links* temporais são iguais.
- ✦ Modelo Estritamente Repetitivo: se também as probabilidades condicionais são idênticas.
- ✦ Modelo Escondido (ou Oculto) de Markov: se também possui a propriedade de Markov, ou seja, o passado não afecta o futuro dado o presente – *Hidden Markov Models*.
- ✦ Filtro de Kalman: se também só uma variável tem pais fora da fatia de tempo.
- ✦ Cadeia de Markov: se também cada fatia de tempo possui só uma variável.

Nos pontos que se seguem será efectuada uma breve abordagem ao Modelo de Markov, dado que este pode ser visto como um caso especial de Redes Bayesianas Dinâmicas.

3.2.1. Premissa de Markov

Markov pressupõe que o estado actual depende de um histórico finito número de estados anteriores:

$$\forall t, P(X_t | X_{0:t-1}) = P(X_t | X_{t-n:t-1}), n \geq 1$$

Processos que seguem este princípio são chamados de Processos ou Cadeias de Markov. Num processo de Markov de primeira ordem, $n=1$.

Assim como uma cadeia de Markov, uma Rede Bayesiana pode ser vista como uma matriz de atributos que possuem uma correlação entre si e que também apresentam influência uns sobre os outros ao longo do tempo.

Exemplificando, considere-se um caso simples de uma Rede Bayesiana composta por duas variáveis: Recomendações e Classificação. A Classificação de uma auditoria efectuada a um balcão depende da quantidade de Recomendações efectuadas. Assume-se também que as auditorias são realizadas em determinada escala de tempo.

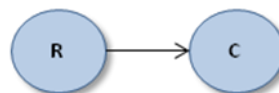


Figura 8 – Rede Bayesiana

Considera-se como valores possíveis para os atributos, os seguintes: **Muitas**, **Médio** e **Poucas** para o número de *Recomendações* e **Bom**, **Satisfatório** e **Com Deficiências** para a *Classificação*. A qualificação das avaliações possíveis dos balcões foi efectuada de acordo com os seguintes critérios:

Tabela 4 – Classificação dos atributos

Classificação	Nº de Recomendações
Bom	0-1 = Poucas
Satisfatório	2-7 = Médio
Com deficiências	8 ou mais = Muitas

Assim sendo, a Rede Bayesiana também apresentaria os dados de probabilidades iniciais e conjuntas apenas para a variável *Classificação* uma vez que é a única que possui uma variável pai, ou seja, uma relação de dependência da *Classificação* dado o número de *Recomendações*. O modelo de dependências e as tabelas de probabilidade denotariam, então, todos os dados que a Rede Bayesiana poderia fornecer.

Ora, seguindo a abordagem Markoviana, o que se pretende obter é o instante de tempo n no qual, a partir de uma inferência, uma determinada configuração de probabilidades de uma variável deve ocorrer.

Se se pretende obter uma nova configuração de uma variável, então é necessário estruturar uma matriz de transição Markoviana vinculada a essa variável, ou seja, é necessário mapear as probabilidades entre os seus estados para a matriz, com base nas

probabilidades condicionais que a referida variável possui em relação às quais mantém um vínculo de dependência.

Considerando o exemplo, devem-se mapear as probabilidades de transição de *Classificação* da seguinte forma: de **Bom** continuar **Bom**, de **Bom** passar para **Satisfatório**, de **Bom** passar para **Com Deficiências** e assim por diante.

Isto significa então a necessidade de se calcular as probabilidades de transição para os estados de uma dada variável, para que se possa determinar a probabilidade de se alcançar um estado C_{t+1} com base em C_t .

A tabela que se segue é resultado da criação da matriz de transição Markoviana, e traduz a transição de um estado x num tempo C_t para um estado y num tempo C_{t+1} , ou seja:

$$P(C_{t+1} = s_y | N_t = s_x) = p_{xy}$$

	Bom	Satisfatório	Com Deficiências
Bom	p_{BB}	p_{BS}	p_{BD}
Satisfatório	p_{SB}	p_{SS}	p_{SD}
Com Deficiências	p_{DB}	p_{DS}	p_{DD}

Tabela 5 – Modelo de matriz de transição de estados a ser construída

Contudo, o raciocínio não termina por aqui. A variável *Recomendações* em relação à variável *Classificação* não é suficiente para se verificar a relação da variável com ela própria e a transição entre os seus estados, uma vez que, desta forma, a matriz de transição Markoviana convergiria imediatamente para o estado estacionário. Assim deve também ser considerado o estado da variável *Classificação* num ponto de tempo anterior, para actuar com a variável *Recomendações* e, desta forma, obter as relações de transição da variável *Classificação*.

Generalizado, a especificação do modelo seria a seguinte:

$$p_{xy} = \frac{\sum_{i=1}^n P(s_y | s_y \cap Pa_i) \times P(Pa_i)}{\sum_{j=1}^m \sum_{k=1}^n P(s_j | s_x \cap Pa_k) \times P(Pa_k)}$$

Onde:

p_{xy} corresponde à probabilidade de se transicionar de um estado x para o estado y .

3.2.2. HMMs

Modelos Ocultos de Markov (*Hidden Markov Models* – HMMs) podem ser considerados casos especiais de Redes Bayesianas Dinâmicas embora tenham sido desenvolvidos antes, em meados da década de 1960, quando alguns importantes artigos de Leonard Baum e outros autores formaram a base da teoria e seus vários algoritmos relacionados (Russell e Norvig, 2003).

O desafio é determinar os parâmetros ocultos a partir dos parâmetros observáveis, por outras palavras, trata-se de representar uma variável aleatória como uma Cadeia de Markov (ou processo de Markov), com a propriedade particular de que tal variável não pode ser medida diretamente, mas sim através de observações possíveis a partir de cada estado da variável.

Trata-se de um modelo onde, para além da função de distribuição de probabilidades associadas aos estados, existe uma função de distribuição de probabilidades para as observações que podem ser realizadas em cada estado (Rabiner e Juang, 1989). Consiste num processo duplamente estocástico composto por um processo oculto (não observável), mas que se manifesta através de um outro processo estocástico que produz a sequência de símbolos observados. Os dois tipos de parâmetros a que um Modelo Oculto de Markov está associado são: probabilidades de emissão dos símbolos e probabilidades de transição de estados.

Um exemplo clássico de modelagem Markoviana é a estimação de um cara-ou-coroa, onde, lançada uma moeda ao ar, a probabilidade de sair cara é igual á probabilidade de sair coroa, quando esta cair, e é igual a 50%. Neste caso, as variáveis são ditas observáveis. Mas se este jogo for efectuado atrás de uma cortina e não for possível saber quantas moedas estão a ser utilizadas, então o modelo irá considerar variáveis ocultas.

3.3. REDES BAYESIANAS DINÂMICAS

As Redes bayesianas modelam as relações entre variáveis num determinado instante do tempo ou durante um intervalo específico. Ainda que a relação causal representada por um arco implique uma relação temporal, as Redes Bayesianas não modelam

explicitamente relações temporais entre as variáveis. A única forma de modelar a relação entre o valor actual de uma variável e o seu valor passado ou futuro é pela adição de uma outra variável que possua nome diferente (Korb e Nicholson, 2003).

A sequência de manipulação de uma rede Bayesiana dinâmica é a mesma apresentada para as redes Bayesianas.

3.3.1. Definição

Redes Bayesianas Dinâmicas são Redes Bayesianas em que algumas variáveis são relativas a um tempo específico, chamado *fatia de tempo* (*time slice*), e são conectadas entre si por *links temporais*. Representam pois, modelos probabilísticos temporais.

São processos estocásticos, na forma de sequência de dados ou séries temporais assumindo bastante expressão em muitas aplicações práticas. Uma rede Bayesiana que modela um processo estocástico é chamada de Rede Bayesiana Dinâmica, que é uma rede bayesiana em que uma célula básica (rede bayesiana simples) se repete ao longo do tempo. O termo dinâmico significa que o processo é gerado por um sistema dinâmico, e não que a rede sofra alguma alteração ao longo do tempo.

Um processo que incorpore uma dinâmica de alteração periódica dos valores das suas variáveis, pode então ser visto como esta série de *fatias de tempo*, onde cada instante contém um conjunto de variáveis aleatórias, algumas observáveis, outras não. Assim, cada fatia pode ter qualquer número de variáveis de estado X_t e de evidência E_t .

Rede Bayesiana Dinâmica é pois, uma rede Bayesiana, que representa um modelo de probabilidades temporal. Considerando que as ligações entre variáveis são replicadas em cada fatia, então a Rede Bayesiana Dinâmica representa um processo de Markov de primeira ordem. O Princípio de Markov enuncia que estados passados e futuros são independentes considerando-se o estado actual. Processos que seguem este princípio são os chamados Processos de Markov ou Cadeias de Markov. $P(X_t|X_{t-1})$ é o chamado modelo de transição para processos de primeira ordem.

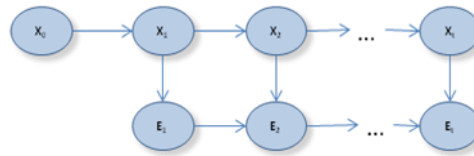


Figura 9 – Exemplo de uma Rede Bayesiana Dinâmica

Neste caso, X_t e E_t correspondem às variáveis observáveis e não observáveis do modelo, respectivamente. As Redes Bayesianas Dinâmicas assumem a proposição de Markov, pois, de outra forma, novas tabelas de probabilidade teriam que ser montadas a cada fatia de tempo.

3.3.2. Construção

Para construir uma Rede Bayesiana Dinâmica é necessário considerar a seguinte informação:

1. Distribuição *a priori* das variáveis iniciais: $P(X_0)$
2. Modelo de transição: $P(X_{t+1}|X_t)$
3. Modelo sensor: $P(E_t|X_t)$

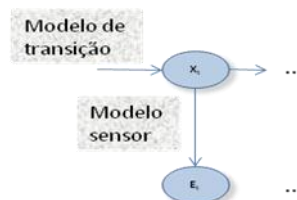


Figura 10 – Rede Bayesiana

Para a rede Bayesiana da figura 10 o modelo de transição é $P(X_t|X_{t-1})$ e o modelo sensor é $P(E_t|X_t)$. A estrutura da figura assume um processo de *Markov* de ordem um, porque assume-se que a probabilidade do momento depende apenas do momento anterior.

Para especificar os modelos de transição de sensor, é necessário especificar a topologia das conexões entre fatias sucessivas e entre as variáveis de estado e de evidência. Ora, como esses modelos são assumidos como sendo estacionários – só é necessário especificá-los para a fatia de tempo inicial.

A figura que se segue foi retirada de (Kuenzer, *et al.*, 2002) – e ilustra algumas das diferentes topologias de redes.

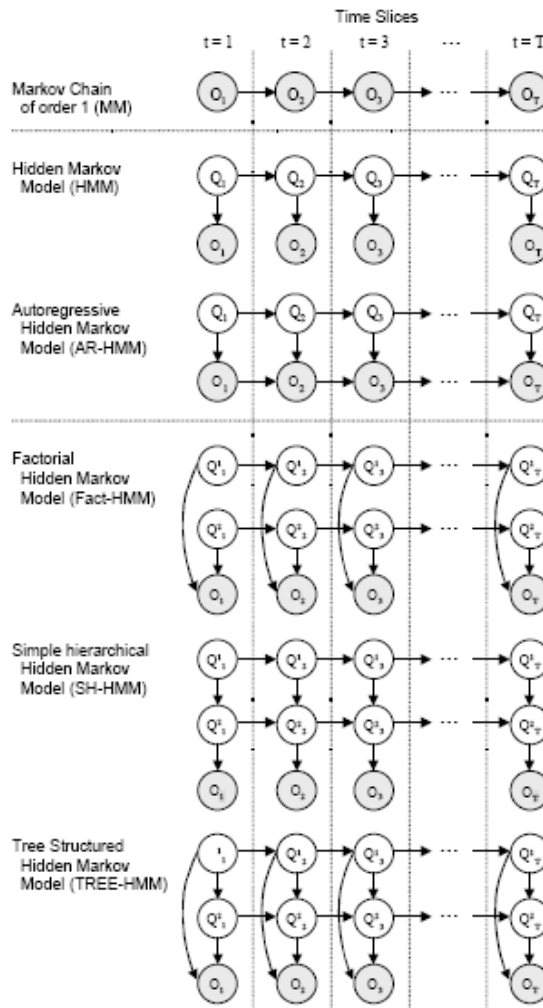


Figura 11 – Diferentes Topologias de Redes Bayesianas Dinâmicas

3.3.3. Premissas adotadas

A necessidade de se determinar tabelas de probabilidade condicionais para cada unidade de discretização de tempo que seja analisada pode ser um problema quando se trabalha com Redes Bayesianas Dinâmicas. Desta forma, assume-se, trabalhar com um processo estocástico estacionário.

Pressupõe-se então que:

- As ligações entre as variáveis são replicadas a cada camada;
- A Rede Bayesiana Dinâmica representa um processo de Markov de 1ª ordem.

3.3.4. Inferência exacta em Redes Bayesianas Dinâmicas

Ainda que seja possível, através de inferências, verificar o comportamento futuro dos seus atributos, as Redes Bayesianas não dispõem de meios que permitam descobrir o quão próximo ou distante esses eventos estariam de acontecer, isto é, não permitem quantificar e apontar, no tempo, o momento de ocorrência dessas inferências.

Como já foi dito anteriormanete, Redes Bayesianas Dinâmicas são, antes de mais, Redes Bayesianas pelo que o processo de inferência é basicamente o mesmo nos dois casos. Ou seja, para um conjunto de observações, pode-se construir uma representação completa da Rede Bayesiana Dinâmica usando uma Rede Bayesiana efectuando o desdobramento da primeira até que a rede fique grande o suficiente para acomodar as observações. Uma vez desdobrada, pode-se usar qualquer algoritmo de inferência probabilística em redes Bayesianas.

Nos pontos que se seguem, é efectuada uma descrição de quatro tipos de inferência utilizados em redes Bayesianas Dinâmicas: o *Filtering* (filtragem), a Previsão, o *Smoothing* (suavização) e a Sequência Mais Provável.

3.3.4.1. Filtering

Trata-se de construir um processo recursivo a partir das probabilidades *a priori* e das probabilidades condicionais, ou seja, obter a distribuição de probabilidades do estado actual dadas as evidências até ao momento, isto é:

$$P(X_t | e_{1:t})$$

Pode ser realizado da seguinte forma: dado o resultado da filtragem no instante t , é possível obter o resultado de $t+1$ a partir da nova evidência e_{t+1} . Ou seja,

$$P(X_{t+1} | e_{1:t+1}) = f(e_{t+1}, P(X_t | e_{1:t}))$$

para uma função f . Este processo é chamado de estimação recursiva. É composto por duas partes: primeiro o estado actual da distribuição é projectado de t para $t+1$; depois é actualizado utilizando uma nova evidência e_{t+1} . Ora, dividindo pela evidência:

$$P(X_{t+1}|e_{1:t+1}) = P(X_{t+1}|e_{1:t}, e_{t+1})$$

Utilizando a Regra de Bayes:

$$= \alpha P(e_{t+1}|X_{t+1}, e_{1:t}) P(X_{t+1}|e_{1:t})$$

E a evidência de Markov:

$$= \alpha P(e_{t+1}|X_{t+1}) P(X_{t+1}|e_{1:t})$$

O α é uma constante normalizadora utilizado para as probabilidades somarem até dar um. O segundo termo $P(X_{t+1}|e_{1:t})$, representa uma previsão, de um passo apenas, do próximo estado e o primeiro termo $P(e_{t+1}|X_{t+1})$, que é obtido directamente a partir do modelo sensor, actualiza-o com a nova evidência. Agora pode ser obtida uma previsão de apenas um passo para o próximo estado condicionando o estado actual X_t :

$$P(X_{t+1}|e_{1:t+1}) = \alpha P(e_{t+1}|X_{t+1}) \sum_{x_t} P(X_{t+1}|x_t, e_{1:t}) P(x_t|e_{1:t})$$

O mesmo que, usando a propriedade de Markov:

$$P(X_{t+1}|e_{1:t+1}) = \alpha P(e_{t+1}|X_{t+1}) \sum_{x_t} P(X_{t+1}|x_t) P(x_t|e_{1:t})$$

Naturalmente, esta equação envolve apenas o modelo de transição e não o modelo sensor. Se se considerar o que acontece quando se tenta prever cada vez mais no futuro, normalmente a previsão converge para um ponto fixo que depois se mantém constante. Isso é a distribuição estacionária do processo de *Markov* definido pelo modelo de transição.

Dentro do somatório, o primeiro factor é o modelo de transição, e o segundo factor é o estado actual da distribuição, pelo que se chega à formulação recursiva desejada.

Assim, a estimativa de filtragem, $P(x_t|e_{1:t})$, pode ser vista como uma mensagem $f_{1:t}$, que é propagada para a frente ao longo da sequência, modificada por cada transição e actualizada por cada nova observação. O processo é o seguinte:

$$f_{1:t+1} = \alpha FORWARD(f_{1:t}, e_{t+1})$$

onde *FORWARD* implementa a actualização descrita na equação anterior.

Quando todas as variáveis de estado são discretas, o tempo de actualização é constante, ou seja, independente de t e o espaço necessário é também constante.

3.3.4.2. Previsão

Trata-se de calcular a distribuição de probabilidades de um estado futuro dadas as evidências até o presente. É semelhante ao *Filtering* mas sem a adição de novas evidências a partir do momento t . À medida que se tenta prever mais longe no futuro, a tendência é que a distribuição de probabilidades convirja para um ponto fixo.

A notação, neste caso é dada por:

$P(X_{t+k}|e_{1:t})$ para algum $k > 0$

Uma vez que pode ser visto simplesmente como a filtragem, mas sem a adição duma nova evidência, então é possível derivar a seguinte actualização:

$$P(X_{t+k+1}|e_{1:t}) = \sum_{x_{t+k}} P(X_{t+k+1}|X_{t+k})P(X_{t+k}|e_{1:t})$$

Naturalmente, esta equação envolve apenas o modelo de transição e não o modelo sensor. Considerando o que acontece quando se tenta prever cada vez mais no futuro, normalmente a previsão converge para um ponto fixo que depois matem-se constante. Também neste caso, isso é a distribuição estacionária do processo de *Markov* definido pelo modelo de transição. A previsão do estado actual está condenada para tempos futuros maiores que uma pequena fracção do tempo.

3.3.4.3. Smoothing

Smoothing significa calcular a distribuição *a posteriori* sobre algum estado passado, dada toda a evidência até ao momento, ou seja,

$P(X_k|e_{1:t})$, para $1 \leq k < t$

É realizado em duas partes, isto é, dividindo a evidência $e_{1:t}$ em $e_{1:k}$, $e_{k+1:t}$:

$$P(X_k | e_{1:t}) = P(X_k | e_{1:k}, e_{k+1:t})$$

Utilizando a Regra de Bayes:

$$= \alpha P(X_k | e_{1:k}) P(e_{k+1:t} | X_k, e_{k+1:t})$$

E a independência condicional:

$$= \alpha P(X_k | e_{1:k}) P(e_{k+1:t} | X_k)$$

Ou seja:

$$= \alpha f_{1:k} b_{k+1:t}$$

onde se define uma mensagem *backward* $b_{k+1:t} = P(e_{k+1:t} | X_k)$. A mensagem *forward* pode ser obtida por filtragem de 1 até k, tal como é dado pela equação alcançada com o *Filtering*. A mensagem *backward* pode ser obtida por um processo recursivo que anda para trás a partir de t:

$$\begin{aligned} P(e_{k+1:t} | X_k) &= \sum_{x_{k+1}} P(e_{k+1:t} | X_k, x_{k+1}) P(x_{k+1} | X_k) \\ &= \sum_{x_{k+1}} P(e_{k+1} | x_{k+1}) P(x_{k+2:t} | X_{k+1}) P(x_{k+1} | X_k) \end{aligned}$$

Dos três factores do somatório, o primeiro e o terceiro são obtidos directamente do modelo e o segundo é uma chamada recursiva. Usando a notação de mensagem, obtém-se:

$$b_{k+1:t} = \text{BACKWARD}(b_{k+2:t}, e_{k+1:t})$$

onde *BACKWARD* implementa a actualização descrita na equação anterior. Tal como para a recursividade *forward*, o tempo e o espaço necessários para cada actualização são constantes, ou seja, independentes de t.

As recursões *backward* e *forward* utilizam um pedaço de tempo constante para cada passo, então o tempo de complexidade para suavizar $e_{1:t}$ é $O(t)$. Se se pretende suavizar a sequência inteira, o método óbvio seria simplesmente correr o processo todo de suavização para cada passo, o que resultaria num tempo de complexidade de $O(t^2)$. Uma melhor aproximação seria a de usar programação dinâmica para reduzir o tempo de complexidade para $O(t)$. A ideia chave aqui é guardar o resultados da filtragem

forward sobre a sequência inteira. Este algoritmo designa-se por algoritmo *forward-backward* (Nilsson, 1998).

3.3.4.4. Sequência mais provável

Um algoritmo recursivo similar ao de filtragem, chamado *Algoritmo de Viterbi* pode ser utilizado para obtenção da Sequência mais provável. Os estados possíveis são representados por um grafo, e procura-se o caminho mais provável nesse grafo.

Assim, dada uma sequência de observações, trata-se de encontrar a sequência de estados que mais provavelmente geraram essas observações, ou seja,

$$\operatorname{argmax}_{X_{1:t}} P(X_{1:t} | e_{1:t})$$

Existe uma relação recursiva entre os caminhos mais prováveis para cada estado X_{t+1} e os caminhos mais prováveis para cada estado x_t :

$$\begin{aligned} \max_{x_1 \dots x_t} P(x_1 \dots x_t, X_{t+1} | e_{1:t+1}) &= \\ &= \alpha P(e_{t+1} | X_{t+1}) \max_{x_t} (P(X_{t+1} | x_t) \max_{x_1 \dots x_{t-1}} P(x_1 \dots x_{t-1}, x_t | e_{1:t})) \end{aligned}$$

Esta equação é semelhante à equação da filtragem, excepto que $f_{1:t}$ é substituído por $m_{1:t}$ e $\sum$ é substituído por $\max$:

$$m_{1:t} = \max_{x_1 \dots x_{t-1}} P(x_1 \dots x_{t-1}, X_t | e_{1:t})$$

que é a probabilidade do caminho mais provável para cada estado x_t .

3.3.5. Inferência aproximada nas Redes Bayesianas Dinâmicas

Dado que a inferência exacta é intratável para redes grandes e muito conectadas, muitas vezes é necessário recorrer a aproximações que permitam obter uma solução mais eficiente, sendo porém, como referido, aproximada. Neste caso, a complexidade em tempo pode custar a qualidade da solução.

Como Redes Bayesianas Dinâmicas são Redes Bayesianas, podem ser utilizados os mesmos algoritmos anteriormente descritos para inferência probabilística nestas

últimas. Por questões de eficiência, geralmente utiliza-se o algoritmo aproximado para inferência conhecido como *Particle Filtering* (Russell e Norvig, 2003), que é uma modificação do algoritmo *Likelihood Weighting*.

O algoritmo *Likelihood Weighting* puro gera amostras dos nós de estado da rede em ordem topológica, dando pesos a cada amostra de acordo com a probabilidade de ocorrência, dadas as variáveis de evidência observadas. Aplicar *Likelihood Weighting* directamente na Rede Bayesiana Dinâmica desdobrada apresentaria os mesmos problemas de tempo e espaço que os apresentados na Rede Normal.

O algoritmo modificado pode ser visto como um algoritmo de *Filtering* onde o conjunto de N amostras corresponde à mensagem propagada (*forward message*). Este algoritmo fornece as probabilidades correctas à medida em que N tende para infinito.

Entretanto, embora o algoritmo pareça manter uma boa aproximação da realidade com um número constante de amostras, não há garantias teóricas disso.

4. TRABALHOS RELACIONADOS

Nos últimos tempos a investigação e algumas implementações eficientes de Redes Bayesianas, têm evidenciado a sua utilização para a representação gráfica de conhecimento incerto nas mais diversas áreas, tais como:

- ✦ Medicina, onde se destaca a sua aplicação para a solução de problemas de diagnóstico. (Milho e Fred, 2002) apresentam um exemplo de aplicação para diagnóstico na área de doenças do sono oferecendo aos especialistas a possibilidade de desenharem as suas aplicações específicas sem lhes ser exigido o conhecimento profundo da teoria de Redes Bayesianas.

Ainda no complexo e vasto domínio da Medicina, (Gerven, *et al.*, 2008) constroem Redes Bayesianas Dinâmicas para prognósticos de carcinomas considerando não apenas a sobrevivência do paciente mas também a progressão da doença, efeito de tratamentos e evolução de possíveis complicações, evidenciando a utilidade deste tipo de abordagem na prática, comparando-a com métodos tradicionais e experimentando-a em três exemplos concretos de pacientes. (Charitos, *et al.*, 2009), para auxílio a médicos que pretendem diagnosticar pneumonia associada ao processo de ventilação mecânica, reportam resultados promissores em experiências efectuadas em pacientes reais no que respeita à captação do desenvolvimento da doença ao longo do tempo, considerando a dinâmica que envolve a participação de peritos e informações associadas aos doentes uma boa receita para diminuir a carga computacional da inferência. (Burge e Lane, 2005), em domínios da neurociência, através de um exemplo que utiliza a informação de ressonâncias magnéticas para o diagnóstico de doenças mentais, mostram como Redes Bayesianas dinâmicas que aprendem a estrutura a partir dos dados se comportam mediante diferentes metodologias, por forma a identificarem classes discriminativas. (Zhang, *et al.*, 2006) utilizam Redes Bayesianas Dinâmicas com o objectivo de encontrar padrões de activação cerebral em termos de conectividade e interactividade.

- ✦ Educação / Ensino: (Bet, *et al.*, 2004) utilizam Redes Bayesianas para o desempenho do papel de um professor que analisa as dificuldades do aluno orientando-o conforme as suas necessidades e respeitando o seu processo de

aprendizagem; (Abbasi, 2007) utiliza este tipo de redes para interpretação dos gestos faciais e corporais de estudantes em prol de um ensino mais eficiente; (Schreiber, *et al.*, 2002) apresentam um projecto de desenvolvimento de um sistema educacional *on-line* com adaptação proactiva para a Internet, por forma a compreender o processo de interacção entre sistema e utilizadores, o tempo de permanência em cada página e qual a ordem de acesso. (Silva e Brito, 2008) constroem um modelo de adaptação relativo a uma instituição de ensino de forma a proporcionar aos alunos diferentes métodos de aquisição de conhecimento e interagir de a forma que o seu conhecimento seja compartilhado e amadurecido com/por demais alunos e professores.

- ✦ Sistemas de Suporte à Decisão: (Bertotti, *et al.*, 2003) utilizam Redes Bayesianas para prever a proporção da primeira estimativa de venda no lançamento de produtos numa indústria de confecção, como ferramenta de apoio à definição de colecções futuras.
- ✦ Internet: Actualmente a utilização das Redes Bayesianas tem crescido bastante para a resolução de problemas relacionados a Internet. A filtragem de SPAM, usualmente feita através da utilização de filtros que não contêm inteligência, tem vindo a utilizar mecanismos de Inteligência Artificial para aumentar a sua eficiência. (Lermen e Bruno, 2008) propõem uma estrutura que utiliza Redes Bayesianas como um método de classificação de alertas por forma a identificar ataques conhecidos e desconhecidos, gerar regras e diminuir o volume dos falsos alertas.
- ✦ Risco Operacional: a utilização de Redes Bayesianas é abordada como forma de trazer transparência às relações de causa e efeito entre as actividades do banco e a volatilidade dos diversos tipos de lucro do banco. Trata-se de uma abordagem que pode ser feita tanto por matemáticos (análise das complexas relações de negócios) como pelos gestores (tomada de decisões e estratégias) e que simplifica a manutenção do modelo pela visualização gráfica das relações, ao invés de equações matemáticas.
- ✦ Segurança: Neste âmbito, (Frigault, *et al.*, 2008) propõem um modelo de Redes Bayesianas dinâmicas que capte vulnerabilidades associadas ao funcionamento

de uma rede de um computador procurando medir continuamente a segurança de uma rede num ambiente dinâmico que incorpore factores temporais. (Zajdel, *et al.*, 2006) trabalharam um sistema de vigilância, através do uso de câmaras distribuídas por uma dada área, por forma a obter padrões de movimentação complexos, que atraíram a atenção de parceiros industriais interessados em otimizar os seus sistemas de segurança. (Bouillaut, *et al.*, 2008) consideram as Redes Bayesianas Dinâmicas estruturas interessantes no âmbito da prevenção rodoviária, mais concretamente no que diz respeito à manutenção dos trilhos dos comboios e úteis para a determinação de parâmetros óptimos de manutenção. (Feng, *et al.*, 2009) utilizaram Redes Bayesianas Dinâmicas para prever se o objectivo das sequencias de chamadas de um sistema é o não cumprido, através de pré-avisos que vão sendo lançados de forma a garantir que sejam tomadas com antecedência medidas adequadas.

- ✦ Reconhecimento de Voz: (Choudhury, *et al.*, 2002), consideram Redes Bayesianas uma ferramenta bastante atractiva no que toca a “sensações humanas” e combinam a representação gráfica desta estrutura com a apresentação de algoritmos de inferência e aprendizagem. A ideia é mostrar um método que possibilite a alteração quer da estrutura da rede quer dos seus parâmetros de forma a aumentar a exactidão dos resultados. (Frankel, *et al.*, 2007) mostram que modelando as dependências entre um conjunto de seis variáveis relacionadas com a articulação da voz, a precisão de reconhecimento melhora relativamente a um sistema que as considere independentes. Para tal utilizam Redes Bayesianas Dinâmica.
- ✦ Sistemas Industriais: (Salem, *et al.*, 2006) utilizam Redes Bayesianas Dinâmicas para construção de um modelo de análise da fiabilidade de sistemas complexos, associadas a ferramentas de simulação, já que estas permitem considerar a influência do tempo nas falhas de um sistema, bem como ilustrar uma estrutura de representação compacta do modelo.

Para além das mencionadas, existem outras aplicações que abrangem um leque bastante diversificado de áreas como são engenharia, análise de texto, agricultura, tratamento de águas, previsões meteorológicas etc.

Muito embora actualmente este tema possa despertar grande interesse entre os Auditores Internos dada a actual procura por mais eficiência e eficácia nas auditorias, não foram encontrados muitos trabalhos desenvolvidos no âmbito da aplicação de Redes Bayesianas aos processos de auditoria bancária.

(Sarkar e Sriram, 2001) procuram demonstrar como modelos probabilísticos automáticos que incorporam um mecanismo formal de revisão de crenças podem ser úteis em trabalhos de auditoria de forma a apoiar o julgamentos dos auditores na difícil tarefa de avaliação da viabilidade de um banco. Neste contexto, analisaram dois modelos probabilísticos distintos um dos quais o modelo de classificação de *Naive Bayes*. O interesse do artigo prende-se no enquadramento que é feito no que toca à área em questão – Banca e Auditoria – mais do que na abordagem metodológica que é alvo deste estudo e que concretamente não é trabalhada neste artigo.

O trabalho de (Laws e O'Hagan, 2003) propõem uma análise Bayesiana no âmbito de Auditoria em termos de aplicação em múltiplas áreas de uma organização. Importa referir que a análise considera a informação prévia dos especialistas – auditores – e utiliza como método de inferência a simulação de Monte Carlo. Mais uma vez, enquadramento na temática em detrimento de aplicação da técnica propriamente dita.

(Fenton e Neil, 2007) aplicam Redes Bayesianas à gestão do risco na actualidade, adoptando esta abordagem como uma boa técnica para a obtenção de bons resultados de investimento. A análise de risco financeiro é referida e a exposição conceptual do raciocínio foram úteis para a estruturação dos casos práticos aqui analisados.

5. AUDITORIA NA BANCA – APLICAÇÕES

Os objectivos da profissão do auditor são os de trabalhar para os mais altos padrões de profissionalismo, para atingir os mais altos níveis de desempenho e, de uma forma geral, para atingir os requisitos de interesse da organização. Estes objectivos exigem que sejam atingidas metas capazes de gerar altos níveis de credibilidade, confiança e qualidade de serviço, metas essas quase sempre, de difícil quantificação e evidência empírica. Se facto é, ser possível atribuir uma classificação a uma auditoria efectuada a um determinado balcão, mais difícil se apresenta medir o seu impacto em termos de desempenho futuro do mesmo.

Este capítulo mostra como é possível gerar Redes Bayesianas no âmbito do impacto da actividade de auditoria na banca, fornecendo modelos gráficos de representação de conhecimento que se apresentam como um poderoso instrumento de apoio quer para utilizadores, auditores ou especialistas. Isto deve-se essencialmente ao facto de as relações (de dependência ou independência) entre as variáveis do domínio poderem ser visualizadas graficamente, bem como à possibilidade de quantificar, em termos probabilísticos, os efeitos dessas relações.

Assim, primeiro, é apresentada a ferramenta utilizada – o GeNIe – com a descrição de características e potencialidades que podem ser exploradas pelo utilizador profissionalmente, academicamente ou até simplesmente por lazer, não sem deixar de sintetizar as vantagens e inconvenientes que se lhe encontram associados.

Segue-se a descrição do funcionamento desta ferramenta, colocando em prática fundamentos teóricos e utilizando um cenário simples de um caso teórico, a análise de falhas de crédito em contas com indicadores de risco. O caso prático apresentado procura representar, através de Redes Bayesianas Simples, o impacto da auditoria em operações de crédito, com o intuito de determinar se a evolução de um cenário que foi alvo de intervenção por parte de agentes auditores é favorável ou não, em termos de evolução de crédito vencido. Aqui distingue-se a capacidade desta ferramenta ser capaz de descobrir o conhecimento inerente a um conjunto de dados fornecidos como “entrada”, da capacidade de “receber” os conhecimentos de um especialista sobre o domínio estudado, e permitir ainda a conjugação destes dois níveis de contribuição para a geração de uma Rede Bayesiana. O caso prático segue os seguintes passos: tratamento

dos dados, caracterização das variáveis, métodos e algoritmos de aprendizagem utilizados, interface gráfico do modelo, descrição da topologia / arquitectura da rede, simulação, propagação e discussão dos resultados.

Finalmente, introduz-se a componente temporal na análise através da construção de uma Rede Bayesiana Dinâmica que permita avaliar, ao longo do tempo, o impacto que as recomendações efectuadas pela equipa de auditoria têm na classificação do perfil associado às operações de crédito efectuadas. Esta análise reúne o tratamento dos dados, a caracterização das variáveis, a metodologia utilizada, o interface gráfico do modelo e a simulação, propagação e discussão dos resultados.

5.1.O SOFTWARE – GENIE 2.0

GeNIe é um *software* que dá suporte à criação, manipulação e avaliação de Redes Bayesianas (Ver Figura 12 – Interface gráfico do GeNIe 2.0), permitindo a construção de modelos de decisão gráficos. É, neste projecto, utilizado como ferramenta de apoio à construção de Redes Bayesianas.

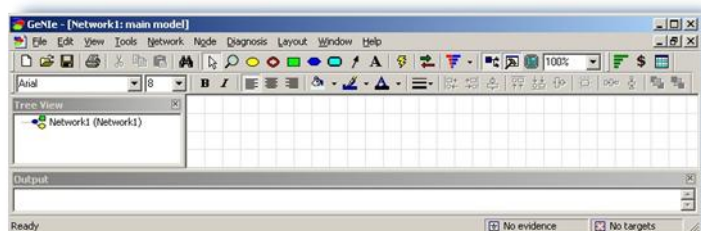


Figura 12 – Interface gráfico do GeNIe 2.0

O software representa a rede graficamente através de um diagrama onde as variáveis são exibidas como nós e as dependências condicionadas exibidas como arcos entre as variáveis. A cada nó é atribuída uma tabela de probabilidades assim como definidas as evidências que podem ser observadas no cenário que está a ser modelado.

As Redes Bayesianas podem ainda ser utilizadas para realizar vários tipos de inferência recorrendo ao uso de algoritmos disponibilizados pelo aplicativo.

O interface gráfico do GeNIe é também uma boa ferramenta de apoio à criação intuitiva de modelos de decisão teóricos com a intervenção de peritos a quem facilmente se permite o desenho do modelo com a simples prática do *click-and-drop*.

O GeNIe incorpora exemplos de Redes Bayesianas, entre os quais o ilustrado na Figura 13 relativo a um problema de concessão de crédito:

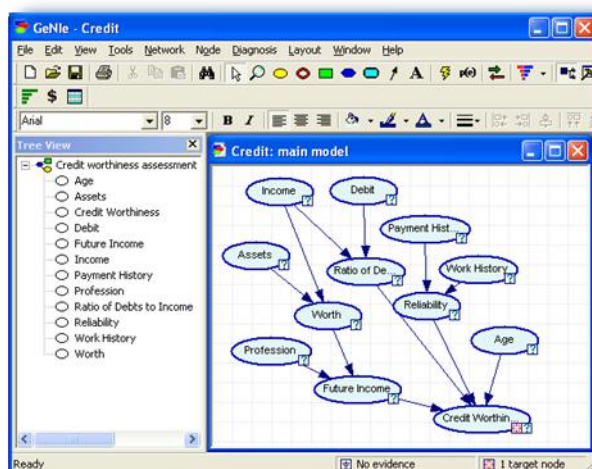


Figura 13 - Exemplo de uma Rede Bayesiana no GeNIe

O GeNIe permite definir ou verificar as propriedades dos nós bem como as propriedades da rede, tal como demonstra a Figura 14. Esta gama de opções é a base do raciocínio que irá ser concretizado.

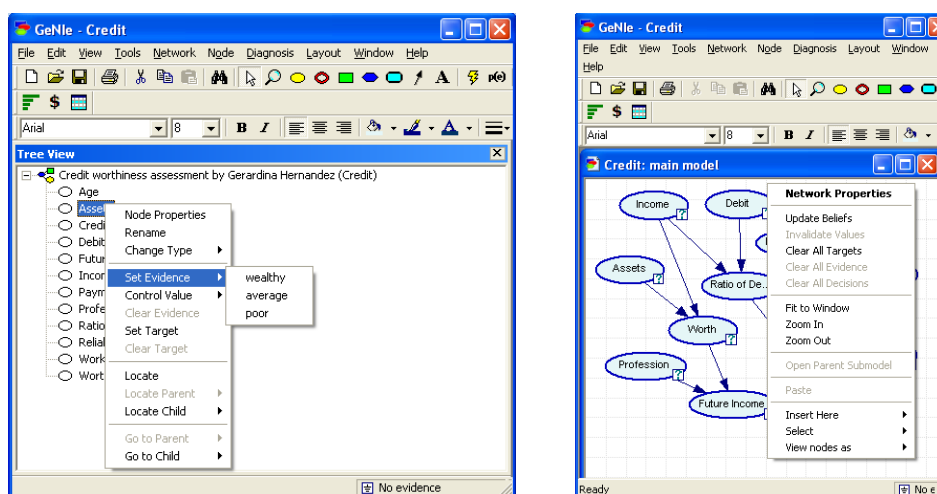


Figura 14 – Propriedades dos nós e da rede no GeNIe

Cada nó da rede tem uma tabela de probabilidades associada. A Figura 15 ilustra as tabelas de probabilidade associadas às variáveis *Work History* e *Reliability*, onde se pode ler estados e associações que geram as condicionadas:

Node properties: Work History

	Value
Stable	0.25
Unstable	0.25
Justified_no_work	0.25
Unjustified_no_work	0.25

Node properties: Reliability

Payment History	Excellent				Acceptable				NoAcceptable				Without Reference			
	Stable	Unstable	Justified	Unjustified	Stable	Unstable	Justified	Unjustified	Stable	Unstable	Justified	Unjustified	Stable	Unstable	Justified	Unjustified
Work History																
Reliable	0.99	0.7	0.7	0.5	0.7	0.55	0.6	0.4	0.196429	0.01	0.1	0.01	0.7	0.3	0.5	0.2
Unreliable	0.01	0.3	0.3	0.5	0.3	0.45	0.4	0.6	0.803571	0.99	0.9	0.99	0.3	0.7	0.5	0.8

Figura 15 – Tabelas de Probabilidades no GeNIe

5.1.1. Vantagens e Inconvenientes do GeNIe 2.0

Como principais vantagens do GeNIe:

- Trata-se de um *software* livre;
- Interface gráfica
- Algoritmos para Aquisição, representação e utilização de conhecimento;
- O mesmo modelo pode ser utilizado para avaliar, prever, diagnosticar ou otimizar decisões, o que contribui para a rentabilização do esforço investido na construção da rede;
- Possibilidade de combinar conhecimento de peritos com o conhecimento incorporado nos dados, isto é, permite intervenção parcial ou total do perito na construção das redes, estabelecendo dependências ou ainda atribuindo valores às tabelas de probabilidades condicionais.

Como principais inconvenientes do GeNIe:

- O espaço necessário para armazenar todas as probabilidades de problemas mais complexos é muito grande, bem como o tempo exigido para calcular essas probabilidades;
- Exigente em termos de conhecimentos teóricos que se encontram na base de selecção dos métodos e algoritmos a utilizar;
- Trabalha apenas com variáveis discretas, muito embora permita a discretização;
- Não permite o manuseamento / alteração da base de dados.

5.2.AUDITORIA NA BANCA USANDO REDES BAYESIANAS NORMAIS

A auditoria interna de um banco deve acompanhar a evolução dos principais processos que tem a seu cargo, não apenas em termos de evolução da aplicação do Acordo de Basileia II e das Directivas comunitárias e orientações do Banco de Portugal sobre a matéria, mas também dos modelos de medição de risco adoptados internamente. Na perspectiva deste estudo, a auditoria enquadra-se num processo de recolha e avaliação de evidências para determinar até que ponto objectivos organizacionais são alcançados. Se bem que a opinião de um auditor aumente a credibilidade afecta a uma organização, essa opinião pode não transmitir segurança suficiente quanto à futura viabilidade de um balcão, nem quanto à eficiência ou eficácia com que a gerência irá conduzir os negócios desse balcão. A credibilidade instaurada pode ser reforçada se medido o impacto dos trabalhos realizados, nomeadamente através da análise de indicadores do grau de execução das recomendações efectuadas.

Neste capítulo introduz-se a ferramenta utilizada na análise e aduz-se a descrição de dois casos que se servem de Redes Bayesianas Normais para estudo dos problemas abordados.

5.2.1. Estudo de um caso teórico – Falhas de Crédito em contas com indicadores de risco

Neste trabalho são analisadas as técnicas para aprendizagem da topologia (relacionamento causal) e dos parâmetros (distribuições de probabilidades condicionais) da uma Rede Bayesiana, a partir de dados históricos de certos aspectos das operações de crédito, fornecidos por uma instituição bancária. Os aspectos observados nesse domínio, para fins deste trabalho, serão detalhados no capítulo 3.3, que trata de um estudo prático de caso. Antes, importa descrever e perceber o funcionamento da ferramenta que se adoptou para a análise, colocando em prática os fundamentos teóricos já abordados e utilizando para tal um cenário mais simples.

Ora, considerando um modelo cujas variáveis são as descritas na Tabela 6, pretende-se, a partir daqui, trabalhar o aplicativo GeNIe e explorar o seu conteúdo tirando partido dos exercícios, testes e resultados obtidos para realização de análises mais complexas.

Tabela 6 – Variáveis do modelo

Variáveis	Estados possíveis	Observações
SRC	Sim / Não	Conta ao serviço da recuperação de crédito
CDEVOLV	Sim / Não	Conta com cheque(s) devolvidos
CCRED_LN	Sim / Não	Conta com cartão de crédito em lista negra
Falha_Crédito	Sim / Não	Deteção de falha por parte da auditoria

A base de dados possui 2.162 registos relativos à análise das contas de uma agência bancária. Não existem variáveis contínuas pelo que a discretização das mesmas não será alvo de análise. O GeNIe impõe a discretização das variáveis e o tratamento dos dados em falta para que seja possível construir a rede automaticamente através dos dados. A Figura 16 ilustra a base de dados importada pelo GeNIe de um ficheiro de texto previamente construído.

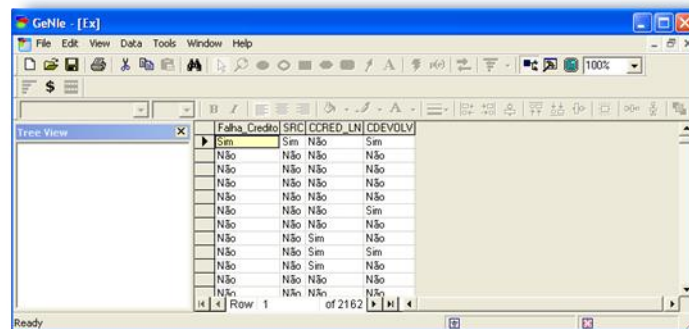


Figura 16 – Importação dos dados

Este exemplo, pela simplicidade que reveste, foi utilizado para mais facilmente se perceber as diferenças existentes entre os vários métodos de aprendizagem da estrutura da rede, o que facilmente se evidencia graficamente. Para isso, utilizando a mesma base de dados, foram utilizados os vários métodos que o GeNIe disponibiliza para aprendizagem da rede (*Learn New Network*), conforme evidenciado na Figura 17. O GeNIe permite ainda seleccionar, de entre todas as variáveis de uma base de dados, aquelas que se pretende que sejam consideradas para a construção da rede.

5.2. Auditoria na Banca usando Redes Bayesianas Normais

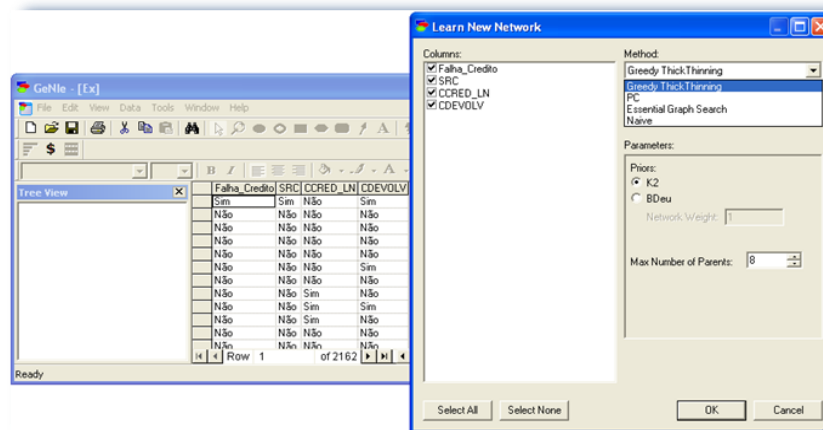


Figura 17 – Aprendizagem da estrutura da rede no GeNIe

A Figura 18 ilustra os resultados obtidos com através da aplicação dos diferentes métodos que constam do aplicativo. Note-se que uma das características deste processo de aprendizagem é a que se prende com o tratamento e/ou inferência realizada sobre “classes”, o GeNIe disponibiliza métodos quer de aprendizagem supervisionada (*Naive*), quer de aprendizagem não supervisionada (*Greedy Thick Thinning*, *PC* e *Essencial Graph Search*) o que possibilita efectuar qualquer tipo de inferência sobre as redes. Essa inferência será posteriormente realizada por intermédio de um algoritmo de propagação que, dada uma ou mais evidências, actualizará as probabilidades de cada uma das variáveis (nós) restantes da rede.

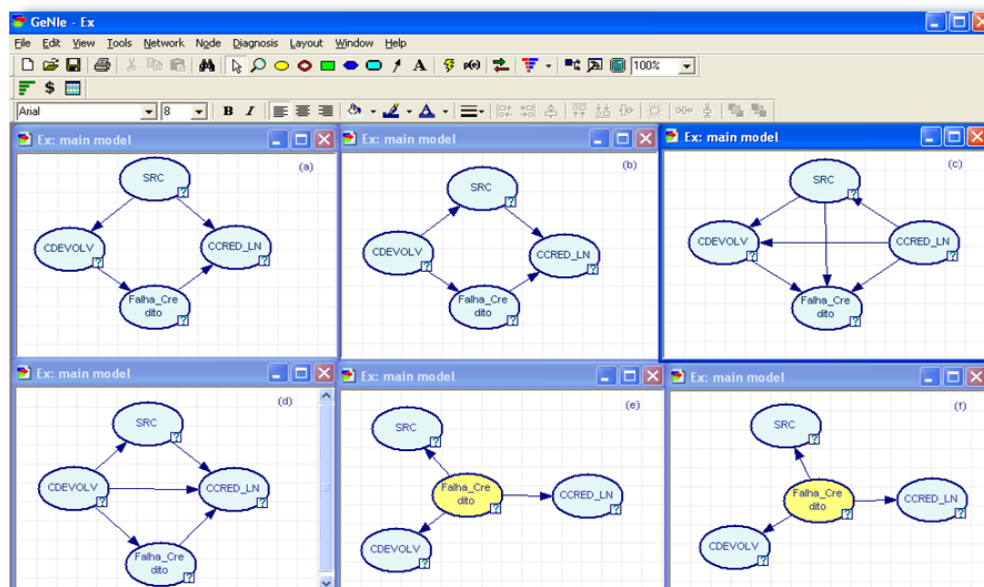


Figura 18 – Aplicação de vários métodos de aprendizagem no GeNIe

Os métodos utilizados para a construção da rede foram, respeitando a legenda atribuída a cada caso:

- a) Método: *Greedy Thick Thinning*; Parâmetros: K2
 - b) Método: *Greedy Thick Thinning*; Parâmetros: BDeu
 - c) Método: PC
 - d) Método: *Essencial Graph Search*
 - e) Método: Naive; Parâmetros: K2
 - f) Método: Naive; Parâmetros: BDeu
- Parâmetros: BDeu

Tal como já referido, uma vez gerada a Rede Bayesiana e as respectivas tabelas de probabilidades, que se podem visualizar no GeNIe através de duplo clique em cada um dos nós, é possível efectuar inferências por intermédio dos algoritmos de propagação que este aplicativo disponibiliza e que se permitem consultar tal como demonstra a Figura 19.

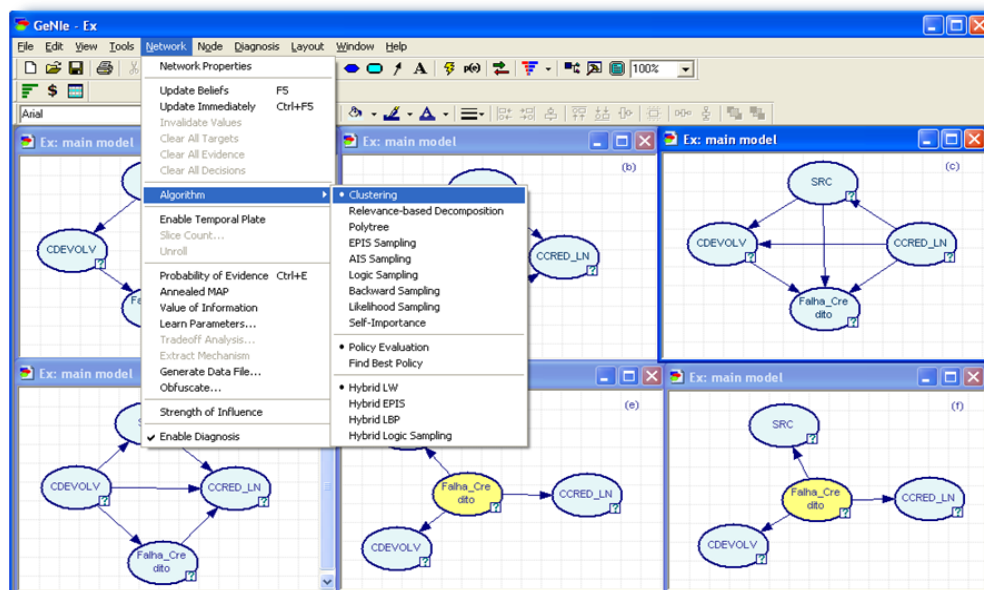


Figura 19 – Algoritmos de propagação dos dados no GeNIe

Como se pode verificar, a aplicação de diferentes métodos origina topologias de rede significativamente distintas que, só por si, podem levar um leigo na matéria a conclusões diferentes orientadas para a resposta a um problema idêntico. Por aqui se realça a importância que um especialista assume, senão directamente, então pelo menos

pelo papel que deve desempenhar na condução das abordagens efectuadas, quer em teoria, quer no que à intuição empírica lhes diga respeito.

A aplicação dos algoritmos de aprendizagem deve ter em consideração a arquitectura da rede, na medida em que deve ser exigente quanto a determinados pressupostos que podem não estar a ser cumpridos. As *polytrees* são um exemplo destes casos, que exigem – Figura 20 – redes simplesmente conectadas, conforme já mencionado na prévia abordagem teórica deste trabalho.

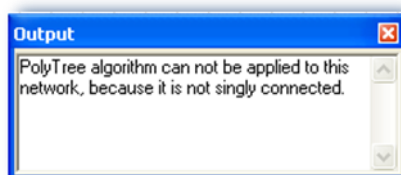


Figura 20 – Alertas no GeNIe

Relativamente à concepção gráfica, os nós podem ser ilustrados como ícones, como aconteceu até aqui, ou como forma de gráfico de barras (ver Figura 21), o que possibilita uma visualização imediata mais completa das informações em causa, isto é, para além da variável, permite visualizar os estados possíveis da mesma e respectivos valores das probabilidades associadas. As figuras que se seguem optam por este tipo de apresentação.

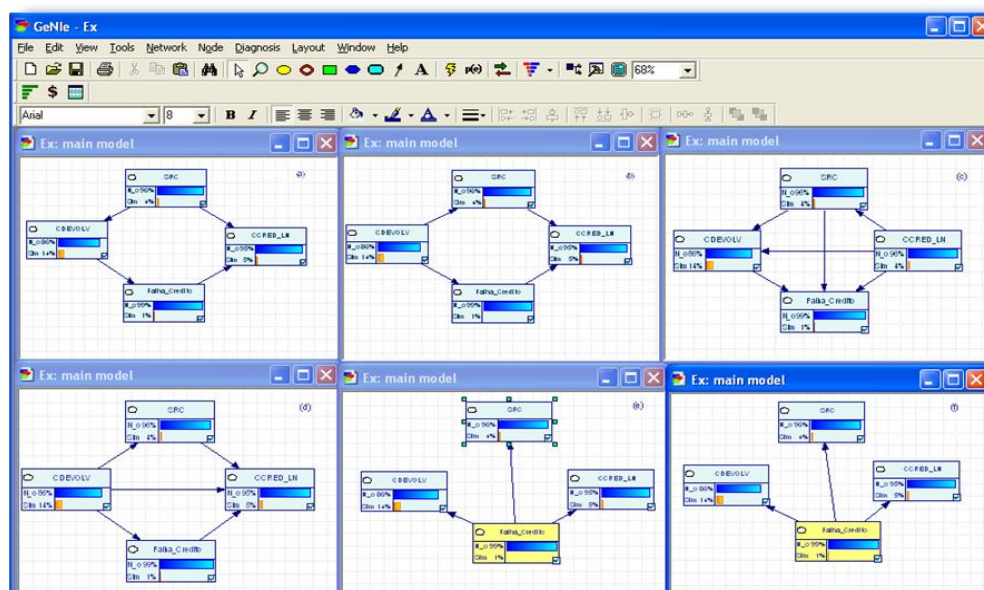


Figura 21 – Gráficos de Barras no GeNIe

A capacidade de realizar simulações, utilizando o conhecimento adquirido e incorporado pelas Redes Bayesianas, é uma das características mais relevantes de ferramentas como o GeNIe. A ideia consiste em definir quais os efeitos da definição de estados específicos em determinadas variáveis, nas restantes variáveis, tal como ilustra a Figura 22.

No caso da Figura 22, a totalidade de falhas de crédito detectadas propaga-se pelas restantes variáveis do modelo, permitindo averiguar quanto ao âmbito de incidência das mesmas. Seria de esperar que as probabilidades associadas aos estados *Sim* das restantes variáveis, o que de facto, aconteceu. Isto porque é mais provável detectar falhas em operações com essas características, do que na ausência delas. O facto de o estado *Não* assumir ainda assim, em alguns casos, probabilidades mais elevadas prende-se com o facto de a auditoria trabalhar com amostras, e desta forma a variável *Falhas_Credito* não implicar a análise integral de todas as observações em causa.

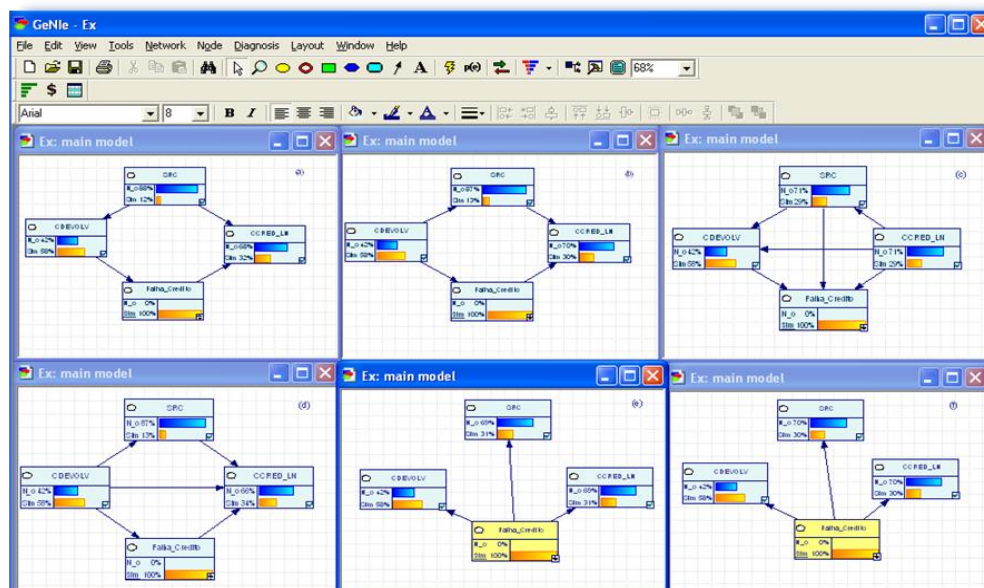


Figura 22 – Simulações no GeNIe

5.2.2. Estudo de um caso prático – Impacto da Auditoria em Operações de Crédito

As autoridades financeiras impõem fortes pressões sobre os bancos, para que estes conduzam os seus negócios de uma forma firme e eficiente. Deve-se garantir que as operações realizadas sejam conduzidas prudentemente e de acordo com as normas, políticas e estratégias da instituição, que as transacções efectuadas cumpram os regulamentos e sejam controladas e que o controlo produza resultados efectivos.

A auditoria de crédito em empresas da área financeira é uma actividade pró-activa, ou seja, procura descobrir e apontar falhas que futuramente poderão comprometer a qualidade do crédito. Neste domínio, pretende-se avaliar o impacto / influência da auditoria no âmbito de operações de crédito realizadas por uma instituição, quer numa vertente qualitativa – existência ou não desse impacto e o que influencia – quer numa vertente quantitativa – com que intensidade ocorre.

Esta abordagem assume o intuito de determinar se a evolução de um cenário que foi alvo de intervenção por parte de agentes auditores é favorável, ainda que esteja exposta a alguns tipos de fragilidades, nomeadamente o não confronto com a evolução de situações que não foram auditadas. Ainda assim, propõe-se fundamentar uma actuação nem sempre possível de quantificar de forma exacta, não fosse a existência de conjunturas internas e externas a interferir no processo, impedindo a assunção totalmente isolada das conclusões que possam vir a ser retiradas. Não é de desprezar, no entanto, a necessidade real da análise das estratégias implementadas e da averiguação quanto à sua adequação ou melhor forma de agir, tendo como base um estudo concreto de casos.

Dado o vasto âmbito de actuação de um departamento como o de auditoria e a diversidade de temáticas a áreas envolvidas, não seria viável avaliar o seu desempenho na globalidade. Uma rede construída com todas as principais variáveis em causa não se revelou sustentável nem tão pouco legível. Por conseguinte a modelação da análise foi efectuada através do estudo de uma das várias áreas possíveis de actuação por parte da auditoria – operações de crédito.

Este trabalho serve-se de uma ferramenta capaz de descobrir conhecimento contido num conjunto de dados e gerar uma Rede Bayesiana que o evidencie.

O caso em estudo segue os seguintes passos: tratamento dos dados, caracterização das variáveis, métodos e algoritmos de aprendizagem utilizados, interface gráfico do modelo, descrição da topologia / arquitectura da rede, simulação, propagação e discussão dos resultados.

5.2.2.1. Tratamento dos dados

Foi extraída uma amostra aleatória de 4 balcões de uma instituição bancária, considerando que estes foram alvo de auditoria 4 vezes e cuja última auditoria efectuada ocorreu entre 2007 e 2008. Foi ainda realizada uma análise preliminar para determinar as variáveis representativas da análise e respectivos estados que apresentavam indícios de serem relevantes, dada a existência de um elevado número de estados para essas variáveis. Contou-se desta forma com a intervenção de um perito na selecção das variáveis relevantes para o estudo.

A base de dados contem 849 registos relativos a operações de crédito e toda a informação dos campos se reporta à data da última auditoria realizada, excepto a informação que descreve o estado actual dessas operações que se reporta a Março de 2009.

Foram ainda eliminados, e porque não interferem na análise, números de contrato, dados pessoais ou identificadores, tais como Nome, Contribuinte, Códigos de Agência, entre outros, com o objectivo de preservar a confidencialidade dos dados e o sigilo bancário.

5.2.2.2. Caracterização das variáveis

Assim como aconteceria na maior parte dos *softwares* integrados neste contexto, a implementação de uma Rede Bayesiana no GeNIe que utilize um determinado processo de aprendizagem para a modelação das variáveis de um problema, não é a tarefa mais difícil no processo. A tarefa mais complicada é a de mapear as variáveis correctas para poder atingir a respostas desejadas. A Tabela 7 caracteriza as variáveis envolvidas nesta análise.

5.2. Auditoria na Banca usando Redes Bayesianas Normais

Tabela 7 – Caracterização das variáveis do caso em estudo

Nome da variável	Estados possíveis	Descrição
Balcao	A B C D	Identificação do balcão gestor da conta associada à operação
Classe_credito	I II III IV	Montantes de crédito atribuídos a cada operação, por ordenação crescente de valores, ou seja, Valores da Classe I são inferiores aos da Classe II, que são inferiores aos da Classe III, que são inferiores aos da classe IV
Endividamento	sim nao	Se a operação evidencia valores em mora
SRC	sim nao	Se o processo se encontra no serviço de recuperação de dívidas
Analisado	sim nao	Identifica as operações analisadas pela auditoria
Recomend	sim nao	Identifica as operações que foram alvo de recomendações ao balcão
Sit_ult_aud	N O V	Situação da operação aquando da realização da última auditoria, N significa normal, V significa com crédito vencido (valor em dívida) e O significa outra situação.
Situacao_Actual	E N O V	Situação actual da operação, N significa normal, V significa com crédito vencido (valor em dívida) e O significa outra situação. E traduz as contas que actualmente se encontram encerradas.

5.2.2.3. Métodos e algoritmos de aprendizagem utilizados

O método utilizado foi o *Greedy Thick Thinning*, algoritmo *BDeu* com máximo número de pais 3 e peso 1. O algoritmo de propagação utilizado foi o assumido por defeito pelo GeNIe, o *Clustering*, no entanto experiências efectuadas permitiram concluir que os resultados obtidos com a utilização de outros algoritmos não diferem consideravelmente.

5.2.2.4. Interface gráfico do modelo

A Figura 23 ilustra a Rede Bayesiana gerada pelo GeNIe para o problema em estudo sem recurso a qualquer tipo de intervenção de peritos no que se refere a probabilidades associadas ou relações estabelecidas:

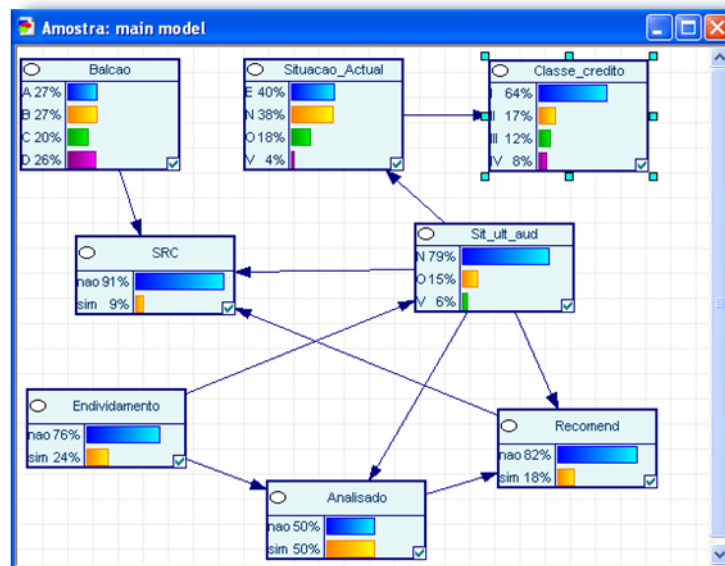


Figura 23 – Rede Bayesiana gerada pelo GeNIe

Cada nó tem, conforme mencionado anteriormente, uma tabela de probabilidades associada. Ilustre-se – Figura 24 – como exemplo a variável **Situacao_Actual**, que depende condicionalmente apenas da variável **Sit_ult_aud**. Esta tabela, tal como as relações estabelecidas entre as variáveis, podem ser alteradas manualmente.

Sit_ult_aud	N	O	V
E	0.44138697	0.19854881	0.39482759
N	0.45915597	0.016490765	0.29137931
O	0.059353406	0.76846966	0.29137931
V	0.040103653	0.016490765	0.022413793

Figura 24 – Tabela de Probabilidades da variável Situacao_Actual

5.2.2.5. Descrição da topologia / arquitetura da rede automática

A rede construída nos moldes descritos, na ausência de intervenção por parte de peritos produziu resultados interessantes na medida em que se verifica:

- O **Balcao** apresenta-se como uma variável ligada à variável **SRC**, no entanto, o Balcão, por si só, não é causa directa de nenhuma das outras variáveis;
- A variável **Analizado** tem como pais as variáveis **Endividamento** e **Sit_ult_aud**, sendo que também faria sentido ter também como pais, as variáveis **SRC** e **Classe_credito**;
- A variável **Recomend** tem como pais as variáveis **Analizado**, e **Sit_ult_aud**; de facto, só se efectuam recomendações relativamente às operações que foram analisadas, a variável **Sit_ult_aud** também pode ser considerada causa para recomendação, ainda que não suficiente;
- A variável **Endividamento** não tem pais; surge como pai das variáveis **Sit_ult_aud** e **Analizado**, o que seria de esperar;
- A variável **SRC** tem como pais as variáveis **Recomend**, **Balcao** e **Sit_ult_aud**; se por um lado, esta está, de facto, dependente da situação das operações, por outro lado as relações estabelecidas com as variáveis **Recomend** e **Balcao** não seriam tão lineares;
- A variável **Classe_credito** surge relacionada apenas com a **Situacao_actual**, o que permite relacionar a situação actual das operações com o montante financiado;
- A variável **Sit_ult_aud** tem como pai a variável **Endividamento**, variável esta que, de facto, define a situação do contrato;
- A variável **Situacao_actual** tem como pai a variável **Sit_ult_aud**, muito embora a sua definição não devesse depender apenas desta variável, mas também, para efeitos deste estudo, da variável **Recomend**.

No que respeita aos valores das probabilidades associadas, verifica-se que, como seria de esperar as probabilidades associadas à variável **Classe_credito** diminuem com o aumento do montante financiado, ou seja, é natural que existam mais operações de montantes menos elevados. Também faz sentido a maior parte das operações se encontrarem em situação normal, não se encontrarem com valores em dívida nem no serviço de recuperação de dívidas e não serem alvo de recomendações, e que parte das operações já se encontrem encerradas.

5.2.2.6. Topologia da rede com intervenção manual

Tecidas as considerações relativamente ao método automático de construção da rede, recorreu-se à intervenção de um perito no sentido de averiguar que tipo de estrutura representaria melhor o problema em estudo. O resultado encontra-se evidenciado pela Figura 25.

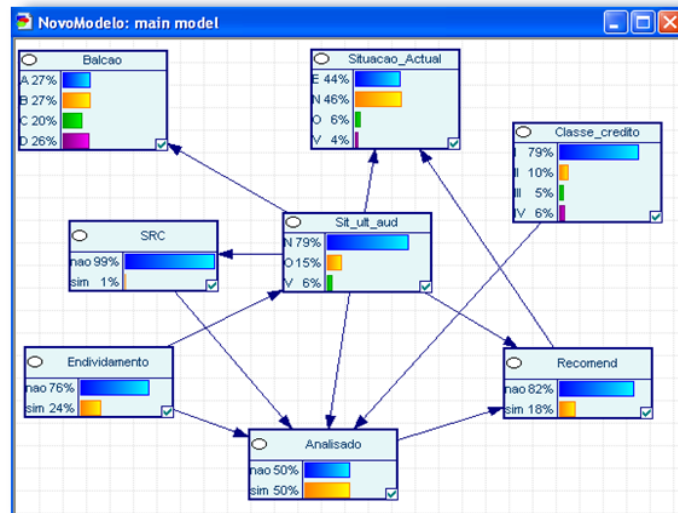


Figura 25 – Nova rede com intervenção do perito

As alterações efectuadas na estrutura fazem com que as tabelas de probabilidades se alterem perante diferentes relações de dependência. Neste caso, não houve intervenção directa por parte do perito nas tabelas, o GeNIe incumbiu-se de calcular a nova distribuição de probabilidades.

A variável **Situacao_Actual** ilustrada na Figura 24, depende agora, para além da **Si_ult_aud**, da variável **Recomend**, pelo que a respectiva tabela de probabilidades sofre alterações, tal como se pode verificar na Figura que se segue:

		N		O		V	
		nao	sim	nao	sim	nao	sim
Sit_ult_aud	E	0.44138697	0.44138697	0.44138697	0.44138697	0.44138697	0.44138697
	N	0.45915597	0.45915597	0.45915597	0.45915597	0.45915597	0.45915597
	O	0.059353406	0.059353406	0.059353406	0.059353406	0.059353406	0.059353406
	V	0.040103653	0.040103653	0.040103653	0.040103653	0.040103653	0.040103653
Recomend	E	0.44138697	0.44138697	0.44138697	0.44138697	0.44138697	0.44138697
	N	0.45915597	0.45915597	0.45915597	0.45915597	0.45915597	0.45915597
	O	0.059353406	0.059353406	0.059353406	0.059353406	0.059353406	0.059353406
	V	0.040103653	0.040103653	0.040103653	0.040103653	0.040103653	0.040103653

Figura 26 – Tabela de Probabilidades da variável **Situacao_Actual** após a intervenção do perito

5.2.2.7. Simulação, propagação e discussão dos resultados

Uma vez gerada a Rede Bayesiana e as respectivas tabelas de probabilidades condicionais, importa efectuar inferências por intermédio do algoritmo de propagação utilizado, de forma a actualizar todas as probabilidades de acordo com as evidências dadas como entrada. Serão de seguida analisadas duas situações.

- Situação 1 – Variável: *Analisado*; Estado: **nao** e **sim**.

Os resultados obtidos pelo GeNIe encontram-se evidenciados na Figura 27.

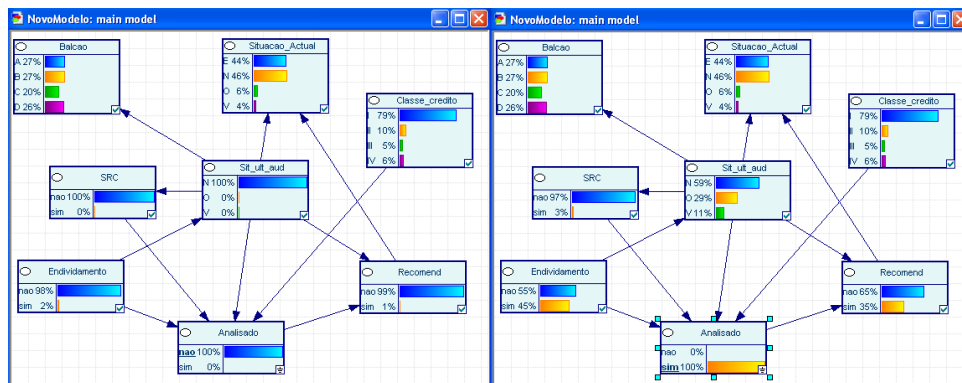


Figura 27 – Actualização Bayesiana para novas evidências da variável *Analisado*

Aparentemente parece não resultar prejuízo da não análise das operações em causa uma vez que se tratavam de operações em situação normal, sem indícios de risco de incumprimento associado. Ainda assim, verifica-se que o estado **V** da situação actual aumentou, o que significa um aumento do número de operações com valores em dívida. Pelo contrário, as operações que foram objecto de análise apresentam actualmente uma diminuição da percentagem afecta a esse estado.

- Situação 2 – Variável: *Recomend* ; Estado: **nao** e **sim**.

A concretização da análise das operações é a elaboração de recomendações. A Figura 28 evidencia os resultados obtidos considerando como evidências os dois estados possíveis da variável *Recomend*.

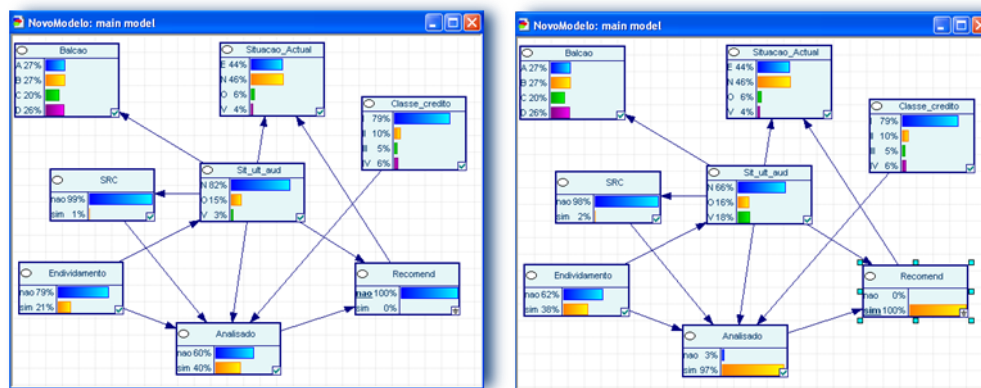


Figura 28 - Actualização Bayesiana para novas evidências da variável Recomend

Como se pode verificar, o impacto das recomendações efectuadas em termos do estado **V** da variável *Situacao_Actual* é o desejável, na medida em que se verifica uma diminuição, face ao valor relativo ao estado **V** da variável *Sit_ult_aud*, o que corrobora o impacto positivo da estratégia levada a cabo pela equipa da auditoria, ou pelo menos, não contraria os objectivos estratégicos da organização.

5.3.AUDITORIA NA BANCA USANDO REDES BAYESIANAS DINÂMICAS

O sucesso da auditoria interna depende muito da interacção com os gestores, de modo que as recomendações formuladas no relatório sejam facilmente entendíveis e haja um sucesso na sua implementação. Assim, as recomendações formuladas devem criar expectativas fundadas que vão contribuir para a mitigação dos riscos e acrescentar valor à empresa. Este processo não é instantâneo, obedece a intervenções que ocorrem ao longo do tempo, muito embora a auditoria tenha como ponto fundamental promover a implementação de recomendações identificadas no mais curto prazo possível. Importa, pois introduzir neste estudo uma dinâmica temporal, que será feita através da construção de uma Rede Bayesiana Dinâmica.

Neste capítulo utiliza-se o GeNIe para representar uma Rede Bayesiana Dinâmica com o objectivo de ilustrar em que medida as recomendações efectuadas por uma equipa de auditoria têm impacto na situação actual em termos de perfil de um conjunto de operações de crédito.

5.3.1. Estudo de um caso prático – Impacto da Auditoria em Operações de Crédito

A concessão de empréstimos é a actividade básica da maioria dos bancos. Para desenvolverem suas actividades de empréstimo, os bancos precisam fazer avaliações da capacidade de crédito dos seus clientes ou potenciais clientes. Consequentemente, um risco importante que os bancos enfrentam é o de crédito ou a falha por parte do mutuário ou proponente de crédito em honrar os seus compromissos contratuais. Este risco aplica-se não apenas a empréstimos, mas também a outras operações acessórias, tais como garantias, aceites, etc. As operações de crédito contratadas recebem, no momento de sua contratação, uma classificação de risco que, em síntese, representa a expectativa de recebimento do capital emprestado mediante a classificação em termos de um perfil apurado. Essa classificação é efectuada com base em critérios consistentes e verificáveis, amparada por informações internas e externas, contemplando, pelo menos, os aspectos relativos aos clientes e garantes (tais como a sua situação económico-financeira, idade, grau de endividamento, pontualidade e atrasos nos pagamentos, limite de crédito atribuído, sector onde actua...) bem como aspectos relativos à operação (tais como a natureza e finalidade da transacção, prazo, características, suficiência e liquidez das garantias, valor do crédito concedido...).

A finalidade do controlo interno da instituição financeira é assegurar que os negócios sejam conduzidos de maneira prudente e de acordo com políticas e estratégias estabelecidas pelo Conselho de Administração; que as transacções somente sejam efectuadas mediante autorização competente; que os activos sejam protegidos e os passivos controlados; que a contabilidade e outros registos forneçam informações completas, precisas e oportunas; e que a administração seja capaz de identificar, avaliar, administrar e controlar os riscos do negócio. A ocorrência de falhas, fraudes ou não cumprimento de normativos legais pelos bancos, são alvo de preocupação das autoridades financeiras que impõem fortes pressões sobre os bancos para que conduzam os seus negócios de modo firme e eficiente.

O auditor interno assume então o papel de um profissional que pretende responder às expectativas da Administração sobre os maiores riscos da empresa, observando, aconselhando e esclarecendo os responsáveis envolvidos, persuadindo-os a implementar as acções correctivas necessárias. Assim, e no que se consubstancia apenas na vertente

de crédito, o auditor estabelece um conjunto de recomendações, quer de carácter específico, quer de carácter geral, tais como: que o balcão cumpra com o que se encontra estipulado no Regulamento Geral de Crédito, que o balcão não autorize o pagamento de cheques e/ou não permita a regularização de contas que apresentam descobertos, que o Balcão cumpra escrupulosamente com o estipulado nos despachos de aprovação das operações de crédito e que proceda ao acompanhamento rigoroso dos seus clientes de forma a evitar situações de incumprimento, regularização imediata do descoberto e do crédito em situação irregular. Importa que estas recomendações, mais do que destinadas a situações pontuais encontradas, produzam efeitos em operações futuras, que os seus destinatários percebam, assimilem e as coloquem em prática de forma a melhorar a sua carteira de clientes e, paralelamente, trabalharem em termos da melhoria do seu desempenho.

O caso apresentado neste capítulo consiste em averiguar se, mediante a intervenção dos auditores, consubstanciada nas recomendações que vão realizando no decurso das suas auditorias, a classificação global do perfil atribuído às operações sofre ou não alterações, e em que sentido tal acontece.

5.3.1.1. Tratamento dos dados

Foi extraída uma amostra relativa a operações de crédito realizadas em 10 balcões de uma instituição bancária, considerando que estes foram alvo de auditoria 3 vezes e cuja última auditoria efectuada ocorreu em 2008. O ciclo de auditoria considerado foi de 2 anos, para todos os balcões, independentemente da classificação atribuída a cada um deles.

A base de dados contem 4.820 registos relativos a operações de crédito e, para cada operação, foi recolhida informação quanto ao seu perfil e quanto ao facto de ter sido ou não alvo de recomendação por parte da auditoria.

Foram ainda eliminados, e porque não interferem na análise, números de contrato, dados pessoais ou identificadores, tais como Nome, Contribuinte, Códigos de Agência, bem como todos os que contribuíram para o apuramento da classificação do perfil da

operação, mais uma vez, com o objectivo de preservar a confidencialidade dos dados e o sigilo bancário.

5.3.1.2. Caracterização das variáveis

As variáveis mapeadas para introduzir a análise do problema proposto através da representação de uma Rede Bayesiana Dinâmica foram duas e encontram-se caracterizadas na tabela 8:

Tabela 8 – Caracterização das variáveis em estudo

Nome da variável	Estados possíveis	Descrição
Perfil	Bom	Classificação de cada operação. Reflecte o cálculo de uma pontuação calculada com base em indicadores económico-financeiros e sócio-demográficos associada a políticas de risco adoptadas internamente pela instituição.
	Fraco	
	Mau	
Recom	S	s – se foi efectuada recomendação, n – caso contrário
	N	

5.3.1.3. Metodologia utilizada

A metodologia utilizada nesta análise difere da anteriormente apresentada na construção de Redes Bayesianas ditas normais uma vez que, introduzida a dinâmica temporal, o GeNIe não sugere uma reprodução automática da rede. Assim sendo, envolve os seguintes passos:

- Importação da base de dados com informação referente às variáveis em estudo mais os dados referentes à variável **rec1** (recomendação ou não aquando da auditoria efectuada dois anos depois ao balcão que contratualizou a operação);
- Determinação da tabela de probabilidades que irá ser utilizada na construção da Rede Bayesiana Dinâmica;
- Definição e imposição do relacionamento das variáveis **perfil** e **rec**;
- Introdução do painel temporal –*Temporal Plate* – e definição do número de fatias de tempo – *slices* – a considerar;
- Alteração da tabela de probabilidades relativa ao período $t=1$;

- Simulação e estudo dos resultados obtidos.

O método considerado na construção da rede foi o *Greedy Thick Thinning* e o algoritmo, o *BDeu* com máximo número de pais 2 e peso 1. Os algoritmos de propagação utilizados foram, tal como descrito na discussão dos resultados, o *Clustering* (assumido por defeito pelo GeNIe) representando o tipo de inferência exacta e o *Likelihood Sampling* em representação do tipo de inferência aproximada.

5.3.1.4. Interface gráfico do modelo

Tal como já referido anteriormente, a construção das Redes Bayesianas pode ser realizada utilizando apenas o conhecimento prévio do domínio, a partir dos dados ou da combinação de ambos. No estudo deste caso, a construção da rede será efectuada utilizando dados históricos da instituição.

Os dados foram importados no GeNIe, numa primeira fase para determinação da tabela de probabilidades, e posteriormente, para construção da Rede Bayesiana Dinâmica.

Na Figura 29 pode visualizar-se a amostra que servirá de base à construção do modelo e que foi importada após recolha e cruzamento da informação, uma vez que a mesma se encontra armazenada em diferentes arquivos e plataformas.

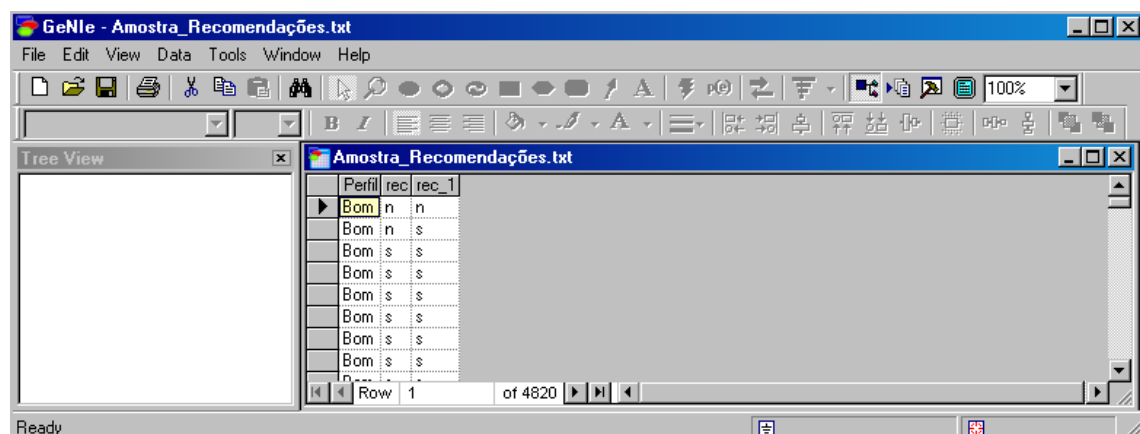


Figura 29 – Importação de dados pelo GeNIe

Segue-se a escolha do método e do algoritmo utilizado para “aprender” a rede. O menu *Data\Learn New Network...* abre a janela que permite efectuar essa selecção:

5.3. Auditoria na Banca usando Redes Bayesianas Dinâmicas

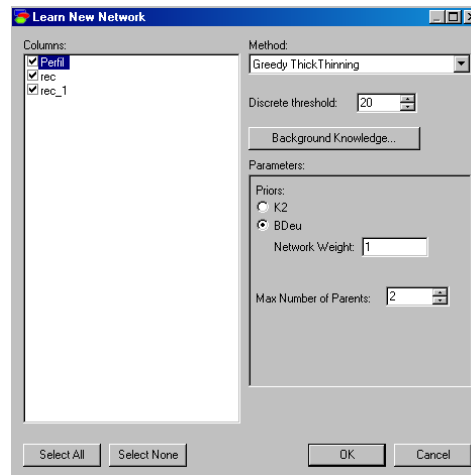


Figura 30- Janela *Learn New Network*

É importante salientar que, neste caso o sentido das relações entre as variáveis do modelo é conhecida e não é flexível, daí a necessidade de se forçar essas relações. O botão *Background Knowledge*, visível na Figura 30, abre o *Knowledge Editor* – Figura 31 – onde se pode definir o posicionamento em termos temporais e o sentido da relação das variáveis em estudo. Os dados referentes à variável *rec_1* referem-se a um período posterior pelo que o respectivo nó se posiciona na primeira fila de tempo – Temporal Tier 1:

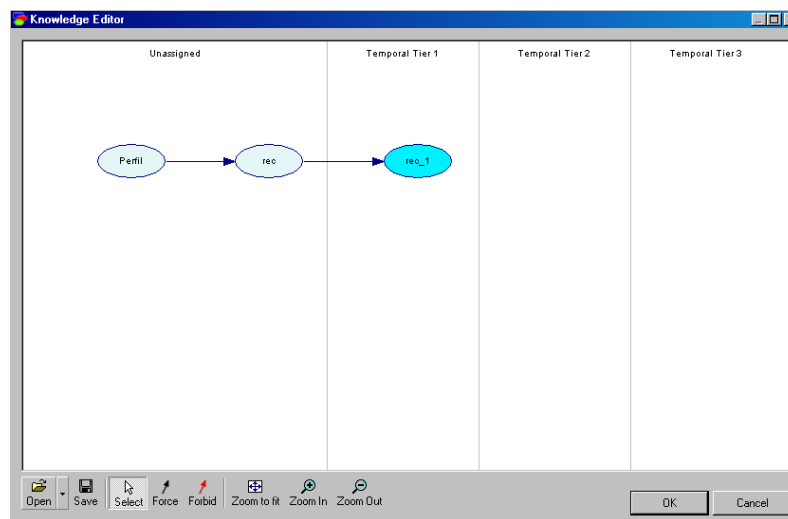


Figura 31 – Janela *Knowledge Editor*

No processo de construção de Redes Bayesianas, tal como já referido, primeiramente é gerada a Rede Bayesiana que representa as relações entre as variáveis do problema, seguindo-se a especificação de como a distribuição de probabilidade de cada nó será

5.3. Auditoria na Banca usando Redes Bayesianas Dinâmicas

representada. Assim sendo, a Figura 32 exibe os cálculos apurados pelo GeNIe para a $P(\text{rec} \mid \text{perfil})$, ou seja, as probabilidades associadas ao facto de se efectuarem ou não recomendações dado um determinado perfil:

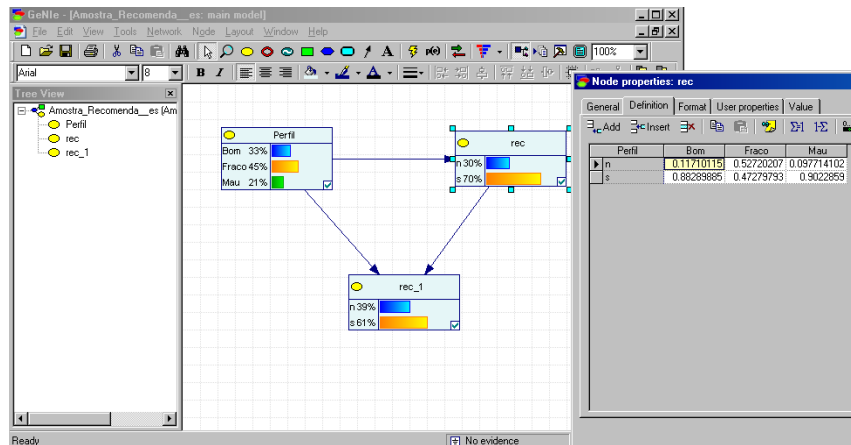


Figura 32 – Tabela de probabilidades da variável rec

Analogamente, clicando no nó **rec_1**, é possível observar, na Figura 33, a tabela de probabilidades – $P(\text{Rec1} \mid \text{Rec} \wedge \text{Perfil})$ – que servirá de base para a construção da Rede Bayesiana Dinâmica e que reflecte o cálculo das probabilidades temporais, no sentido em que traduz a probabilidade de se efectuar recomendações num determinado período dado que já haviam ou não sido efectuadas antes e dado que se está perante a análise de determinado perfil:

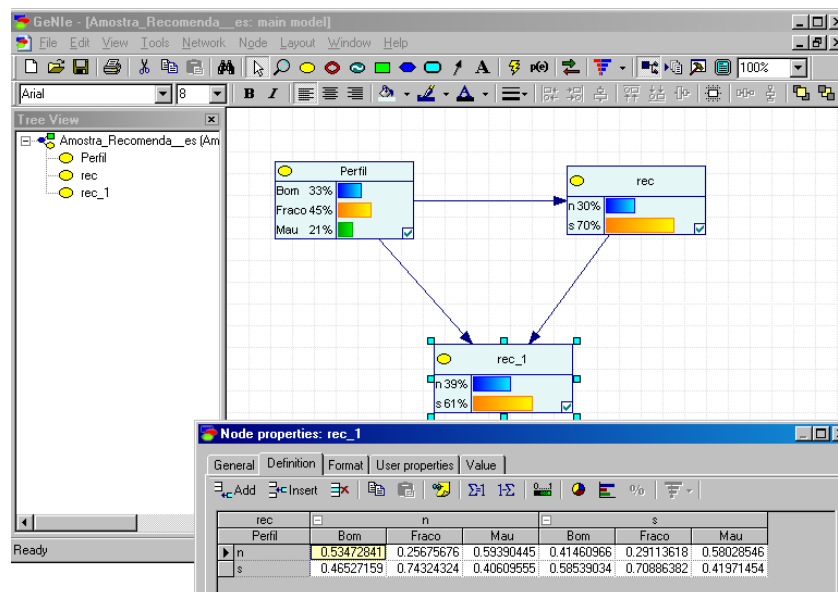


Figura 33 – Tabela de probabilidades da variável rec_1

5.3. Auditoria na Banca usando Redes Bayesianas Dinâmicas

A partir deste momento, inicia-se a construção da Rede Bayesiana Dinâmica propriamente dita, voltando a aprender a rede, agora apenas considerando as variáveis **Perfil** e **rec**, mantendo o método e o algoritmo seleccionado, bem como o *Background Knowledge* anteriormente definido. Note-se que não houve necessidade de importar novamente os dados uma vez que o GeNIe permite a selecção de variáveis para a construção da rede, tal como se pode observar na Figura 34 na área conferida a *Columns*:

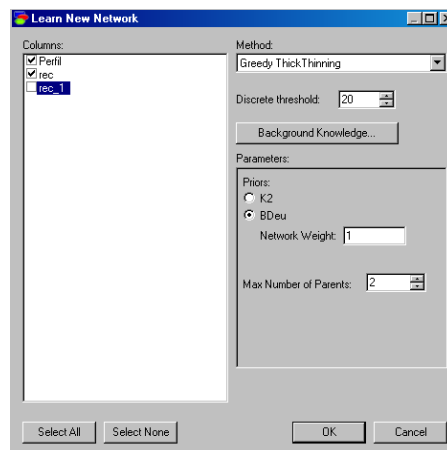


Figura 34 – Janela *Learn New Network*

A selecção da opção *Enable Temporal Template* no menu *Network* introduz a componente temporal na construção da rede. Agora, para além da área de construção da rede que alberga todas as variáveis estáticas do modelo, introduz-se este painel, constituído por:

Initial conditions: área da rede onde nós de raiz – *anchor nodes* – são armazenados;

Temporal plate: área da rede onde os “nós temporais” são armazenados. Só aqui é possível atribuir arcos temporais às variáveis seleccionadas. É também nesta área que se irá definir o número de períodos – *time-slices* – em análise que irão ser alvo de inferência;

Terminal conditions: área da rede onde os nós terminais – *terminal nodes* – são armazenados.

5.3. Auditoria na Banca usando Redes Bayesianas Dinâmicas

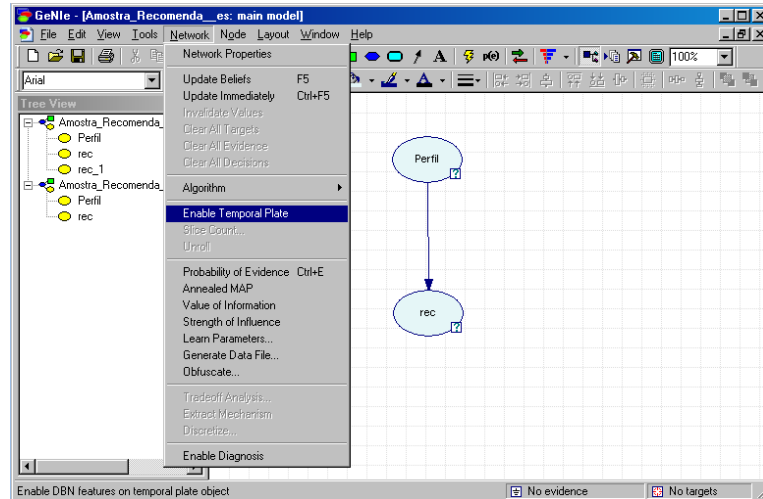


Figura 35 – Menu *Enable Temporal Plate*

Uma vez activado o painel, o nó temporal em análise – **rec** – deverá ser arrastado para o *Temporal plate* onde se adiciona o arco temporal. A selecção da opção *Order 1* significa dizer que a Rede Bayesiana Dinâmica em estudo representa um processo de Markov de 1ª ordem.

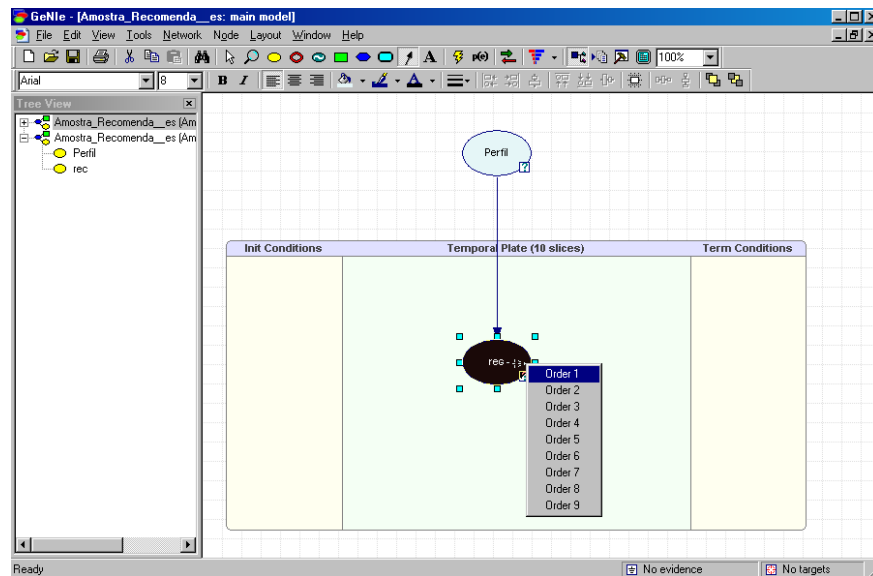


Figura 36 – Desenho do arco temporal

Falta indicar qual o número de períodos considerados na análise. Este número vai reflectir o número de auditorias efectuadas em ciclos temporais de 2 anos, que no caso em análise é igual a 3:

.

5.3. Auditoria na Banca usando Redes Bayesianas Dinâmicas

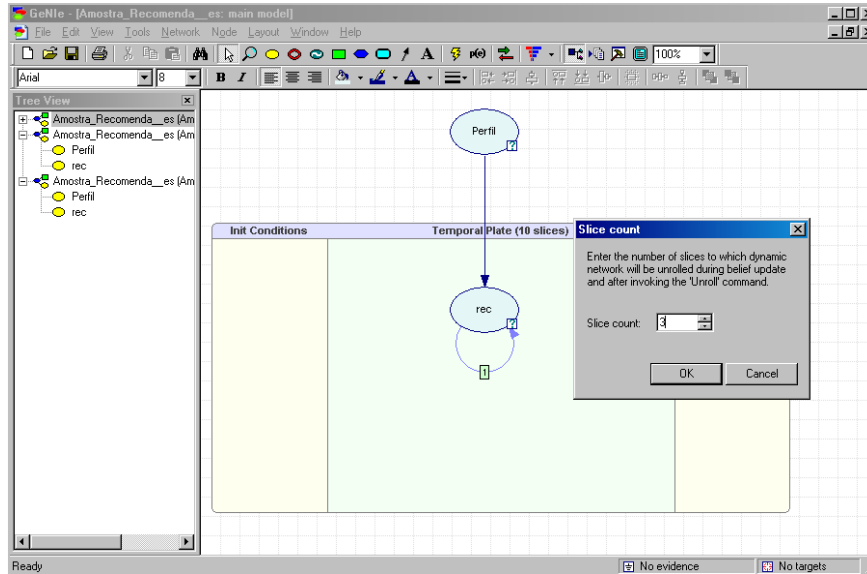


Figura 37 – Definição do número de fatias de tempo

Nesta fase, a estrutura da rede está definida, tal como ilustrada na Figura 38:

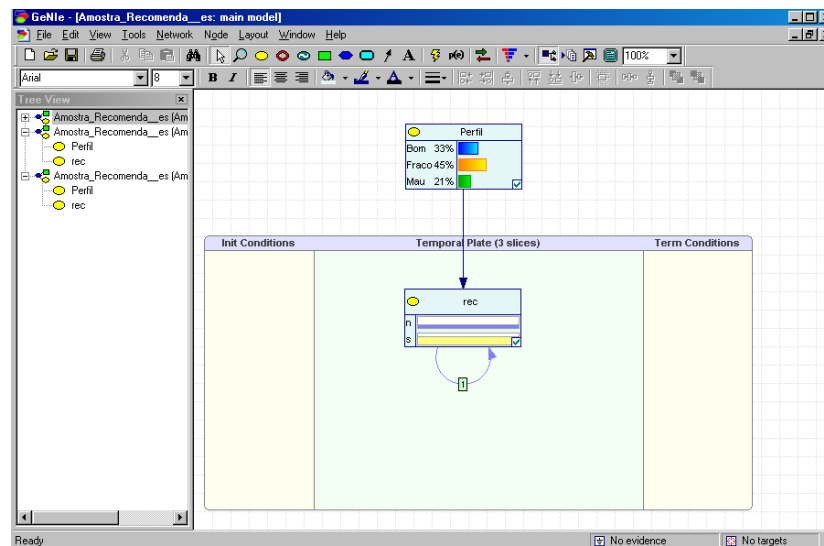
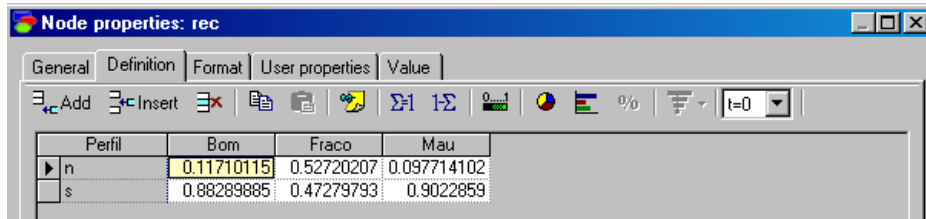


Figura 38 – Rede Bayesiana Dinâmica

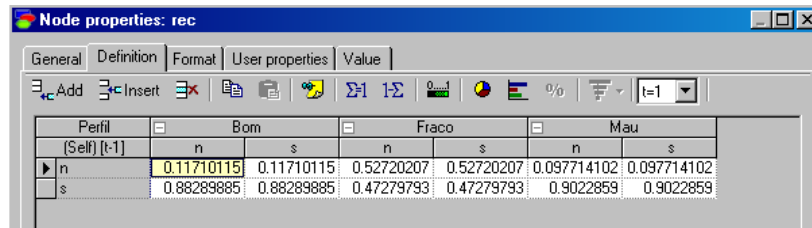
A tabela de probabilidades da variável Perfil mantém-se inalterada relativamente à anterior construção da Rede Bayesiana Normal, no entanto, as tabelas de probabilidades associadas aos nós temporais são agora duas, uma para $t=0$ e outra para $t=1$. As probabilidades calculadas por defeito pelo GeNIe são as que constam nas Figuras 39 e 40:

5.3. Auditoria na Banca usando Redes Bayesianas Dinâmicas



Perfil	Bom	Fraco	Mau
n	0.11710115	0.52720207	0.097714102
s	0.88289885	0.47279793	0.9022859

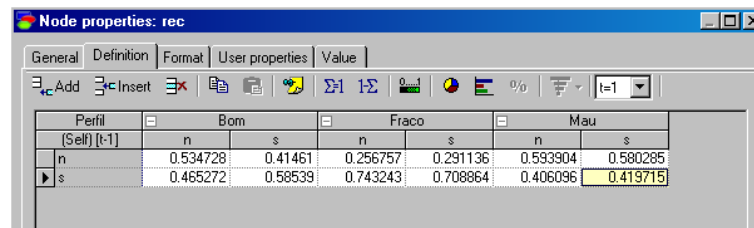
Figura 39 – Tabela de probabilidades para $t=0$



Perfil	Bom	Fraco	Mau
(Self) [t-1]			
n	0.11710115	0.52720207	0.097714102
s	0.88289885	0.47279793	0.9022859

Figura 40 – Tabela de probabilidades para $t=1$

Importa agora alterar a tabela de probabilidades de acordo com as obtidas na Figura 33, para ajustar o modelo à aprendizagem já efectuada através de dados históricos:



Perfil	Bom	Fraco	Mau
(Self) [t-1]			
n	0.534728	0.256757	0.593904
s	0.465272	0.743243	0.406096

Figura 41 – Nova tabela de probabilidades para $t=1$

Após alteração da tabela de probabilidades, é possível consultar os resultados obtidos em termos de distribuição de probabilidades temporais. A Figura 42 evidencia estes números bem como a sua ilustração gráfica:

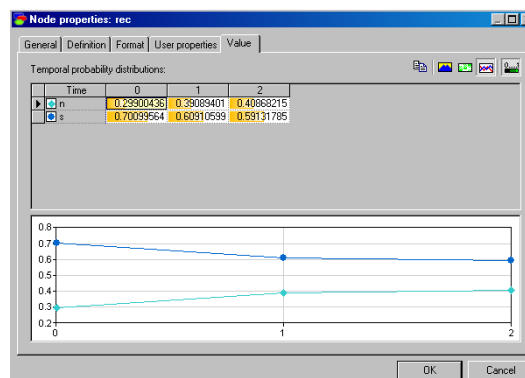


Figura 42 – Probabilidades temporais

Finalmente, a Rede Bayesiana está construída e pronta a permitir a simulação de diferentes cenários.

5.3.1.5. Simulação, propagação e discussão dos resultados

À medida que a auditoria interna vem consolidando o seu papel dentro das organizações e ampliando a sua actuação estratégica, aumenta também a procura por resultados. Obter de forma clara os indicadores chave dos processos fundamentais e avaliar o impacto dos trabalhos de auditoria torna-se quase obrigatório.

É importante que a auditoria compreenda com maior amplitude as melhores estratégias, utilizadas para inserir a área de forma alinhada dentro do plano estratégico da organização, maximizando desta forma a obtenção de resultados.

Neste âmbito, medir não é fácil. Optou-se nesta exposição de caso, simplificar a análise no que diz respeito ao número de variáveis consideradas. A variável **Perfil**, por si só, incorpora o resultado da combinação de um elevado número de outras variáveis, resumido a três níveis de classificação: Bom, Fraco e Mau.

A simulação dos vários cenários possíveis é efectuada seleccionando as evidências para cada fatia de tempo, conforme ilustrado na Figura 43:

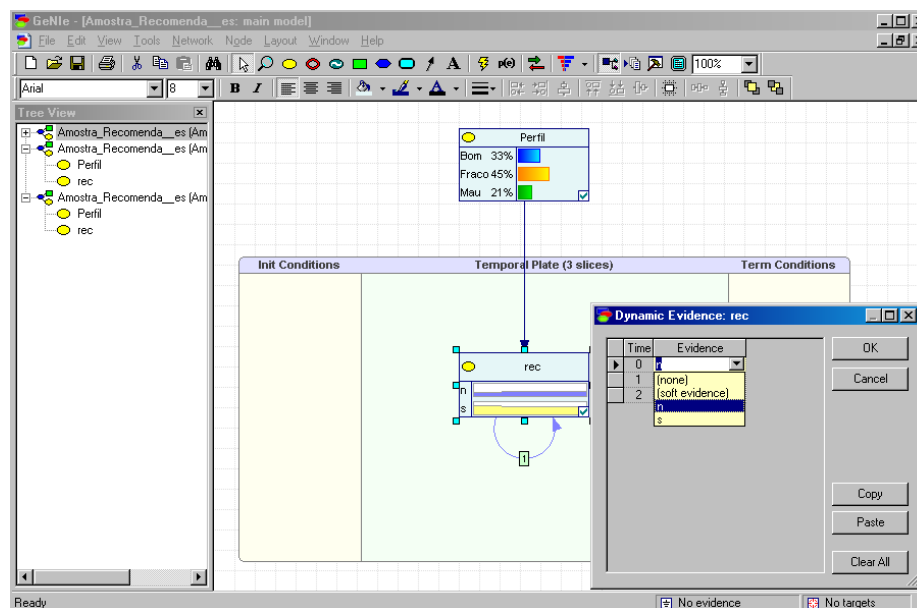


Figura 43 – Simulação de evidências na Rede Bayesiana Dinâmica

Os resultados obtidos simulando os dois cenários extremos, ou seja, considerando que nunca foram efectuadas recomendações e que foram efectuadas recomendações em todos os períodos encontram-se evidenciados na Figura 44:

5.3. Auditoria na Banca usando Redes Bayesianas Dinâmicas

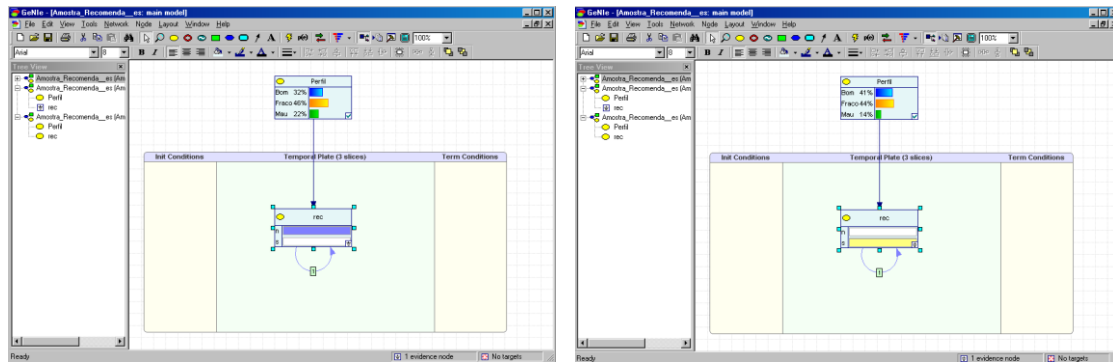


Figura 44 – Análise de situações extremas

No primeiro cenário – nunca foram efectuadas recomendações – observa-se que as probabilidades, em termos de perfil, apresentam valores muito semelhantes aos que se apresentam na tabela de probabilidades *a priori* obtida na Figura 38. Este comportamento altera-se mediante a consideração do cenário oposto, ou seja, considerando que foram efectuadas recomendações em todos os períodos, isto é, considerada uma intervenção intensa por parte da auditoria na área de crédito dos balcões, verifica-se que uma clara tendência para a melhoria da carteira. Aumenta a probabilidade do perfil ser bom e reduz a probabilidade de ser mau.

Outra forma de visualização gráfica desta dinâmica temporal, e porque se entende mais ilustrativo em termos de percepção visual é proceder ao “desenrolar da rede” – *Unroll* – através do menu *Network* – Figura 45:

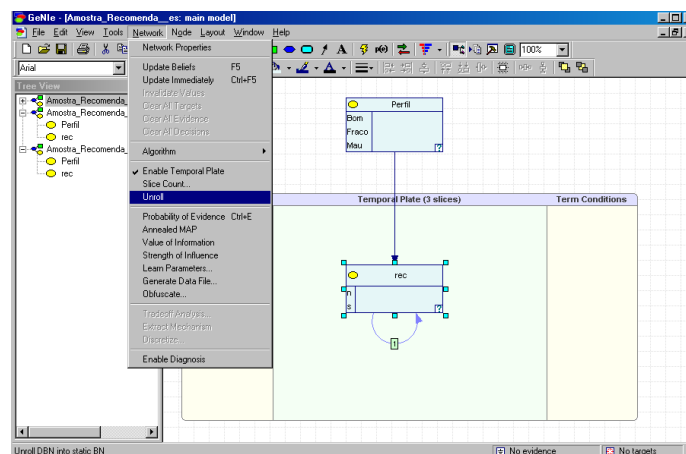


Figura 45 – Unroll da rede

Assim, as simulações que se seguem serão ilustradas tal como na Figura 46, onde se encontra representada a Rede Bayesiana Dinâmica como que convertida a uma Rede

5.3. Auditoria na Banca usando Redes Bayesianas Dinâmicas

Bayesiana Normal agora com mais duas variáveis, uma correspondente a $t=1$ e a outra correspondente a $t=2$.

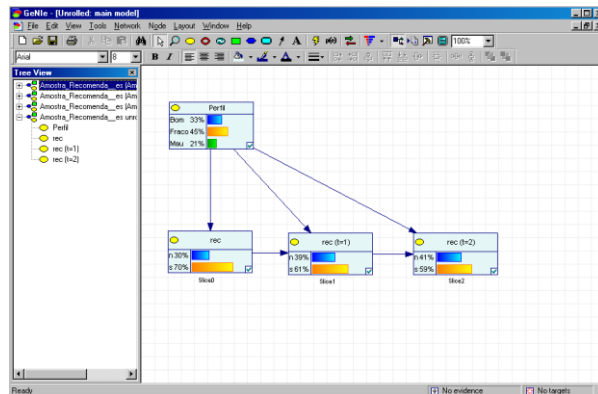
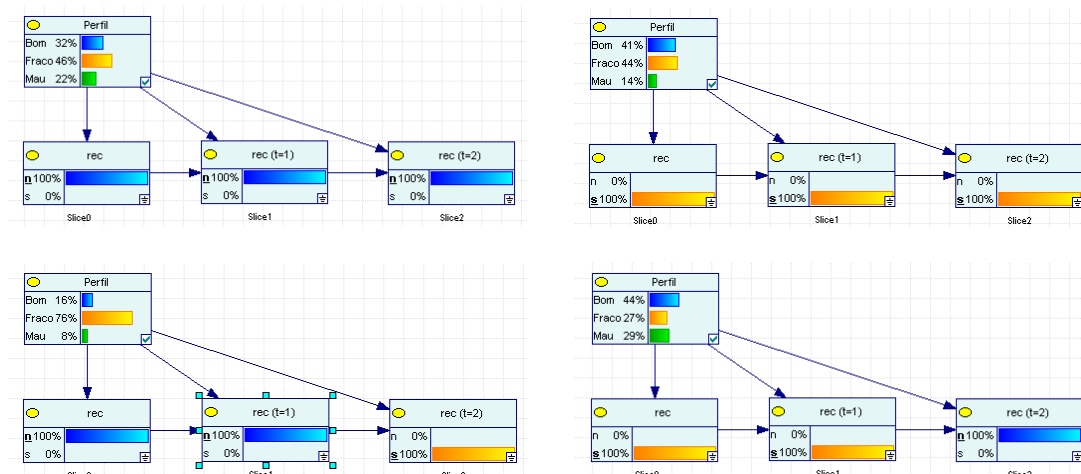


Figura 46 – Outra forma de representar a Rede Bayesiana Dinâmica

Após a construção das redes e o fornecimento das respectivas probabilidades condicionais, pode-se efectuar inferências sobre as mesmas.

Uma Rede Bayesiana Dinâmica, tal como a representada na Figura anterior, é uma Rede Bayesiana Normal que pode ser utilizada para representar um modelo probabilístico temporal, e pode, por isso, utilizar os mesmos algoritmos para inferência probabilística.

A Figura 47 consubstancia os diferentes cenários possíveis considerando o algoritmo de propagação assumido por defeito pelo GeNIe, que é um algoritmo exacto, conforme já referido na abordagem teórica relativa a esta matéria. Este algoritmo de propagação é o Clustering. Segue-se a apresentação dos diferentes cenários possíveis – Figura 48 – se considerado um algoritmo aproximado – o Likelihood Sampling.



5.3. Auditoria na Banca usando Redes Bayesianas Dinâmicas

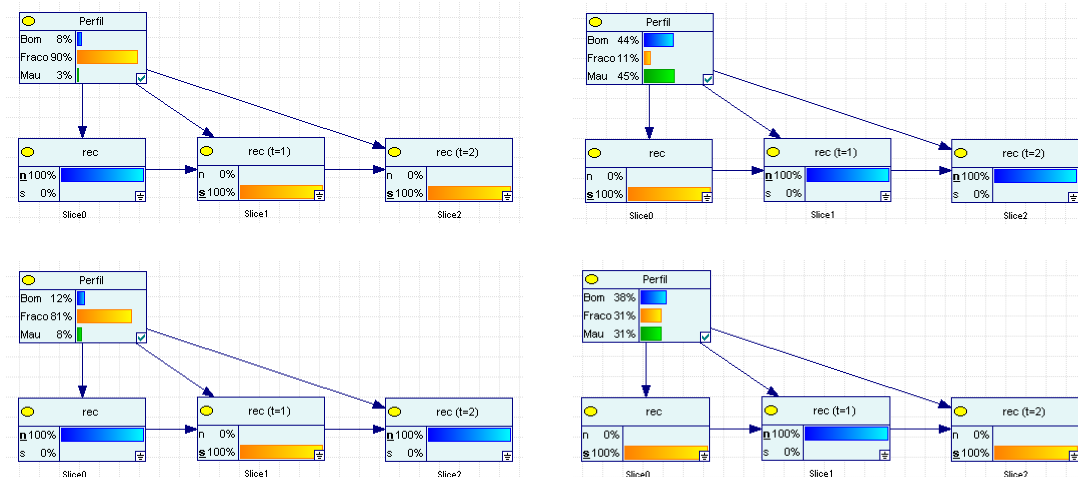


Figura 47 – Simulação de cenários utilizando o *Clustering*

É possível observar que os resultados diferem consoante os padrões de actuação ao longo do tempo. Ainda assim, verifica-se que, sempre que foram efectuadas recomendações na primeira auditoria, e independentemente de se reiterarem na segunda e / ou na terceira, a probabilidade de alcançar operações com bom perfil é sempre maior do que na ausência total de recomendações.

Por outro lado, também os resultados em termos de probabilidade associada a um mau perfil apresentam valores mais elevados. Não recomendar na primeira auditoria e recomendar na segunda e / ou na terceira tem forte impacto no perfil mediano, isto é, classificado como fraco. As expectativas em termos de performance, alcançadas com base neste indicador, vão no sentido de considerar que os padrões de actuação que vão sendo assumidos ao longo do tempo em termos de gestão da carteira não tendem para melhorias significativas nem tão pouco para resultados propriamente drásticos. Como que se a auditoria actuasse como um sinal, mais do que factor fortemente interventivo, e provocasse comportamentos menos audazes ou arriscados em termos de medidas adoptadas para a gestão da carteira e, com a evolução do tempo, tal conduzissem para padrões mais normais e acomodados, do que propriamente exemplares ou proactivos.

Não parece ser indiferente o impacto das auditorias, quando traduzido pelo registo das recomendações que se estabelecem nos balcões. Se os resultados forem interpretados em termos de primeiro impacto, o que pode ser alvo de reflexão importante é o alvo das recomendações. Obviamente, o auditor efectua

recomendações quando encontra não conformidades, mas o seu papel não se esgota no registo dos problemas que encontra e nas providências que toma para a sua resolução. Se o impacto das suas recomendações se estender a eventuais ocorrências futuras, tudo indica que a actuação da equipa na primeira auditoria assume uma importância acrescida.

Importa ressaltar ainda a possibilidade que o auditor interno tem, como especialista, de aplicar os conhecimentos, experiência e expectativas que assume *a priori*, de intervir na análise, neste caso atribuindo valores diferentes às tabelas de probabilidades (uma vez que que faria sentido alterar a única relação de dependência existente).

A Figura 48 ilustra os resultados obtidos nos diferentes cenários possíveis, com a utilização de um algoritmo de propagação aproximado, o *Likelihood Sampling*:

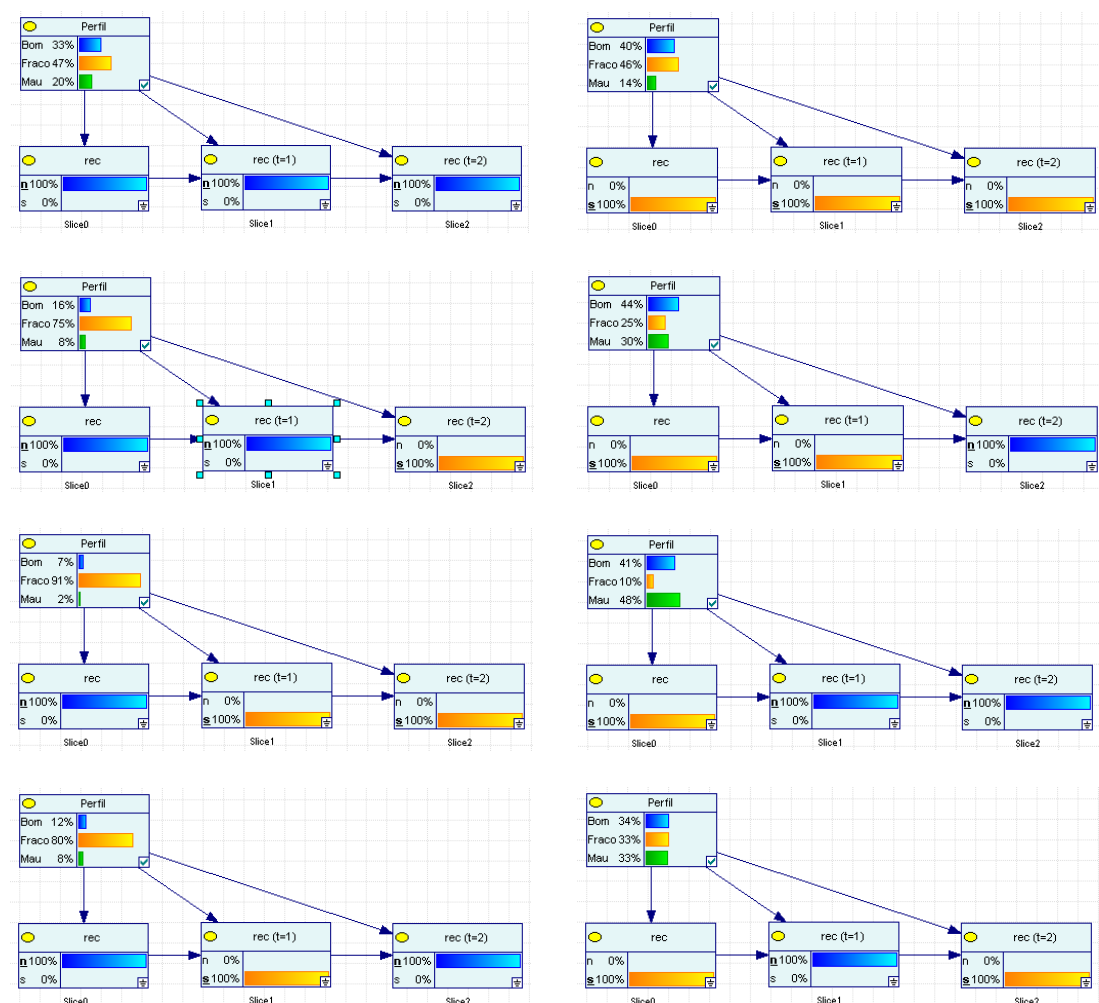


Figura 48 – Simulação de cenários utilizando o *Likelihood Sampling*

Em termos de resultados, os diferentes cenários simulados com o algoritmo de propagação *Likelihood Sampling* não evidenciam diferenças substanciais relativamente ao *Clustering*, o que significa dizer que, neste caso, a escolha do algoritmo de propagação não afectou a performance da rede considerada.

6. CONSIDERAÇÕES FINAIS E TRABALHO FUTURO

Um dos objectivos da auditoria é contribuir para a melhoria das práticas e qualidade dos processos estabelecidos, através da elaboração, quando se justifique, de recomendações de alterações e/ou novos procedimentos organizativos. Ora, para qualquer acção, o especialista, neste âmbito, auditor, tem expectativas que umas vezes são realizadas, outras não. Independentemente do resultado, o auditor aprenderá a partir de experiências que o ajudam a entender o seu universo, o que passa por avaliar o impacto das acções levadas a cabo, neste caso avaliar o impacto das recomendações efectuadas.

Este trabalho não teve a pretensão de provar a aplicabilidade empírica da utilização de Redes Bayesianas numa actividade tão vasta e complexa como a auditoria bancária. Aliás, considerando o também vasto âmbito de actuação e aplicabilidade das Redes Bayesianas, tema alvo de crescentes e impetuosos estudos académicos, com especial ênfase na área da medicina, a contribuição deste trabalho surge como um teste à abrangência do tema a uma área ainda não tão estudada – auditoria à área de crédito de uma instituição bancária – e que poderia eventualmente tirar algum proveito dos resultados alcançados.

Para tal, foram efectuadas várias experiências, através da análise de casos, primeiro utilizando Redes Bayesianas Normais, depois introduzindo uma vertente temporal na análise, através da apresentação de uma Rede Bayesiana Dinâmica.

No primeiro caso prático apresentado utilizou-se uma Rede Bayesiana Simples (ou Estática, em contraponto com Dinâmica utilizada mais tarde), para procurar observar o impacto de recomendações efectuadas nas auditorias no crédito vencido (valor em dívida) associado às operações de crédito de uma instituição bancária. A situação dos contratos de crédito foi utilizada como indicador da “sanidade” da instituição, ainda que esta escolha não seja muito robusta, dada a dependência que assume relativamente a factores externos tais como conjuntura económico-financeira, capacidade creditícia dos titulares e até mesmo estratégias de actuação no interior da empresa. Seguindo o modelo proposto, relativo a uma área específica relacionada com a análise de operações de crédito, confundem-se os resultados obtidos com os que seriam desejados, pelo que não se apresenta premente alterar a estratégia de actuação levada a cabo pela direcção de auditoria. Ainda assim, a base de dados utilizada referente à informação relativa a

quatro balcões num universo de mais de cem, não garante conclusões a nível global, que possam fundamentar ou justificar a actuação de todas as equipas de auditoria em toda a instituição. Para tal, impõe-se a utilização de uma base de dados mais vasta.

No segundo caso prático analisado foi considerado um indicador de análise mais abrangente – perfil de risco associado às operações de crédito formalizadas – e apresentada a abordagem dinâmica destes modelos. Mais uma vez se verificou que as técnicas de mensuração causal e a modelação do risco de crédito permanecem num estágio inicial de desenvolvimento na maioria das instituições e, como os demais aspectos de outros riscos, são ainda difíceis de medir. Os resultados obtidos, reflectem não só a importância que pode ser atribuída à primeira intervenção efectuada por um auditor, como também à possibilidade de, ao longo do tempo, essa actuação produzir resultados. Mais uma vez, daqui se pode concluir que os resultados obtidos apontam para uma avaliação positiva da estratégia levada a cabo pela direcção de auditoria da instituição em causa.

Em suma, as Redes Bayesianas possuem uma semântica clara, que permite processar diagnóstico, aprendizagem, explanação e muitas outras tarefas de inferência necessárias para sistemas inteligentes de apoio à decisão, pelo que se apresentam como uma ferramenta útil para trabalhar modelos que se relacionem com análises deste tipo.

Neste trabalho, foram introduzidos conceitos teóricos, foi apresentada uma ferramenta de apoio e foi sugerida uma análise empírica de casos que operam com dados reais.

Fica como sugestão de trabalhos futuros o aprofundar o conhecimento dos algoritmos, com a realização de mais análises com os diferentes algoritmos de aprendizagem de estrutura, através de experiências com dados e comparações de resultados, bem como o estudo, em profundidade, das propriedades e características dos algoritmos de aprendizagem de estrutura que permitam uma maior contribuição para a área.

Seria também interessante, alargar a estrutura da Rede Bayesiana Dinâmica construída, introduzindo as variáveis que se entendessem pertinentes para o análise de forma a apreender mais informação e enriquecer a trabalho.

O estudo de abordagens de aprendizagem de estrutura que trabalhem melhor quando a base de dados contem dados em falta, a avaliação do impacto do processo de discretização de variáveis e a comparação do GeNIe com a utilização de outras

ferramentas, são outras sugestões a ser consideradas em trabalhos futuros neste contexto.

Os modelos criados, bem como as conclusões preliminares, serão submetidos à instituição financeira fornecedora da base de dados, para avaliação da continuidade do processo de análise em termos profissionais, em continuidade dos trabalho acadêmico já realizado.

REFERÊNCIAS BIBLIOGRÁFICAS

- Abbasi, A. R. (2007), "A Bayesian Network Approach to Interpret Affective States from Human Gestures-An Application to Affective Tutoring", *Proceedings of Doctoral Consortium in 2nd Conference on Affective Computing and Intelligent Interaction*, Lisboa.
- Bertotti, S., M. L. S. Caetano e P. S. S. Borges (2003), "Previsão de Vendas para Artigos de Confecção utilizando Redes Bayesianas – Um caso prático", *Anais do V Encontro de Estudantes de Informática do Tocantins*, pp. 221-226, Palmas.
- Bet, S., A. A. Frozza e A. A. C. Arruda (2004), "Banco de Desafios: Uma Aplicação de Hipermídia Adaptativa e Redes Bayesianas", *Simpósio Brasileiro de Informática na Educação*, pp. 599-602, Manaus.
- Bouillaut, L., O. François, P. Leray, P. Akinin e S. Dubois (2008), "Dynamic Bayesian Network modelling Maintenance Strategies : Prevention of Broken Rails", *World Congress on Railway Research*, Seoul, South Korea.
- Burge, J. e T. Lane (2005), "Learning Class-Discriminative Dynamic Bayesian Networks", *ACM International Conference Proceeding Series; Proceedings of the 22nd international conference on Machine learning*, pp. 97-104, Bonn, Germany.
- Castillo, E., J. M. Gutiérrez e A. S. Hadi (1997), *Expert systems and probabilistic network models*, New York, Springer-Verlag.
- Charitos, T., L. C. Gaag, S. Visscher, K. A. M. Schurink e P. J. F. Lucas (2009), "A Dynamic Bayesian Network for Diagnosing Ventilator-Associated Pneumonia in ICU Patients", *Expert Systems with Applications: An International Journal*, Vol. 36, Nº 2, pp. 1249-1258.
- Charniak, E. (1991), "Bayesian networks without tears: making Bayesian networks more accessible to the probabilistically unsophisticated", *AI Magazine*, Vol. 12, Nº 4, pp. 50-60.
- Cheng, J. e M. J. Druzdzel (2000), "AIS-BN: An Adaptive Importance Sampling Algorithm for Evidential Reasoning in Large Bayesian Networks", *Journal of Artificial Intelligence Research*, Vol. 13, Nº pp. 155-188.
- Chickering, D. M. (2002), "Optimal structure identification with greedy search", *The Journal of Machine Learning Research*, Vol. 3, Nº pp. 507-554.
- Choudhury, T., J. M. Rehg, V. Pavlović e A. Pentland (2002), "Boosting and structure learning in dynamic Bayesian networks for audio-visual speaker detection", *International Conference on Pattern Recognition* pp. 789-794.

Cowell, R. G., S. L. Lauritzen, A. P. David e D. J. Spiegelhalter (1999), *Probabilistic Networks and Expert Systems*, 1st edition, Springer-Verlag New York, Inc.

Feng, L., W. Wang, L. Zhu e Y. Zhang (2009), "Predicting intrusion goal using dynamic Bayesian network with transfer probability estimation", *J. Netw. Comput. Appl.*, Vol. 32, Nº 3, pp. 721-732.

Fenton, N. e M. Neil (2007), *Managing Risk in the Modern World - Applications of Bayesian Networks*.

Frankel, J., M. Wester e S. King (2007), "Articulatory feature recognition using dynamic Bayesian networks", *Comput. Speech Lang.*, Vol. 21, Nº 4, pp. 620-640.

Friedman, N. e M. Goldszmidt (1996), "Discretizing continuous attributes while learning bayesian networks", *International Conference on Machine Learning*, pp. 157-165.

Frigault, M., L. Wang, A. Singhal e S. Jajodia (2008), "Measuring Network Security Using Dynamic Bayesian Network", *Conference on Computer and Communications Security - Proceedings of the 4th ACM workshop on Quality of protection*, pp. 23-30, Virginia.

Fung, R. e B. Favero (1994), "Backward simulation in Bayesian networks.", *Proceedings of the Tenth Annual Conference on Uncertainty in Artificial Intelligence (UAI-94)*, pp. 227-234, San Francisco.

GeNIe. <http://genie.sis.pitt.edu>. Editado por Decision Systems Laboratory. University of Pittsburgh .

Gerven, M. A. J., B. G. Taal e P. J. F. Lucas (2008), "Dynamic Bayesian networks as prognostic models for clinical patient management", *Journal of Biomedical Informatics* Vol. 41, Nº 4, pp. 515-529.

Heckerman, D. (1995), *A tutorial on learning with Bayesian networks*, Technical Report MSR-TR-95-06.

Henrion, M. (1988), "Propagating uncertainty in Bayesian networks by probabilistic logic sampling", *Uncertainty in Artificial Intelligence 2*, pp. 149-163.

Jensen, F. V. (2001), *Bayesian Networks and Decision Graphs*, New York: Springer.

Johnstone, D. J. (1994), *Audit risk in terms of probabilities: The AUP24 model*.

Kalisch, M. (2006), *A Tutorial on Learning With Bayesian Networks*.

Kanazawa, K., D. Koller e S. Russell (1995), *Stochastic simulation algorithms for dynamic probabilistic networks*, San Francisco, CA, Morgan Kaufmann.

- Korb, K. B. e A. E. Nicholson (2003), *Bayesian artificial intelligence*, Florida: Chapman & Hall/CRC.
- Kuenzer, A., C. Schlick, F. Ohmann, L. Schmidt e H. Luczak (2002), "An empirical study of dynamic bayesian networks for user modeling", *Workshop on Machine Learning for User Modeling*.
- Laws, D. J. e A. O'Hagan (2003), "A hierarchical Bayes model for multilocation auditing", *Journal of the Royal Statistical Society: Series D (The Statistician)*, Vol. 51, Nº 4, pp. 431 - 450.
- Lermen, E. G. e G. G. E. Bruno (2008), "Framework para Detecção e Filtragem de Alertas de Intrusão utilizando Redes Bayesianas", *VIII Simpósio Brasileiro em Segurança da Informação e de Sistemas Computacionais*.
- Milho, I. e A. Fred (2002), "Sistema de Apoio ao Diagnóstico Médico Baseado em Redes Bayesianas", *Conferência Científica e Tecnológica em Engenharia - CCTE2002*, Lisboa.
- Nilsson, N. J. (1998), *Artificial intelligence: a new synthesis*, Morgan Kaufmann Publishers Inc.
- Pearl, J. (1988), *Probabilistic reasoning in intelligent systems: networks of plausible inference*, San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.
- Rabiner, L. R. e B. H. Juang (1989), *An Introduction to Hidden Markov models*.
- Rebane, G. e J. Pearl (1987), "The recovery of causal poly-trees from statistical data", *AAAI Workshop on Uncertainty in AI*, UCLA Cognitive System Laboratory. Seattle.
- Russell, S. e P. Norvig (2003), *Artificial Intelligence: A Modern Approach*, Prentice Hall Series in Artificial Intelligence.
- Salem, A. B., A. Muller e P. Weber (2006), "Dynamic Bayesian Networks in System Reliability Analysis", *6th IFAC Symposium on Fault Detection, Supervision and Safety of technical processes*, pp. 481-486, Beijing : China.
- Sarkar, S. e R. S. Sriram (2001), "Bayesian Models for Early Warning of Bank Failures", *Management Science* Vol. 47, Nº 11, pp. 1457 - 1475.
- Schreiber, J. N. C., R. S. Wazlawick e P. S. S. Borges (2002), *Uma proposta de navegação adaptativa na WEB utilizando Redes Bayesianas*, Vigo, Servicio de Publicación Teleco Vigo.
- Shachter, R. D. e M. A. Peot (1990), "Simulation Approaches to General Probabilistic Inference on Belief Networks", *Proceedings of the Fifth Annual Conference on Uncertainty in Artificial Intelligence* pp. 221 - 234.

Silva, E. M. M. e P. F. Brito (2008), "Utilização de Redes Bayesianas em Sistemas Hipermídia Adaptativo para o Ensino", *XIX Simpósio Brasileiro de Informática na Educação*, Palmas.

Wikipedia <http://en.wikipedia.org/wiki/>

Yuan, C. e M. J. Druzdzel (2004), *Importance Sampling Algorithms for Bayesian Networks: Principles and Performance*.

Zajdel, W., A. T. Cemgil e B. J. A. Krose (2006), "Dynamic Bayesian Networks for Visual Surveillance with Distributed Cameras", *European Conference on Smart Sensing and Context No1*, pp. 240-243, Enschede.

Zhang, L., D. Samaras, N. Alia-Klein, N. D. Volkow e R. Z. Goldstein (2006), "Modeling neuronal interactivity using dynamic Bayesian networks", *Advances in Neural Information Processing Systems*, Cambridge.