FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Development and testing of a personalized recommender system based on mobility profile analysis and passenger activity

Luís Miguel Azevedo Duarte

DISSERTATION

Mestrado Integrado em Engenharia Informática e Computação

Supervisor: Prof. Teresa Galvão Dias

Second Supervisor: Prof. Marta Campos Ferreira

July 25, 2018

Development and testing of a personalized recommender system based on mobility profile analysis and passenger activity

Luís Miguel Azevedo Duarte

Mestrado Integrado em Engenharia Informática e Computação

July 25, 2018

Abstract

Sustainable Mobility has been rising as one of the most prominent concerns of everyday life. The need for awareness of climate changes and the overpopulation of metropolitan and urban areas has created a problem, urgently in need of a solution.

Some previous approaches have handled solutions regarding only the transportation component, while some conceptual models have tried to increase the attractiveness of public transportation by matching the services provided by public transportation operators to several services provided by the cities. However, a practical, empirical solution has yet to be presented.

This dissertation proposes to address these concerns and presents a personalized recommender system based not only on mobility profiles, but also on the activities of everyday users of public transportation systems. The cross reference of both information sources, allows for a more user-based experience, making sure that the offers provided are indeed in accordance with customer preferences. Additionally, the system has a built in learning mechanism, that improves it's recommendations based on user feedback, improving the overall experience. With the objective to achieve more accurate results two types of data sources were used: Automated Fare Collection system data from the Metropolitan Area of Porto, Portugal, and city services information extracted from Google Places. Both data sources were thoroughly analysed and filtered increasing result precision and improving the overall recommendation quality.

As a current and widely used tool, a mobile application prototype was developed, in order to correctly showcase the full potential impact of the aforementioned system.

System testing was performed through internal offline data validation, that is by splinting the given dataset in learning and testing data, measuring not only the system's ability to generate tailored recommendations, but also to anticipate user preferences.

The main idea in addressing this issue was to promote public transportation usage, by supplying customers with benefits via discounts, promotions, service offerings, among some others, increasing the use of cleaner and more efficient transportation modes.

The developed work resulted as well in the writing of a scientific paper, whose oral presentation was accepted to be part of the European conference *21st EURO Working Group on Transportation Meeting 2018.*

Resumo

A temática da Mobilidade Sustentável tem ganho grande relevo naquilo que é o panorama atual da sociedade. A necessidade para a consciencialização de alterações climáticas e a sobrepopulação das áreas metropolitanas e urbanas, tem vindo a criar um problema, urgentemente à espera de uma solução.

Algumas abordagens prévias apresentaram soluções focadas somente na componente de transporte, enquanto alguns modelos conceptuais tentaram aumentar o interesse pelos transportes publicos aliando os serviços fornecidos pelas suas operadoras, com serviços locais interentes às respectivas cidades. Todavia, uma solução empírica ainda não foi apresentada.

Esta dissertação propõe abordar estes problemas e apresenta um sistema de recomendação personalizado baseado, não só na análise de perfis de mobilidade, mas também na atividade diária de passageiros. O cruzamento de dados destes tipos de informações permite uma experiência mais orientada ao utilizador, garantindo que a oferta está efectivamente de acordo com as suas preferências. Adicionalmente, o sistema tem embutido um mecânismo de aprendizagem baseado no *feedback* dos utilizadores, melhorando a sua experiência global. Com o objectivo de obter resultados mais precisos, dois tipos de fontes de dados foram utilizados: informações recolhidas pelo sistema *Andante* na Área Metropolitanta do Porto, Portugal e detalhes acerca dos serviços locais da cidade, extraídos da API Google Places. Ambas as fontes de informação foram analizadas e filtradas de forma detalhada, aumentando a precisão dos resultados e melhorando a qualidade global das recomendações.

Sendo uma ferramenta cada vez mais utilizada, um protótipo de uma aplicação móvel foi desenvolvido, com o intuito de demonstrar corretamente o impacto potencial completo do sistema.

Os testes ao sistema foram realizados através de validações *offline* internas, isto é, dividindo o *dataset* disponível numa parte para aprendizagem e outra para teste, avaliando não só a capacidade do sistema gerar recomendações personalizadas, mas também de antecipar as preferências dos utilizadores.

A idea principal na abordagem deste problema foi a promoção da utilização do transporte público, fornecendo ao clientes benefícios, sejam eles através de descontos, promoções, ofertas, entre outros, aumentando a utilização de modos de transporte limpos e eficientes.

Adicionalmente o trabalho desenvolvido resultou na escrita de um artigo cientifico, cuja apresentação oral foi aceite para fazer parte da conferência europeia *21st EURO Working Group on Transportation Meeting 2018*.

Acknowledgements

At this time I would like to take the opportunity to thank everyone involved in the development of this dissertation. To professors Teresa Galvão and Marta Ferreira, for all the knowledge, help and, much needed guidance, for the motivation and for asking more and more of me each week, I thank you. I would like to thank as well, Joana Hora and Miguel Botelho for all the help and input given during the development of this project.

And since this also represents the culmination of what were 5 challenging years of my master's degree, I would like to thank everyone involved in what was an important time for professional and personal growth. To my family a special thanks for all the support, encouragement and understanding. To my close friends André, André, Edgar, Gustavo and Leonardo, that shared with me this amazing journey, and to whom without this would not be possible, I thank you.

Luís Miguel Azevedo Duarte

"I think the essence of hard work is one that's pretty straightforward. You'll never be the best looking, you'll never be the tallest, the most talented, most capable, you'll never have the most money - there will always be someone better at whatever you're doing than you are. But you can always be the hardest working person in the room."

Casey Neistat

Contents

List of Figures

LIST OF FIGURES

List of Tables

LIST OF TABLES

Abbreviations

AFC	Automated Fare Collection
NFC	Near Field Communication
BLE	Bluetooth Low Energy
MAP	Metropolitan Area of Porto
API	Application Programming Interface
POI	Point of interest
STCP	Sociedade de Transportes Colectivos do Porto
ES	Espírito Santo
CP	Comboios de Portugal
MP	Metro do Porto
TIP	Transportes Intermodais do Porto
START	Sustainable Transportation Awareness Recommender Tool

Chapter 1

Introduction

1.1 Context

Sustainable mobility as been rising has one of the most prominent concerns of everyday life. This can be seen, in the unwillingness to decrease the use of transportation of private usage [DoT16], which not only aggravates the inefficiency of transportation itself, but also compromises the everyday life of every citizen. It results in an increase of gas emissions and air pollution, causing even more climate changes to those that already have occurred in prior years. In that sense, promoting the use of public transportation could appease or even aim to solve these problems. And while technology continues to rapidly grow in an astounding pace, transport efficiency and environmental awareness and knowledge, seem to not be able to follow the same rhythm [KA02].

Currently, due to technological devices of private usage, such as smartphones, and to artificial intelligence systems being developed and improved upon, users are now able to access, not only large amounts of information, in a few simple clicks, but have that information tailored to their preferences and needs.

Nowadays, it seems that almost every single person has indeed a smartphone device. Not only this, but the enormous dependency and addiction on these devices [BJR16] has made them the perfect tool to present or manage information. People are now more and more accustomed to send emails, check their social media profiles and do specific or everyday searches in these devices. Additionally, it is now more common for users to watch videos or television and even manage important personal information such as baking and finance account assets through these devices. Some cities, such as Amsterdam, are even using mobile applications to reroute tourists, in order to decrease the congestion that some touristic attractions face, using a creative approach and way of thinking to tackle this issue [Ind17]. Recommender systems are also an important component in these products. They are becoming more accurate on data gathering and the information they present due to mechanisms of self learning and feedback, playing a major role in user oriented information retrieval software tools.

Yet, in a society that continuously keeps overpopulating urban and metropolitan areas, looking for better life opportunities, and that is more than aware of the need to protect the environment and put a stop to climate changes, not many solutions regarding sustainable mobility and transportation, at least with a technological component, have been presented or discussed. Trying to solve this problem in a legislative way, several major cities worldwide have had to implement strict private usage transportation circulation rules [CdM04], that are crucial to the maintenance and sustainability of the cities themselves.

1.2 Motivation and Goals

Combining recommender systems that have a one-to-one approach with its users, via a mobile application, with the opportunity to address an up to date problem that is sustainable mobility, in the form of increasing public transportation usage, is an exciting, yet very challenging task.

This work tries to answer these concerns by combining recommender systems that have a one-to-one approach with its users, via a mobile application, with the opportunity to address an up to date problem that is sustainable mobility, by increasing public transportation usage. In an attempt to take sustainable mobility a step forward, the proposal of a recommender system based on mobility profiles and passenger activity is presented. The idea behind it is to, after an extensive analysis and categorization of different validation datasets, generate user profiles with specific needs and preferences and, afterwards, combine them with a recommender system, available via a mobile application that, taking in consideration these elements, presents users with cultural and local business options for them to visit or consult [FG13]. The main idea in the development of this platform is to supply public transport users with personalized benefits via discounts, promotions, service offerings, among others, increasing the use of cleaner and more efficient transportation modes. By combining marketing strategies such as discounts and other offers upon each trip, it is expected to increase awareness of both services, attract new customers and retain the existing ones and promote a sustainable mobile. Even more, the opportunity to actually experiment with real data from the *Andante* system, currently in place in the Metropolitan Area of the city of Porto, creates extra motivation and caution upon its development.

The main goals of this solution are threefold. First, developing a system that is user oriented, and that supplies users with personalized offers based on their mobility profiles. By evaluating and cross-referencing data of location and activity of each user, it is possible to create offers that not only match their preferences but also their needs. Secondly, this will result in local business promotion, increasing customer loyalty and public transportation usage, via the need to accumulate discounts and offers, gathered with the number of trips each user does. Finally, this will allow to showcase the companies of public transportation of the Metropolitan Area of the city of Porto a new perspective in user recommendation and city culture promotion.

1.3 Dissertation Structure

In addition to this introductory chapter, in which the context, motivation and goals are stated, this dissertation has an additional 6 chapters. The next chapter, chapter 2, presents the state of the art and literature review on this subject, exploring exiting technologies and related work, as well as different approaches to the same, or similar, problems. Afterwards, chapter 3 goes into detail on the mobility profile analysis conducted, exploring both the data used to perform this analysis and the conditions, restrictions and variants that were used in order to achieve the most accurate outcome, and an overview on passenger activities that was crucial to conduct user profiling. Chapter 4 addresses the recommendation items, in this case, the points of interest deemed fit to be recommended and their extraction process was developed. Chapter 5 addresses the recommender system implementation and the mobile application development, describing the system architecture an algorithms used. Afterwards, chapter 6 reviews the results gathered from the usage of the developed system, how they were obtained and what conclusions can be extracted from them. Finally, as a closing chapter, Chapter 7 will explore what can be concluded from the developed work, it's implications and what future work can and needs to be done regarding this subject.

Chapter 2

Literature Review

In this chapter, the literature review conducted prior to the system's development will be presented and discussed, describing the state of the art on the subjects at hand, the main technologies used, and other approaches that have presented solutions to address the same or similar problems.

Sequentially, from broader to more specific subjects, an analysis will be performed, beginning with the study of mobility itself, it's definition, the importance of sustainable mobility, and what is the current paradigm and worldwide stance on this topic. Afterwards, automated fare collection will be discussed, explaining its underling mechanisms and what has been done until now with this type of gathered data. In section 3, recommender systems will be reviewed, describing and examining feedback categorization methods and filtering techniques, and their importance for today's society. Narrowing the scope even further, since the working data available originates from the metropolitan area of the city of Porto's public transportation system, a study of this subject will be presented, detailing its current mobility system, future proposals and metrics, and how this can be used in the development of this project. Finally, an overview of different approaches will be described, alongside with an analysis of these proposals, culminating in an overview closing section, that evaluates the literature review conducted and its importance in the development of this project.

2.1 Mobility and Transportation

While being defined in *Mirriam-Webster*'s dictionary [1] simply as "the ability or capacity to move" [(n.17], nowadays, mobility includes much more. As a matter of fact, the terms mobility and transportation come hand in hand, more often than not. Both are a requirement and a necessity of our society, crucial to maintain every citizen's quality of life, since businesses, institutions and services are deeply spread throughout a city's geographical area. Therefore, it is of the utmost importance that mobility is continuously improved upon, presenting newer, cleaner, and more efficient versions of itself, safeguarding not only the present, but also the future generations of our society [Wac10]. This is where the concept of sustainable mobility appears.

[1] https://www.merriam-webster.com/

Due to environmental awareness, being discussed and debated on a regular basis in today's world leaders gatherings and conventions, this need for a more sustainable mobility and clean public transportation usage is quickly becoming an important issue embraced by many countries worldwide [EW09].

2.1.1 Singapore - From an Area Licensing Scheme to a Car Circulation Limit

After experiencing a massive growth in private usage vehicles in the decade of the 1960's, and as a means of promoting a more sustainable approach to their circulation problem, in 1975, Singapore proposed an area licensing scheme that made it mandatory for drivers to purchase a license, in order to access specific metropolitan areas during peak traffic hours [Sei98]. While effective in reducing car circulation and increasing public transportation usage [CdM04], the system required police patrol in the designated access areas and has since been automatized with a toll mechanism that uses technologies like smart-cards and video analysis, with the objective of safeguarding it's purpose [Goh02].

Singapore's approach was the first of it's kind, but after proving it's effectiveness has, since then, paved the way for other cities to adopt similar mechanisms, like London or Trondheim [CdM04]. Nevertheless, Singapore was also the pioneer in acknowledging that these measurements are insufficient [Reu17]. Due to the large increase in the number of cars in circulation, as of February 2018, a vehicle circulation limit will be imposed, meaning that "aspiring car owners will have to wait for other drivers to give up their certificates in order to get permission" to legally drive in the city [kn:c].

2.1.2 São Paulo's Vehicle Circulation Restrictions

In an attempt to decrease the number of vehicles circulating in São Paulo, Brazil, a restriction protocol was implemented, regarding car circulation, first state wide, but now, comprising only its inner city area [CdM04]. This vehicle restriction paradigm states that during traffic peak hours (07:00-10:00 and 17:00-20:00), of weak days, car circulation is restricted based on an altercation of the last digit of the license plate being an odd or even number.

2001 studies indicate that these measures had a positive effect on the city's sustainable mobility, reducing traffic by 14% [Mah08]. Similar to this scenario, other Southern America cities, such has Mexico City, Bogotá and Santiago have also in motion a circulation restriction plan.

2.1.3 Oslo's Parking Ban Plan

Initially considering a more drastic approach then the aforementioned measures, the city of Oslo, in Norway, proposed a car ban policy in the city centre in 2015. The idea was to have a car free zone by the year 2019, but after facing enormous backlash from citizens, the car ban was replaced with a new two step policy. With the same deadline, Oslo is set to impose a city wide parking ban and pedestrian network extension, in an effort to reach sustainable mobility in a more progressive, graduate pace [CK17].

In 2019 an evaluation of the plan will be performed, assessing its effectiveness in the city's global environmental metrics. If the results are not satisfactory, the car ban proposal will be reassessed and possibly reimplemented.

2.1.4 Tokyo's Energy Producing Metro Stations

Opened since 1927, the Tokyo's metro system is one of the highest densely populated subway systems worldwide, with an estimated 14 billion passengers each year [Fal17]. Highly efficient, and holding two distinct underground systems [Wri15], Tokyo has, for a long time, been held as one of the best subway systems worldwide.

In an effort to promote a more sustainable transportation method, in December 2008, Tokyo metro stations took it one step further by implementing energy harnessing flooring tiles, that through the footsteps of the thousand daily commuters were able to produce clean energy [Rya08]. Implemented at the ticket gates, these piezoelectric sheets, have since been helping to power the subway stations themselves.

2.1.5 London's Intelligent System Management

In order to achieve a more sustainable and efficient transportation, the city of London has adopted three techniques with an innovative component, focused on revolutionizing it's transportation policy [GG06].

The first measure, similar to so many other cities, is a road pricing scheme that enforces drivers to pay a fee if they wish to circulate in central London on week days. This standard was implemented in 2003 and has it's earnings redirected to transit and transportation improvement spending. The second is a bus priority network that helps passengers avoid traffic jams, and navigate throughout the city more quickly [kn:16b]. Implemented in 2016, it has since helped the city's congestion problem and improved it's transportation efficiency. Finally, London has implemented an automated traffic enforcement measure, that though the usage of cameras helps to record illegal usage of the priority circulation networks [GG06].

2.2 Automated Fare Collection

Public transportation service providers are currently able to access and collect travel and passenger data automatically, in an effective, inexpensive, efficient and secure process entitled Automated Fare Collection (AFC) [Gor06]. This allows for a more productive payment method, that is not only faster, improving passenger circulation, that is critical in peak traffic hours [Adv18], but also reduces the need for a larger base of operation staff.

Despite AFC technologies originating with magnetic striped printed tickets, in the 1960's [Gor06], currently, and since the introduction of Hong Kong's *Octopus Card* [2], in 1997 [CP03],

[2] https://www.octopus.com.hk/en/consumer/index.html

and popularization with the London's *Oyster Card* [3] five years later [fL], contactless smart cards seem to have dominated both public transport service providers and user preferences [Qui08]. As a matter of fact, almost every capital and major city worldwide, have largely embraced this type of AFC technology and other electronic payment services. For example, the city of Madrid, Spain, as of the year 2018, has fully converted transportation AFC technologies exclusively to contactless smart cards [Mad18]. Other cities like Porto, Portugal, are even taking a step further setting to replace the already in vigour aforementioned system, with a mobile application that replaces physical cards by June 2018, entitled *Anda* [Cmp18]. Berliner Verkehrsbetriebe [4], also known as the transportation Company of Berlin, in association with the brand *Adidas* [5] are proposing an out-of-the-box idea, by incorporating stamped ticketing codes in shoes, and a walking validation system, increasing the speed of the ticket validation process [Viv18]. Owners of these shoes will be able to travel for free in zones A and B of the city's transportation grid. As a final example, the *East Japan Railway Company* [6] has developed an innovative business model idea called the *SUICA Super Urban Intelligent Card* [7][kn:16a] that uses a contactless smart cards, not only for mobility purposes, but also for shopping in specific stores [kn:a].

These mechanisms are made possible mainly due to two technologies: Bluetooth, in particular BLE communications, and NFC systems. While BLE uses a low consumption data transfer mechanism, taking advantage of Bluetooth beacons [Ban17a], NFC uses a wireless chip read-write short communication method, in order to, transmit information [CLCG15]. The effectiveness and simplicity of both mechanisms are the primary reasons for their extensive usage.

Nevertheless, as mentioned before, automated fare collection technologies offer more than just an innovative payment method, as they supply an enormous amount of important user and travel information that can be used in a variety of ways.

2.2.1 AFC's Data Usage

As mentioned before, the AFC's data can be used in many ways. One of these various applications is to make assumptions about transportation planning on a strategical, tactical and operational level [PTM11]. Using passenger validation data such as, travel time, location and card type, among other types of information, several studies have been able to develop predictions about transit patterns and withdraw well founded conclusions. While some studies have shown that an analysis of the AFC's data can lead to a more deeper understanding of user behaviour [MTA06], helping to identify potential marketing segments [AMT06], others focus on the effectiveness of the transit flow itself, and on finding ways to improve it [MCBO14]. These studies are not only crucial in improving the perceived attractiveness of public transportation among users [Bly04], but also in the adaptation of the transportation network, reallocation of passenger flows and trips estimation statistics.

[3] https://oyster.tfl.gov.uk/oyster/entry.do
[4] http://www.bvg.de/en/
[5] https://www.adidas.com
[6] http://www.jreast.co.jp/e/index.html
[7] http://www.jreast.co.jp/e/pass/suica.html

2.3 Recommender Systems

With everyday users being granted access to such large amounts of information via the Internet, it is now, more than ever, necessary for that information to be filtered, prioritized and efficiently tailored to each subject and personal needs or preferences [IFO15]. In this regard, recommender systems assume a huge role in this process.

Recommender systems are defined as software tools, algorithms and, above all, assistants in the social process of providing the knowledge "necessary to make choices without sufficient personal experience of the alternatives" [RVE97]. The recommendation process itself is divided in to three distinct phases. The first, an information collection phase that gathers data via implicit, explicit or hybrid feedback, a second regarding the learning component of the process, and a third and final phase that address the prediction and recommendations themselves [IFO15].

Currently, these systems are abundant, most noticeably in leisure, e-commerce, retail and scientific areas [AT05] and are making a breakthrough to almost every market.

2.3.1 Information Collection Phase

As mentioned before, regarding the information collection phase, there are three types of feedback categories when dealing with recommender systems. These address the information needed to make the recommendation itself, making sure that the data gathered is properly captured.

- Implicit feedback: this type of information collection occurs when the data regarding a user's preferences is gathered or inferred by the motorization of different actions the user might take part [IFO15]. These actions can origin from mainly three types of observable behaviour: examination, retention and referencing [OK98], as seen in Table 2.1

Table 2.1: Observable behaviour for implicit feedback [OK98]

Category	**Description**
Examination	This type of observable behaviour regards information that can be withdrawn when a user either selects an item, spends a given amount of time reading or observing an item, has repeated interactions with it, or purchases or subscribes to certain types of content with relation to said item.
Retention	This type of observable behaviour regards information that can be deduced when a user saves or prints an object that intends to use in the future.
Reference	This type of observable behaviour regards information that can be collected with the association between two objects. Common examples are forward replies, cut & paste functions and citations or quotations.

- Explicit feedback: this type of feedback is when the information collected is obtained through a rating system used by the user itself, in order to continuously improve the recommendations. This results in the accuracy of the recommendation being directly proportional to the number of ratings provided [IFO15].

- Hybrid feedback: as the name itself indicates, this type of feedback combines the strengths of both the aforementioned methods, with the objective of increasing the accuracy and efficiency of each recommendation.

While explicit feedback is still considered more reliable than its implicit counterpart, resulting in more confident recommendations, since it involves the intervention of the user itself, in the process, implicit feedback is set to be better in terms of performance and independence of the results themselves. New studies show that, in order to reduce biased input and the cognitive effort to process the ratings provided [OK98], implicit feedback could rise as a more effective solution. Nevertheless, combining the best from each method, a hybrid approach is set to decrease weakness and improve the overall performance of the system, resulting trustworthy and efficient solution.

2.3.2 Learning Phase

In order to perform an analysis of the information provided by enormous data sets collected through AFC technologies and properly use them in recommender systems, different data mining methodologies and algorithms must first be implemented. Only then, can actual knowledge be extracted and assumptions validated. These data mining techniques, through the combination of various methods in different scientific areas, such as machine learning and database theory, among others, help to categorize and find patterns in large amounts of information, otherwise nearly impossible to identify [BAH97].

Throughout the years, different classifications of data mining algorithms and function types have been defined [AMT06]. Currently, and as seen in Table 2.2 five distinct types emerge: classification, regression, segmentation, association and sequence analysis.

Table 2.2: Data mining algorithm types [Mic18]

Function Type	**Description**
Classification	Consideration of previous information knowledge to label the new data set (prediction of discrete variables).
Regression	Consideration of previous data to create a function to fill in missing values in new records (prediction of continuous variables).
Segmentation	Division of the population in different segments of similar behaviour according to a predefined metric (clustering).
Association	Establishment of correlations between dataset attributes.
Sequence Analysis	Summary of different sequences in data, that occur on a regular basis.

According to a study conducted in 2008 [WKR+], some the most commonly used data mining algorithms are described as follows:

- **C4.5 (classification algorithm):** This algorithm produces classifiers, intended to predict the class in which a new case belongs, by expressing them through decision trees or rulesets.
  The decision trees are constructed through a training set and follow, most commonly, two heuristics. The first, entitled information gain, aims to minimize the total entropy (unpredictability) of the subsets, while the second, default gain ratio, "divides information gain by the information provided by the test outcomes" [WKR+].
  Generically, this algorithm works as follows: first, it is necessary to search for the existence of base cases. Afterwards, for each attribute, it is necessary to discover the normalized information gain obtained by the partition (split) of this same attribute. Then, a decision node is created, splitting on the "attribute with the highest normalized information gain". Finally, a recursive method is implemented on the sub lists originated from the partition, and added those new nodes as children of aforementioned node [Nan11].
  On the other hand, ruleset classifiers try to simplify the understanding of complex decision trees, introducing "an alternative formalism consisting of a list of rules of the form *if A and B and C and ... then class X*, where rules for each class are grouped together" [WKR+].

- **k-Means (segmentation algorithm):** This iterative algorithm splits a new given dataset into a k number of clusters that are previously specified by the user [WKR+]. The algorithm begins by choosing k objects randomly. Then, for each one of them, assigns the n/k (n being the number of samples in the new dataset) closest attributes, being the proximity obtained by the Euclidean distance. Finally, in an iterative method, for each cluster defines the *centroid* (object that is the average of all of the objects in the cluster) and assigns the n/k with the highest proximity.

- **AdaBoost (classification algorithm):** In this algorithm, a base classifier is appointed using the original training set. Then, for each object of this same set, its weight is measured in accordance to how well the aforementioned classifier can predict the class. "The weight of an object define its probability of being selected for the training of the next base classifier" [Mor].

- **k-Nearest Neighbour (classification algorithm):** This "instance-based learning algorithm" [Nan11], works by discovering a group of objects, entitled k objects, in it's training set that are very similar to the new object, and uses the predominance of a specific class in it's neighbourhood to assign a label [WKR+].

Other algorithms supported by many of the existing tools, are, among others, the Naive Bayes classification algorithm and the classification and regression algorithms Random Forest and CART.

2.3.2.1 Technologies

Currently there are many technologies that through the implementation of data mining algorithms offer a construction ground stone for the implementation and development of recommender systems. Since the proposed solution is inserted in an academic environment, rather than the exploration of commercial tools, this section will focus only on the overview and analysis of open-source technologies.

Table 2.3: Overview of Data Mining Tools - *RapidMiner*, *R*, *Weka* and *Apache Spark*

Name	Description
RapidMiner [8]	Data science unified software platform, with a visual workflow design, broad connectivity and high scalability [kn:18]. Its primary focus is data mining computing processes, mainly supporting feature tables and time series [MR11].
R [9]	Highly extensible language and a statistical, and graphical, environment [kn:e], that supports feature tables and time series. Requires programming and script coding knowledge, highly appealing to programmers due two the easy implementation of data mining algorithms with pre developed extensions [MR11].
Weka [10]	Data mining platform and library developed in *Java* that holds a collection of machine learning algorithms, oriented to solve real-world problems [kn:d]. Commonly used in "algorithm development and applied research, for embedding data mining software into larger data mining software tools or specific solutions for narrow applications" [MR11].
Apache Spark [11]	Programming oriented "engine for large-scale data processing" [kn:b], highly efficient and scalable, with an easy deployment system. Usable with various languages such as *R*, *Java*, *Python* and *Scala*.

2.3.3 Prediction Phase

When considering the prediction/recommendation phase of recommender systems and their most commonly used filtering paradigms, three methods stand out.

- **Content-based**: This approach, which consists in producing a recommendation based on the comparison of items, rated previously by a user. This allows for the creation of a model that categorizes and profiles a user and his respective preferences according to the features

[8] https://rapidminer.com/
[9] https://www.r-project.org/
[10] https://www.cs.waikato.ac.nz/ml/weka/
[11] https://spark.apache.org/

of the items that were previously rated by that user [LGS11]. It uses as a algorithmic resources, most commonly, probabilistic models, decision trees and neural networks, being the recommendations underlined with machine learning techniques and statistical analysis [IFO15].

Nevertheless, this approach has some disadvantages mainly regarding the difficulty in providing a full system coverage analysis, the over-specialization of item recommendations, and the quantity of feedback vs. performance problem [BS97].

An example of a content-based recommender system is *LIBRA* [12], a book recommending tool that uses learning algorithms for text categorization [MR99].

- **Collaborative-filtering**: This paradigm presents a different approach, by recommending items that users with a similar profile have expressed interest in. Instead of computing similarities between items, this approach computes the similarities between users, and the correlation they present [BS97]. In order to do so, two different techniques can be used when dealing with collaborative-filtering. The first, entitled memory-based is built upon algorithms that use the entire user-item database, with the intent of generate a recommendation. The underlying premise is that "every user is a part of a group of people with similar interests" [SK09] and, as such, stored information about other individuals, and the ratings they gave to each item can be useful and should be taken in to account when a recommendation is formed. The second technique, model-based, is the use of data mining and machine learning algorithms to find patterns in previously rated items, and taking in account trained and learned data to create a well founded and intelligent prediction. Additionally a hybrid solution can also be used by combining the strongest aspects of both algorithms. Figure 2.1 presents a more thorough overview and comparison of these mentioned techniques.

  While solving some of the shortcomings of the content-based paradigm, collaborative-filtering presents some difficulties of its own. These include the addition of a new item to the database that cannot be recommended until another user has shown interest in it and the "unique user problem" that happens when a user has unusual preferences, not similar to those of any other user [BS97].

  GroupLens [13] is an example of this type of system, that uses the collaborative-filtering paradigm to filter the wide broad spectrum of available articles found in the Internet in to those the user actually has interest in consulting [RIS+94].

- **Hybrid**: as a combination of both of the aforementioned techniques, this paradigm explores the strongest features of both content-based and collaborative-filtering, to produce an optimal recommendation. It manipulates each algorithm, in order to, fix or improve the weaknesses the other might present, and as such improve the overall performance of the recommendation [IFO15].

[12] http://libra.inf.usi.ch/
[13] https://grouplens.org/

P-Tango is an example of this type of system, combining the speed of content-based methods, with the depth of the collaborative-filtering techniques in the information presentation of an online newspaper [CGM+99].

CF categories	Representative techniques	Main advantages	Main shortcomings
Memory-based CF	•Neighbor-based CF (item-based/user-based CF algorithms with Pearson/vector cosine correlation) •Item-based/user-based top-N recommendations	•easy implementation •new data can be added easily and incrementally •need not consider the content of the items being recommended •scale well with co-rated items	•are dependent on human ratings •performance decrease when data are sparse •cannot recommend for new users and items •have limited scalability for large datasets
Model-based CF	•Bayesian belief nets CF •clustering CF •MDP-based CF •latent semantic CF •sparse factor analysis •CF using dimensionality reduction techniques, for example, *SVD, PCA*	•better address the sparsity, scalability and other problems •improve prediction performance •give an intuitive rationale for recommendations	•expensive model-building •have trade-off between prediction performance and scalability •lose useful information for dimensionality reduction techniques
Hybrid recommenders	•content-based CF recommender, for example, *Fab* •content-boosted CF •hybrid CF combining memory-based and model-based *CF* algorithms, for example, Personality Diagnosis	•overcome limitations of CF and content-based or other recommenders •improve prediction performance •overcome CF problems such as sparsity and gray sheep	•have increased complexity and expense for implementation •need external information that usually not available

Figure 2.1: Overview of collaborative filtering techniques [SK09]

2.4 Public Transportation in the Metropolitan Area of Porto

After considering mobility as a whole, and being it the primary subject of studying, it is important to understand, on a more restricted scope, the Metropolitan Area of the city of Porto (MAP) [14]. This region includes 17 municipal territories and 1.700.000 inhabitants spread throughout an area of 2.040 square kilometres. 3 public transport operators (*STCP* [15] - Sociedade de Transportes Colectivos do Porto; *Metro do Porto* [16]; *CP* [17] - Comboios de Portugal) alongside with 29 private road service operators [MM18] supply around 700.000 trips per day in public transportation, across 653 route lines through trains, metro, buses, pluvial connections, funiculars and cable cars [AMP18].

[14] http://portal.amp.pt/
[15] http://www.stcp.pt/
[16] http://www.metrodoporto.pt/
[17] https://www.cp.pt

2.4.1 The *Andante* System

The main system behind MAP's public transportation services is the *Andante* [18] platform [Lea15], a contactless smart card AFC system, developed in 2002 [dP05b], that requires validation before a passenger initiates a trip, or when a transshipment occurs, but not when the trip is finished. As seen in table 2.4, it is currently accessible to users with two types of cards: *Cartão Andante Azul* and *Cartão Andante PVC* [dP05a].

Table 2.4: *Andante* system card comparison [dP05a]

Property	***Cartão Andante Azul***	***Cartão Andante PVC***
Rechargeable	Yes	Yes
Title Types	Single Trip and Daily	Single Trip, Daily and Monthly Subscription
Multiple Titles	Yes	Yes
Mixed Titles	No	No
Expiration Date	1 year	5 years
Usage	One person at a time	Owner only
Cost (in euros)	0.60	6.00

In addition to the title types presented in the table above, an *Andante Tour* is also available for tourists in a single or three day usage format [Seab]. Since the MAP region is extensive, fare prices for the *Andante* system vary according to the transition between specific zones (Figure 2.2). These transitions are ranked from Z2 to Z12 and specify the number of neighbour rings crossed between entry and exit stations [Seaa].

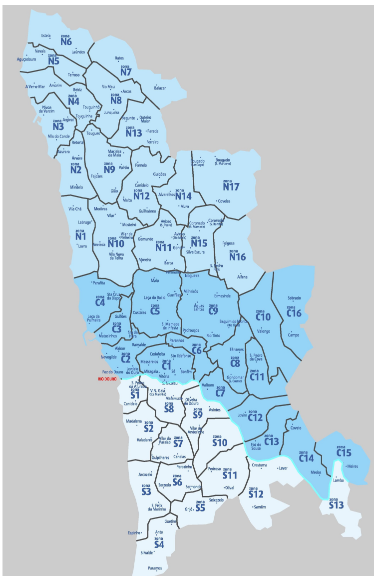

Figure 2.2: MAP Zone Division [Seaa]

[18] http://www.linhandante.com/

2.4.2 The *Anda* System

Currently undergoing a testing phase and set to launch in June of 2018 [Cmp18], the *Anda* system is a pilot-project developed as a partnership between the *Transportes Intermodais do Porto* and the *Faculdade de Engenharia da Universidade do Porto* [dP17], that presents a "mobile ticketing solution" through the use of smartphone devices [Ban17b]. Using both BLE and NFC technologies [dP18], this solution not only allows for a simplified usage of public transportation in the MAP region, but is also pioneer in passenger monthly fare optimization. It uses a post-billing mechanism, as well as, a favourable fare calculation system, that doesn't require prior knowledge from the user regarding fares and transportation zones [dP17].

2.4.3 AMP2020 Strategy

Aiming for a more sustainable mobility, Porto is currently implementing a strategic planning protocol entitled *AMP2020*, with a 6 year duration, that started in 2014 [AMP13]. This plan presents an intelligent, sustainable and inclusive growth paradigm for the MAP region, that includes not only, the promotion of sustainable transportation and reduction of overcrowding and network infrastructural blockages, but also promotes public transportation and its usage.

Regarding mobility itself, several steps have already been taken. These include the aforementioned usage of a mobile application in alternative to the *Andante* smartcards [Cmp18], the proposal of a new metro line, between *São Bento*'s and *Casa da Música*'s stations and a 3.2 kilometre extension of the yellow line [Not18a] and a metro wagon remodelling system, undergoing currently a testing phase, allowing more users to travel together simultaneously [Not18b].

2.5 Other Approaches

After consideration of the various disciplines relevant to the sustainable mobility problem and after analysing the current paradigms of today's society regarding this subject, it is important to understand and revise previous approaches regarding this, or similar challenges.

2.5.1 Mobile Technologies Multiservice Approach

Presenting an innovative approach on the encouragement of public transportation as a means to advance society's paradigm and stance on sustainable mobility, the mobile technology based multiservice solution was proposed in 2015, defining a conceptual model that combined public transportation service providers with local businesses in an everyday user-oriented mobile platform [FG13].

The conceptual model defined four distinct stages sequentially linked. First, a mobility profile analysis and user preference evaluation, that would need both explicit data elicitation methods, as well as system recording techniques, in order to accurately retrieve customer choices. Afterwards, a data mining phase that would categorize users into different target clusters and an analysis of the possible delivery service offerings. Third, as any project with a practical intent, a measurement stage of the effectiveness and efficiency of the system would be required. Lastly, a phase of improvement of the service offerings would set the tone for another iteration of the model, now updated and enhanced.

While being an innovative approach on the subject, considering an holistic view of transportation, this solution presents an opportunity for an empiric testing and development of a mobile system that, not only considers the aforementioned aspects, but also is built upon an intelligent user-centred recommendation system.

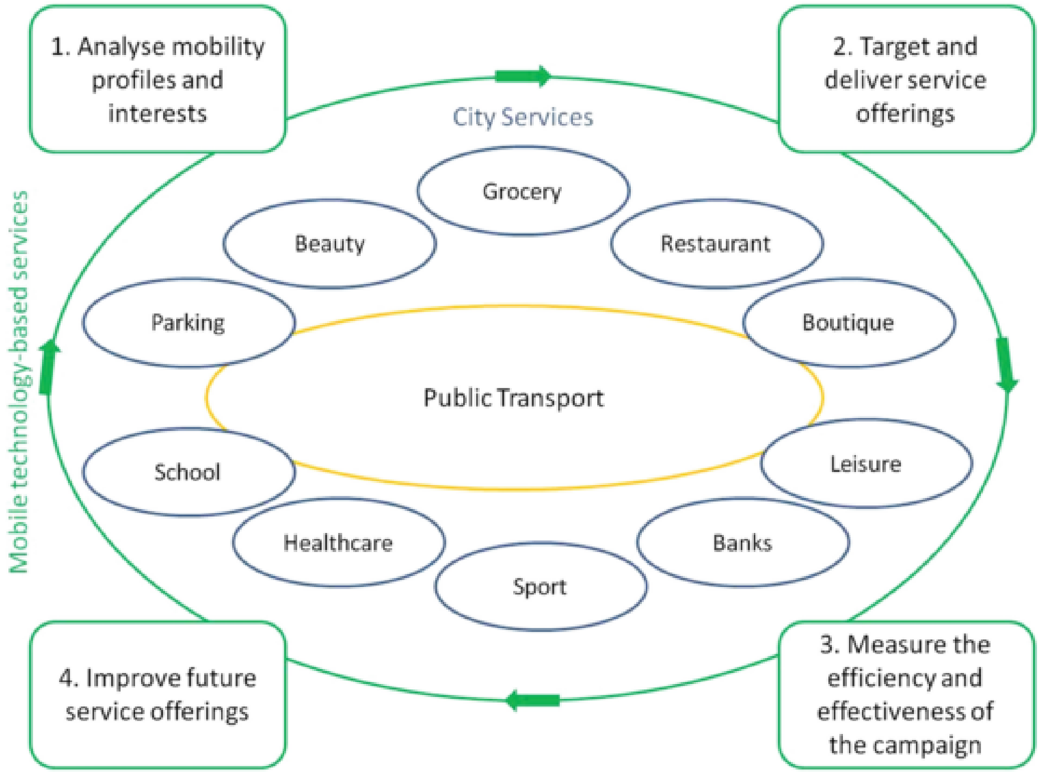

Figure 2.3: Conceptual model of the multiservice approach [FG13]

2.5.2 MARS

The Mobility Aware Recommender System or MARS is a project based on an a master's dissertation of the Faculty of Engineering of the University of Porto, that approached the mobility awareness problem through the development of a recommender system that integrated location and frequency as its main variables. It produces recommendations based on user preferences, public transportation mobility patterns and service characteristics and location [LCG15].

The project's methodology included data gathering and analysis of the public transportation network of the Metropolitan Area of the City of Porto via smart cards, user profile categorization and algorithmic implementation using the variables mentioned before.

The overall recommender system uses a content-based approach [LGS11] and was able to produce promising results regarding recommendations in highly frequented areas. Nevertheless, the system could be improved with a hybrid recommendation approach, taking in consideration user input as well, by using a clustering algorithm and considering the time factor, that drastically shifts the transportation patterns throughout a year. Finally, it also presents the opportunity for an one-to-one implementation of a recommender system, through an user available mobile application.

2.5.3 PILGRIM

PILGRIM is a mobile location broker and mobility awareness recommender system that not only filters Internet links based on a user's location, but also maintains an historic database of previous locations and corresponding links, as a means to help identify patterns and calculate user preferences [BB03].

The main idea behind it is to promote local services, that have their websites buried in the Internet's vast amounts of information, by presenting a mobile solution that has fewer, more specific links than other search engines.

Despite being an interesting solution for the local cultural service's promotion, it targets a very specific problem and presents some scalability and user security issues due to the algorithms used. Additionally, no empirical testing data was collected, and the system lacked an experimentation in a real life scenario validating it's overall performance.

2.6 Conclusions

The literature review allowed, not only to perceive the problem inserted in a more wider environment, but also, to understand how to act and build upon existing knowledge that helps the development of a fuller and more robust solution. The study helped to understand the current paradigm regarding mobility and how cities and nations worldwide are implementing mechanisms to aim for a more sustainable future. It was well underlined the importance of automated fare collection systems and technologies and how data gathering can be used to improve sustainable transportation. The analysis of the state of the art of recommender systems provided the knowledge for the development of the solution regarding both algorithms and technologies, as well as, methodologies to use. Overseeing the Metropolitan Area of Porto was also important to understand the target environment of the solution and what approaches are currently being implemented.

Finally, revising other approaches to the same or similar problems narrowed the field of action, while enabling the withdraw previous experiences that could help create an effective, efficient and innovative solution.

Chapter 3

Mobility Profiling and Passenger Activity Analysis

Given the conclusions found when revising both the current context and paradigm, as well as the state of the art and other projects on this matter, three main questions arose when defining the problem statement of this project. The first problem at hands would be how to promote sustainable mobility through the increase usage of public transportation, the second, how to approach this problem through an holistic perspective, contextualizing the purpose of each trip, and finally, how to combine public transportation with local city services. In order to address the aforementioned problem, a solution consisting of a personalized recommender system based on mobility profile analysis and passenger activity was developed. Despite being a highly complex system, the methodology of the developed project can be divided into three different stages. This chapter presents the first, a detailed analysis of the available datasets, that allowed to withdraw conclusions about passenger activity allowing to map where, when and how users travel, and a profiling component that grouped/categorized different users according to different mobility profiles.

3.1 Available Datasets

For the development of this project, two distinct sets of data were supplied by *Transportes Intermodais do Porto* (TIP), in order to generate the mobility profiles required for the recommendation process and for the analysis of passenger activity. The first, as described in subsection 3.1.1, is a stations dataset, mainly used as a connection point between the places to be recommended and the passenger activity. This activity is present in the second dataset, as seen in subsection 3.1.2, and is comprised of passenger validations throughout the year 2013 in the *Andante* system. Both datasets were crucial in the development of the proposed solution, as they allowed for a more accurate and close to real life portrait of the applicability of the system, given the authenticity of their information.

3.1.1 Stations

Allowing to map the city's local services to public transportation trips, the first of the available datasets consists of a 4075 entry file of stations in the MAP's area, regarding the bus operators *STCP - Sociedade de Transportes Colectivos do Porto* and *ES - Espírito Santo*, the metro operator *MP - Metro do Porto* and the Portuguese train service *CP - Comboios de Portugal*. In table 3.1, the known fields of each of the datasets entries are presented:

Table 3.1: Station dataset fields

Field	**Type**	**Description**
name	String	Station name
type	String	Station operator type
zone	String	Corresponding *Andante* zone
latitude	Double	Location latitude
longitude	Double	Location longitude
newCode	Integer	Station identifier
oldCode	Integer	Previously used station identifier
realCode	Integer	Corresponding identifier used by TIP

The positioning information included in this dataset (latitude and longitude variables), allowed for a showcase of the stations distribution within the MAP's area, as seen in figure 3.1, where it is undoubtedly noticeable the higher density of stations in the central region, as opposed to the southern and northern regions. Complementing this information, table 3.2 allowed for a comparison of the number of stations per transportation service operator supplying a much needed understanding of how these stations are distributed.

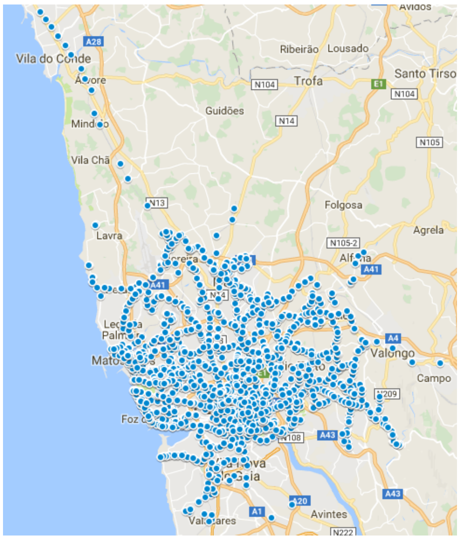

Figure 3.1: Station distribution over the MAP's area

Table 3.2: Station distribution per service operator

Operator	Number of Stations
CP	26
Espirito Santo	61
Metro do Porto	82
STCP	3906

3.1.2 Validations

As mentioned before, the validations dataset corresponds to the approximately 130 million validations of passengers of the *Andante* system, throughout the year 2013. These validations are for trip entry or transshipment, providing no information about a passenger's exit location. Table 3.3 presents the different fields in each of the dataset's entries.

Table 3.3: Validation dataset fields

Field	Type	Description
id	Integer	Validation identifier
op	String	Transportation service operator
cod_cartao	Integer	Card code
Paragem	Integer	Station identifier
Linha	Integer	Transportation Line
Sentido	Integer	Line Direction
Variante	Integer	Unknown variable
Veiculo	Integer	Vehicle identifier
zonamento	Integer	Unknown variable
dhiniviagem	DateTime	Trip initial datetime
datahoravalidacao	DateTime	Validation datetime
Tip_Valid	Integer	Unknown variable
Grupo_Cod	Integer	Group Identifier
Grupo	String	Group name
Perfil_Cod	Integer	Profile identifier
Perfil	String	Profile name
zonas	String	Valid circulation zones
Zona	String	Initial Stop Zone
NO. of Validations	Integer	Number of validations of the card within the valid time
Entry Stop	Integer	Initial station identifier
Exit Stop	Integer	Final station identifier
Type	Integer	Trip type (specifies if the trip is a transshipment or final destination)

For the development of the proposed recommender system, this data was filtered in order to consider only validations regarding the month of January of the aforementioned year, increasing computer calculation time and processing capacity.

The month of January accounts for a total of 21.953.417 validations and can be categorized into 12 different groups according to the type of card validated. Table 3.4 and figure 3.2 present these results, comparing the number of validations in this month, according to the respective car type.

Table 3.4: Number of January validations per card type

Card Type	**Number of Validations**
Título de Viagem	3956664
Ocasionais STCP	468142
Clikz Título de Viagem	43598
Clikz Andante 24	62
Bilhete Euro (3 Dias)	7
Bilhete Euro (1 Dia)	7
Assinatura STCP	2776022
Assinatura Fim de Semana STCP	21991
Assinatura	14617382
Andante Tour3	14843
Andante Tour1	27002
Andante 24	27697

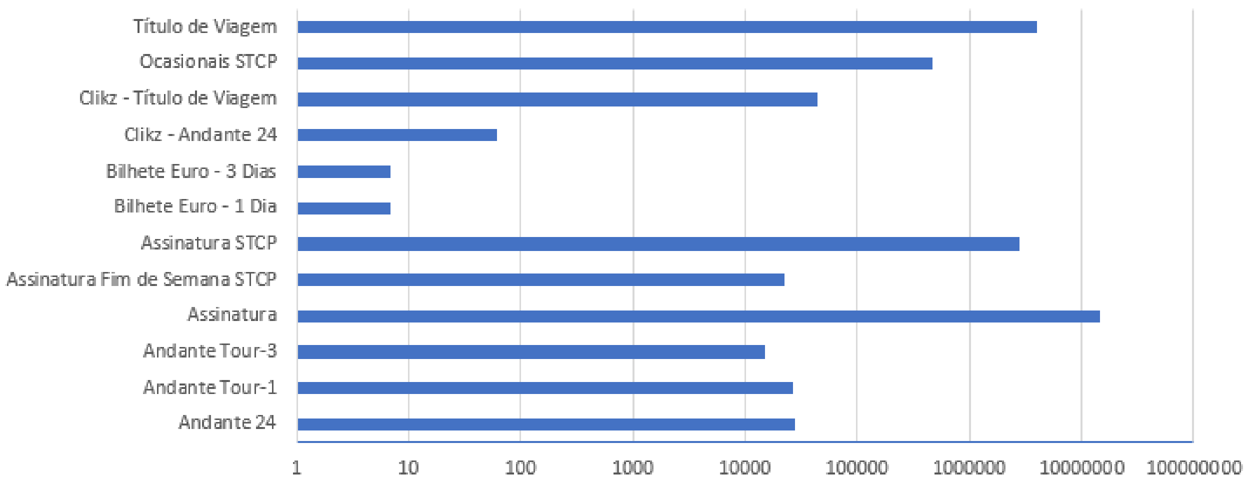

Figure 3.2: Number of January validations (Logarithmic scaling)

Upon analysis, it is visible that both monthly subscription cards (*Assinatura* and *Assinatura STCP*). and single unmarked tickets (*Titulo de Viagem* and *Ocasionais STCP*), make up almost the entirety of the validations for the studied month.

As monthly subscriptions (*Assinatura, Assinatura STCP and Assinatura Fim de Semana STCP*) combined outnumber validations by any other type of card, and since they are correctly labelled by the *Andante* system, they serve as the better study sample for the development of this project. Monthly subscription validations can be divided according to predefined profiles, as seen in the following table (table 3.5). Described in greater detail in section 3.2 these available profile types served as the ground stone for passenger profiling.

Table 3.5: Number January validations per profile

Profile Type	**Number of Validations**
13_18_escola.tp	212
13_18_escola.tp A	370929
13_18_escola.tp B	230752
13_18_escola.tp F	35689
3 Idade	3769061
4_12_escola.tp	61
4_12_escola.tp A	115069
4_12_escola.tp B	48404
4_12_escola.tp F	5905
Combinado STCP/ CP	57379
Combinado STCP/ Operador Privado	36027
Especial	7542
Estudante	2149648
Menor13	68062
Normal	7327024
Normal M	1145
Ref./ Pens.	166928
Social+	2155452
Social+ A	317193
Social+ D	43190
Social+ R	152245
sub23_superior.pt	496
sub23_superior.pt A	337019
sub23_superior.pt F	19963

Transportation operator usage analysis was also performed as seen in figure 3.3. For the month of January, and taking in consideration only monthly subscription cards, *STCP* and *Metro do Porto* public transportation service operators account for all validations, being the first responsible for 81% of the entries (14.122.837) and the latter for the remaining 19% (3.292.558).

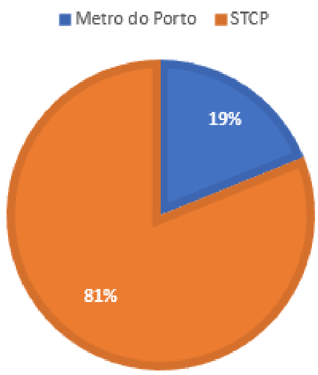

Figure 3.3: Transportation operator usage

When analysing the distribution of passenger validations per *Andante* zone, it is visible the influence the central (Zone name started with C) area of the MAP's region, accounting for the majority of the validations of the month in question. Table 3.6 presents the mentioned distribution, matching each zone with the respective number of validations.

Table 3.6: January validations distribution per *Andante* zone

Zone	**Number of Validations**
N2	19.723
N3	54.048
N10	69.634
N11	54.123
N14	8.150
N16	47.624
C1	6.916.045
C2	1.841.781
C3	399.327
C4	41.507
C5	674.557
C6	1.626.585
C8	496.353
C9	738.319
C10	81.233
C11	62.671
C16	10.021
S1	43.777
S2	125.364
S8	1.214.422
S9	92.118

Figure 3.4 presents the distribution of validations per *Andante* zone area, dividing validations according to a northern, central or southern area of the MAP's region.

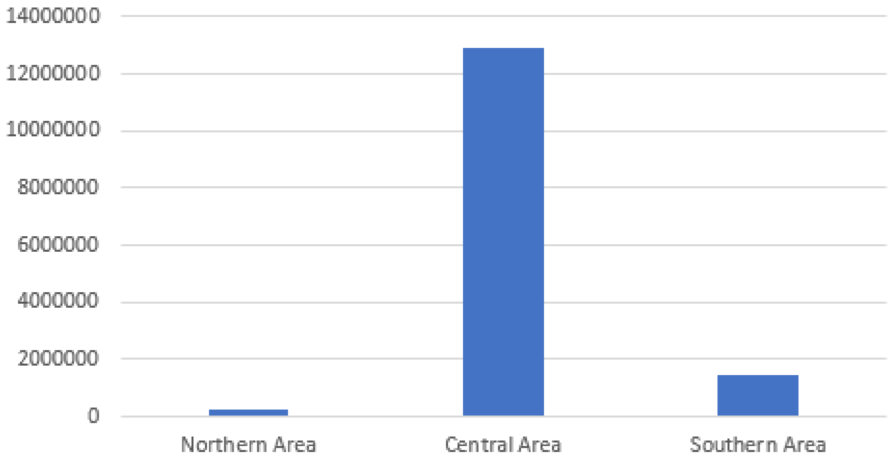

Figure 3.4: Validation distribution per zone area

Finally, a temporal analysis was also conducted, and was divided into 2 distinct sets of approaches. The first, as shown in figure 3.5, presents validation distribution divided in weeks [1] throughout the month of January, giving a better understanding of how dataset partitioning could be used, dividing it into subsets of training and testing.

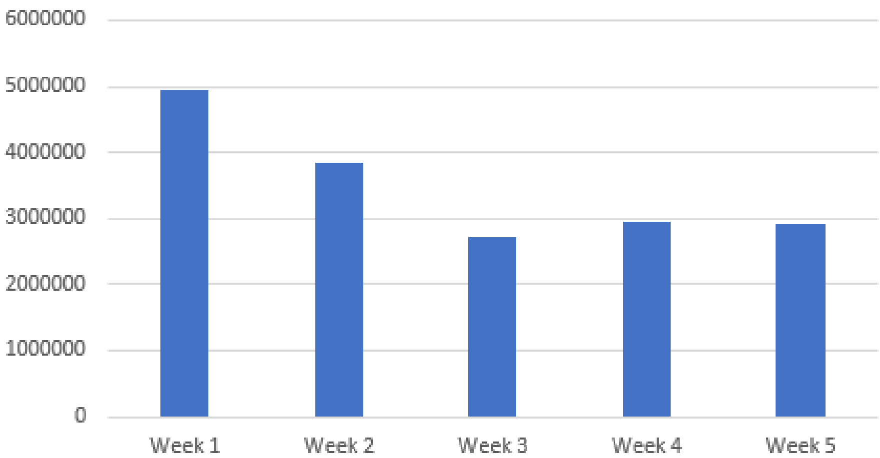

Figure 3.5: Validation distribution per week over the month of January

Table A.2 presents the absolute values of the temporal analysis with the corresponding week.

The second analysis was executed in order to better understand which hours of a day, hold the highest number of validations (figure 3.6).

[1] Week 1 - 01/01/2013 to 05/01/2013; Week 2 - 06/01/2013 to 12/01/2013; Week 3 - 13/01/2013 to 20/01/2013; Week 4 - 20/01/2013 to 26/01/2013; Week 5 - 27/01/2013 to 31/01/2013;

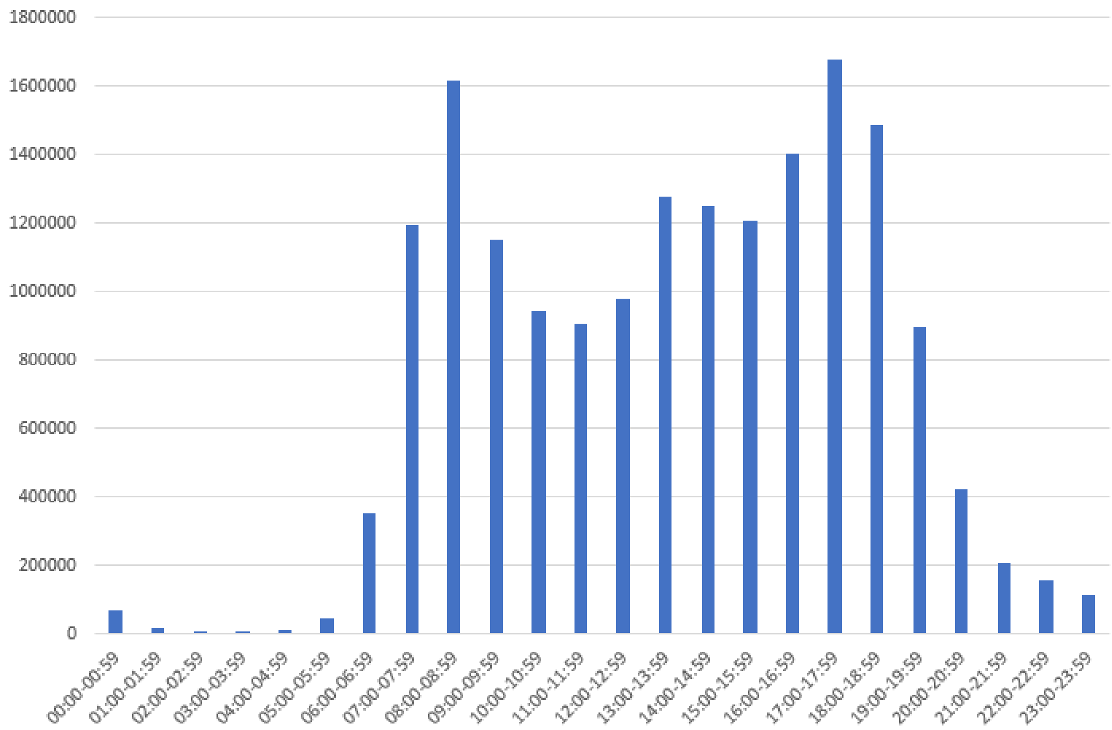

Figure 3.6: Validation distribution per hourly interval

Table A.1 presents the absolute values of the temporal analysis with the corresponding hourly interval.

3.2 Profiling

As previously mentioned, the *Andante* monthly subscription card labelling was vital in the formation of different profiling groups, that were used as a cold start clustering mechanism used by the recommender system. It is the corner stone for the collaborative-filtering technique proposed by the developed system. In that sense, it is first required to understand what the labelling meant and what assumptions would be valid to withdraw from that data. Tables 3.7 and 3.8 present a more through comprehension of each of the data profile types.

Table 3.7: Profile type description (Part 1) [kn:12]

Profile Type	**Description**
13_18_escola.tp	*Andante* monthly subscription fare for students between the ages of 13 and 18 years old that do not attend universities and do not have public transportation assured by the municipal counties. Can be divided in type A (50% discount), type B (25% discount) or F (Family plan - 25% discount).

Table 3.8: Profile type description (Part 2) [kn:12]

Profile Type	**Description**
4_12_escola.tp	*Andante* monthly subscription fare for students between the ages of 4 and 12 years old that do not have public transportation assured by the municipal counties. Can be divided in type A (50% discount), type B (25% discount) or F (Family plan - 25% discount).
Combinado STCP/ CP	Fare discount card for STCP and CP usage only.
Combinado STCP/ Operador Privado	Fare discount card for STCP and private transportation operators usage only.
Especial	Unknown
Estudante	*Andante* monthly subscription fare for all students until the age of 25, with no zone restrictions holding a 25% discount on the origin value.
Menor13	*Andante* monthly subscription fare for people under the age of 13, holding a 25% discount on the origin value.
Normal	Normal *Andante* fare.
Normal M	Unknown
Ref./ Pens.	Senior fare for retired or pensioner holding a 25% discount.
Social+	*Andante* monthly subscription fare for people with state financial support. Can be divided in type A (50% discount), type D (25% discount) or R (Elderly only - 25% discount).
sub23_superior.pt	*Andante* monthly subscription fare for students under the age of 23. Can be divided in type A (50% discount) or type F (25% discount).

Given this information, two distinct groups of profiles were created: the first, according to age stratus (Young, University Student, Adult and Elderly) and the second according to the existence of financial aid or not (Normal and Beneficiary). Table 3.9 presents how each group was created, taking in consideration the labelling of the *Andante* system [2].In other words, the Young mobility profile, represents people between the ages of 4 and 18, enrolled in a scholastic environment, while the University Student profiles, represents students enrolled in in Universities, between the ages of 18 and 25. The Adult profile takes into consideration all other non studying young adults and people between the ages of 25 and 65, excluding retired people, who fall in the last category of Elderly profiles, representing not only them but all elderly people above the age of 65 and pensioners. The financial distribution is more straightforward, dividing between Beneficiaries

[2]Label types *Combinado STCP/CP*, *Combinado STCP/Operador Privado*, *Especial* and *Normal M* were not used in the grouping of the mobility profiles either due to the wideness of their scope or lack of information.

those who get state well fare support or other benefits, and those who don't (Normal mobility profile).

Table 3.9: Profiling

	Income		Age			
	Normal	**Beneficiary**	**Young**	**University Student**	**Adult**	**Elderly**
13_18_escola.tp	x		x			
13_18_escola.tp A		x	x			
13_18_escola.tp B		x	x			
13_18_escola.tp F		x	x			
3 Idade	x					x
4_12_escola.tp	x		x			
4_12_escola.tp A		x	x			
4_12_escola.tp B		x	x			
4_12_escola.tp F		x	x			
Estudante	x			x		
Menor13	x		x			
Normal	x				x	
Ref./ Pens.		x				x
Social+		x			x	
Social+ A		x			x	
Social+ D		x			x	
Social+ R		x				x
sub23_superior.pt	x			x		
sub23_superior.pt A		x		x		
sub23_superior.pt F		x		x		

Figures 3.7 and 3.8 present a graphical overview of the card distribution of the month of January according to each profile.

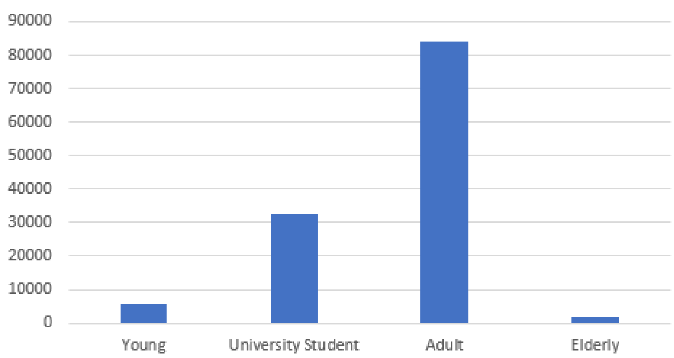

Figure 3.7: Profiling by age stratus card distribution

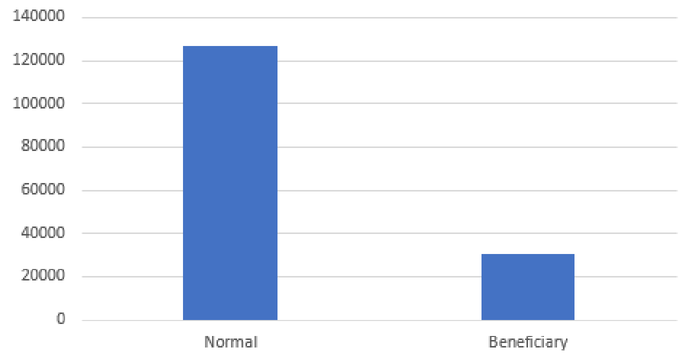

Figure 3.8: Profiling by financial stratus card distribution

Nevertheless, it is important to consider that while this grouping methodology is a great starting point for the definition of the groups, it holds a certain amount of uncertainty and error margin associated with it, due to possible overlaps of labelling, mainly present in the Young and University Student profiles.

Chapter 4

MAP's Local Services

In a follow up component of the work methodology, local services of the Metropolitan Area of Porto, were extracted with the Google Places API, and were used has the recommendation POI's, or points of interest, of the developed solution.

4.1 Extraction Process

The extraction process of the system's points of interest is an iterative process, where for each of the stations available in the dataset, its latitude and longitude is retrieved and used as the centre of the search circle of local services. This area is defined by a search radius of 1000 meters and, among others, retrieves the following information fields, that were used in the recommender system:

Table 4.1: Retrieved fields from the Google Places API

Field	Type	Description
GoogleID	String	Unique place identifier used by the Google platform for each place
Name	String	Place name
Latitude	Double	Location latitude
Longitude	Double	Location longitude
Rating	Long	User based rating of the POI
Address	String	Place address location
Price Level	Double	Service price level (range 0-5)
Type	String	Place type
Photo	String	Place photograph identifier
Open Now	Boolean	Returns true if the place is open when the request is made

Furthermore, POI's were also filtered by type, in order to improve computing processing time, and to adequately fit the problem at hands, since not all services would be able to supply discounts or other types of offers. In that sense, the following table presents the 20 used place types:

Table 4.2: Place types used for recommendation

Types				
Amusement Park	Aquarium	Art Gallery	Bakery	Bar
Beauty Salon	Bicycle Store	Book Store	Cafe	Clothing Store
Convenience Store	Florist	Gym	Museum	Night Club
Restaurant	Spa	Stadium	Store	Zoo

4.2 Obtained Results and Analysis

After performing the aforementioned process, specified in table 4.2 to the 4075 available stations, that is, by only considering POI's with a place type tag equal to one of the preselected types, 1013 points of interest were selected. These POI's account for 16970 associations between stations and places, and their distribution can be seen in figure 4.1.

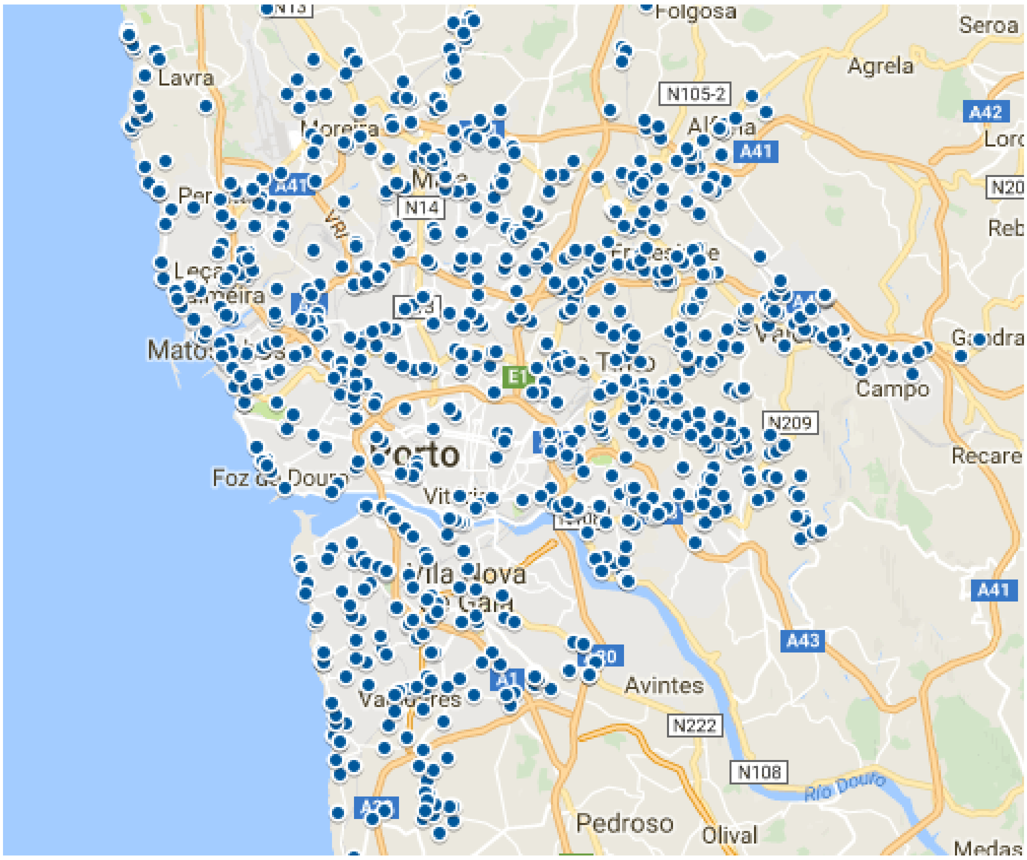

Figure 4.1: Partial POI distribution over the MAP's area

The analysis of the number of POI's per *Andante* zone [1] resulted in the following tables:

[1]Note that only stations with POI's are presented.

Table 4.3: POI's per *Andante* zone (Northern Area)

Zone	N1	N2	N3	N10	N11	N15	N16	N17	N18
Number of POI's	24	33	13	17	41	8	23	6	6

Table 4.4: POI's per *Andante* zone (Central Area)

Zone	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C16
Number of POI's	29	59	51	80	72	42	29	72	147	53	28	20

Table 4.5: POI's per *Andante* zone (Southern Area)

Zone	S1	S2	S3	S7	S8	S9
Number of POI's	16	50	15	22	35	22

The following chart in figure 4.2 presents the collective information presented in the tables 4.3, 4.4 and 4.5 where it is notorious the amount of POI's in the central region when compared to other regions of the MAP's area [2].

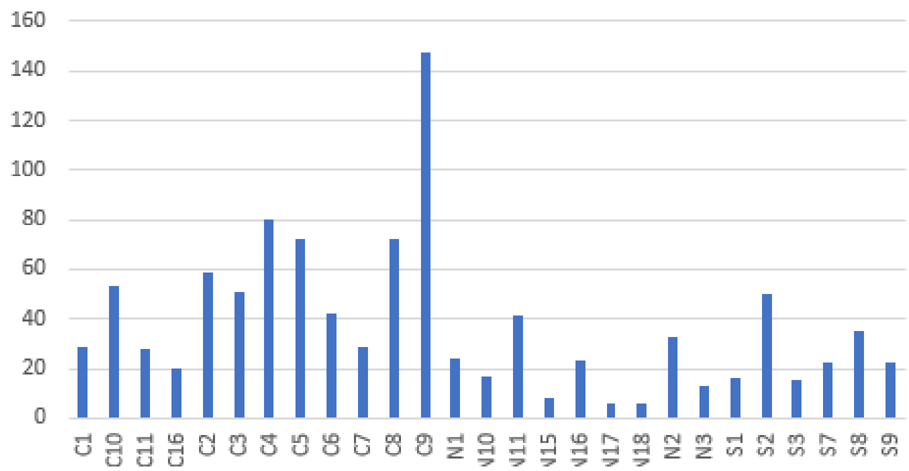

Figure 4.2: POI distribution over the MAP's area

This analysis allowed to perceive that, in order to achieve more and different POI recommendations, given the performance time limitations, the *Andante* zones C4, C5, C8 and C9, present the best study target, having the highest count of POI's in a single zone, despite not having necessarily the highest number of passenger validations.

[2] Zone C9 presents an abnormally high POI concentration. Despite not being able to withdraw any conclusions about the cause of this disparity, it could be a result of the wide distribution of stations in this area.

MAP's Local Services

Chapter 5

Recommender System

As previously mentioned, the proposed solution is a recommender system that generates tailored, user-centred recommendations based on a user's mobility profile and public transportation activity. In that sense, given that mobility profiles were already covered in Chapter 3, as were the recommendation items as well (Chapter 4), it is now important to define what is the core of the proposed solution: the recommender system.

This component is responsible for both connecting all other components and engineering results, through mobility profile assignment, passenger activity analysis, POI recommendation probability calculation and feedback handling. Being a complex structure system, this chapter will be devoted to detail an overview of the recommender system's architecture, database and methodology, as well as present an analysis of the developed user mobile system and its usability.

5.1 System Architecture

The architecture of the developed solution is mainly composed of three distinct blocks, as portrayed in figure 5.1.

The first block, detailed more thoroughly in section 5.2, is composed by the system's database, a local MySQL database, in charge of not only holding relevant entities and information, but also in charge of the associations that those entities have with each other. In similarity to all of the other components, this database runs locally, given the amount of data it holds, maximizing the system's overall response time.

The second block corresponds to the *RapidMiner* analysis, critical in processing performance and profile definition, that allowed a proper categorization and validation processes. Despite having a standalone application, its use is enabled by a built in plugin in the system's server.

The main and final block is the Java Server. It holds the modules for POI extraction, that directly communicates with the Google Places API, for profiling definition and station parsing, communicating both with the *RapidMiner* block and the local database connector, and for the recommender system module, detailed in section 5.3, in charge of producing the recommendations themselves.

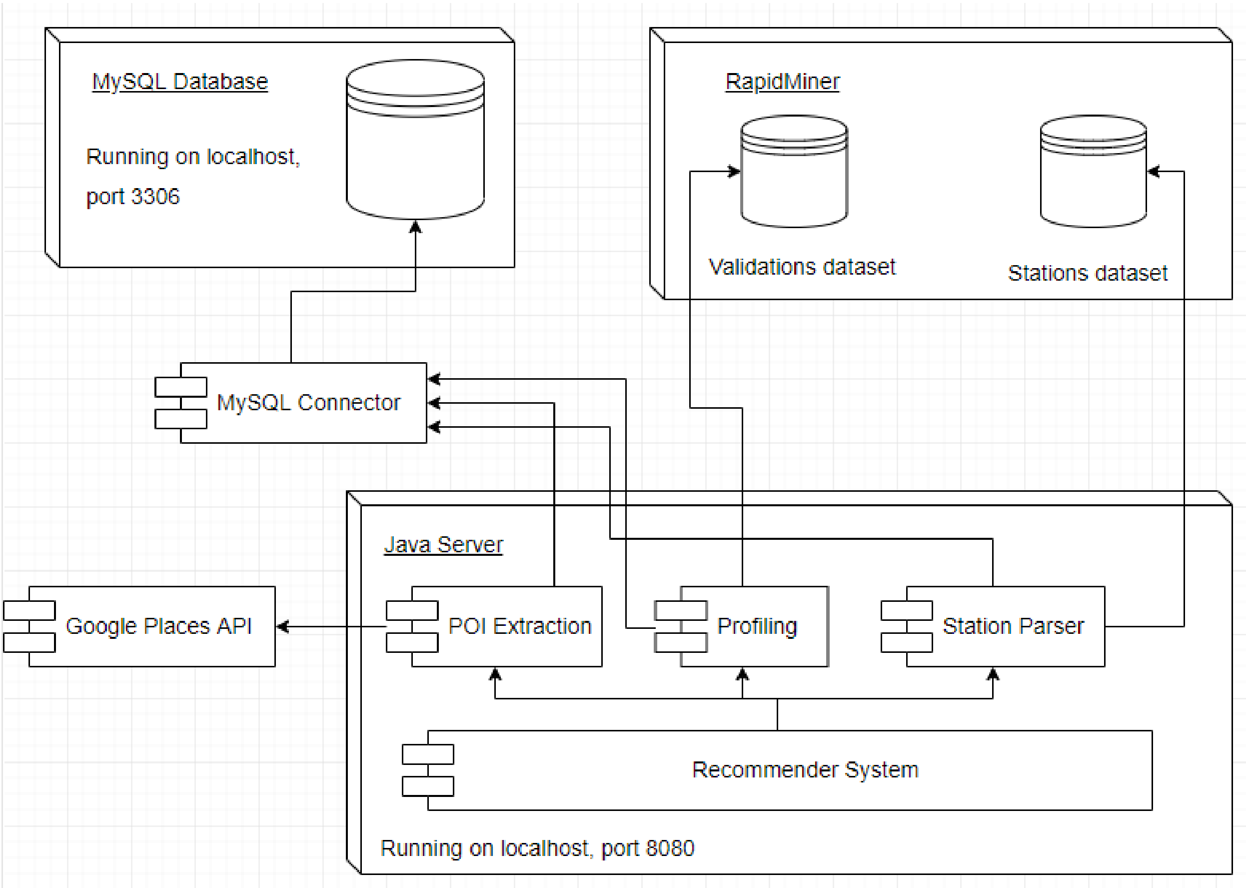

Figure 5.1: System Architecture Diagram

Additionally, a mobile prototype was also developed, and is detailed in section 5.5

5.2 System Database

Apart from the datasets specified in section 3.1 the recommender system uses a local *MySQL* database comprised of 10 tables representing each of the entities and relationships between them.

The idea behind this database is that each Recommendation is a correspondence between a Card and a Place with a given probability, user feedback and novelty status. Each Card has one or more Profiles associated with it and establishes this association with another table named Profile-Card. The Distribution class connects a Profile with a PlaceType, while each Place is associated with a Station via a StationPlaces table. Finally, each Station has one Zone to which it belongs, being that a zone can have multiple stations.

Figure 5.2 presents a diagram of the architecture of the database, previously described.

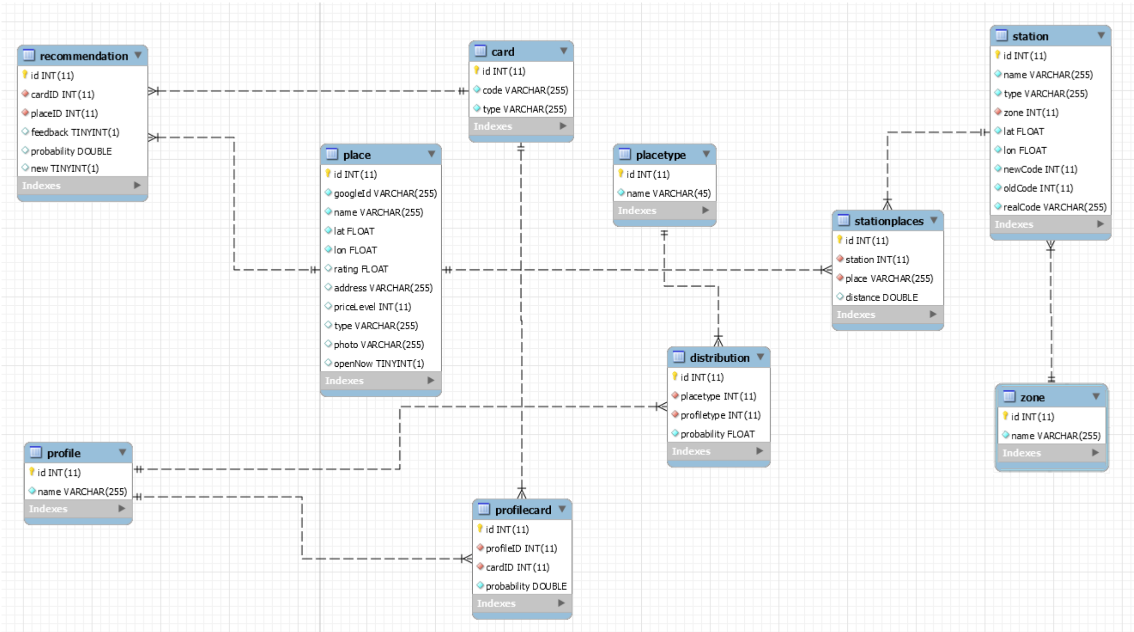

Figure 5.2: Database Architecture Diagram

5.3 Methodology

The development methodology of the recommender system underwent a series of phases, both regarding development and testing. The development methodology can be divided into 3 great groups: Mobility profiling and passenger activity analysis, POI extraction and the recommendation process itself. While the first two were already covered in previous chapters (Chapter 3 and Chapter 4 respectively), the recommendation process has yet to be discussed. On the other hand, the testing methodology is divided into 4 distinct components: a first stage where recommendations are generated from the training dataset, a second where persona feedback is applied to the aforementioned dataset, a third where new validations are introduced to the original recommendations dataset and a final stage where new validations are introduced to the already feedback influenced recommendations. This section will cover these subjects presenting a detailed analysis of each process as well as the data scope decisions that were made.

5.3.1 Data Scope Decisions

Despite already being loosely covered in the previous chapters, given the large amount of available data, scope decisions were taken throughout the development of this project, in order to achieve valid results in an acceptable time frame. In that sense, this section will cover these three major decisions that heavily influence the outcome of the results and the structure of the recommender system.

As mentioned in subsection 3.1.2, the original validations dataset accounted for all the validations registered during the year of 2013 (approximately 130 million validations). While some interesting factors could be extracted when addressing the entirety of the dataset, mainly regarding seasonal passenger flow distribution, for the development of a prototyping recommender system, capable of producing a valid starting point for a mutual benefit business model that addresses sustainable mobility, a portion of this dataset is considered to be sufficient. In that sense, the month of January of 2013 (21.953.417 validations) was selected as the training and testing dataset of the recommender system.

The second major decision was to filter validations that only matched monthly subscription card types. This allowed to reduce the error margin associated with the profiling component, as these cards are already labelled by the *Andante* system, as seen in section 3.2 (17.415.395 validations).

Finally, and given the information collected in chapter 4, it was important to validate the generated recommendations with a large sample of POI's. Thus, for the result analysis component of this solution, 4 distinct *Andante* zones were selected, guaranteeing a high POI to station ratio. The selected zones chosen were C4, C5, C8 and C9.

Figure 5.3 showcases these decisions in a diagram with a complete overview.

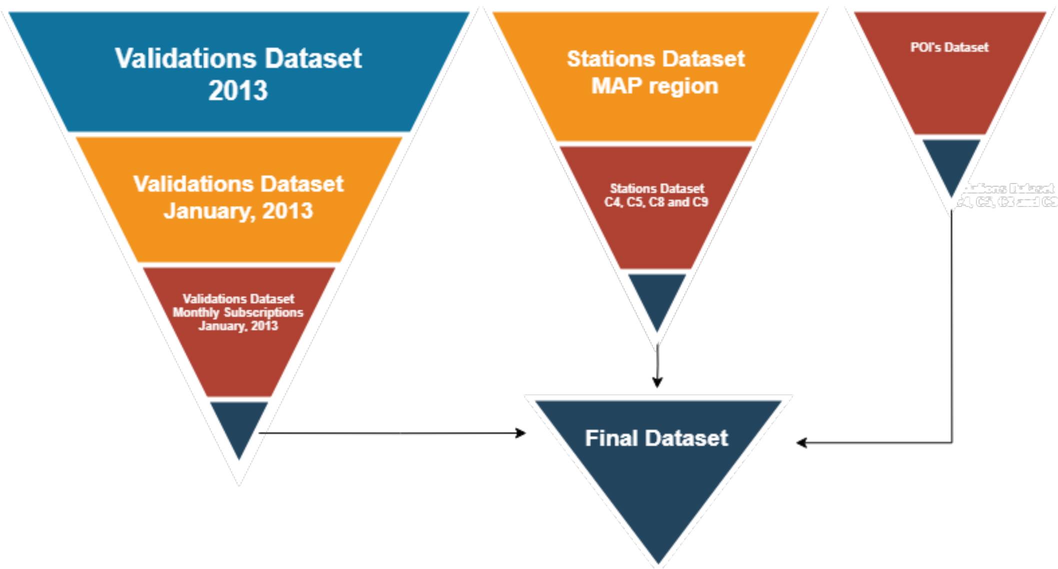

Figure 5.3: Data Scope Decisions Diagram

5.3.2 Recommendation Process Overview

The recommendation process begins with a request from the *Java* server to the *MySQL* database, in order to withdraw the information of the card in question. Once this data is retrieved, the respective profile validations file is scanned, and associates a card with all the stations it visited (validated). Afterwards, given that the stations are already mapped to the card, for each station POI's are withdrawn from the database and ordered according to a recommendation function detailed in subsection 5.3.3. Once this information is collected, a recommendation is formed, associating a card with a place, according to a specific probability. If the recommendation is new, it is immediately added to the database, otherwise, it's probability value is updated and later on added. Figure 5.4 presents a diagram of the aforementioned process.

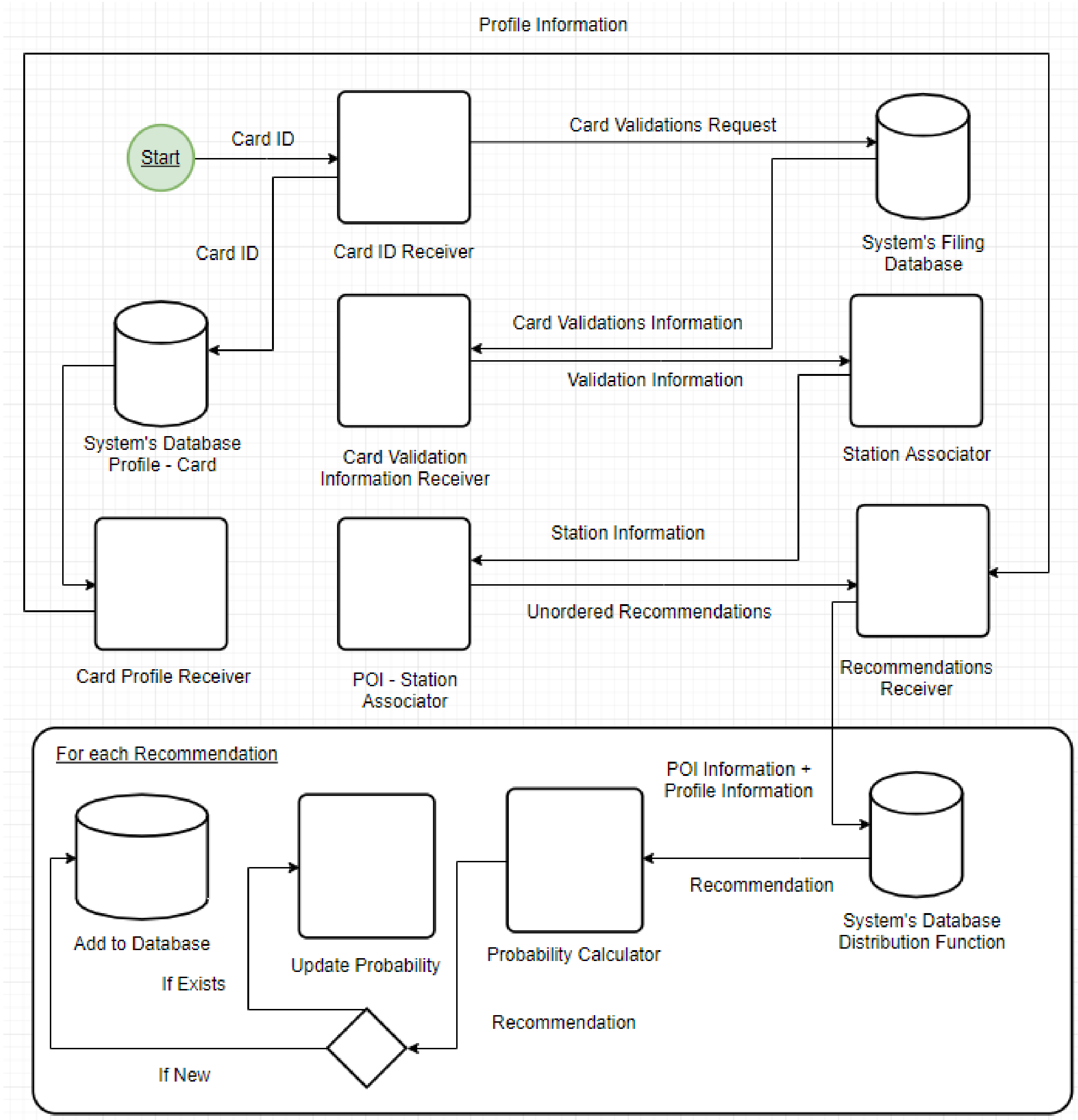

Figure 5.4: Recommendation Process Overview

Users are also able to give feedback to the recommender system once they receive a recommendation. This triggers an update request to the database, altering the feedback property of the given recommendation, and updating the distribution with a +50% or -90% variance on the type depending on it being positive or negative. This is will be relevant to future recommendations as the system is being continuously improved with feedback provided by the users.

5.3.3 Recommendation Function

The recommender system uses a recommendation function to map each type of place to a given card profile. The probability of recommending a place p to a card c, as seen in equation 5.1. can be divided into the sum of 2 auxiliary functions. The first, (5.2) consists in the sum of the probability of recommending a given place type knowing the probability of the corresponding card having a specific mobility profile. The second (5.3) is the sum of the probability of 2 fractions: the product of the probability of the given card being of the normal profile type with the place rating divided by 5, and the product of the probability of the given card being of the beneficiary profile type with the place priceLevel divided by 5.

$$P_c(p) = F_1(p.type|C) + F_2(p|C) \tag{5.1}$$

$$F_1(t|C) = D(t|C.young) + D(t|C.university) + D(t|C.adult) + D(t|C.elderly) \tag{5.2}$$

$$F_2(p|C) = \frac{(P(c.normal) * p.rating)}{5} + (c.beneficiary) * \frac{1}{p.priceLevel} \tag{5.3}$$

Since in the beginning of the developed recommender system there is no user feedback, the probability of a place type being recommended to a specific card profile was performed according to a distribution function that follows a negative exponential function when rearranging place types. Figure 5.5 presents a visualization of each profile distribution, while table 5.1 states the values initially used. It is important to understand that these values are arbitrary and do not reflect any previously analysis or study foundation.

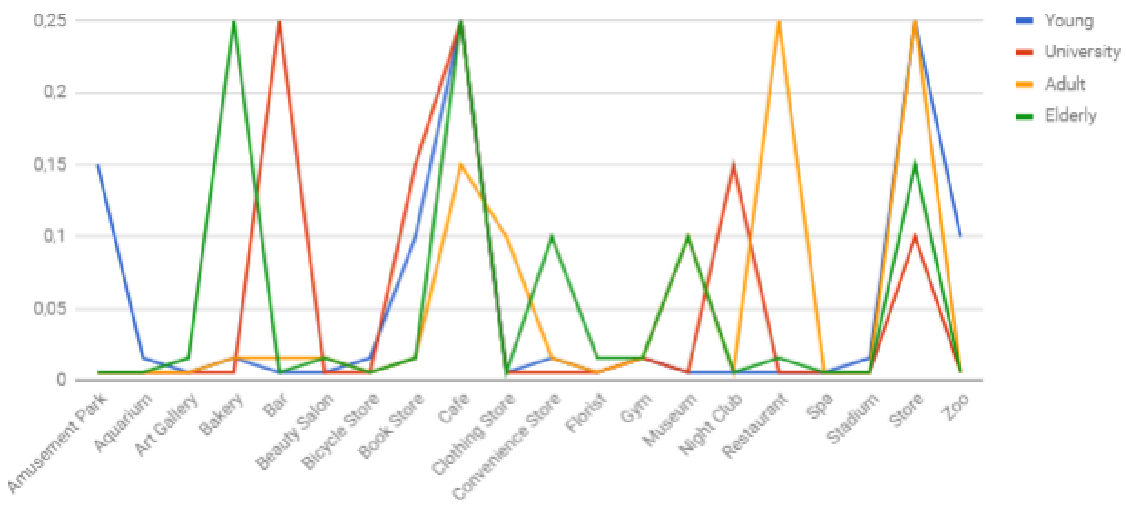

Figure 5.5: Distribution used for place ordering according to type and profile

Table 5.1: Distribution values used for place ordering according to type and profile

	Young	**University**	**Adult**	**Elderly**
Amusement Park	0.15	0.006	0.006	0.006
Aquarium	0.016	0.006	0.006	0.006
Art Gallery	0.006	0.006	0.006	0.016
Bakery	0.016	0.006	0.016	0.25
Bar	0.006	0.25	0.016	0.006
Beauty Salon	0.006	0.006	0.016	0.016
Bicycle Store	0.016	0.006	0.006	0.006
Book Store	0.1	0.15	0.016	0.016
Cafe	0.25	0.25	0.15	0.25
Clothing Store	0.006	0.006	0.1	0.006
Convenience Store	0.016	0.006	0.016	0.1
Florist	0.006	0.006	0.006	0.016
Gym	0.016	0.016	0.016	0.016
Museum	0.006	0.006	0.1	0.1
Night Club	0.006	0.15	0.006	0.006
Restaurant	0.006	0.006	0.25	0.016
Spa	0.006	0.006	0.006	0.006
Stadium	0.016	0.006	0.006	0.006
Store	0.25	0.1	0.25	0.15
Zoo	0.1	0.006	0.006	0.006

5.4 System Validation and Testing Methodologies

When revising recommender system's evaluation literature, studies found that it is important that these systems both predict accurately user interests, but also anticipate their tastes. In that sense, for the validation and testing of the presented solution, 4 distinct stages of offline experimentations were conducted, simulating user behaviour interacting with the system [RRS+].

Stage 1, also designated Training Recommendations is the most straightforward component. It applies a 3 week filter to the validations that are considered in order to withdraw stations visited in the first 3 weeks of the month. With no feedback applied, the recommendation process is as described in the first stages of subsection 5.3.2. The main focus of this evaluation is making sure that the recommendation's probabilities are indeed in accordance with the function detailed in subsection 5.3.3, analysing value disparity and precision.

Stage 2, Feedback Application, uses the personas created, described in subsection 5.4.1 and applies their feedback to the recommendations generated in stage 1. Here, it is expected that the probability weights of the distribution function shift in accordance with the given positive or negative feedback.

The next stage, stage 3, entitled Testing Recommendations applies the recommendation process to the remaining weeks of the month, considering recommendations uninfluenced by the feedback in stage 2.

Finally stage 4, Testing Recommendations with Feedback, applies the process described in stage 3 but to the feedback altered dataset. By comparing the results between stages 3 and 4 it is expected to be able to withdraw conclusions about the validity of the recommender system and if it presents a valid business model for sustainable transportation and city local services promotion.

5.4.1 Persona Feedback

In order to generate usable feedback for the evaluation of the recommender system, different personas were created, combining exiting card knowledge with predefined assumptions regarding user interest. Since financial variances are given due to place data and not card data, 8 different types of personas were generated, two for each of the representative age mobility profiles (Young, University Student, Adult and Elderly). For studying purposes, each persona is simplified, having only one like and one dislike. Table 5.2 presents the persona information relevant for the system's evaluation.

Table 5.2: Persona Information

Card ID	Age Profile Type	Financial Profile Type	Likes	Dislikes
86090	Young	Normal	Restaurants	Stores
86063	Young	Beneficiary	Gyms	Cafes
94147	University	Normal	Stadiums	Bars
98942	University	Beneficiary	Restaurants	Cafes
4150	Adult	Normal	Bars	Restaurants
2285	Adult	Beneficiary	Gyms	Stores
1755	Elderly	Beneficiary	Clothing Stores	Cafes
19	Elderly	Beneficiary	Stores	Gyms

These cardID numbers while picked randomly, were also subject to the precondition of having validations in the selected areas. The following table, table 5.3, presents the activity of these cards throughout the predefined time period. In it is it visible that the randomly chosen dataset also presents a high variation in terms of number of validations and average daily validation, accounting for a heterogeneous sample.

Table 5.3: Persona Card Activity

Card ID	Nº of validations (January 2013)	Daily validation average
86090	51	1.65
86063	107	3.45
94147	30	0.97
98942	19	0.61
4150	143	4.61
2285	45	1.45
1755	2	0.06
19	139	4.48

5.5 Mobile Application Prototype

With the objective of being available to different users on an everyday use basis, the START's (Sustainable Transportation Awareness Recommender Tool) mobile application was developed, using the information generated by the recommender system in a cloud hosted server. The mobile application uses a *Node.js* server hosted in the *Heroku* platform and uses a *JSON* based database present in the platform *Firebase* to store and handle data.

START allows users to, upon *NFC* card validation, access user oriented recommendations and the corresponding information, giving them also the ability to supply feedback in their mobile devices. The following figures illustrate the different interfaces and information types the mobile application presents to users.

Figure 5.6: Landing Interfaces

Figure 5.6 presents the different landing pages of the mobile application, depending on the NFC's activation status. They either instruct the user to activate his NFC connection (left) or scan his *Andante* card for reading (right).

Figure 5.7: Recommendations Main Page Interface Example

Figure 5.7 presents a recommendation listing example, where it is visible the POI's name, type (indicated by an icon in the upper right corner of the recommendation), photograph, address, rating, open status, novelty status, feedback (depending if the *thumbs up* or *thumbs down* buttons are pressed) and the available discounts (5.8).

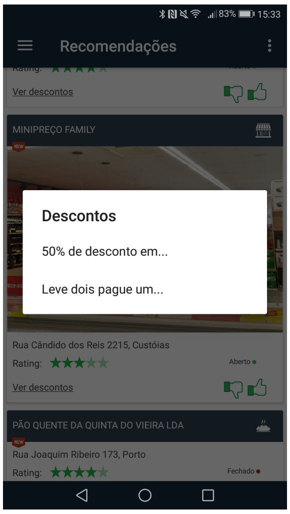

Figure 5.8: Discounts Interface Example

Finally, under a format of a side menu, figure 5.9, the app presents the types of recommendation filters a user can activate, being them, unfiltered in the *Recomendações* section, new recommendations in *Novidades* or upvoted in the *Gostos* section.

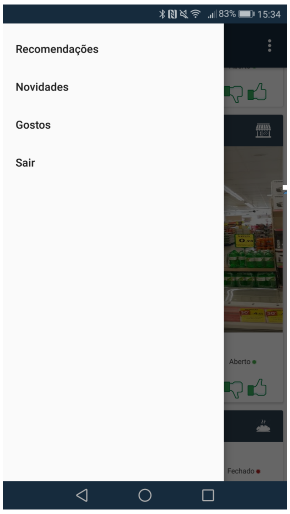

Figure 5.9: Menu Interface Example

Recommender System

Chapter 6

Results

Based on the methodology described in section 5.3, this chapter presents an overview of the obtained results, analysing the findings that were gathered with experimentation, by following a 20 card sample (tables 6.1 and 6.2), that was randomly selected and kept throughout the different stages of testing [1].

Table 6.1: Card Testing Sample (Part 1)

Card ID	Age Profile	Financial Profile	Persona
86090	Young	Normal	Yes
86063	Young	Beneficiary	Yes
89402	Young	Beneficiary	No
86075	Young	Beneficiary	No
86096	Young	Normal	No
94147	University	Normal	Yes
98942	University	Beneficiary	Yes
98920	University	Normal	No
98918	University	Normal	No
99068	University	Beneficiary	No
4150	Adult	Normal	Yes
2285	Adult	Beneficiary	Yes
2286	Adult	Normal	No
2287	Adult	Beneficiary	No
2288	Adult	Normal	No
1755	Elderly	Beneficiary	Yes
19	Elderly	Beneficiary	Yes

[1] The obtained sample results for each stage are partially presented in the chapter, while the remaining appear in the Annex section of this dissertation.

Table 6.2: Card Testing Sample (Part 2)

Card ID	Age Profile	Financial Profile	Persona
25	Elderly	Beneficiary	No
27	Elderly	Beneficiary	No
30	Elderly	Beneficiary	No

6.1 Stage 1 - Training Recommendations

This section details the obtained results of the first stage of the recommender system's testing and validation. As mentioned before, the objective of this experimentation is to validate the recommendations, making sure that they are indeed in accordance with the distribution function defined in the project's methodology.

In order to do so, for each of the cards presented in the previously described chosen sample, a table containing the recommendation results, describing POI's name, type, rating, price level and probability, will be collected, ordering the recommendations according to their chosen probability. Not only this, but to allow a more graphic perception of each of the sample results, a boxplot chart will be presented as well as a brief commentary of the obtained results, when pertinent for the study.

Table 6.3: Stage 1 recommendations - Card ID 86090

Name	Type	Rating	PriceLevel	Probability
Invescorte - Computadores e Sistemas, SA	Store	5/5	N/A	0.625
APBS Interiores	Store	5/5	N/A	0.625
Artur Lagoela & Filhos - Industria e Comercio De Materiais De Construcao, Lda	Store	4.7/5	N/A	0.595
Farmacia Sousa Oliveira	Store	4.7/5	N/A	0.595
Tipografia Lessa	Store	4.7/5	N/A	0.595
Farmacia Lima Coutinho	Store	4.5/5	N/A	0.575
Catassol Central Pharmacy	Store	4.2/5	N/A	0.545
Minipreco Family	Store	3.8/5	N/A	0.505
Pao Quente da Quinta do Vieira, Lda	Bakery	4.3/5	N/A	0.438
MaiaFit - Espacos Desportivos, SA	Gym	4.2/5	N/A	0.428
McDonald's	Restaurant	4.1/5	1/5	0.413
Churrasqueira Sao Tiago	Restaurant	4/5	N/A	0.403
Doce Maia Gourmet - Restauracao, Lda	Restaurant	3.9/5	N/A	0.393

The obtained results for the sample with CardID equal to 86090, presented in table 6.3, are indeed in accordance with the previously defined distribution function for a Young-Normal profile. Not only this, but POI type variation (figure 6.1) confirms the rating's importance in a Normal profile recommendation, as well. Standard deviation for the overall recommendations is 3.216, with standard deviation according to place type being 2.331 for Store types, 0.281 for Bakery, 0.286 for Gym and 1.126 for Restaurant.

Figure 6.1: Stage 1 boxplot chart results - CardID 86090

Table 6.4: Stage 1 recommendations - Card ID 89402 (Part 1)

Name	Type	Rating	PriceLevel	Probability
Sportsdirect Leca PT	Bicycle Store	4.2/5	1/5	0.503
Artur Lagoela & Filhos - Industria e Comercio De Materiais De Construcao Lda	Store	4.7/5	N/A	0.125
Atelier Joaquim Pombal	Store	4/5	N/A	0.125
Farmacia Sousa Oliveira	Store	4.7/5	N/A	0.125
Espaco Saude Lionesa	Store	4.2/5	N/A	0.125
Sabores da Herdade	Store	4.1/5	N/A	0.125
Fnac	Store	4.1/5	N/A	0.125
Jumbo	Store	4.1/5	N/A	0.125
Eureka	Store	3.5/5	N/A	0.125
Smart Cartridge Marshopping	Store	N/A	N/A	0.125
Midas shop MAR Shopping - Matosinhos	Store	4.1/5	N/A	0.125
Expocar Porto - Audi	Store	4.4/5	N/A	0.125

Table 6.5: Stage 1 recommendations - Card ID 89402 (Part 2)

Name	**Type**	**Rating**	**PriceLevel**	**Probability**
Laskasas Interiores - Leça da Palmeira	Store	4.1/5	N/A	0.125
Portabonus Pneus e Acessorios, Lda	Store	4.5/5	N/A	0.125
Pao Quente da Quinta do Vieira, Lda	Bakery	4.3/5	N/A	0.008
Estadio do Leca FC	Stadium	3.8/5	N/A	0.008
Churrasqueira Sao Tiago	Restaurant	4/5	N/A	0.003
TRYP Porto Expo Hotel	Bar	4.1/5	N/A	0.003
Brasileirao	Restaurant	4.5/5	N/A	0.003
ArtiFlor - Florista Decoracao Custoias Matosinhos	Florist	5/5	N/A	0.003

Despite the lack of information regarding the price level of the recommendations of the sample with cardID 89402, which conditioned the type variation, the obtained results present in tables 6.4 and 6.5 showcase the expected outcome of recommendations for a Young-Beneficiary profile type, being indeed in accordance with the original distribution function. Given the lack of variation of probability the boxplot chat was omitted from this study.

Table 6.6: Stage 1 recommendations - Card ID 94147

Name	**Type**	**Rating**	**PriceLevel**	**Probability**
Iberacero Portugal Sociedade de Representacoes, Lda	Store	5/5	N/A	0.55
TRYP Porto Expo Hotel	Bar	4.1/5	N/A	0.535
Alto Cristelo - Pet Shop	Store	4.7/5	N/A	0.520
Expocar Porto - Audi	Store	4.4/5	N/A	0.490
Fnac	Store	4.1/5	N/A	0.460
Jumbo	Store	4.1/5	N/A	0.460
Midas shop MAR Shopping - Matosinhos	Store	4.1/5	N/A	0.460
Laskasas Interiores - Leça da Palmeira	Store	4.1/5	N/A	0.460
Brasileirão	Restaurant	4.5/5	N/A	0.453
Suprides XXI, Lda	Store	3.8/5	N/A	0.430
Sportsdirect Leca PT	Bicycle Store	4.2/5	1/5	0.423
Eureka	Store	3.5/5	N/A	0.400
Estadio do Leça FC	Stadium	3.8/5	N/A	0.383
Jular Madeiras - Porto	Store	3/5	N/A	0.35
Smart Cartridge Marshopping	Store	N/A	N/A	0.05
Hitlife	Beauty Salon	N/A	N/A	0.003

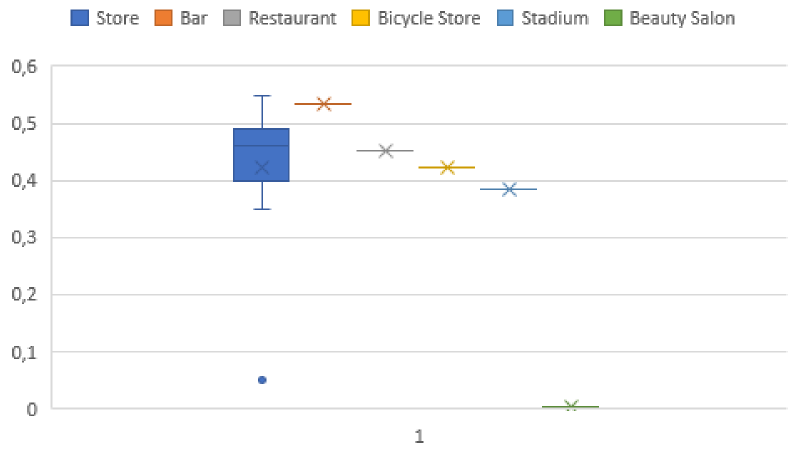

Figure 6.2: Stage 1 boxplot chart results - CardID 94147

The obtained results (Table 6.6 and Figure 6.2) present a standard University Student-Normal profile with no abnormal results without accounting for the POI store with no available rating. Standard deviation for the overall recommendations is 3.673, with standard deviation according to place type being 2.926 for Store types, 0.233 for Bar, 0.274 for Restaurant, 2.885 for Bicycle Store, 0.309 for Stadium and 0.499 for Beauty Salon.

Table 6.7: Stage 1 recommendations - Card ID 99068 (Part 1)

Name	**Type**	**Rating**	**PriceLevel**	**Probability**
O Paparico	Bar	4.5/5	N/A	0.125
Lidl	Store	4/5	N/A	0.05
Lupi - Fotografia - Aguas Santas	Store	4.5/5	N/A	0.05
Pharmacy Sousa Reis	Store	4.5/5	N/A	0.05
Data Rio Servicos de Informatica, Lda	Store	4.6/5	N/A	0.05
BDR - Bandeiras e Mastros	Store	4.3/5	N/A	0.05
Telecao	Store	4.3/5	N/A	0.05
Casa Carvalho - Moveis, Eletrodomesticos e Decoracoes	Store	4.7/5	N/A	0.05
Farmacia Central Rio Tinto	Store	4.5/5	N/A	0.05
Koket	Store	4.5/5	N/A	0.05
Materiais de Construcao Dias	Store	2.7/5	N/A	0.05
MCoutinho Parts and Repair Automotive, SA	Store	4.2/5	N/A	0.05
Inauto	Store	3.8/5	N/A	0.05

Table 6.8: Stage 1 recommendations - Card ID 99068 (Part 2)

Name	Type	Rating	PriceLevel	Probability
Invescorte - Computadores e Sistemas, SA	Store	5/5	N/A	0.05
APBS Interiores	Store	5/5	N/A	0.05
Tipografia Lessa	Store	4.7/5	N/A	0.05
Farmacia Lima Coutinho	Store	4.5/5	N/A	0.05
Sakhti	Store	3/5	N/A	0.05
Ginasio da venda Nova	Gym	4.4/5	N/A	0.008
Centro de Yoga Vaidika	Gym	5/5	N/A	0.008
MaiaFit - Espacos Desportivos SA	Gym	4.2/5	N/A	0.008
Churrasqueira O Grelhador da Giesta	Restaurant	4.2/5	N/A	0.003
Jose Ribeiro Cabeleireiros	Beauty Salon	4.9/5	N/A	0.003
Restaurante Dona Tila	Restaurant	3.7/5	N/A	0.003
Winsowelu - Miguel Garcia	Spa	N/A	N/A	0.003
Puro Equilibrio Day Spa	Spa	5	N/A	0.003
Churrasqueira da Estacao	Restaurant	4.2/5	N/A	0.003
Doce Alto	Bakery	4.5/5	N/A	0.003
Casa do Lopes	Restaurant	4.2/5	N/A	0.003
Doce Maia Gourmet - Restauracao, Lda	Restaurant	3.9/5	N/A	0.003

The obtained results (Tables 6.7 and 6.8) present a standard University Student-Beneficiary profile with no abnormal results. Given the lack of variation of probability the boxplot chat was omitted from this study.

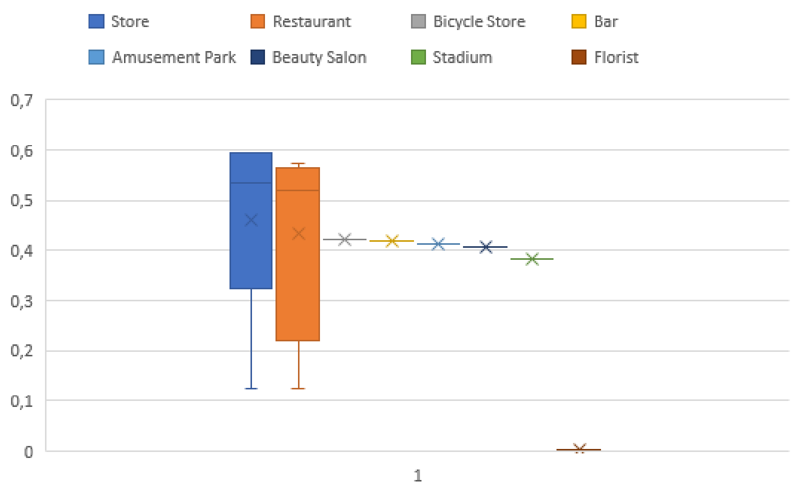

Figure 6.3: Stage 1 boxplot chart results - CardID 4150

Table 6.9: Stage 1 recommendations - Card ID 4150

Name	Type	Rating	PriceLevel	Probability
Bompiso - Trade Tires, SA	Store	4.7/5	N/A	0.595
Moldacril - Acrilicos	Store	4.7/5	N/A	0.595
Brasileirao	Restaurant	4.5/5	N/A	0.575
Expocar Porto - Audi	Store	4.4/5	N/A	0.565
Grade Restaurante, Lda	Restaurant	4.1/5	N/A	0.535
Laskasas Interiores - Leça da Palmeira	Store	4.1/5	N/A	0.535
NewStaff Restaurante	Restaurant	3.8/5	N/A	0.505
Maxmat	Store	3.6/5	N/A	0.485
Sportsdirect Leca PT	Bicycle Store	4.2/5	1/5	0.423
TRYP Porto Expo Hotel	Bar	4.1/5	N/A	0.418
Maria Rapaz	Amusement Park	4.1/5	N/A	0.413
Donacorpus - Centro de Estetica e Bem Estar	Beauty Salon	4/5	N/A	0.408
Estádio do Leça FC	Stadium	3.8/5	N/A	0.383
Biozoo - Plantas e Animais de Companhia, Lda	Store	2/5	N/A	0.325
Armando Moreira Comercio de Automoveis	Store	N/A	N/A	0.125
Take away Veleiros	Restaurant	N/A	N/A	0.125
Mania das Plantas, Lda	Florist	N/A	N/A	0.003

The obtained results for card 4150 demonstrate an accurate Adult-Normal profile, showcasing a high variant of result probabilities due to the higher amount of recommendations that were generated (Table 6.9 and Figure 6.3). Standard deviation for the overall recommendations is 3.803, with standard deviation according to place type being 5.472 for Store types, 1.435 for Restaurant, 0.289 for Bicycle Store, 0.291 for Bar, 0.294 for Amusement Park, 0.296 for Beauty Salon, 0.389 for Stadium and 0.499 for Florist.

Table 6.10: Stage 1 recommendations - Card ID 2287 (Part 1)

Name	Type	Rating	PriceLevel	Probability
Electro Rayd	Store	4.3/5	N/A	0.125
Data Rio Servicos de Informatica, Lda	Store	4.6/5	N/A	0.125
Rota das Regioes	Store	3.5/5	N/A	0.125
Restaurante Charco	Restaurant	4.2/5	N/A	0.125
Garagem Cavada Nova	Store	4.5/5	N/A	0.125
Jumbo	Store	4.2/5	N/A	0.125
Central Churrasco	Restaurant	4.5/5	N/A	0.125

Table 6.11: Stage 1 recommendations - Card ID 2287 (Part 2)

Name	Type	Rating	PriceLevel	Probability
H3 Parque Nascente	Restaurant	5/5	N/A	0.125
CPCdi - Companhia Portuguesa de Computadores e Distribuicao de Produtos Informaticos	Store	4.3/5	N/A	0.125
Restaurante Toca da Formiga	Restaurant	4.2/5	N/A	0.125
Afontec, Lda	Store	3.9/5	N/A	0.125
Ribeiro & Tavares LDA	Store	4/5	N/A	0.125
Decorarte - Bernardino Cardoso	Store	5/5	N/A	0.125
Auto - Industrial Lda	Store	4.1/5	N/A	0.125
Perfumes & Companhia	Clothing Store	4.5/5	N/A	0.05
Sabores Magnolia	Bakery	4/5	N/A	0.008
Ginasio da Venda Nova	Gym	4.4/5	N/A	0.003
Winsowelu - Miguel Garcia	Spa	N/A	N/A	0.003
FitSpot	Gym	3.3/5	N/A	0.003
Clube Zupper	Gym	3.3/5	N/A	0.003

CardID 2287 values (Tables 6.10 and 6.11) while presenting a clear use of the distribution function on Adult profiles, do not allow for any conclusions regarding Financial profiling, due to the lack of information, and thus the boxplot chart for these results is omitted.

Table 6.12: Stage 1 recommendations - Card ID 1755

Name	Type	Rating	PriceLevel	Probability
Doce Alto (Ermesinde)	Cafe	4.1/5	N/A	0.125
Copidouro	Store	3.7/5	N/A	0.075
Fotosport	Store	4/5	N/A	0.075
ACMA	Store	5/5	N/A	0.075
Boutique dos Tecidos	Store	N/A	N/A	0.075
Imagine Foto	Store	N/A	N/A	0.075
ZUNI Personalbit, Lda	Store	3/5	N/A	0.075
Cascata Maia Shopping Snack-bar, Lda	Restaurant	3.1/5	N/A	0.008
Churrasqueira da Estacao - Ermesinde	Restaurant	4.1/5	N/A	0.008
Perfumes & Companhia	Clothing Store	3.5/5	N/A	0.003

Similar to previous observations, cardID 1755 values (Table 6.12) while presenting a clear use of the distribution function on Elderly profiles, does not allow for any conclusions regarding Financial profiling, due to the lack of information, and thus the boxplot chart for these results is omitted.

Table 6.13: Stage 1 recommendations - Card ID 19

Name	Type	Rating	PriceLevel	Probability
Sportsdirect Maia PT	Bicycle Store	3.8/5	1/5	0.503
Tintas Robbialac - Maia	Store	5/5	N/A	0.075
Auto Maiamotor, Lda	Store	3.4/5	N/A	0.075
Maxmat	Store	3.6/5	N/A	0.075
Radio Popular	Store	3.8/5	N/A	0.075
Damaceno & Antunes - Tecidos de Decoracao, Lda	Store	4.4/5	N/A	0.075
Farmacia Agra	Store	4.4/5	N/A	0.075
LandParts	Store	N/A	N/A	0.075
Reciclacar - Importacao, Exportacao e Reciclagem de Veiculos Automoveis, Lda	Store	3.6/5	N/A	0.075
Lopes dos Piglets	Store	4.3/5	N/A	0.75
Stand Pinto - Comercio de Automoveis	Store	4.1/5	N/A	0.75
Armando Moreira Comercio de Automoveis	Store	N/A	N/A	0.75
Copidouro	Store	3.7/5	N/A	0.075
Fotosport	Store	4/5	N/A	0.075
ACMA	Store	5/5	N/A	0.75
Imagine Foto	Store	N/A	N/A	0.075
ZUNI Personalbit, Lda	Store	3/5	N/A	0.075
Mobiliario Mani Design	Store	N/A	N/A	0.075
Moldacril - Acrilicos	Store	4.7/5	N/A	0.075
Churrascaria Portuguesa da Maia	Restaurant	4.3/5	N/A	0.008
Body Space	Gym	1/5	N/A	0.008
Pavilhao Municipal de Gueifaes II	Gym	4.4/5	N/A	0.008
Restaurante Sabores de Prata	Restaurant	4.2/5	N/A	0.008
Just 4 Fit, Lda	Gym	4.7/5	N/A	0.008
Quinta Casa do Arco	Restaurant	4.2/5	N/A	0.008
Mania das Plantas, Lda	Florist	N/A	N/A	0.008
Zen Gym	Gym	4.5/5	N/A	0.008
Cascata Maia Shopping Snack bar, Lda	Restaurant	3.1/5	N/A	0.008
Churrasqueira da Estacao	Restaurant	4.1/5	N/A	0.008
Campo Longo Industria e Comercio de Vestuario, Lda	Clothing Store	N/A	N/A	0.003
Perfumes & Companhia	Clothing Store	3.5/5	N/A	0.003

As stated in the previous figure and table, the conclusions gathered from card 17 regarding Financial status are invalid due to lack of information, and thus the boxplot chart for these results is omitted. Nevertheless, table 6.13 demonstrates the influence the distribution function has on recommendations of the age profile Elderly.

Given the obtained results in this section, it is visible to see that each place probability is indeed in accordance with the initial values of the distribution function present in subsection 5.3.3. It is mainly visible the influence of the POI's rating in the probability value of Normal profiles, while the influence of the POI's price level in the Beneficiary profiles due the lack of information compromises conclusions about the probability shifting in these profile types.

6.2 Stage 2 - Feedback Application

As detailed in subsection 5.4.1, for each of the predefined personas, recommendation feedback was provided, based on their likes and dislikes. Given the results obtained in the previous stage, for each POI type of interest to the user, a positive feedback was given, increasing the overall place type - mobility profile distribution probability by 50%. Whenever the recommended POI type was disliked by the user, a negative feedback was provided, decreasing the distribution function recommendation probability of that same type - profile association by 90%. For each of the positive or negative variance, a redistribution of the probability's weights was given to the other POI types of the same mobility profile.

When executing this methodology to all the predefined personas, the distribution function originated the results present in tables 6.14 and 6.15. While these values are a correct calculation of the distribution function when subject to the aforementioned feedback, all values are rounded up to 3 decimal places and, in some cases, may not portrait the existing redistribution of the weight, since the changes could be very small.

Table 6.14: Distribution Function Values Comparison (Part 1)

	Young	**University**	**Adult**	**Elderly**
Amusement Park	0.167	0.016	0.033	0.007
Aquarium	0.033	0.016	0.033	0.007
Art Gallery	0.023	0.016	0.033	0.017
Bakery	0.033	0.016	0.043	0.251
Bar	0.023	0.155	0.047	0.007
Beauty Salon	0.023	0.016	0.042	0.017
Bicycle Store	0.033	0.016	0.032	0.007

Table 6.15: Distribution Function Values Comparison (Part 2)

	Young	**University**	**Adult**	**Elderly**
Book Store	0.117	0.160	0.042	0.017
Cafe	0.162	0.016	0.176	0.146
Clothing Store	0.023	0.15	0.126	0.012
Convenience Store	0.016	0.016	0.042	0.101
Florist	0.023	0.033	0.032	0.017
Gym	0.034	0.026	0.047	0.001
Museum	0.023	0.016	0.126	0.101
Night Club	0.023	0.16	0.032	0.007
Restaurant	0.039	0.021	0.018	0.017
Spa	0.023	0.033	0.032	0.007
Stadium	0.033	0.031	0.032	0.007
Store	0.016	0.11	0.002	0.246
Zoo	0.117	0.016	0.032	0.007

After analysis it is concluded that the distribution function was updated as expected, correctly altering values in accordance to the specified methodology of subsection 5.4. Figure 6.4 presents a visual comparison between the original distribution values and the feedback influenced ones. The obtained expectations were met, as the following results were obtained:

- Probability increment for Restaurants and Gyms in Young profiles.
- Probability decrement for Stores and Cafes in Young profiles.
- Probability increment for Stadiums and Restaurants in University Student profiles.
- Probability decrement for Bars and Cafes in University Student profiles.
- Probability increment for Bars and Gyms in Adult profiles.
- Probability decrement for Restaurant and Stores in Adult profiles.
- Probability increment for Clothing Stores and Stores in Elderly profiles.
- Probability decrement for Cafes and Gyms in Elderly profiles.

As previously mentioned, when focusing in only one of the POI's types, Restaurant for instance, it is quite visible the influence the negative feedback of Adult profiles as distribution weights percentages dropped from 25% to 1.8%, as seen in figure 6.5.

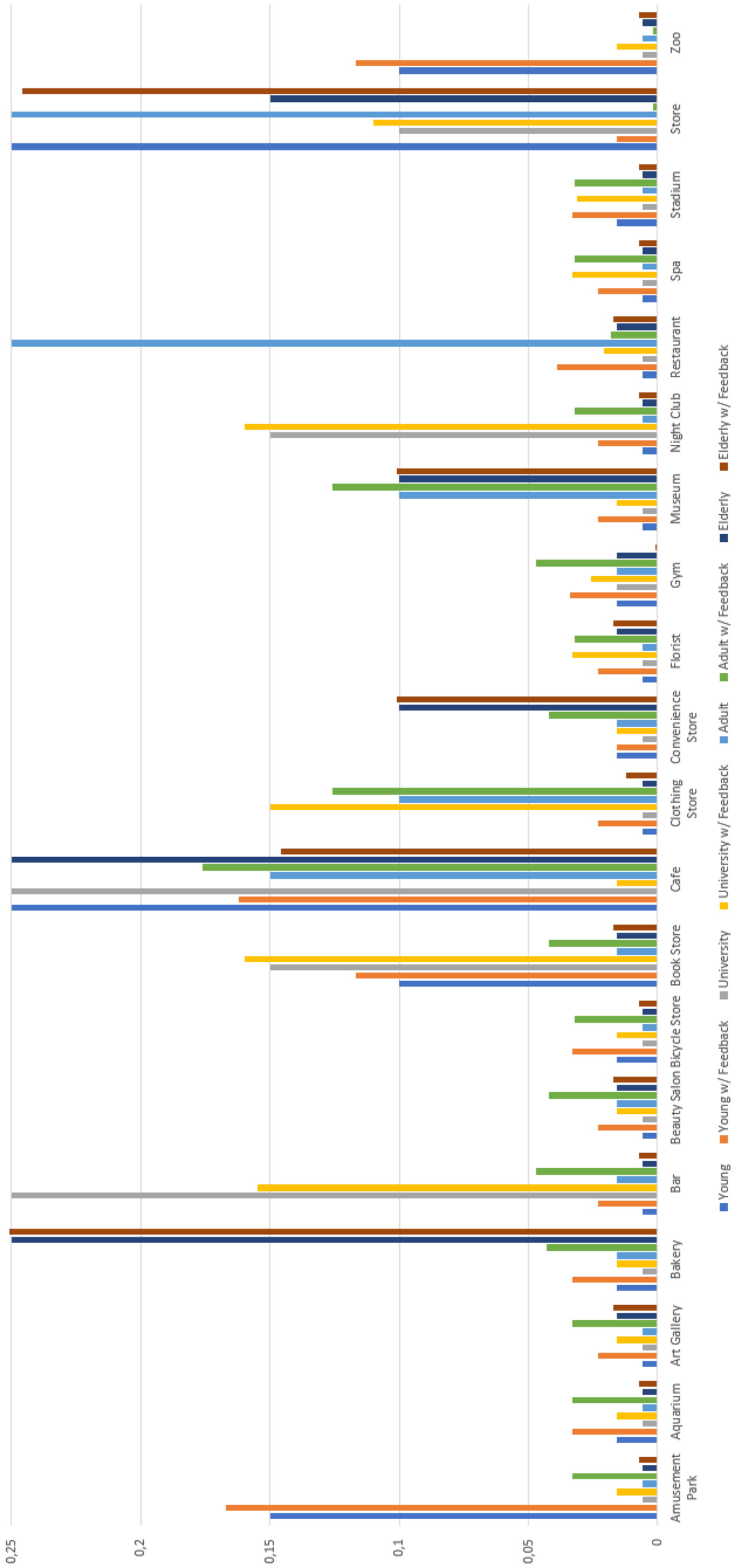

Figure 6.4: Distribution Values Comparison - No Feedback vs. Feedback

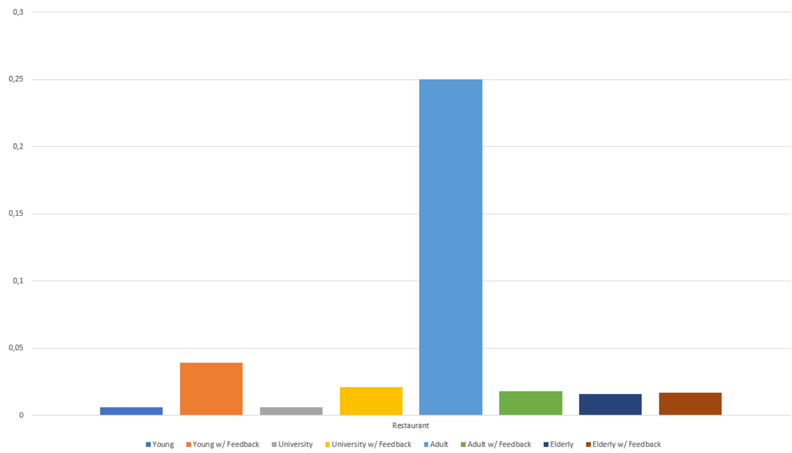

Figure 6.5: Distribution Values Comparison Restaurant Type - No Feedback vs. Feedback

In order to avoid the repetition of the recommendation tables, only the results of card 86090 will be presented in this section, while all the other persona [2] results will be available in the Annex A portion of this dissertation.

Table 6.16: Stage 2 recommendations - Card ID 86090

Name	**Type**	**Rating**	**PriceLevel**	**Probability**	**Feedback**
McDonald's	Restaurant	4.1/5	1/5	0.620	+
Churrasqueira Sao Tiago	Restaurant	4/5	N/A	0.605	+
Doce Maia Gourmet - Restauracao, Lda	Restaurant	3.9/5	N/A	0.590	+
Pao Quente da Quinta do Vieira, Lda	Bakery	4.3/5	N/A	0.446	
MaiaFit - Espacos Desportivos, SA	Gym	4.2/5	N/A	0.428	
Invescorte - Computadores e Sistemas, SA	Store	5/5	N/A	0.063	-
APBS Interiores	Store	5/5	N/A	0.063	-
Artur Lagoela & Filhos - Industria E Comercio De Materiais De Construcao, Lda	Store	4.7/5	N/A	0.060	-
Farmacia Sousa Oliveira	Store	4.7/5	N/A	0.060	-
Tipografia Lessa	Store	4.7/5	N/A	0.060	-
Farmacia Lima Coutinho	Store	4.5/5	N/A	0.058	-
Catassol Central Pharmacy	Store	4.2/5	N/A	0.055	-
Minipreco Family	Store	3.8/5	N/A	0.051	-

[2] Note that only the predefined persona's results will be shown as they are the only ones from the sample that have given feedback.

The obtained results present in table 6.16 illustrate the importance of the persona feedback on his own recommendations, where a negative opinion towards stores and positive towards restaurants, noticeably change the probability percentage of the POI's of those types, when in comparison to the results of table 6.3.

Nevertheless, not all recommendations that are influenced by feedback, present results validating the profiling previously done. In some cases, when the a user's preferences are so off put with his own group, not even feedback is able to immediately and drastically shift probabilities. For example, when comparing table B.5 to table B.22 it is visible the impact the dislike of Cafe types and fondness of Restaurant types had on the overall probability distribution. Nevertheless, these results constitute the perfect example that even with negative feedback, if it is given to a type of place with high probability, can still surpass the probability of a recommendation of a positive feedback POI. A valid assumption would be that in this scenario, the persona is likely to be misplaced in it's respective age cluster.

6.3 Stage 3 - Testing Recommendations

Following the same principle as the previous stage, this section will only present the new recommendations for the user with cardID 86090. Additionally, in order to better understand the changes this stage of testing originated in comparison to Stage 1, only new recommendations will be present.

Table 6.17: Stage 3 recommendations - Card ID 86090

Name	Type	Rating	PriceLevel	Probability
Restaurante Paraiso	Restaurant	4.6/5	N/A	0.463
Restaurante Veleiros	Restaurant	4.5/5	N/A	0.453
Casa Velha	Restaurant	4.4/5	N/A	0.443
Restaurant Mari-Ze	Restaurant	3.7/5	N/A	0.373
Restaurante Campo Alegre Mar	Restaurant	N/A	N/A	0.003

As seen in table 6.17, when considering the testing dataset, corresponding to a similar passenger validations dataset, containing only the last 2 weeks of the month of January, 2013, this particular user, after travelling to new locations, received 5 new recommendations, all of the same type, Restaurant, when comparing to the results of table 6.3. These results were calculated using the same recommendation process as stage 1, but considering now the new stations the user visited.

This example presents the system's ability to generate new POI's based on new passenger activity. Nevertheless, not all sample user's had new recommendations, as trip patterns may remain the same throughout the entirety of the month in question. One example of this type of occurrence is the user with cardID 89402, that given the fact that he did not travel to new stations and that all of the available POI's had already been recommended, no new recommendations were generated.

6.4 Stage 4 - Testing Recommendations with Feedback

In similarity with previous stages, this section will only present the new recommendations for the user with cardID 86090.

Table 6.18: Stage 4 recommendations - Card ID 86090

Name	Type	Rating	PriceLevel	Probability
Restaurante Paraiso	Restaurant	4.6/5	N/A	0.479
Restaurante Veleiros	Restaurant	4.5/5	N/A	0.469
Casa Velha	Restaurant	4.4/5	N/A	0.459
Restaurant Mari-Ze	Restaurant	3.7/5	N/A	0.389
Restaurante Campo Alegre Mar	Restaurant	N/A	N/A	0.019

The obtained results seen in table 6.18 present the expected outcome, producing the same recommendations as table 6.17, but with new probabilities, now according to the probabilities presented in tables 6.14 and 6.15. Given the positive feedback Restaurant type had by Young profiles, recommendation probability for this type increased, as seen in figure 6.6.

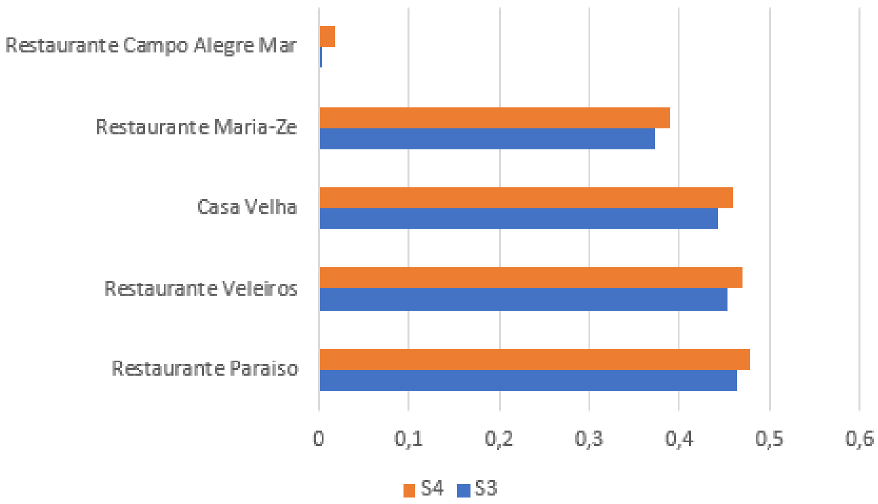

Figure 6.6: Card ID 86090 Stage 3 - Stage 4 Result Comparison

6.5 Results Overview

Given the obtained results it is possible to state that the developed system is able to generate user-centred recommendations, presenting users with local city services of the MAP's area, given their mobility profile and public transportation activity.

Stage 1 results, present in section 6.1, demonstrate accurately that for Normal financial profile types, POI rating is considered the key factor in probability variation, with probabilities shifting according to age profile - place type distribution definition. And while, priceLevel values are few, Beneficiary profile results seem to follow in the same footsteps. When comparing with stage 2 results (section 6.2) feedback influence is notorious, demonstrating the full capabilities of what is a collaborative-filtering recommender system.

More interesting is the comparison between stages 3 and 4, present in sections 6.3 and 6.4, respectively. Not only do they demonstrate the capabilities of the recommender system to generate new recommendations based on new activity, but also how the system learned from the user input, slowly improving recommendation anticipation precision, and calibrating the recommendations themselves.

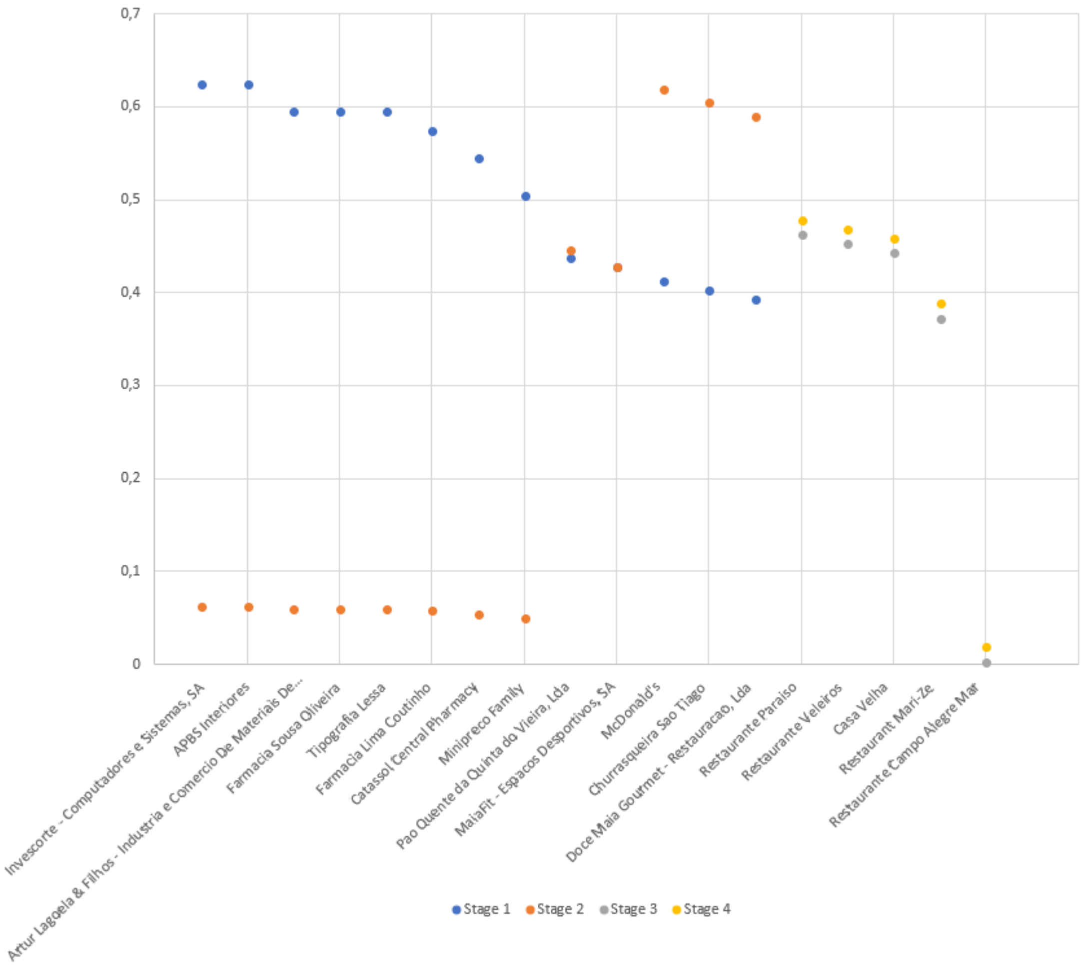

Figure 6.7: Card ID 86090 Overall Result Comparison

Considering the aforementioned example of cardID 86090, and as portrayed in figure 6.7, given a Normal-Young profile, stage 1 recommendations follow the distributed function initially specified. Nevertheless, when subject to feedback in stage 2, POI's that disinterest the user, decrease drastically their recommendation probability (-90%), while positive feedback POI's increase somewhat their probability (+50%), as specified in subsection 5.3.2. In this stage it is also visible that other POI types suffer minor variations in their recommendation probability as a result of the distribution function readjustment.

Finally, new recommendations specified by stages 3 and 4 take in consideration the previous feedback given, personalizing their recommendations to user interests and anticipating it's tastes, resulting in an adjustment of the recommendation probability.

Chapter 7

Conclusions

Throughout the development of this dissertation, a detailed and thorough analysis of passenger data and activity was performed, in order to guarantee highly tailored and user-centred recommendations. This analysis resulted in some conclusions. First regarding station distribution, which allowed to understand which areas have the higher density of public transportation stations and how they distribute themselves among operators. Secondly, and very important for the developed work, regarding passenger validations, where and when they most often occur, which types of card are most often used, and what information can be extracted and used from these cards, in order to create profiles.

In addition, a study of points of interest of the MAP's area was performed, given the data withdraw from the Google Palces API, as these POI's serve as the items the recommender system will recommend. This allowed to perceive how they are distributed in the MAP's area, and what types of information this API supplies, that can be used to present the recommendations.

Afterwards, the recommendation process was design, as well as the system's architecture . Probability calculation mathematical formulas were defined and the mobile prototype was created. This allowed for the development and testing of a fully functional recommender system, capable of providing users with recommendations based on their mobility profiles and activity history. Despite the system's cold-start mechanism, due to the feedback self-learning mechanism that was implemented, it emerges not only as an innovative potential sustainable mobility promoter solution, but also as an innovative new business model with a practical application.

Through experimentation, the developed solution produced not only user-centred recommendations, but also anticipated user preferences given it's collaborative-filtering methodology, making for a complete broad solution.

Finally, the different stages of testing validated the system's recommendations, allowing to analyse each recommendation based on the respective user profile and activity, as well as measure the feedback's influence on new recommendations.

7.1 Future Work

The proposed solution presents itself as a starting point for a scalability process, in order to expand, not only the target population, but also the intrinsic capacities of the system, that, given the available time scope, were not able to be implemented.

First and foremost, the system with benefit for the automatic integration of both databases, allowing for real-time responses to user feedback. This would improve the overall usability of the system, allowing to test it in a real environment, provided the update of the available validations dataset.

An analysis of the entirety of the validation's dataset could also provide additional conclusions about seasonal patterns and passenger activity, resulting in more precise recommendations.

Another, and possibly the most interesting improvement to the developed solution, would be to not only use the information provided by monthly subscription cards, to generate mobility profiles, but also use single trip cards as well, categorizing them through clustering algorithms, expanding the applicability scope and evaluating if the presented group mobility profiles make sense or should differ in any way.

An integration with the *Anda*'s, system could also be mutual beneficial for both business models, providing an all-in-one application for different purposes.

Finally, performance enhancements could be conducted, increasing sample testing and the recommender's system precision.

References

[Adv18] Advantech. Available at http://www.advantech.eu/intelligent-transportation/afc, last accessed 12th of January 2018, 2018.

[AMP13] AMP. Amp2020. pages 1–38, 2013.

[AMP18] AMPorto. Portal da autoridade de transportes da amp. Available at http://autoridade.amp.pt/pt-pt/, last accessed 8th of January 2018, 2018.

[AMT06] Bruno Agard, Catherine Morency, and Martin Trépanier. Mining public transport user behaviour from smart card data. *IFAC Proceedings Volumes*, 39(3):399–404, 2006.

[AT05] Gediminas Adomavicius and Alexander Tuzhilin. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, 17(6):734–749, 2005.

[BAH97] Alex Bochner, Sarabjot Anand, and John Hugues. Data mining in manufacturing environments: Goals, techniques and applications. *Studies in Informatics and Control*, 6(4):319–328, 1997.

[Ban17a] João Bandeira. Mobile ticketing using bluetooth low energy technology. 2017.

[Ban17b] João Bandeira. Mobile ticketing using bluetooth low energy technology. 2017.

[BB03] M. Brunato and R. Battiti. Pilgrim: A location broker and mobility-aware recommendation system. *Proceedings of the First IEEE International Conference on Pervasive Computing and Communications, 2003. (PerCom 2003).*, pages 265–272, 2003.

[BJR16] Amhmed A. Bhih, Princy Johnson, and Martin Randles. Diversity in smartphone usage. *Proceedings of the 17th International Conference on Computer Systems and Technologies 2016 - CompSysTech '16*, pages 81–88, 2016.

[Bly04] Philip Blythe. Improving public transport ticketing through smart cards. *Proceedings of the ICE - Municipal Engineer*, 157(1):47–54, 2004.

[BS97] Marko Balabanović and Yoav Shoham. Fab: content-based, collaborative recommendation. *Communications of the ACM*, 40(3):66–72, 1997.

[CdM04] Paulo Câmara and Laura Valente de Macedo. Restrição veicular e qualidade de vida: O pedágio urbano em londres e o rodízio' em são paulo. pages 1–12, 2004.

[CGM+99] Mark Claypool, Anuja Gokhale, Tim Miranda, Pavel Murnikov, Dmitry Netes, and Matthew Sartin. Combining content-based and collaborative filters in an online newspaper. *Computer Science Technical Report Series*, (June), 1999.

[CK17] Athlyn Cathcart-Keays. Oslo's car ban sounded simple enough. then the backlash began. Available at https://www.theguardian.com/cities/2017/jun/13/oslo-ban-cars-backlash-parking, last accessed 12th of January 2018, Jun 2017.

[CLCG15] Rui Couto, Joao Leal, Pedro Mauricio Costa, and Teresa Galvao. Exploring ticketing approaches using mobile technologies: Qr codes, nfc and ble. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, pages 7–12, 2015.

[Cmp18] Cmp. Andante vai ter uma aplicação móvel em abril - notícias. Available at http://www.porto.pt/noticias/andante-vai-ter-uma-aplicacao-movel-em-abril, last accessed 17th of January 2018, 2018.

[CP03] Patrick Y. K. Chau and Simpson Poon. Octopus: an e-cash payment system success story. *Communications of the ACM*, 46(9):129, 2003.

[DoT16] United States Department of Transportation. Transportation statistics annual report 2015. 2016.

[dP05a] Transportes Intermodais do Porto. Cartões. Available at http://www.linhandante.com/cartoes.asp, last accessed 18th of January 2018, 2005.

[dP05b] Transportes Intermodais do Porto. Quem somos. Available at http://www.linhandante.com/quemsomos.asp, last accessed 18th of January 2018, 2005.

[dP17] Tranportes Intermodais do Porto. Piloto anda : Simplificação e melhoria do andante. 2017.

[dP18] Transportes Intermodais do Porto. Vem aí o andante no telemóvel. Available at =http://www.linhandante.com/noticias-det.asp?noticiaid=141, last accessed 7th of Febreuary 2018, 2018.

[EW09] Larry Elliott and Patrick Wintour. G8 agrees to climate targets despite differences with developing nations. Available at https://www.theguardian.com/world/2009/jul/08/g8-climate-carbon-emission-targets, last accessed 9th of January 2018, Jul 2009.

[Fal17] Edward Falzon. What are the world's best metro systems? Available at https://edition.cnn.com/travel/article/world-best-metro-systems/index.html, last accessed 29th of January 2018, Jul 2017.

[FG13] Marta Campos Ferreira and Teresa Galvão. How to encourage the use of public transport? a multiservice approach based on mobile technologies. volume 143 of *Lecture Notes in Business Information Processing*, pages 314–325. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.

[fL] Transport for London. What is oyster? Available at https://tfl.gov.uk/fares-and-payments/oyster/what-is-oyster, last accessed 30th of January 2018.

[GG06] Todd Goldman and Roger Gorham. Sustainable urban transport: Four innovative directions. *Technology in Society*, 28(1-2):261–273, 2006.

[Goh02] Mark Goh. Congestion management and electronic road pricing in singapore. *Journal of Transport Geography*, 10(1):29–38, 2002.

[Gor06] Fabio Gordillo. The value of automated fare collection data for transit planning : an example of rail transit od matrix estimation. 2006.

[IFO15] F. O. Isinkaye, Y. O. Folajimi, and B. A. Ojokoh. Recommendation systems: Principles, methods and evaluation. *Egyptian Informatics Journal*, 16(3):261–273, 2015.

[Ind17] The Independent. Amsterdam has a new solution for overtourism. Available at https://www.independent.co.uk/travel/news-and-advice/amsterdam-overtourism-solution-tourists-technology-van-gogh/museum-canal-boat-rides-a8015811.html, last accessed 26th of December 2017, 2017.

[KA02] Anja Kollmuss and Julian Agyeman. Mind the gap: Why do people behave environmentally and what are the barriers to pro-environmental behaviour. *Environmental Education Research*, 8(3):239–260, 2002.

[kn:a] Available at =http://www.jreast.co.jp/e/pass/suica.html, last accessed 30th of January 2018.

[kn:b] Apache spark™ - lightning-fast cluster computing. Available at https://spark.apache.org/, last accessed 25th of January 2018.

[kn:c] Singapore slaps limit on the number of cars on its roads. Available at http://money.cnn.com/2017/10/24/news/singapore-car-numbers-limit/index.html, last accessed 29th of January 2018.

[kn:d] Weka. Available at https://sourceforge.net/projects/weka/?source=navbar, last accessed 25th of January 2018.

[kn:e] What is r? Available at https://www.r-project.org/about.html, last accessed 25th of January 2018.

[kn:12] Tarifario social andante, 2012.

[kn:16a] All you need to know about smart cards and how to use them in japan! Available at =http://jpninfo.com/53077, last accessed 30th of January 2018, 2016.

[kn:16b] £200m being invested in bus priority schemes. Available at https://tfl.gov.uk/info-for/media/press-releases/2016/january/-200m-being-invested-in-bus-priority-schemes, last accessed 29th of January 2018, 2016.

[kn:18] Lightning fast unified data science platform. Available at https://rapidminer.com/products/, last accessed 25th of January 2018, 2018.

[LCG15] Ricardo Leal, Pedro Mauricio Costa, and Teresa Galvão. Discovering mars: A mobility aware recommender system. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, 2015-Octob:1–6, 2015.

[Lea15] Ricardo Leal. Voyage to mars: A mobility aware recommender system. 2015.

[LGS11] P. Lops, M. Gemmis, and G. Semeraro. Content-based recommender systems: State of the art and trends. In *Recommender Systems Handbook*, pages 73–105. 2011.

[Mad18] Metro De Madrid. Available at https://www.metromadrid.es/en/viaja_en_metro/Tarifasybilletes/TTP/, last accessed 15th of January 2018, 2018.

[Mah08] Anjali Mahendra. Vehicle restrictions in four latin american cities: Is congestion pricing possible? *Transport Reviews*, 28(1):105–133, 2008.

[MCBO14] Mohamed K El Mahrsi, Etienne Côme, Johanna Baro, and Latifa Oukhellou. Understanding passenger patterns in public transit through smart card and socioeconomic data. *The 3rd International Workshop on Urban Computing (UrbComp 2014)*, 2014.

[Mic18] Microsoft. Data mining algorithms. Available at https://technet.microsoft.com/pt-br/library/ms175595(v=sql.90).aspx, last accessed 18th of January 2018, 2018.

[MM18] MOVE-ME. Ficha técnica. Available at http://mobilidade.amp.pt/pt/, last accessed 17th of January 2018, 2018.

[Mor] João Moreira. Advanced predictive topics.

[MR99] Raymond J. Mooney and Loriene Roy. Content-based book recommending using learning for text categorization. pages 195–204, 1999.

[MR11] Ralf Mikut and Markus Reischl. Data mining tools. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 1(5):431–443, 2011.

[MTA06] C Morency, M Trepanier, and B Agard. Analysing the variability of transit users behaviour with smart card data. *2006 IEEE Intelligent Transportation Systems Conference*, pages 44–49, 2006.

[(n.17] (n.d.). Mobility. Available at https://www.merriam-webster.com/dictionary/mobility, last accessed 8th of January 2018, 2017.

[Nan11] P Nancy. A comparison on performance of data mining algorithms in classification of social network data. *October*, 32(8):47–53, 2011.

[Not18a] Jornal de Notícias. Metro do porto terá mais 5,7 quilómetros em 2022. Available at https://www.jn.pt/nacional/interior/metro-do-porto-tera-mais-57-quilometros-em-2022-9049041.html, last accessed 23th of January 2018, Jan 2018.

[Not18b] Jornal de Notícias. Novo metro no porto com espaço para mais passageiros. Available at https://www.jn.pt/local/noticias/porto/porto/interior/novo-metro-no-porto-com-espaco-para-mais-passageiros-9062944.html, last accessed 23th of January 2018, Jan 2018.

[OK98] Douglas W Oard and Jinmook Kim. Implicit feedback for recommender systems. *Proceedings of the AAAI workshop on recommender systems*, pages 81–83, 1998.

[PTM11] Marie-Pier Pelletier, Martin Trépanier, and Catherine Morency. Smart card data use in public transit: A literature review. *Transportation Research Part C: Emerging Technologies*, 19(4):557–568, 2011.

[Qui08] Nasreen Quibria. The contactless wave : A case study in transit payments. pages 1–33, 2008.

[Reu17] Reuters. Singapore will stop adding cars to the roads in 2018. Available at http://www.businessinsider.com/singapore-ban-cars-february-2018-2017-10, last accessed 29th of January 2018, Oct 2017.

[RIS+94] Paul Resnick, Neophytos Iacovou, Mitesh Suchak, Peter Bergstrom, and John Riedl. Grouplens : An open architecture for collaborative filtering of netnews. *Proceedings of the 1994 ACM conference on Computer supported cooperative work*, pages 175–186, 1994.

[RRS+] Francesco Ricci, Lior Rokach, Bracha Shapira, Paul B Kantor, and Francesco Ricci. *Recommender Systems Handbook*.

[RVE97] Paul Resnick, Hal R. Varian, and Guest Editors. Recommender systems. *Communications of the ACM*, 40(3):56–58, 1997.

[Rya08] Julian Ryall. Japan harnesses energy from footsteps. Available at http://www.telegraph.co.uk/news/earth/energy/3721841/Japan-harnesses-energy-from-footsteps.html, last accessed 29th of January 2018, Dec 2008.

[Seaa] Seara.com. Ocasionais andante. Available at http://www.stcp.pt/pt/viajar/tarifas/titulos-intermodais-andante/ocasionais-andante/, last accessed 18th of January 2018.

[Seab] Seara.com. Títulos turísticos. Available at http://www.stcp.pt/pt/viajar/tarifas/titulos-turisticos/, last accessed 18th of January 2018.

[Sei98] Foo Tuan Seik. A unique demand management instrument in urban transport: the vehicle quota system in singapore. *Cities*, 15(1):27–39, 1998.

[SK09] Xiaoyuan Su and Taghi M. Khoshgoftaar. A survey of collaborative filtering techniques. *Advances in Artificial Intelligence*, (Section 3):1–19, 2009.

[Viv18] Dinheiro Vivo. Adidas lança ténis que são passe de transportes público. Available at https://www.dinheirovivo.pt/outras/adidas-lanca-tenis-que-sao-passe-de-transportes-publicos/, last accessed 23th of January 2018, Jan 2018.

[Wac10] Monica Wachowicz. *Movement-aware applications for sustainable mobility : technologies and approaches*. Information Science Reference, Hershey, PA, 2010.

[WKR+] Xindong Wu, Vipin Kumar, Quinlan J. Ross, Joydeep Ghosh, Qiang Yang, Hiroshi Motoda, Geoffrey J. McLachlan, Angus Ng, Bing Liu, Philip S. Yu, Zhi Hua Zhou, Michael Steinbach, David J. Hand, and Dan Steinberg. Number 1.

[Wri15] Sarah Wright. 9 unique subway and metro maps from around the world. Available at https://www.smartrailworld.com/9_of_the_best_subway_and_metro_maps_from_around_the_world, last accessed 29th of January 2018, 2015.

REFERENCES

Appendix A

Passenger Activity Analysis

Table A.1: Validation distribution per hour interval

Interval	Number of Validations
00:00-00:59	67930
01:00-01:59	18773
02:00-02:59	10229
03:00-03:59	9930
04:00-04:59	11002
05:00-05:59	43617
06:00-06:59	352841
07:00-07:59	1195555
08:00-08:59	1618446
09:00-09:59	1154452
10:00-10:59	940985
11:00-11:59	905335
12:00-12:59	980209
13:00-13:59	1277035
14:00-14:59	1249999
15:00-15:59	1209260
16:00-16:59	1401432
17:00-17:59	1678871
18:00-18:59	1485889
19:00-19:59	898652
20:00-20:59	424289
21:00-21:59	207730
22:00-22:59	155298
23:00-23:59	117636

Table A.2: Validation distribution per weekly interval

Week	**Number of Validations**
Week 1	4957378
Week 2	3844974
Week 3	2732956
Week 4	2949275
Week 5	2930812

Appendix B

Obtained Results

Table B.1: Stage 1 recommendations - Card ID 86063 (Part 1)

Name	Type	Rating	PriceLevel	Probability
Pharmacy Sousa Reis	Store	4.5/5	N/A	0.125
CVA Electronica - Fabrica de Equipamento Som e Luz, Lda	Store	4.6/5	N/A	0.125
Casa Carvalho - Moveis, Electrodomesticos e Decoracoes, Lda	Store	4.7/5	N/A	0.125
L. Pinto Monteiro, Lda	Store	5/5	N/A	0.125
Tintal - Company Inks Manufacutring, Ltd	Store	4.5/5	N/A	0.125
Koket	Store	4.5/5	N/A	0.125
CPCdi - Companhia Portuguesa de Computadores e Distribuicao de Produtos Informaticos	Store	4.3/5	N/A	0.125
Artur Lagoela & Filhos - Industria e Comercio de Materiais de Construcao, Lda	Store	4.6/5	N/A	0.125
Rendibor	Store	N/A	N/A	0.125
Copidouro	Store	3.7/5	N/A	0.125
Fotosport	Store	4/5	N/A	0.125
ACMA	Store	5/5	N/A	0.125
Boutique dos Tecidos	Store	N/A	N/A	0.125
Jular Madeira - Porto	Store	3/5	N/A	0.125
Iberacero Portugal Sociedade de Representacoes, Lda	Store	5/5	N/A	0.125
Briel	Store	3.7/5	N/A	0.125

Table B.2: Stage 1 recommendations - Card ID 86063 (Part 2)

Name	Type	Rating	PriceLevel	Probability
Doce Alto (Ermesinde)	Cafe	4.1/5	N/A	0.125
Imagine Foto	Store	N/A	N/A	0.125
Zuni Personalbit, Lda	Store	3/5	N/A	0.125
Ginasio Venda Nova	Gym	4.4/5	N/A	0.008
Zen Gym	Gym	4.5/5	N/A	0.008
Winsowelu - Miguel Garcia	Spa	N/A	N/A	0.003
Restaurante Charco	Restaurant	4.2/5	N/A	0.003
Choupal dos Melros	Restaurant	4.1/5	N/A	0.003
Gare Caffe	Restaurant	4.1/5	N/A	0.003
Belissima -Fashion Accessories, Lda	Clothing Store	4.2/5	N/A	0.003
Taberninha Repelao	Bar	4.5/5	N/A	0.003
Puro Equilibrio Day Spa	Spa	5/5	N/A	0.003
Churrasqueira da Estacao	Restaurant	4.2/5	N/A	0.003
Mania das Plantas, Lda	Florist	N/A	N/A	0.003
Perfumes & Companhia	Clothing Store	3.5/5	N/A	0.003
Cascata Maia Shopping Snack-bar, Lda	Restaurant	3.1/5	N/A	0.003
Churrasqueira da Estacão - Ermesinde	Restaurant	4.1/5	N/A	0.003

Table B.3: Stage 1 recommendations - Card ID 86075

Name	Type	Rating	PriceLevel	Probability
Cestaria Cunha	Florist	4.3/5	N/A	0.003
Central Churrasco - São Mamede de Infesta	Restaurant	4.4/5	N/A	0.003
Ana Moura Boutique	Clothing Store	5/5	N/A	0.003

Table B.4: Stage 1 recommendations - Card ID 86096

Name	Type	Rating	PriceLevel	Probability
Bompiso-trade Tires SA	Store	4.7/5	N/A	0.595
Moldacril - Acrilicos	Store	4.7/5	N/A	0.595
MaxMat	Store	3.6/5	N/A	0.485
Donacorpus - Centro de Estetica e Bem Estar	Beauty Salon	4/5	N/A	0.403
Armando Moreira Comercio de Automoveis	Store	N/A	N/A	0.125
Mania das Plantas, Lda	Florist	N/A	N/A	0.003

Table B.5: Stage 1 recommendations - Card ID 98942

Name	Type	Rating	PriceLevel	Probability
Doce Alto (Ermesinde)	Cafe	4.1/5	N/A	0.125
MaxMat	Store	3.6/5	N/A	0.05
Armando Moreira Comercio de Automoveis	Store	N/A	N/A	0.05
ACMA	Store	5/5	N/A	0.05
A.J. Pinto - Distribution, Ltd	Store	4/5	N/A	0.05
House of Lamps, SA	Store	4.3/5	N/A	0.05
Garantia da Quintas - Sociedade Agrcola e Comercial, Lda	Store	N/A	N/A	0.05
Imagine Foto	Store	N/A	N/A	0.05
Zuni Personalbit, Lda	Store	3/5	N/A	0.05
Moldacril - Acrilicos	Store	4.7/5	N/A	0.05
Mania das Plantas, Lda	Florist	N/A	N/A	0.003
Churrasqueira da Estacao - Ermesinde	Restaurant	4.1/5	N/A	0.003

Table B.6: Stage 1 recommendations - Card ID 98920 (Part 1)

Name	Type	Rating	PriceLevel	Probability
O Paparico	Bar	4.5/5	N/A	0.575
Tintas Robbialac - Maia	Store	5/5	N/A	0.55
Investcorte - Computadores e Sistemas SA	Store	5/5	N/A	0.55
APBS Interiores	Store	5/5	N/A	0.55
Pharmacy Portas da Maia	Store	4.8/5	N/A	0.53
Tipografia Lessa	Store	4.7/5	N/A	0.52
Centro de Yoga Vaidika	Gym	5/5	N/A	0.508
Centro de Treino e Avaliacao Desportiva	Gym	5/5	N/A	0.508
Farmacia Lima Coutinho	Store	4.5/5	N/A	0.500
Farmacia Santana	Store	4.5/5	N/A	0.500
MCoutinho Parts and Repair Automotive, SA	Store	4.2/5	N/A	0.470
Catassol Central Pharmacy	Store	4.2/5	N/A	0.470
Farmacia Bastos	Store	4.1/5	N/A	0.460
Sabores da Herdade	Store	4.1/5	N/A	0.460
Doce Alto	Bakery	4.5/5	N/A	0.453
Sportzone Maia Jardim	Store	3.9/5	N/A	0.440

Table B.7: Stage 1 recommendations - Card ID 98920 (Part 2)

Name	Type	Rating	PriceLevel	Probability
Churrasqueira Portuguesa da Maia	Restaurant	4.3/5	N/A	0.43
Inauto	Store	3.8/5	N/A	0.430
Radio Popular	Store	3.8/5	N/A	0.430
MaiaFit - Espacos Desportivos	Gym	4.2/5	N/A	0.428
Casa do Lopes	Restaurant	4.2/5	N/A	0.423
McDonald’s	Restaurant	4.1/5	1/5	0.413
Furusato	Restaurant	4.1/5	N/A	0.413
Perfumaria Barreiros Faria	Clothing Store	4/5	N/A	0.403
Doce Maia Gourmet - Restauracao, Lda	Restaurant	3.9/5	N/A	0.393
Restaurante Dona Lurdes	Restaurant	3.9/5	N/A	0.393
Auto Maiamotor Lda	Store	3.4/5	N/A	0.39
Sportsdirect Maia PT	Bicycle Store	3.8/5	1/5	0.383
Materiais de Construcao Dias	Store	2.7/5	N/A	0.320
Belcol - Sociedade de Representacoes, Lda	Store	1/5	N/A	0.32
Fonseca, Lda	Store	N/A	N/A	0.05
Emilio Colombo - Comercio de Artigos de Pesca, Lda	Store	N/A	N/A	0.05

Table B.8: Stage 1 recommendations - Card ID 98918 (Part 1)

Name	Type	Rating	PriceLevel	Probability
Filotipo, Lda	Store	4.9/5	N/A	0.54
TRYP Porto Expo Hotel	Bar	4.1/5	N/A	0.535
Alflora - Flowers and Decoration	Store	4.8/5	N/A	0.530
Artur Lacerda & Filhos - Industria e Comercio de Materiais de Construcao, Lda	Bar	4.7/5	N/A	0.520
Farmacia Sousa Oliveira	Store	4.7/5	N/A	0.520
Alto Cristelo - Pet Shop	Store	4.7/5	N/A	0.520
Data Rio Servicos de Informatica, Lda	Store	4.6/5	N/A	0.510
Lupi Fotografia	Store	4.5/5	N/A	0.499
Pharmacy Sousa Reis	Store	4.5/5	N/A	0.5
VirtualLeds	Store	4.5/5	N/A	0.5
Jose Ribeiro Cabeleireiros	Beauty Salon	4.9/5	N/A	0.493
Expocar Porto - Audi	Store	4.4/5	N/A	0.490

Table B.9: Stage 1 recommendations - Card ID 98918 (Part 2)

Name	**Type**	**Rating**	**PriceLevel**	**Probability**
BDR - Bandeiras e Mastros	Store	4.3/5	N/A	0.480
Telecao	Store	4.3/5	N/A	0.480
Jumbo	Store	4.1/5	N/A	0.460
Midas shop MAR Shopping - Matosinhos	Store	4.1/5	N/A	0.460
Decathlon Matosinhos	Bicycle Store	4.5/5	N/A	0.452
Brasileirao	Restaurant	4.5/5	N/A	0.453
Lidl	Store	4/5	N/A	0.45
Ginasio Venda Nova	Gym	4.4/5	N/A	0.448
Pao Quente Da Quinta do Veira, Lda	Bakery	4.3/5	N/A	0.433
Minipreco Family	Store	3.8/5	N/A	0.430
Suprides XXI, Lda	Store	3.8/5	N/A	0.430
Churrascaria o Grelhador da Giesta	Restaurant	4.2/5	N/A	0.423
Sportsdirect Leca PT	Bicycle Store	4.2/5	1/5	0.423
Briel	Store	3.7/5	N/A	0.420
Grade Restaurante, Lda	Restaurant	4.1/5	N/A	0.413
Maria Rapaz	Amusement Park	4.1/5	N/A	0.413
Churrasqueira Sao Tiago	Restaurant	4/5	N/A	0.403
Estádio do Leça FC	Stadium	3.8/5	N/A	0.383
Restaurante Dona Tila	Restaurant	3.7/5	N/A	0.373
Jular Madeiras - Porto	Store	3/5	N/A	0.35
Winsowelu - Miguel Garcia	Spa	N/A	N/A	0.003
Hitlife	Beauty Salon	N/A	N/A	0.003

Table B.10: Stage 1 recommendations - Card ID 2286

Name	**Type**	**Rating**	**PriceLevel**	**Probability**
Decorarte - Bernardino Cardoso	Store	5/5	N/A	0.625
Restaurante Paraiso	Restaurant	4.6/5	N/A	0.585
Restaurante Veleiros	Restaurant	4.5/5	N/A	0.575
Casa Velha	Restaurant	4.4/5	N/A	0.565
Restaurante Toca da Formiga	Restaurant	4.2/5	N/A	0.545
Ribeiro & Tavares LDA	Store	4/5	N/A	0.525
Afontec, Lda	Store	3.9/5	N/A	0.515
Restaurante Mari-Ze	Restaurant	3.7/5	N/A	0.495
Clube Zupper	Gym	3.3/5	N/A	0.338
Restaurante Campo Alegre Mar	Restaurant	N/A	N/A	0.125

Table B.11: Stage 1 recommendations - Card ID 2285

Name	**Type**	**Rating**	**PriceLevel**	**Probability**
CDPCdi - Companhia Portuguesa de Computadores e Disribuicao de Produtos Informatios	Store	4.3/5	N/A	0.125
Copidouro	Store	3.7/5	N/A	0.125
Fotosport	Store	4/5	N/A	0.125
Cascata Maia Shopping Snack-bar LDA	Restaurant	3.1/5	N/A	0.125
ACMA	Store	5/5	N/A	0.125
Boutique dos Tecidos	Store	N/A	N/A	0.125
Churrasqueira da Estacao - Ermesinde	Restaurant	4.1/5	N/A	0.125
Imagine Foto	Store	N/A	N/A	0.125
Doce Alto (Ermesinde)	Cafe	4.1/5	N/A	0.075
Perfumes & Companhia	Clothing Store	3.5/5	N/A	0.05
Sabores Magnolia	Bakery Store	4/5	N/A	0.008
Zen Gym	Gym	4.5/5	N/A	0.008

Table B.12: Stage 1 recommendations - Card ID 2288 (Part 1)

Name	**Type**	**Rating**	**PriceLevel**	**Probability**
clinicabiologica.com	Store	5/5	N/A	0.625
Domingos Manuel M Freitas	Store	5/5	N/A	0.625
Decorarte - Bernardino Cardoso	Store	5/5	N/A	0.625
Farmacia Sousa Oliveira	Store	4.7/5	N/A	0.595
Luci Fotografia - Aguas Santas	Store	4.5/5	N/A	0.575
Vales & Vales - Acessorios Auto, Lda	Store	4.5/5	N/A	0.575
Restaurant Alcaide	Restaurant	4.5/5	N/A	0.565
Leitao e Coisas	Restaurant	4.5/5	N/A	0.565
MHR Computer and Appliances	Store	4.4/5	N/A	0.565
Churrasqueira POrtuguesa	Restaurant	4.5/5	N/A	0.555
Restaurante Toca da Formiga	Restaurant	4.2/5	N/A	0.545
Jardiland	Store	4.1/5	N/A	0.535
Pao Quente Sangemil	Cafe	4.6/5	N/A	0.535
Lidl	Store	4/5	N/A	0.525
A.J. Pinto - Distribution, Ltd	Store	4/5	N/A	0.525
Ribeiro & Tavares LDA	Store	4/5	N/A	0.525
Hotel Ibis Porto Sao Joao	Restaurant	3.9/5	N/A	0.515
Afontec, Lda	Store	3.9/5	N/A	0.515

Table B.13: Stage 1 recommendations - Card ID 2288 (Part 2)

Name	Type	Rating	PriceLevel	Probability
Quinta D'As Raparigas - Eventos, Lda	Restaurant	3.5/5	N/A	0.475
Restaurante Marito	Restaurant	3/5	N/A	0.425
Sentir Bem Estar, Massagem e Terapias	Spa	4/5	N/A	0.403
Clube Zupper	Gym	3.3/5	N/A	0.338
Parabichos Petshopt	Store	N/A	N/A	0.125

Table B.14: Stage 1 recommendations - Card ID 25 (Part 1)

Name	Type	Rating	PriceLevel	Probability
McDonald's	Restaurant	4.1/5	1/5	0.508
Neta I padaria e confeitaria	Bakery	4.6/5	N/A	0.125
Hora da Ribalta	Cafe	4.2/5	N/A	0.125
Data Rio Servicos de Informatica, Lda	Store	4.6/5	N/A	0.125
Jumbo	Store	4.2/5	N/A	0.075
Tintas Robbialac - Maia	Store	5/5	N/A	0.075
Auto Maiamotor, Lda	Store	3.4/5	N/A	0.075
Pharmacy Porta da Maia	Store	4.8/5	N/A	0.075
ACMA	Store	5/5	N/A	0.075
Jular Madeiras	Store	3/5	N/A	0.075
Iberacero Portugal Sociedade de Representacoes, Lda	Store	5/5	N/A	0.075
Briel	Store	3.7/5	N/A	0.075
Imagine Foto	Store	N/A	N/A	0.075
Zuni Personalbit, Lda	Store	3/5	N/A	0.075
Mobiliario Mani Design	Store	N/A	N/A	0.075
Helena Pires Sociedade Farmaceutica unipessoal, Lda	Store	4.7/5	N/A	0.075
Pinto & Sousa	Store	4.3/5	N/A	0.075
Parque Soccer	Store	3/5	N/A	0.075
Ginasio da Venda Nova	Gym	4.4/5	N/A	0.008
FitSpot	Gym	3.3/5	N/A	0.008
Central Churrasco	Restaurant	4.5/5	N/A	0.008
Churrasqueira Portuguesa da Maia	Restaurant	4.3/5	N/A	0.008
Mania das Plantas, Lda	Florist	N/A	N/A	0.008
Zen Gym	Gym	4.5/5	N/A	0.008

Table B.15: Stage 1 recommendations - Card ID 25 (Part 2)

Name	Type	Rating	PriceLevel	Probability
Churrasqueira da Estacao - Ermesinde	Restaurant	4.1/5	N/A	0.008
Rei do Churrasco	Restaurant	4/5	N/A	0.008
Winsowelu - Miguel Garcia	Spa	N/A	N/A	0.003
Perfumes & Companhia	Clothing Store	4.5/5	N/A	0.003
Maia Zoo	Zoo	4.2/5	N/A	0.003

Table B.16: Stage 1 recommendations - Card ID 27

Name	Type	Rating	PriceLevel	Probability
Sportsdirect Maia PT	Bicycle Store	3.8/5	1/5	0.503
Doce Alto	Cafe	4.1/5	1/5	0.125
Casa Carvalho - Moveis, Electrodomesticos e Decoracoes, Lda	Store	4.7/5	N/A	0.075
Koket	Store	4.5/5	N/A	0.075
Tintas Robbialac - Maia	Store	5/5	N/A	0.075
Auto Maiamotor LDA	Store	3.4/5	N/A	0.075
MaxMat	Store	3.6/5	N/A	0.075
Radio Popular	Store	3.8/5	N/A	0.075
Armando Moreira Comercio de Automoveis	Store	N/A	N/A	0.075
ACMA	Store	5/5	N/A	0.075
Boutique dos Tecidos	Store	N/A	N/A	0.075
Zuni Personalbit, Lda	Store	3/5	N/A	0.075
Moldacril - Acrilicos	Store	4.7/5	N/A	0.075
Churrasqueira da Estacao	Restaurant	4.2/5	N/A	0.008
Churrasqueira Portuguesa da Maia	Restaurant	4.3/5	N/A	0.008
Mania das Plantas, Lda	Florist	N/A	N/A	0.008
Churrasqueira da Estacao - Ermesinde	Restaurant	4.1/5	N/A	0.008
Puro Equilibrio Day Spa	Spa	5/5	N/A	0.003

Table B.17: Stage 1 recommendations - Card ID 30 (Part 1)

Name	Type	Rating	PriceLevel	Probability
Neta Padaria e Confeitaria 1	Bakery	4.6/5	N/A	0.125
Hora da Ribalta	Cafe	4.2/5	N/A	0.125
Lidl	Store	4/5	N/A	0.075
Lupi Fotografia	Store	4.5/5	N/A	0.075

Table B.18: Stage 1 recommendations - Card ID 30 (Part 2)

Name	**Type**	**Rating**	**PriceLevel**	**Probability**
Pharmacy Sousa Reis	Store	4.5/5	N/A	0.075
Data Rio Servicos de Informatica, Lda	Store	4.6/5	N/A	0.075
BDR - Bandeiras e Mastros	Store	4.3/5	N/A	0.075
Jumbo	Store	4.2/5	N/A	0.075
Casa Carvalho - Moveis, Electrodomesticos e Decoracoes, Lda	Store	4.7/5	N/A	0.075
L. Pinto Monteido, Lda	Store	5/5	N/A	0.075
Sociedade de Importacao Enrique Thumann, Lda	Store	N/A	N/A	0.075
Tintal - Company Inks Manufacturinf, Ltd	Store	4.5/5	N/A	0.075
Farmacia Esposade	Store	4.8/5	N/A	0.075
Carta Aberta, Lda	Store	N/A	N/A	0.075
Kancela - Ilustracao e Design	Store	5/5	N/A	0.075
Arcitel	Store	N/A	N/A	0.075
Expo Laranjeiras	Store	4.3/5	N/A	0.075
Casotas&Companhia!	Store	5/5	N/A	0.075
Flor e Arte	Store	4/5	N/A	0.075
Ibertec - Armando Ferreira Freire & Filho, Lda	Store	N/A	N/A	0.075
Irmaos Pinto Cardoso, Lda	Store	N/A	N/A	0.075
Ibermola	Store	4.2/5	N/A	0.075
MaxMat	Store	3.6/5	N/A	0.075
Bompiso - Trade Tires, SA	Store	4.7/5	N/A	0.075
Grafislab - Packaging, Design	Store	5/5	N/A	0.075
Carimbex - Carimbos e Gravuras, Lda	Store	4.3/5	N/A	0.075
Imagine Foto	Store	N/A	N/A	0.075
Mobiliario Mani Design	Store	N/A	N/A	0.075
Helena Pires Sociedade Farmaceutica Unipessoal LDA	Store	4.7/5	N/A	0.075
Pinto & Sousa	Store	4.3/5	N/A	0.75
Parque Soccer	Store	3/5	N/A	0.075
Churrascaria O Grelhador da Giesta	Restaurant	4.2/5	N/A	0.008
Jose Ribeiro Cabeleireiros	Beauty Salon	4.9/5	N/A	0.008
FitSpot	Gym	3.3/5	N/A	0.008

Table B.19: Stage 1 recommendations - Card ID 30 (Part 3)

Name	Type	Rating	PriceLevel	Probability
Central Churrasco	Restaurant	4.5/5	N/A	0.008
O Cardeal	Restaurant	4.2/5	N/A	0.008
Kim Kim	Restaurant	4.2/5	N/A	0.008
O Ze Pacheco	Restaurant	4.2/5	N/A	0.008
Academia 7éight	Gym	4.1/5	N/A	0.008
Perfektus Gym	Gym	3.9/5	N/A	0.008
Zen Gym	Gym	4.5/5	N/A	0.008
Rei do Churrasco	Restaurant	4/5	N/A	0.008
Perfumes & Companhia	Clothing Store	4/5	N/A	0.003
Puro Equilibrio Day Spa	Spa	5/5	N/A	0.003
Seara Producoes Audiovisuais	Clothing Store	5/5	N/A	0.003

Table B.20: Stage 2 recommendations - Card ID 86063 (Part 1)

Name	Type	Rating	PriceLevel	Probability	Feedback
Pharmacy Sousa Reis	Store	4.5/5	N/A	0.125	
CVA Electronica - Fabrica de Equipamento Som e Luz, Lda	Store	4.6/5	N/A	0.125	
Casa Carvalho - Moveis, Electrodomesticos e Decoracoes, Lda	Store	4.7/5	N/A	0.125	
L. Pinto Monteiro, Lda	Store	5/5	N/A	0.125	
Tintal - Company Inks Manufacutring, Ltd	Store	4.5/5	N/A	0.125	
Koket	Store	4.5/5	N/A	0.125	
CPCdi - Companhia Portuguesa de Computadores e Distribuicao de Produtos Informaticos	Store	4.3/5	N/A	0.125	
Artur Lagoela & Filhos - Industria e Comercio de Materiais de Construcao, Lda	Store	4.6/5	N/A	0.125	

Table B.21: Stage 2 recommendations - Card ID 86063 (Part 2)

Name	Type	Rating	PriceLevel	Probability	Feedback
Rendibor	Store	N/A	N/A	0.125	
Copidouro	Store	3.7/5	N/A	0.125	
Fotosport	Store	4/5	N/A	0.125	
ACMA	Store	5/5	N/A	0.125	
Boutique dos Tecidos	Store	N/A	N/A	0.125	
Jular Madeira - Porto	Store	3/5	N/A	0.125	
Iberacero Portugal Sociedade de Representacoes, Lda	Store	5/5	N/A	0.125	
Briel	Store	3.7/5	N/A	0.125	
Imagine Foto	Store	N/A	N/A	0.125	
Zuni Personalbit, Lda	Store	3/5	N/A	0.125	
Ginasio Venda Nova	Gym	4.4/5	N/A	0.012	+
Zen Gym	Gym	4.5/5	N/A	0.012	+
Winsowelu - Miguel Garcia	Spa	N/A	N/A	0.003	
Restaurante Charco	Restaurant	4.2/5	N/A	0.003	
Choupal dos Melros	Restaurant	4.1/5	N/A	0.003	
Gare Caffe	Restaurant	4.1/5	N/A	0.003	
Belissima -Fashion Accessories, Lda	Clothing Store	4.2/5	N/A	0.003	
Taberninha Repelao	Bar	4.5/5	N/A	0.003	
Puro Equilibrio Day Spa	Spa	5/5	N/A	0.003	
Churrasqueira da Estacao	Restaurant	4.2/5	N/A	0.003	
Mania das Plantas, Lda	Florist	N/A	N/A	0.003	
Perfumes & Companhia	Clothing Store	3.5/5	N/A	0.003	
Cascata Maia Shopping Snackbar, Lda	Restaurant	3.1/5	N/A	0.003	
Churrasqueira da Estacao - Ermesinde	Restaurant	4.1/5	N/A	0.003	
Doce Alto (Ermesinde)	Cafe	4.1/5	N/A	0.0125	-

Table B.22: Stage 2 recommendations - Card ID 98942

Name	Type	Rating	PriceLevel	Probability	Feedback
MaxMat	Store	3.6/5	N/A	0.05	
Armando Moreira Comercio de Automoveis	Store	N/A	N/A	0.05	
ACMA	Store	5/5	N/A	0.05	
A.J. Pinto - Distribution, Ltd	Store	4/5	N/A	0.05	
House of Lamps, SA	Store	4.3/5	N/A	0.05	
Garantia da Quintas - Sociedade Agrcola e Comercial, Lda	Store	N/A	N/A	0.05	
Imagine Foto	Store	N/A	N/A	0.05	
Zuni Personalbit, Lda	Store	3/5	N/A	0.05	
Moldacril - Acrilicos	Store	4.7/5	N/A	0.05	
Doce Alto (Ermesinde)	Cafe	4.1/5	N/A	0.013	-
Churrasqueira da Estacao - Ermesinde	Restaurant	4.1/5	N/A	0.005	+
Mania das Plantas, Lda	Florist	N/A	N/A	0.003	

Table B.23: Stage 2 recommendations - Card ID 94147

Name	Type	Rating	PriceLevel	Probability	Feedback
Estadio do Leça FC	Stadium	3.8/5	N/A	0.582	+
Iberacero Portugal Sociedade de Representacoes, Lda	Store	5/5	N/A	0.55	
Alto Cristelo - Pet Shop	Store	4.7/5	N/A	0.520	
Expocar Porto - Audi	Store	4.4/5	N/A	0.490	
Fnac	Store	4.1/5	N/A	0.460	
Jumbo	Store	4.1/5	N/A	0.460	
Midas shop MAR Shopping - Matosinhos	Store	4.1/5	N/A	0.460	
Laskasas Interiores - Leça da Palmeira	Store	4.1/5	N/A	0.460	
Brasileirão	Restaurant	4.5/5	N/A	0.453	
Suprides XXI, Lda	Store	3.8/5	N/A	0.430	
Sportsdirect Leca PT	Bicycle Store	4.2/5	1/5	0.423	
Eureka	Store	3.5/5	N/A	0.400	
Jular Madeiras - Porto	Store	3/5	N/A	0.35	
Smart Cartridge Marshopping	Store	N/A	N/A	0.05	
TRYP Porto Expo Hotel	Bar	4.1/5	N/A	0.054	-
Hitlife	Beauty Salon	N/A	N/A	0.003	

Table B.24: Stage 2 recommendations - Card ID 4150

Name	**Type**	**Rating**	**PriceLevel**	**Probability**	**Feedback**
TRYP Porto Expo Hotel	Bar	4.1/5	N/A	0.627	+
Bompiso - Trade Tires, SA	Store	4.7/5	N/A	0.595	
Moldacril - Acrilicos	Store	4.7/5	N/A	0.595	
Expocar Porto - Audi	Store	4.4/5	N/A	0.565	
Laskasas Interiores - Leça da Palmeira	Store	4.1/5	N/A	0.535	
Maxmat	Store	3.6/5	N/A	0.485	
Sportsdirect Leca PT	Bicycle Store	4.2/5	1/5	0.423	
Maria Rapaz	Amusement Park	4.1/5	N/A	0.413	
Donacorpus - Centro de Estetica e Bem Estar	Beauty Salon	4/5	N/A	0.408	
Estádio do Leça FC	Stadium	3.8/5	N/A	0.383	
Biozoo - Plantas e Animais de Companhia, Lda	Store	2/5	N/A	0.325	
Armando Moreira Comercio de Automoveis	Store	N/A	N/A	0.125	
Brasileirao	Restaurant	4.5/5	N/A	0.058	-
Grade Restaurante, Lda	Restaurant	4.1/5	N/A	0.054	-
NewStaff Restaurante	Restaurant	3.8/5	N/A	0.051	-
Take away Veleiros	Restaurant	N/A	N/A	0.013	-
Mania das Plantas, Lda	Florist	N/A	N/A	0.003	

Table B.25: Stage 2 recommendations - Card ID 2285 (Part 1)

Name	**Type**	**Rating**	**PriceLevel**	**Probability**	**Feedback**
Cascata Maia Shopping Snack-bar, Lda	Restaurant	3.1/5	N/A	0.125	
Churrasqueira da Estacao - Ermesinde	Restaurant	4.1/5	N/A	0.125	
Doce Alto (Ermesinde)	Cafe	4.1/5	N/A	0.075	
Perfumes & Companhia	Clothing Store	3.5/5	N/A	0.05	
CDPCdi - Companhia Portuguesa de Computadores e Disribuicao de Produtos Informatios	Store	4.3/5	N/A	0.013	-

Table B.26: Stage 2 recommendations - Card ID 2285 (Part 2)

Name	**Type**	**Rating**	**PriceLevel**	**Probability**	**Feedback**
Copidouro	Store	3.7/5	N/A	0.013	-
Fotosport	Store	4/5	N/A	0.013	-
ACMA	Store	5/5	N/A	0.013	-
Boutique dos Tecidos	Store	N/A	N/A	0.013	-
Imagine Foto	Store	N/A	N/A	0.013	-
Zen Gym	Gym	4.5/5	N/A	0.012	+
Sabores Magnolia	Bakery Store	4/5	N/A	0.008	

Table B.27: Stage 2 recommendations - Card ID 1755

Name	**Type**	**Rating**	**PriceLevel**	**Probability**	**Feedback**
Copidouro	Store	3.7/5	N/A	0.075	
Fotosport	Store	4/5	N/A	0.075	
ACMA	Store	5/5	N/A	0.075	
Boutique dos Tecidos	Store	N/A	N/A	0.075	
Imagine Foto	Store	N/A	N/A	0.075	
ZUNI Personalbit, Lda	Store	3/5	N/A	0.075	
Doce Alto (Ermesinde)	Cafe	4.1/5	N/A	0.0125	-
Cascata Maia Shopping Snack-bar, Lda	Restaurant	3.1/5	N/A	0.008	
Churrasqueira da Estacao - Ermesinde	Restaurant	4.1/5	N/A	0.008	
Perfumes & Companhia	Clothing Store	3.5/5	N/A	0.005	+

Table B.28: Stage 2 recommendations - Card ID 19 (Part 1)

Name	**Type**	**Rating**	**PriceLevel**	**Probability**	**Feedback**
Sportsdirect Maia PT	Bicycle Store	3.8/5	1/5	0.503	
Tintas Robbialac - Maia	Store	5/5	N/A	0.113	+
Auto Maiamotor, Lda	Store	3.4/5	N/A	0.113	+
Maxmat	Store	3.6/5	N/A	0.113	+
Radio Popular	Store	3.8/5	N/A	0.113	+
Damaceno & Antunes - Tecidos de Decoracao, Lda	Store	4.4/5	N/A	0.113	+
Farmacia Agra	Store	4.4/5	N/A	0.113	+
LandParts	Store	N/A	N/A	0.113	+

Table B.29: Stage 2 recommendations - Card ID 19 (Part 2)

Name	Type	Rating	PriceLevel	Probability	Feedback
Reciclacar - Importacao, Exportacao e Reciclagem de Veiculos Automoveis, Lda	Store	3.6/5	N/A	0.113	+
Lopes dos Piglets	Store	4.3/5	N/A	0.113	+
Stand Pinto - Comercio de Automoveis	Store	4.1/5	N/A	0.113	+
Armando Moreira Comercio de Automoveis	Store	N/A	N/A	0.113	+
Copidouro	Store	3.7/5	N/A	0.113	+
Fotosport	Store	4/5	N/A	0.113	+
ACMA	Store	5/5	N/A	0.113	+
Imagine Foto	Store	N/A	N/A	0.113	+
ZUNI Personalbit, Lda	Store	3/5	N/A	0.113	+
Mobiliario Mani Design	Store	N/A	N/A	0.113	+
Moldacril - Acrilicos	Store	4.7/5	N/A	0.113	+
Churrascaria Portuguesa da Maia	Restaurant	4.3/5	N/A	0.008	
Restaurante Sabores de Prata	Restaurant	4.2/5	N/A	0.008	
Quinta Casa do Arco	Restaurant	4.2/5	N/A	0.008	
Mania das Plantas, Lda	Florist	N/A	N/A	0.008	
Cascata Maia Shopping Snack bar, Lda	Restaurant	3.1/5	N/A	0.008	
Churrasqueira da Estacao	Restaurant	4.1/5	N/A	0.008	
Campo Longo Industria e Comercio de Vestuario, Lda	Clothing Store	N/A	N/A	0.003	
Perfumes & Companhia	Clothing Store	3.5/5	N/A	0.003	
Body Space	Gym	1/5	N/A	0.001	-
Pavilhao Municipal de Gueifaes II	Gym	4.4/5	N/A	0.001	-
Just 4 Fit, Lda	Gym	4.7/5	N/A	0.001	-
Zen Gym	Gym	4.5/5	N/A	0.001	-

Table B.30: Stage 3 recommendations - Card ID 86063

Name	**Type**	**Rating**	**PriceLevel**	**Probability**
A.J. Pinto II - Distribution, Ltd	Store	4/5	N/A	0.125
House of Lamps, SA	Store	4.3/5	N/A	0.125
Garantia das Quintas - Sociedade Agricola Comercial, Lda	Store	N/A	N/A	0.125

Table B.31: Stage 3 recommendations - Card ID 94147

Name	**Type**	**Rating**	**PriceLevel**	**Probability**
O2 Fitness	Gym	5/5	N/A	0.508
Leitao & Coisas	Restaurant	4.4/5	N/A	0.443
Casa Velha	Restaurant	4.4/5	N/A	0.443
Central Churrasco - Sao Mamede Infesta	Restaurant	4.4/5	N/A	0.443
Tropical Burguer - Circunvalacao	Restaurant	4/5	N/A	0.403

Table B.32: Stage 3 recommendations - Card ID 98920

Name	**Type**	**Rating**	**PriceLevel**	**Probability**
PiuBelle - Clothing Industria e Comercio, SA	Store	4.6/5	N/A	0.51
Damaceno & Antunes - Tecidos de Decoracao, Lda	Store	4.4/5	N/A	0.49
Farmacia Agra	Store	4.4/5	N/A	0.49
Solemai - Acessorios Agricolas, Lda	Store	4/5	N/A	0.45
Pavilhao Municipal Gueifaes II	Store	4.4/5	N/A	0.448
Restaurante Sabores de Prata	Restaurant	4.2/5	N/A	0.423

Table B.33: Stage 3 recommendations - Card ID 98918

Name	**Type**	**Rating**	**PriceLevel**	**Probability**
Mega Ensaio - Producao de Audiovisuais, Lda	Store	4.5/5	N/A	0.5
Fnac	Store	4.1/5	N/A	0.46
Michaela Larisch	Clothing Store	3.5/5	N/A	0.353

Table B.34: Stage 3 recommendations - Card ID 99068

Name	Type	Rating	PriceLevel	Probability
Biozoo - Plantas e Animais de Companhia, Lda	Store	2/5	N/A	0.05

Table B.35: Stage 3 recommendations - Card ID 2285

Name	Type	Rating	PriceLevel	Probability
Materiais de Construcao Dias	Store	2.7/5	N/A	0.125
Mecinor - Metalurgica Civil e Campismo Lda	Store	N/A	N/A	0.125
Centro de Yoga Vaidika	Gym	5/5	N/A	0.008
Pao Quente Nortenho, Lda	Bakery	4.3/5	N/A	0.008

Table B.36: Stage 3 recommendations - Card ID 2287

Name	Type	Rating	PriceLevel	Probability
Farmacia de Fanzeres	Store	4.4/5	N/A	0.125
Maquinouro	Store	4.2/5	N/A	0.125
Acilio Sousa e Castro, Lda	Store	N/A	N/A	0.125
Masipao	Bakery	4.4/5	N/A	0.0.008

Table B.37: Stage 3 recommendations - Card ID 19

Name	Type	Rating	PriceLevel	Probability
Farmacia Bastos	Store	4.1/5	N/A	0.075
PiuBelle - Clothing Industria e Comercio, SA	Store	4.6/5	N/A	0.075
Solemai - Acessorios Agricolas, Lda	Store	4/5	N/A	0.075

Table B.38: Stage 3 recommendations - Card ID 27

Name	Type	Rating	PriceLevel	Probability
Neta I padaria e confeitaria	Bakery	4.6/5	1/5	0.125
Hora da Ribalta	Cafe	4.2/5	1/5	0.125
Mobiliario Mani Design	Store	N/A	1/5	0.075
Helena Pires Sociedade Farmaceutica Unipessoal	Store	4.7/5	1/5	0.075
Pinto & Sousa	Store	4.3/5	1/5	0.075
Parque Soccer	Store	3/5	1/5	0.075

Table B.39: Stage 4 recommendations - Card ID 94147

Name	Type	Rating	PriceLevel	Probability
O2 Fitness	Gym	5/5	N/A	0.513
Leitao & Coisas	Restaurant	4.4/5	N/A	0.451
Casa Velha	Restaurant	4.4/5	N/A	0.451
Central Churrasco - Sao Mamede Infesta	Restaurant	4.4/5	N/A	0.451
Tropical Burguer - Circunvalacao	Restaurant	4/5	N/A	0.411

Table B.40: Stage 4 recommendations - Card ID 98920

Name	Type	Rating	PriceLevel	Probability
PiuBelle - Clothing Industria e Comercio, SA	Store	4.6/5	N/A	0.515
Damaceno & Antunes - Tecidos de Decoracao, Lda	Store	4.4/5	N/A	0.495
Farmacia Agra	Store	4.4/5	N/A	0.495
Solemai - Acessorios Agricolas, Lda	Store	4/5	N/A	0.455
Pavilhao Municipal Gueifaes II	Store	4.4/5	N/A	0.453
Restaurante Sabores de Prata	Restaurant	4.2/5	N/A	0.431

Table B.41: Stage 4 recommendations - Card ID 98918

Name	Type	Rating	PriceLevel	Probability
Mega Ensaio - Producao de Audiovisuais, Lda	Store	4.5/5	N/A	0.505
Fnac	Store	4.1/5	N/A	0.465
Michaela Larisch	Clothing Store	3.5/5	N/A	0.358

Table B.42: Stage 4 recommendations - Card ID 99068

Name	Type	Rating	PriceLevel	Probability
Biozoo - Plantas e Animais de Companhia, Lda	Store	2/5	N/A	0.055

Table B.43: Stage 4 recommendations - Card ID 2285

Name	Type	Rating	PriceLevel	Probability
Centro de Yoga Vaidika	Gym	5/5	N/A	0.024
Pao Quente Nortenho, Lda	Bakery	4.3/5	N/A	0.021
Materiais de Construcao Dias	Store	2.7/5	N/A	0.001
Mecinor - Metalurgica Civil e Campismo Lda	Store	N/A	N/A	0.001

Table B.44: Stage 4 recommendations - Card ID 2287

Name	Type	Rating	PriceLevel	Probability
Masipao	Bakery	4.4/5	N/A	0.021
Farmacia de Fanzeres	Store	4.4/5	N/A	0.001
Maquinouro	Store	4.2/5	N/A	0.001
Acilio Sousa e Castro, Lda	Store	N/A	N/A	0.001

Table B.45: Stage 4 recommendations - Card ID 19

Name	Type	Rating	PriceLevel	Probability
Farmacia Bastos	Store	4.1/5	N/A	0.123
PiuBelle - Clothing Industria e Comercio, SA	Store	4.6/5	N/A	0.123
Solemai - Acessorios Agricolas, Lda	Store	4/5	N/A	0.123

Table B.46: Stage 4 recommendations - Card ID 27

Name	Type	Rating	PriceLevel	Probability
Neta I padaria e confeitaria	Bakery	4.6/5	1/5	0.126
Mobiliario Mani Design	Store	N/A	1/5	0.123
Helena Pires Sociedade Farmaceutica Unipessoal	Store	4.7/5	1/5	0.123
Pinto & Sousa	Store	4.3/5	1/5	0.123
Parque Soccer	Store	3/5	1/5	0.123
Hora da Ribalta	Cafe	4.2/5	1/5	0.073