UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO

PAUL DANY FLORES ATAUCHI

Broker-RecSys: An Interactive Recommender System for Insurance Brokerage

Thesis presented in partial fulfillment
of the requirements for the degree of
Master of Computer Science

Advisor: Profa. Dra. Renata Galante
Coadvisor: Profa. Dra. Luciana Nedel

Porto Alegre
July 2020

CIP — CATALOGING-IN-PUBLICATION

Atauchi, Paul Dany Flores

Broker-RecSys: An Interactive Recommender System for Insurance Brokerage / Paul Dany Flores Atauchi. – Porto Alegre: PPGC da UFRGS, 2020.

130 f.: il.

Thesis (Master) – Universidade Federal do Rio Grande do Sul. Programa de Pós-Graduação em Computação, Porto Alegre, BR–RS, 2020. Advisor: Renata Galante; Coadvisor: Luciana Nedel.

1. Recommender System. 2. Data Mining. 3. Data Visualization. 4. Insurance Brokerage. I. Galante, Renata. II. Nedel, Luciana. III. Título.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
Reitor: Prof. Rui Vicente Oppermann
Vice-Reitora: Prof[a]. Jane Fraga Tutikian
Pró-Reitor de Pós-Graduação: Prof. Celso Giannetti Loureiro Chaves
Diretora do Instituto de Informática: Prof[a]. Carla Maria Dal Sasso Freitas
Coordenadora do PPGC: Prof[a]. Luciana Salete Buriol
Bibliotecária-chefe do Instituto de Informática: Beatriz Regina Bastos Haro

“ There is a more powerful driving force than steam, electricity and atomic energy: the willpower. ”

— ALBERT EINSTEIN

ACKNOWLEDGEMENTS

I want to thank my advisor, Profa. Renata Galante, and co-advisor, Profa. Luciana Nedel, for their extraordinary support during my master's period. I thank Luís Evers, an insurance broker specialist from "Clic corretores" for the technical support and for providing the insurance portfolio data used in this work. I thank all friends that accompanied me during this period. I thanks my family, my mother Luz Marina, my aunt Magdalena and my siblings Sharmeli and Russel for being my support and motivation. Finally, I want to thank the Brazilian National Council for Scientific and Technological Development (CNPq) to finance my scholarship during the development of this work.

ABSTRACT

Recommender systems are widely used in diverse domains such as e-commerce, tourism, insurance, and so on. However, numerous recommender systems take users into account as an endpoint in the recommendation process and act like a black-box. Therefore, the black-box nature of the recommendation systems limits the understanding and acceptance of the recommendation received by the user. In contrast, interactive recommender systems can solve these drawbacks. Interactive recommender systems combine user interaction, information visualization, and recommender system methods. In the brokerage domain, insurance brokers offer, negotiate, and sell insurance products for their customers. Support brokers into the recommendation process for offering the most relevant insurance products for their customers can improve their loyalty, profit, and marketing campaign in their client portfolio. This work presents Broker-RecSys, an interactive insurance product recommender system framework, to support brokers into the recommendation process for offering insurance products in their client portfolios. The system operates at two levels to provide recommendations: recommendations for a specific customer; and recommendations for a group of customers in the portfolio of clients of a broker. Looking for offering personalized recommendations, Broker-RecSys provides a module to perform customer segmentation based on specific customer characteristics that are interesting for the broker. Broker-RecSys provides two types of recommendations based on popularity and purchase behavior. Several interactions and visualization methods are integrated into Broker-RecSys to support brokers in the recommendation process. Broker-RecSys is evaluated into the usability and usefulness dimensions. Thus, we combine the widely used evaluation method based on questionnaires and the evaluation based on the eye-tracking analysis. Results achieved suggest that data mining methods, while combined with user interaction and data visualization methods, support users in the recommendation process.

Keywords: Recommender System. Data Mining. Data Visualization. Insurance Brokerage.

Broker-RecSys: Um Sistema Interativo de Recomendação para Corretagem de Seguros

RESUMO

Os sistemas de recomendação são amplamente utilizados em diversos domínios, como comércio eletrônico, turismo, seguros, entre outros. No entanto, muitos sistemas não levam em consideração o usuário no processo de recomendação, agindo como uma caixa preta. A natureza da caixa preta limita o entendimento e a aceitação da recomendação recebida pelo usuário. Por outro lado, os sistemas interativos de recomendação podem solucionar essas limitações, pois combinam métodos de interação com o usuário, visualização de informações e sistema de recomendação. No domínio de corretoras, os corretores de seguros oferecem, negociam e vendem produtos de seguro para os seus clientes. Apoiar os corretores no processo de recomendação para oferecer os produtos de seguro mais relevantes para seus clientes pode melhorar sua confiança, lucro e campanha de marketing em seu portfólio de clientes. Este trabalho apresenta o Broker-RecSys, um framework de sistema interativo de recomendação de produtos de seguros para apoiar os corretores no processo de recomendação para oferecer produtos de seguros em seu portfólio. O sistema opera em dois níveis para oferecer recomendações: recomendações para um cliente específico; e recomendações para um grupo de clientes da carteira de clientes de um corretor. Procurando oferecer recomendações personalizadas, o Broker-RecSys fornece um módulo para realizar segmentação de clientes com base em características específicas do cliente que são relevantes para o corretor. Dois tipos de recomendações são fornecidos pelo Broker-RecSys: com base na popularidade e no comportamento de compra. Diversas interações e métodos de visualização são integrados ao Broker-RecSys para oferecer suporte aos corretores no processo de recomendação. O Broker-RecSys é avaliado nas dimensões de usabilidade e utilidade. Assim, para avaliar o Broker-RecSys, combinamos o método de avaliação amplamente usado com base em questionários e a avaliação com base na análise de rastreamento ocular. Os resultados alcançados sugerem que os métodos de mineração de dados, combinados aos métodos de interação e visualização de dados, apoiam os usuários no processo de recomendação.

Palavras-chave: Sistema de Recomendação, Mineração de Dados, Visualização de Dados, Corretagem de Seguros.

LIST OF ABBREVIATIONS AND ACRONYMS

ARM	Association Rule Mining
UI	User Interface
UX	User Experience
D3	Data-Driven Documents
MVT	Model-View-Template
SUS	System Usability Scale
1D	One dimension
2D	Two dimensions
ND	N-dimensions

LIST OF FIGURES

LIST OF TABLES

CONTENTS

1 INTRODUCTION

Recommendation systems are becoming very popular and widely used in many traditional and emerging domains such as e-commerce, news, tourism, and finance (LU et al., 2015; KARIMI; JANNACH; JUGOVAC, 2018; ZIBRICZKY, 2016). These systems help users into the decision making to choose the most relevant and accurate items (products or services) that meet their needs, preferences, and interests (ADOMAVICIUS; TUZHILIN, 2005).

For example, when a user enters into the Netflix platform (television programs and movies streaming service by internet), this platform collects information such as user interaction (choice of genre, category, year of shows and movies or ratings that user gives to these programs), period the user watches programs, and device used to navigate and watch Netflix. Netflix recommender systems use this information to recommend the most personalized shows and movies for the user in order to increase the user acceptance (GOMEZ-URIBE; HUNT, 2016; Netflix, 2019).

In order to perform recommendations, recommender systems use different and varied resources such as rating, transaction data, item descriptions, and demographic information as well as recommendation system methods and techniques (BOBADILLA et al., 2013). In the literature, there are several methods to build a recommender system such as collaborative-based filtering that recommend items to a user based on preferences and interests of similar users, content-based filtering that perform the recommendation based on the content of items chosen by the user in the past, demographic filtering that follows a principle that users have similar preferences and interests while they are similar in personal characteristics, knowledge-based filtering where is consider prior knowledge constraints related to a domain, popularity-based recommendation where is recommend the most popular items for a user and hybrid filtering that combine the diverse recommender system methods in order to suppress their limitations (ADOMAVICIUS; TUZHILIN, 2005; AGGARWAL, 2016b; BRESSAN et al., 2016).

Cold-start is a common problem affecting recommendation systems. The cold-start problem occurs when there is no prior information on user preferences (new user or user-items sparsity) or about an item (without sales record or rating evaluation). This problem is commonly addressed with a hybrid approach where personal and popularity information acts as a complement to recommendation methods to deal with the cold-start problem (SOBHANAM; MARIAPPAN, 2013; LIKA; KOLOMVATSOS; HAD-

JIEFTHYMIADES, 2014; PANDEY; RAJPOOT, 2016). On the other hand, numerous recommender systems take users into account as an endpoint in the recommendation process and act like a black-box. Therefore, the black-box nature of recommender systems limits the understanding and acceptance of the recommendation suggested for the users. In contrast, interactive recommender systems can solve these drawbacks. Interactive recommender systems combine user interaction, information visualization, and recommender system methods. Recent works reported that interactive recommender systems improve the user experience into the recommendation process, increasing the understanding and recommendation acceptance (HE; PARRA; VERBERT, 2016; VALDEZ; ZIEFLE; VERBERT, 2016; JUGOVAC; JANNACH, 2017; DU et al., 2018; LOEPP; ZIEGLER, 2019; TSAI; BRUSILOVSKY, 2019).

In the financial area, users and organizations generate massive and rapid amounts of data every day. This large volume of data is a rich source for recommender systems to recommend suitable products recommendation. Recommender systems are well applied to different financial subdomains such as banks, insurance, loan, stock (YAN; XIE, 2009; ZIBRICZKY, 2016). In the insurance domain, provide a suitable insurance product for the customers is a challenging task due to the existence of some restrictions such as indirect interaction with the client and low frequency of products purchase (ROKACH et al., 2013). In recent years, several approaches are proposed for the insurance products recommendation based on different methods and techniques, some of which are: case-based reasoning (RAHMAN; NORMAN; SOON, 2006), association rule mining (XU et al., 2014), collaborative filtering (ROKACH et al., 2013) and Bayesian networks (QAZI et al., 2017). Although these works are diverse in methods and techniques, they do not support agents/brokers in the recommendation process (customer identification, understanding of the recommendation, and recommendation of products). Knowing which products to recommend to a user is just as important as knowing why a customer receive a certain recommendation and how the recommendation was obtained.

Our work is centered on the insurance brokerage domain in Brazil. An insurance broker is an intermediate between the insurer and the policyholder. The broker offers, negotiates, and finally sell products for their customers. For retain existing policyholders, ideal insurance brokers activities would consist of exploring and identifying potential customers for recommending insurance products (e.g., Gretel, a young girl that lives in the south of Brazil has a car insurance product purchased, she needs more products) as well as identify interesting recommendations to offer (e.g., recommend life insurance for Gretel with 80% of probability of acceptance based on her purchasing behavior considering similar people to her).

Support brokers into the recommendation process for offering insurance products for their customers can improve the loyalty, profit, and marketing campaign in their client portfolio.

In this context, we propose Broker-RecSys, an interactive recommender system framework to support brokers into the recommendation process for offers insurance products in their portfolio at two levels: recommendations for a specific customer; and recommendations for a group of customers. Looking for offering personalized recommendations, Broker-RecSys provides a module to perform customer segmentation based on specific customer characteristics that are interesting for the broker. Broker-RecSys provides two types of recommendations: based on purchase behavior and popularity. Besides, the insurance broker can answer questions such as what products offer to my clients? and why offer these products to my customers?. Several interactions and visualization methods are integrated into Broker-RecSys in order to support brokers in the recommendation process.

To evaluate Broker-RecSys, we combine the widely used evaluation method based on questionnaires and the evaluation based on the eye-tracking analysis. Broker-RecSys is evaluated into the usability and usefulness dimensions. Results show that data mining methods, while combined with interaction and visualization methods, support users into the recommendation process for recommending products.

We aim to answer two research questions related to usability and usefulness dimensions. The research questions are detailed as follow:

1. **Broker-RecSys turn able naive users to perform insurance products recommendation tasks?**. This question allows us to answer what is the Broker-RecSys usability level.
2. **Broker-RecSys supports insurance brokers in the recommendation process for offers insurance products in their client portfolio?**. This question allows us to answer what is the Broker-RecSys usefulness level.

The main contributions of our interactive recommender system are five-folds:

- Dual-level Recommendation System: The proposed interactive recommender system framework allows performing recommend at two levels: for a specific customer and a group of customers. The Broker can answer situations such as: *I want to rec-*

ommend insurance products for Ana or *I want to recommend insurance products for a group of people of the south of Brazil.*

- Controllability (segmentation): Cluster customers based on specific attributes allow them to obtain a diverse variety of recommendations, and it allows offers more personalized recommendations. We look for the broker to answer situations such as *I want to recommend insurance products for my clients based on gender and age*, *I want to recommend insurance products for my clients based on region* and so on.
- User interaction interface: The user interaction and data visualization help into the cognitive process and human visual interpretation. The broker can able to answer situations such as *I want to recommend insurance products for my clients that are between 30 and 50 years old, live in the west of Brazil and have two insurance products purchased* or *I want to perform a market campaign of insurance car and travel; then, I want to identify a group of customers in my portfolio that are younger and live in the west that can receive this recommendation.* All these situations making use of the user interaction interface.
- Explanation: The explanation of the recommendation is a crucial factor that has a strong relationship with the recommendation acceptance. The recommendation is explained based on their measures. The broker can answer questions such as *what insurance product is offered to my client/clients and why?.*
- Cold-start Problem: To address the no-prior sales information and as a complement to the recommendations based on the purchase behavior, the most popular products in a customer group are taken into consideration for recommending. The broker can visualize the *top-n most popular insurance products to offers to an identified customer.*

This dissertation is organized as follows: Chapter 2 briefly describes theoretical backgrounds, Chapter 3 reviews related works, Chapter 4 presents Broker-RecSys, Chapter 5 detail the experimental methodology and results and Chapter 5.5 discusses the experimental results. Finally, Chapter 6 concludes the dissertation and present future directions of the work.

2 BACKGROUND THEORY

This chapter is divided into three parts. In the first part, the fundamental concepts of recommender systems related to this work are presented, and the second part describes several data mining methods that take part in the development of the present work. Finally, the third part describes a brief data visualization fundamentals.

2.1 Recommender System

Recommender systems are widely used in diverse types of applications such as business, e-commerce, tourists, and so on (LU et al., 2015). Recommender systems are techniques and methods that allow alleviating the overhead of information that receives a user to provide only the most relevant items for him/her. These systems use various types of resources, such as transaction data, rating, demographic data, or item content information, as well as diverse methods and techniques from retrieval information, data mining, and machine learning. In the literature, we can find several recommendation systems approaches, such as collaborative filtering, demographic-based filtering, content-based filtering, and hybrid filtering (ADOMAVICIUS; TUZHILIN, 2005; BOBADILLA et al., 2013; RICCI; ROKACH; SHAPIRA, 2015; TARNOWSKA; RAS; JASTREBOFF, 2017). In the following paragraphs, some approaches used in this work are described.

2.1.1 Popularity

The popularity approach is a useful and straightforward method in recommender systems (ZHAO et al., 2010). We can see some real cases where this approach is applied, for example, it is common for people to read the news at the main cover of a portal journal (news trend) (ZIHAYAT et al., 2019), people buy popular products suggested by the social tendency (fashion trend) (HWANGBO; KIM; CHA, 2018) or discover social popularity in social media (WANG; ZHANG; YAMASAKI, 2020). Thus, the popularity approach can perform recommendations based on the frequency of product purchase, the number of views of an online movie, ratings of a post in a social media, or any information that denotes popularity or tendency (TATAR et al., 2014; NATARAJAN; MOH, 2016; MAO et al., 2019). The popularity method, combined with the clusterization method, is a useful strategy to alleviate the cold-start problem in recommender systems (LOH et al., 2009; SHINDE; KULKARNI, 2012; MIAO et al., 2016).

2.1.2 Collaborative Filtering

Collaborative-based filtering is one of the most used methods in recommendation systems (LU et al., 2015; AMIN; PHILIPS; TABRIZI, 2019). This method performs the recommendation of items in a collaborative manner, i.e., to recommend items for a user, it uses the rating history from similar users (ADOMAVICIUS; TUZHILIN, 2005). For example, if two users have purchased two similar products, and one of the users purchase an extra item, it can be probable that the other user will also buy the extra item. In the following paragraph, it technically describes collaborative-based filtering.

Given a set of n users $U = \{u_p\}$ where $p \in \{1, ..., n\}$ and a set of m items $I = \{i_q\}$ where $q \in \{1, ..., m\}$. Then, the users and items are mapped into the index rows and columns of a matrix R respectively. The element of the matrix $R[u_p, i_q]$ represents the rating given by the user u_p for the item i_q. Commonly the rating $R[u_p, i_q]$ vary between 1 to 5, and represent the level of preference of the user u_p for a particular item i_q. For an unrated item i_q by a user u_p, the Collaborative-based filtering calculate $R[u_p, i_q]$ based on an aggregate function and preferences of other users U' for the item i_q (See Equation 2.1).

$$R[u_p, i_q] = \operatorname*{aggr}_{u'_p \in U'} R[u'_p, i_q] \tag{2.1}$$

Collaborative-based filtering is divided into two groups: memory-based methods and model-based methods (BOBADILLA et al., 2013). In the first group, the methods use a rating matrix for performing prediction of ratings. The results are continuously updated based on the change of the entry rating matrix because, for each prediction, it uses the entry rating matrix (See Equation 2.1). On the contrary, the second group includes methods that create models, abstractions about real-world processes. These models are more persistent in the time (low frequency of model updates) but need considerable computational resources at the beginning of the model's creation.

2.1.3 Content-based Filtering

The Content-based filtering approach aims to recommend items for a user based on the similarity of item profile (items rated by the user in the past) of the user with each item profile available in the item catalog (ADOMAVICIUS; TUZHILIN, 2005; AGGARWAL, 2016a). This approach takes into account the item information in order to create the item profile of the user. A feature vector represents the item profile. Due that the content-based filtering works with textual information, the most used method to create the feature vector is based on the term frequency, and inverse document frequency (TF-IDF) measures (SALTON, 1989; BAFNA; PRAMOD; VAIDYA, 2016). To recommend an item for a user, it calculates the similarity between the item profile of the user with each item profile in the item catalog. Finally, the k most similar items profile in the item catalog with high similarity is recommended for the user u. Content-based filtering approach is well applied in several applications domains such as citation recommendation (ZARRINKALAM; KAHANI, 2012), publication recommendation (WANG et al., 2018) , movie recommendation (REDDY et al., 2019) and so on.

2.1.4 Demographic Filtering

Demographic-based filtering approach follows the premise that similar users based on personal information (gender, age, region, and others) are likely to have similar preferences (PAZZANI, 1999). Demographic data is used to predict the personality traits (PARYUDI; ASHARI; TJOA, 2019), product, and user demographic information to perform stratified product recommendation (ZHAO et al., 2016) and so on. This method is commonly used to enhancement the collaborative-based filtering approach (VOZALIS; MARGARITIS, 2004; SRIDEVI; RAO, 2017).

2.1.5 Hybrid Filtering

In order to alleviate the drawbacks in recommender systems methods, the hybrid approach is used. The hybrid approach treats about to combine two or more methods such that they complement each other as well as improve their performance (ÇANO; MORISIO, 2017).

Classical recommender systems combine collaborative filtering, content-based filtering, demographic filtering, or Knowledge-based filtering to address the cold-start problem, data sparsity, accuracy, scalability, diversity, or others (ÇANO; MORISIO, 2017).

2.2 Data Mining

The data is generated for several different sources, such as IoT, online social networks, or search engines (CHEN; MAO; LIU, 2014). Those data contain much information for exploitation. Data mining involves several phases to extract and discover that information and patterns hidden in the data (WITTEN et al., 2016). Data mining provides several methods according to the goal to perform, such as predictive or descriptive analysis in the data. Data mining is well applied to recommender systems as related in (AMATRIAIN et al., 2011; NAJAFABADI; MOHAMED; MAHRIN, 2019). In this work, we use some data mining techniques to build the core of Broker-RecSys; more specifically, we use the association rule mining and clusterization method.

2.2.1 Association Rule Mining

The association rules mining is one of the most used methods of data mining with several applications in recommender systems, medical diagnostics (WANG et al., 2019; RAMASAMY; NIRMALA, 2020), customer market analysis (WANG, 2019), customer behavior analysis (CHENG; WU; CHEN, 2019) and so on. This method aims to find related items in transaction data (AGRAWAL; IMIELIŃSKI; SWAMI, 1993; AGGARWAL; PHILIP, 1999; ZHANG; ZHANG, 2002; OLSON; LAUHOFF, 2019). The association between items measures the co-occurrence, not the causality between them. For example, in the market basket analysis[1], given items purchased by customers, it is detected that a considerable quantity of customers purchases beer and diapers. Then, an interest question can be *what this means?*, and *what can we make with this?*. It can means that beer and diapers have a relationship, and possibly beer purchase is purchased together with a diaper if we presume that a customer is a man that has a baby and went to the supermarket after the work. The supermarket can improve the user experience by move beer closer to the diaper in the late afternoon or weekend. It is a famous beer and diaper story.

[1]technique used to analyze customer purchase behavior by analyzing the shopping baskets.

The paragraph below describes the association rule mining method technically.

Given a set of items $I = \{I_i\}$ where $i \in \{1, .., n\}$ and n denotes the number of items and a set of transactions $T = \{T_j\}$ where $j \in \{1, .., m\}$ and m denote the number of transactions, each transaction contains a subset of items $T_j = \{I_s\}$ where $s \in K$ and $K \subset \{1, .., n\}$ that indicates the active items in the transaction T_j. The rules are generated in the form $X \Rightarrow Y$ where X and Y are itemsets in the transaction T, called antecedent and consequent respectively. The transaction T_j is composed by X_j and Y_j where $X_j \cup Y_j = T_j$ and $X_j \cap Y_j = \Phi$. The importance of the rule $X \Rightarrow Y$ can be measured by the $support(X \Rightarrow Y) = P(X \cap Y)$ and $confidence(X \Rightarrow Y) = P(Y \mid X) = P(X \cap Y)/P(X)$, that means, the probability of appearance together of the itemsets X and Y in the transaction data and the probability of the itemsets X and Y appearance together given the probability of appearance of the itemset X respectively.

Considering exist several methods for association rule mining rules such as the Apriori algorithm, FP-tree Based Model, and OPUS Based Algorithm (ZHANG; ZHANG, 2002). In this work, we use the Apriori algorithm. The Apriori algorithm (AGRAWAL; SRIKANT et al., 1994) is simple, widely used, and a fast algorithm for mining transaction data. This method uses prior knowledge of frequent itemset. The frequent itemset properties consist in that all subset of a frequent itemset must be frequent, and all subset of an infrequent itemset must be infrequent. Besides, for prune itemsets, the apriori algorithm uses a threshold of minimum support. This property improves efficiency in time and reduces the search space in the rule mining process. The Apriori algorithm is detailed in Algorithms 1, 2, 3 and 4.

Algorithm 1: Apriori algorithm

input : T: transaction data
$minsup$: threshold of minimum support
output: F: frequent itemsets

```
T_{i∈{1,..,|T|}} ∈ T where T_i is a transaction and I_{i∈{1,..,|T_i|}} ∈ T_i is an item
frequentItemSet ← FirstPassApriori(T)
F ← F ∪ frequentItemSet
k ← 2
while frequentItemSet ≠ ∅ do
    candidatesItemSet ← AprioriGen(frequentItemSet, k)
    frequentItemSet ← AprioriPrune(candidatesItemSet, T, minsup)
    F ← F ∪ frequentItemSet
    k ← k + 1
end
```

Algorithm 2: FirstPassApriori

input : T: transaction data
output: L: large itemset

$totalItems \leftarrow \sum_{i \in \{1,..,|T|\}} |T_i|$
$itemCounts \leftarrow [0]_{totalItems}$
foreach $t \in T$ **do**
  foreach $i \in t$ **do**
    $itemCounts[i] \leftarrow itemCounts[i] + 1$
  end
  if $itemCounts[i] >= minsup$ **then**
    $L \leftarrow L \cup I_i$
  end
end

Algorithm 3: AprioriPrune

input : L: large itemsets
T: transaction data
$minsup$: threshold of minimum support
output: L^*: large itemsets greather than a minimum support

$L^* \leftarrow \varnothing$
foreach $l \in L$ **do**
  $supportItemset \leftarrow countItemset(l \subset T_{i \in \{1,..,|T|)\}})/|T|$
  if $supportItemset >= minsup$ **then**
    $L^* \leftarrow L^* \cup l$
  end
end

Algorithm 4: AprioriGen

input : L: large itemsets
k: k-itemset
output: L^*: large k-itemsets

$L^* \leftarrow \varnothing$
foreach $l_p, l_q \in L$ **do**
  if $l_p[i] = l_q[i]$,$\forall i \in \{1, .., k-1\}$ **then**
    (k+1)-itemset creation from (k-1)-Itemset
    $l^* \leftarrow \{l_p[1], ...l_p[k-1], l_p[k], l_q[k]\}$
    $l^* \leftarrow sortItemset(l^*)$
    $L^* \leftarrow L^* \cup l^*$
  end
end

2.2.2 Cluster Analysis

Clustering is an unsupervised learning method. The goal of this method is to cluster data into groups that contain similar objects inside the groups and dissimilar objects between groups. The clustering approach allows the exploration of data to perform some analysis. This method is used in a large variety of applications such as computer vision, image processing, customer segmentation, education and so on (JAIN, 2010; DEGELE et al., 2018; KYLVAJA; KUMPULAINEN; KONU, 2019). There are several approaches for clustering, such as Partitional clustering, Hierarchical clustering, Density-based clustering, Grid-based clustering, Spectral Clustering, and others (WONG, 2015; PATEL; THAKRAL, 2016).

Given a set of data points $P = \{p_i\}$ where $i \in \{1, .., n\}$ and $n \in \mathbb{Z}$, $p_i \in \mathbb{R}^d$ where $d \in \mathbb{Z}$ that denote the dimension of the data point p_i and given a set of labels $L = \{l_j\}$ where $j \in \{1, .., m\}$ and $m, l_j \in \mathbb{Z}$, there are a function that maps each data point p_i to a label l_j. This can be summarized by the function $f \colon \mathbb{R}^d_n \mapsto \mathbb{Z}_m$. Thus, the clusterization method treat about the assignation of a label for each data point such that each group of data points with same label be homogeneous.

There are several clusterization algorithms, such as K-means, Mean Shift, Spectral clustering, Agglomerative clustering, DBSCAN, Gaussian Mixture, and others. The paragraph below describes a well-known and used clusterization algorithm, Mean Shift. This algorithm is used in the development of the present work.

Mean Shift (COMANICIU; MEER, 2002) is a partitional clustering algorithm. This nonparametric technique can determine clusters through exploiting the density in the data distribution without the prior knowledge of the number of clusters and does not assume any prior shape in the data distribution. Mean-shift requires a bandwidth (window) parameter that is used to estimate a scale kernel density. The kernel density allows moving data points to increase local density in each iteration until convergence is reached or the number of iteration is finished. Convergence is achieved when there is no significant shift in data points. Considering it is not trivial to choose an appropriate bandwidth parameter, there are alternatives to estimate this parameter in an adaptative mode, such as the sample point estimator (COMANICIU; RAMESH; MEER, 2001). The Mean Shift algorithm is detailed in Algorithm 5 and 6.

Algorithm 5: Mean Shift algorithm

input : P: data points
σ: bandwidth
θ: convergence threshold
n: number of iterations
output: P^* data points labeled

$T \leftarrow P$ (replication of P to simulate data point shift)
Initialization of seeds
repeat
  foreach $t \in T$ **do**
    $neighborsList \leftarrow$ getNeighbors(t)
    $t \leftarrow$ meanShift$(t, neighborsList, bandwidth)$
  end
until *($\Delta < \theta$ or $iteration > n$)(convergence or maximum iteration reached)*
$P* \leftarrow GetLabels(P, T)$

Algorithm 6: meanShift

input : p: data point
N: list of neighbors of data point p
$bandwidth$: scale density kernel
output: p^*: new data point location

$numerator \leftarrow 0$
$denominator \leftarrow 0$
foreach $neighbor \in N$ **do**
  $\mu \leftarrow \|p - neighbor\|_{L2}$
  $\sigma \leftarrow bandwidth$
  $numerator \leftarrow numerator + \mathcal{N}(\mu, \sigma^2) \times neighbor$
  $denominator \leftarrow denominator + \mathcal{N}(\mu, \sigma^2)$
end
$p^* \leftarrow numerator/denominator$

2.3 Data Visualization

Human brains are specialized much more for process visual information compared to touch or hearing information (KLEINSCHMIDT; HANRIEDER, 1992). In this context, data visualization is a research field that aims to represent data in a visual form (WARD; GRINSTEIN; KEIM, 2010). The visual representation turns able humans to understand the information easily. There are some important challenges in data visualizations, such as the visual representation of large or complex data. Data visualization are well applied in several applications such as transport (MONSIVAIS et al., 2018), health (LIU et al., 2016), finance (PERDANA; ROBB; ROHDE, 2019), education

(WILLIAMSON, 2016), identification (YANG et al., 2019) and so on.

There is a variate data type that can be represented visually. These include spatial data, geospatial data, time-oriented data, multivariate data, trees data structures, graphs, and network data structures, and text and document data (WARD; GRINSTEIN; KEIM, 2010). In data visualization, there are several techniques for representing data according to the data type, such as stacked display, dense pixel display, iconic display, geometrically-transformed display, and standard 2D/3D display as well as interaction techniques such as projection, filtering, zoom, distortion, pan, brush and so on (WARD; GRINSTEIN; KEIM, 2010). In the following subsections, we describe some relevant techniques in data visualization used in this work.

2.3.1 Visualization Techniques for Multivariate Data

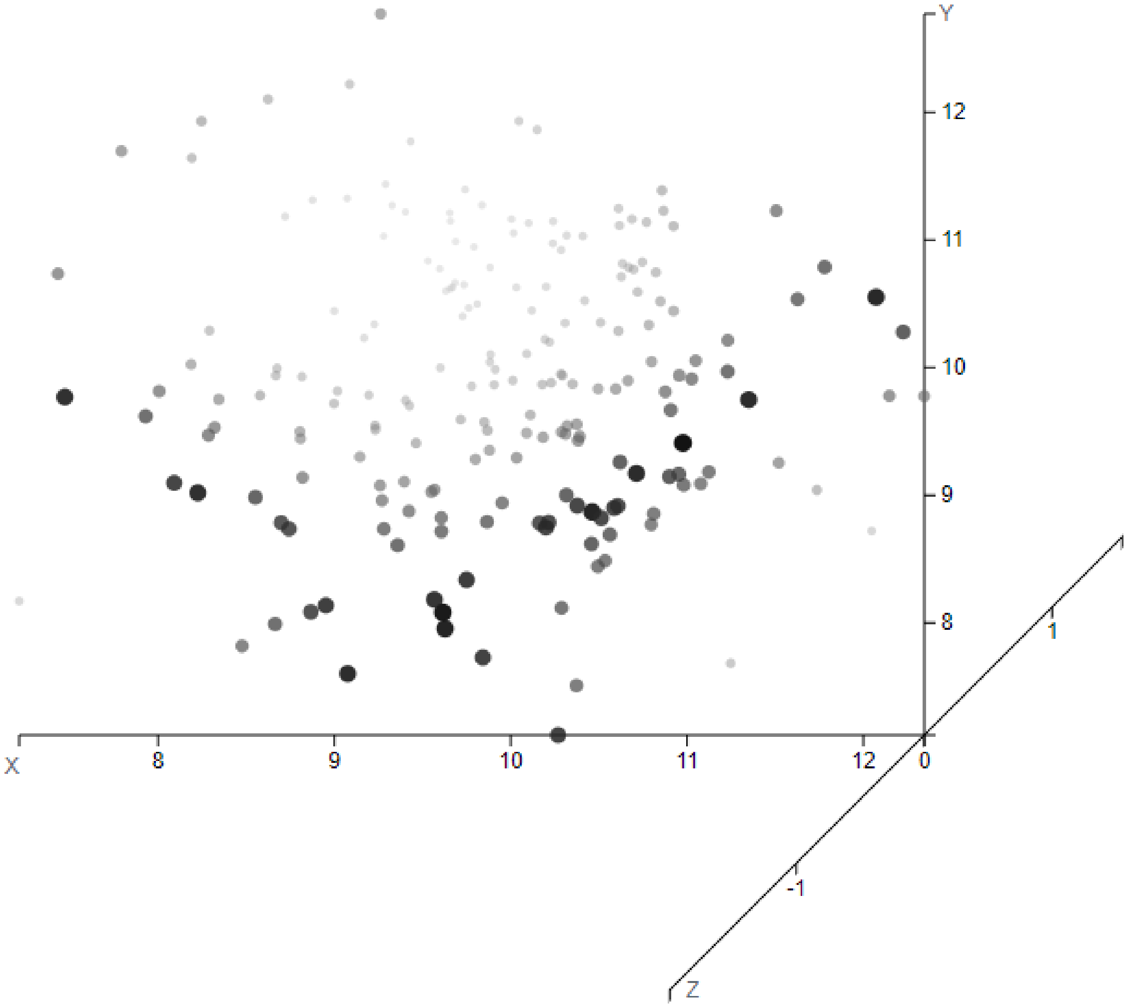
Multivariate data has one or more dimensions, i.e., 1D, 2D, or ND. To represent this data exists several techniques, such as point-based techniques, line-based techniques, and region-based techniques. Below, we briefly described these techniques.

2.3.1.1 Point-based Techniques

The point-based techniques represent n-dimensional data to a p-dimensional visual space using a dot or a mark to represent each data values. Several popular methods represent n-dimensional data in a visual form, such as scatterplots or forced-based methods.

The scatterplot is a widely used and early visualization method for 2D and 3D data (PIRINGER; KOSARA; HAUSER, 2004). This visualization allows use data attributes to encode such as color, size, position, shapes as well as interactivity such as zoom, brush, and pan. Figure 2.1 shows the 2D and 3D Visualization method.

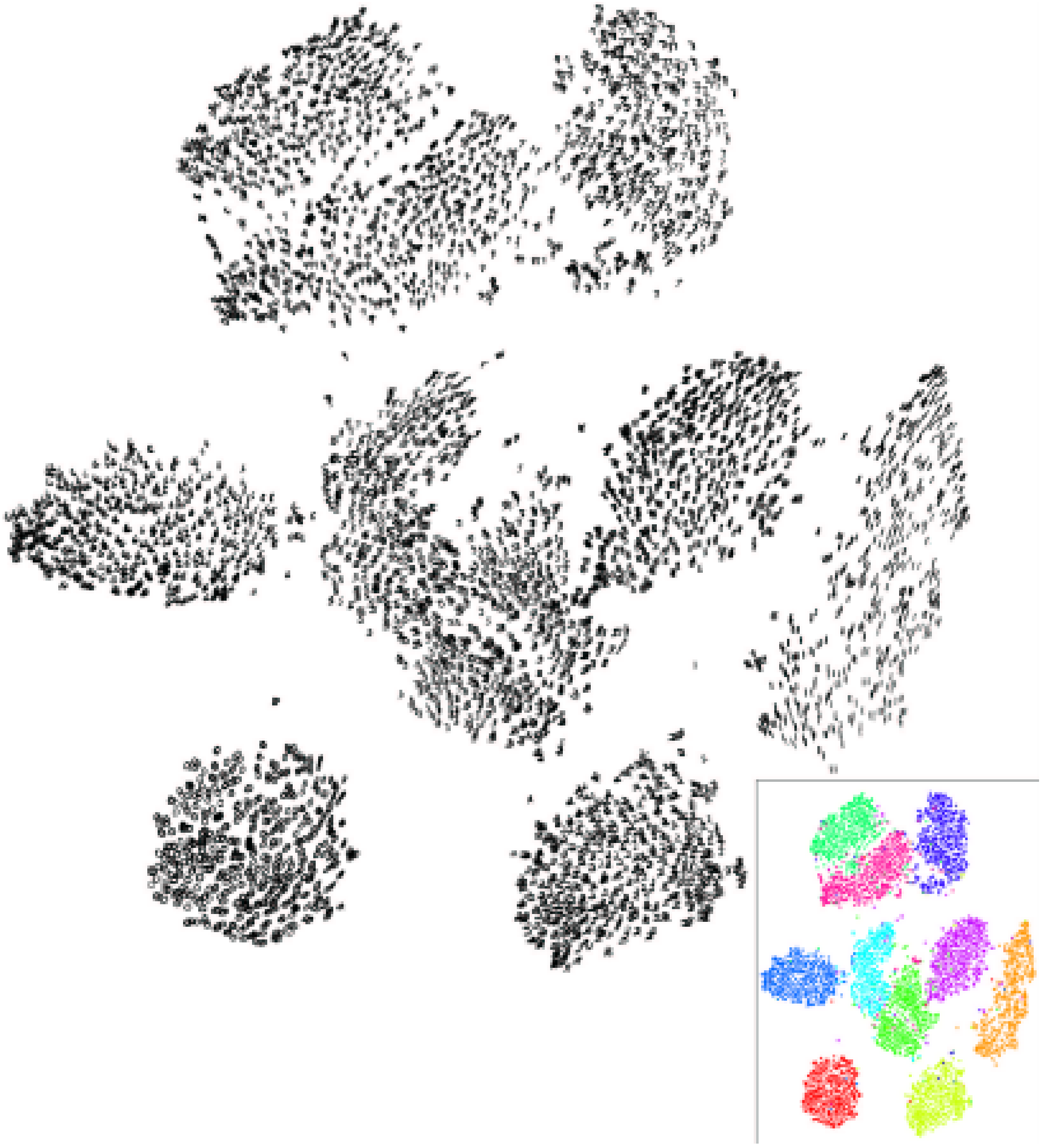
Figure 2.1: 2D and 3D scatterplot visualization



Implemented using the D3 library (BOSTOCK; OGIEVETSKY; HEER, 2011)

Forced-based visualization methods aim to represent N-dimensional data via a projection of data into a 2D or 3D visual space. Relevant methods to project data are MSD (Multidimensional scaling) (COX; COX, 2008), principal components analysis (PCA) (JOLLIFFE; CADIMA, 2016), t-distributed stochastic neighbor embedding (T-SNE) (MAATEN; HINTON, 2008). Figure 2.2 shows the Visualization of the MNIST database (LECUN, 1998) using the T-SNE projection method.

Figure 2.2: Visual projection of the MNIST database using the T-SNE projection method



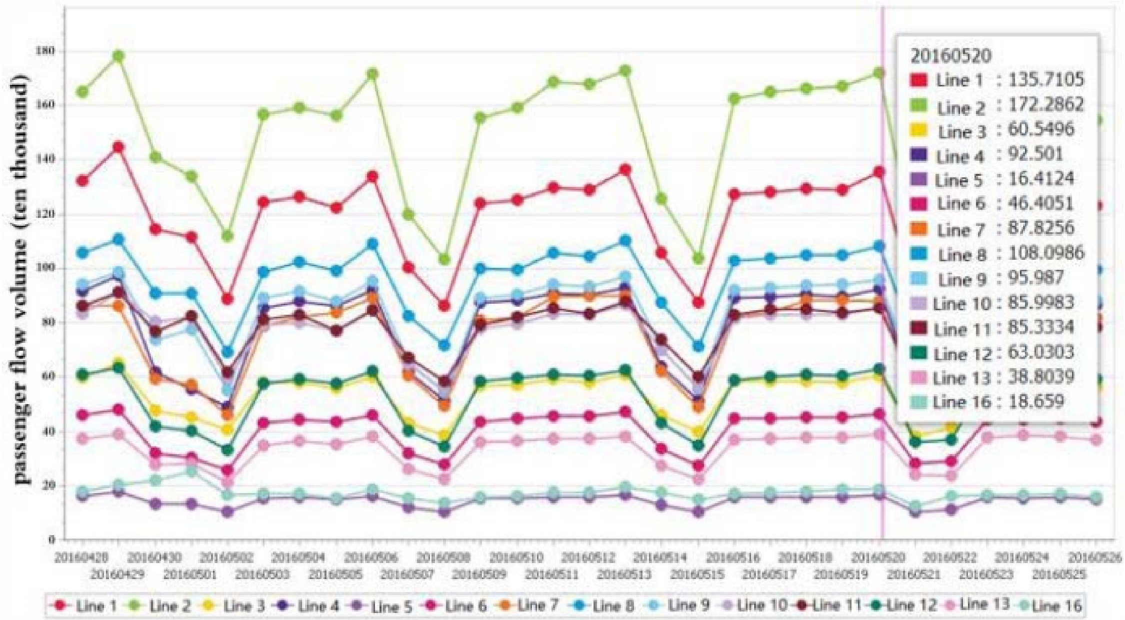
Source: (MAATEN; HINTON, 2008)

2.3.1.2 Line-based Techniques

Contrary to point-based techniques, the line-based techniques represent data values, maintaining a relationship between them in each variable. To represent the relationship is commonly use line patterns such as slopes, curvature, crossings, and others. Some representatives line-based techniques are the line graph and the parallel coordinates.

The line plot is a univariate visualization technique. The y-axis represents the range of values in the data, and the x-axis represents the data in an orderly manner. Also, multivariate data can be represented using the univariate visualization technique where can be overlapping between them. Figure 2.3 shows the visual representation of the daily flow of passengers in a metro network.

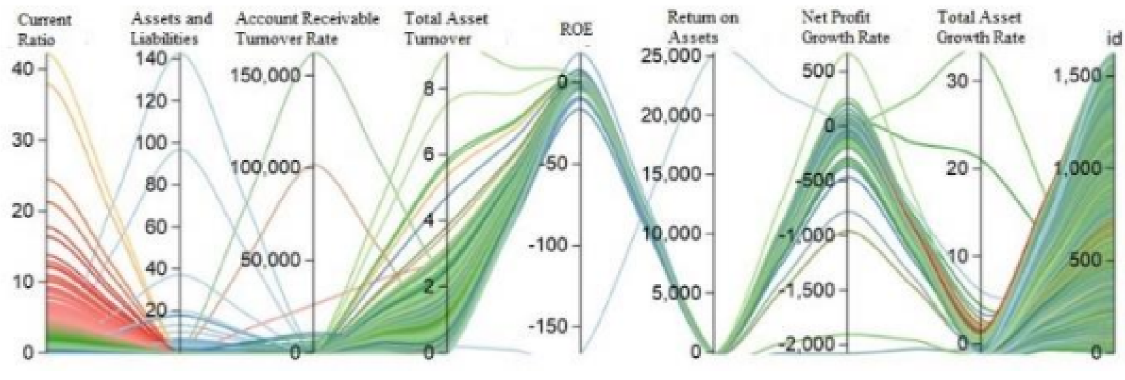
Figure 2.3: Daily passenger flow visualization using line graph



Source: (ZHIYUAN et al., 2017)

The parallel coordinates is a well-used method for representing multivariate data where multiple spaced horizontal parallel axes are used to represent the multi-dimension of the data. This representation allows the visual relationship between variables. Figure 2.4 shows the visual representation of financial data from companies using the parallel coordinates visualization method.

Figure 2.4: Visualization of financial data from companies using parallel coordinate visualization method



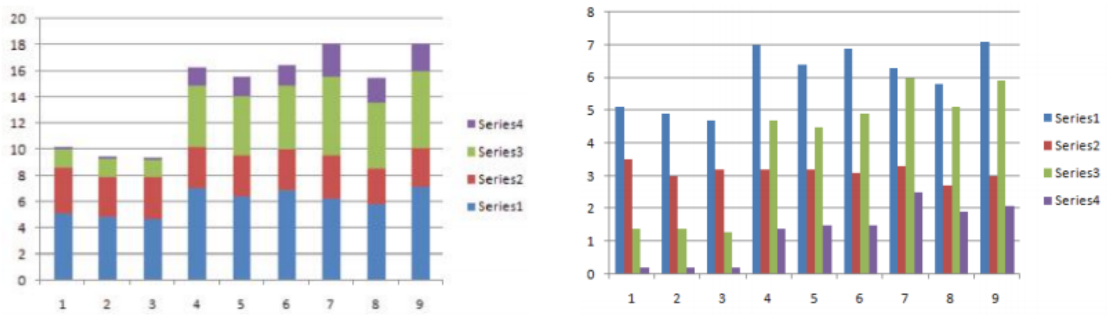
Source: (LI et al., 2019)

2.3.1.3 Region-based Techniques

In region-based techniques, the polygon represents data attributes and values such as shape, color, size, and others. These techniques are used mainly to represent summarization data in a visual form. Some representative methods are the bar chart and tabular display.

Bar chart visualization is a widely used visualization method that allows visualizations of different data types in 2D and 3D. Each axis is easily interpretable. To represent a multivariate data can be used clustering all data dimensions or stacked the data dimensions. Figure 2.5 (a) shows data represented using the stacked bar chart and Figure 2.5 (b) represents data using the clustered bar chart.

Figure 2.5: Bar chart visualization in two forms stacked bar chart and clustered bar chart.



(a) Stacked bar chart. (b) Clustered bar chart.

Source: (WARD; GRINSTEIN; KEIM, 2010)

The tabular display is a visualization technique that allows representing multivariate data. This technique allows several user interactions with data such as filters, searching, sortings. A representative visualization method is a heatmap. Heatmap represents table data values encoded as colors. Figure 2.6 shows a financial data representation using heatmap visualization..

Figure 2.6: Heatmap visualization representing tabular financial data.

Source: (ARBATLI; JOHANSEN, 2017)

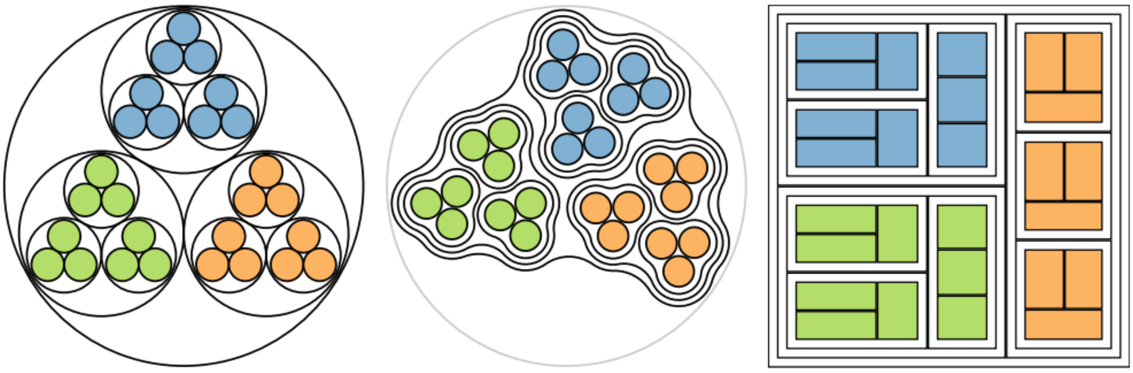
2.3.2 Visualization Techniques for Graphs

The graph visualization techniques, instead of only focuses on representing visual data, also focuses on the relationship between data attributes and values. This relationship can be in several forms, such as hierarchical, connectedness, sequence, similarities temporal or spatial, and others (WARD; GRINSTEIN; KEIM, 2010). This visualization technique is very used in many applications, such as networks (MCGEE et al., 2019) or transportation (ANDRIENKO et al., 2017). Below, we described the hierarchical technique that is used in this work.

The tree is a hierarchical structure that represents relationships. There are two types of methods that apply to hierarchical structure data: space-filling and non-space-filling. In the following paragraphs, it is described as the two techniques.

The space-filling method is a method to represents hierarchical visualization maximizing visual space. Representative visualization methods are treemap and radial visualization. Treemap can use several layouts to encode information in a form that is easy to explore, such as the Venn diagram or rectangular layout. Figure 2.7 shows examples of treemap visualization layouts. The radial visualization is a common visualization used to represent the correlation between data attributes in a hierarchical form. Contrary to the treemap, radial visualization is displayed at a rings levels. Figure 2.8 shows the traffic accidents distributions by hours using a radial visualization.

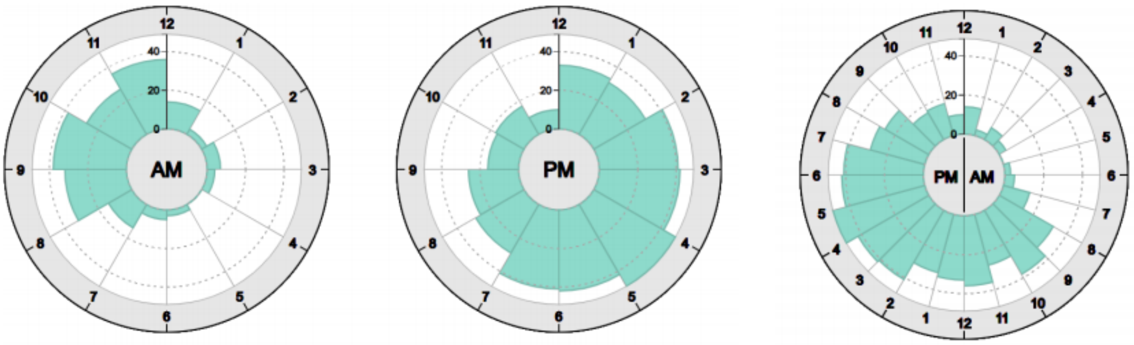
Figure 2.7: Different Treemap visualization layouts.



(a) Circular Treemap (b) Bubble Treemap (c) Nested Treemap

Source: (GÖRTLER et al., 2017)

Figure 2.8: Representation of traffic accidents by hours using radial visualization.

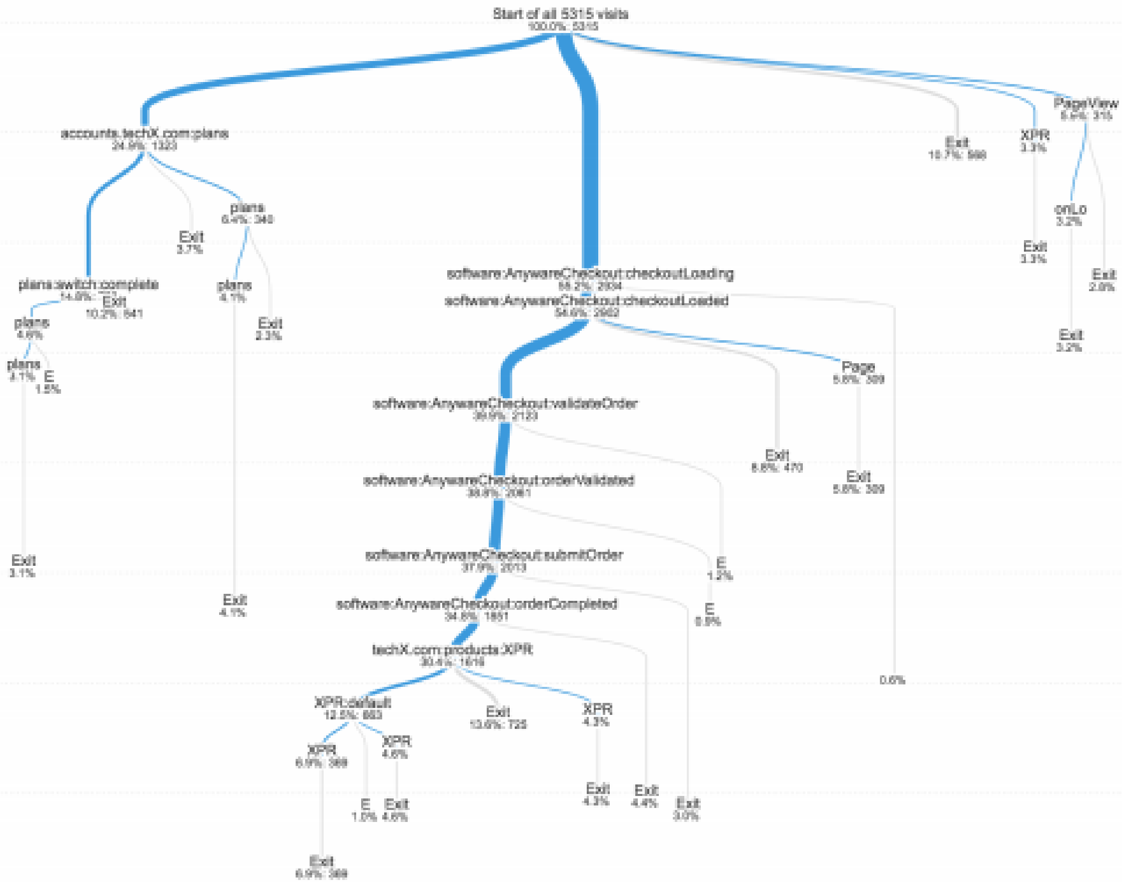


(a) 12-hours radial (12r) (b) 24-hours radial (24r)

Source: (WALDNER et al., 2019)

The most representative non-space-filling method is the node-link diagram. The node-link diagram visualization use nodes to represent entity/object/thing and links to denote the connection between them. Two relevant properties in the node-link diagram visualizations are the degree (number of connections in a node) and depth (level of hierarchical representation). It is important to take into consideration the space between nodes at levels to avoid overlapping. figure 2.9 shows the node-link diagram visualization.

Figure 2.9: Node-link diagram visualization.



Source: (LIU et al., 2017)

3 RELATED WORK

In this chapter, we describe representative works of the literature related to this dissertation. These works are organized into three sections. The first section presents recommendation systems in the insurance domain. The second section gives an overview of interactive recommender systems, and finally, in section three, we detail works related to visualization of recommendations.

3.1 Recommender System in the Insurance Domain

In order to obtain an active Customer Relationship Management (CRM), insurance and brokerage firms have to take advantage of the data generated by the users. Offers the most relevant products for users can improve loyalty and profit (VERHOEF; DONKERS, 2001; ZHANG et al., 2011). In past years, several recommender systems are proposed in the insurance domain. These systems propose a variety of methods to recommend insurance products according to the data and case study.

Case-based reasoning is a method that aims to solve a problem based on past solutions from similar problems. In this line, Rahman et al. (RAHMAN; NORMAN; SOON, 2006) develop a web insurance policies recommender system via a case-based reasoning algorithm where the recommendation is performed based on the similarity of user characteristics and past recommendation solutions as well as expert rules. The method is composed of four components: retrieve, reuse, revise, and retain. The retrieve component allows retrieving past recommendations for a user profile based on similar users profile stored in the case library. If similarity overtakes a threshold, then it retrieves and reuses that recommendations. Otherwise, the input user profile is revised by an insurance expert that determines the new case; then, it retains via the store in the case library. This approach is intuitive and straightforward but robust, which takes advantage of previous solutions and user profiles similarities.

On the other hand, Kumar et al. (KUMAR; SINGH, 2011) propose an insurance life recommender system. In combination with an analytics hierarchy process, data mining methods compound the dual-stage for the product recommendation. They enhance the recommendations taking into account the lifetime value of customers. The lifetime value of customers is given by age and income. Then, the K-means clustering method is applied to segment the customers based on their lifetime value. An analytics hierarchy process, i.e., a multicriteria decision-making method, determines each cluster the best product alternatives for a user. This is performed via the multicriteria flow designed by insurance experts.

Several insurance recommender systems are focused on the cross-selling strategy, i.e., recommend insurance products at the same level for the customer, e.g., life insurance, car insurance, and others. However, in different scenarios, it is used the up-selling strategy, where it is recommended complementary products that add particular benefits to the main product. In the insurance domain, the complementary extension of an insurance product is called an insurance rider. Rokach et al. (ROKACH et al., 2013) introduce an insurance rider recommendation system based on an item-item collaborative filtering approach. This work aims to help inexperienced agents from call-center to provide suitable recommendation rides for a policyholder. The method used in this work is a conditional probability of purchase an item given an item purchased. With this simple approach, they improve performance compare with other methods. Besides, call-center used the recommender systems for several months, and results reported an increment in sales converted in a factor of 3 compared to the traditional recommendations.

Insurance products are policies that involve a set of terms related to requirements for obtaining the policy and benefits that offer if the policy is contracted. Razak et al. (RAZAK; TAN; LIM, 2014) propose an insurance product recommendation system based on expert rules and fuzzy rule generation method. The fuzzy rule generation is put in the form of expert rules. Then, the fuzzy rules and the expert rules are merged to create a rule catalog used by insurance advisors to recommend insurance products for their clients. They use the Wang and Mendel algorithm for generating the fuzzy rules given the rules base generated by insurance experts.

Unlike previous works, Xu et al. (XU et al., 2014) propose a vehicle insurance recommender system based on association rule mining and customer segmentation. In this work, the customers are segmented based on the value of their profit. It is performed to turn the product recommendation more personalized. In each cluster, the transaction data is mining for obtaining association rules that are used as an indicator of recommendation. As a result, they show an increase in the performance of mining the transaction data in each cluster instead of mining in the whole transaction data.

In the real world, the data has a dirty nature. Qazi et al. (QAZI et al., 2017) present an insurance product recommender system based on the Bayesian Networks approach using customer information and portfolio data. In this work, they create multi models for specific groups of customers to offer a most personalized recommendation. They show the Bayesian Network method deal well with missing values because the variable has multiple conditional dependencies where each dependency contributes partially. An online validation shows an increase in the conversion of insurance policies.

Hinduja et al. (HINDUJA; PANDEY, 2017) propose a life insurance recommender system based on user preferences and utility theory. An Intuitionistic Fuzzy Sets method uses the user preference to estimate the importance of insurance features. Multicriteria constraints allow filter products that satisfy the user demographic information. By default, insurance experts assign values to policy features between 0 to 100, where 0 means lack of policy feature, and 100 represents the presence of policy feature in its maximum benefit. Using a Grey Relational Analysis is estimated the degree of utility of insurance policies filtered that is calculated using the importance of policy features and the default values of policy features. Finally, the insurance policies with maximum utility are recommended for the user.

Data mining methods are widely used in recommender systems to alleviate drawbacks or improve recommendation systems approaches. Kaewkiriya et al. (KAEWKIRIYA, 2017) propose a life insurance recommender system based on data mining methods. Firstly the data is prepared and preprocessed. The most relevant features for data analysis are selected using a decision tree algorithm. After identifying the most relevant customer features, the customers are segmented using the k-means algorithm. Finally, it is created a neural network model that is used to recommend insurance products. They obtain an increment in performance using the combination of data mining methods compared to use them independently.

Lesage et al. (LESAGE et al., 2020) propose a car insurance recommender system. They combine the XGBoost algorithm and apriori algorithm to determine the right customer to recommend and which insurance covers recommend. This approach improves the up-selling campaigns. To obtain the recommendation list of items, they follow a sequence of steps. Firstly, several datasets from different sources are integrated into a unified dataset. These datasets contain information about the client, the insurance car characteristics, premium, coverage, income, claims rated. Then, it is performed feature engineering in the unified dataset to build relevant features validated by experts. Consequently, it uses the XGBoost algorithm to determine the potential customers to acquire riders, and the apriori algorithm is used to determine what rider is most likely to be accepted. Before obtaining the final recommendation list, multicriteria filtering is applied based on expert business rules.

Qazi et al. (QAZI et al., 2020) propose an insurance coverage recommender system based on a Bayesian Networks method. Customer characteristics and customer portfolio data are used. Multicriteria-base filtering based on business rules is used before obtain the final recommendation list of products. Additionally, they propose a deep learning model to perform recommendations to potential customers based on external data.

The table 3.1 shows some comparison of the most relevant insurance recommendation systems in terms of the target user, method, sales, and interactivity functionality.

Table 3.1: Recommender system works for insurance domain.

Work	User	Method	Sales	Interactive
Rahman et al. (2006)	Client	Case-based reasoning	Cross selling	Yes
Kumar et al. (2011)	Agent	Multicriteria	Cross selling	No
Rokach et al. (2013)	Agent	Collaborative-based filtering	Up selling	No
Razak et al. (2014)	Agent	Fuzzy rules generation	Cross selling	No
Xu et al. (2014)	Agent	Association rules mining	Cross selling	No
Qazi et al. (2017)	Agent	Bayesian networks	Cross and up selling	No
Hinduja et al. (2017)	Agent	Utility theory and Multicriteria	Cross selling	No
Kaewkiriya et al. (2017)	Agent	Neural network	Cross selling	No
Lesage et al. (2020)	Agent	XGBoost algorithm and Association rules mining	up selling	No

Continued on next page

Table 3.1 – *Continued from previous page*

Work	User	Method	Sales	Interactive
Qazi et al. (2020)	Agent	Bayesian networks and Deep learning	cross and up selling	No
Broker-RecSys	Agent/Broker	Association rules	Cross selling	Yes

3.2 Persuasive and Interactive Recommender System

Recommender systems aim to recommend the best services or products to users based on their preferences, characteristics, or tastes. However, besides the accuracy, other factors such as persuasion and user interaction play a vital role in the recommendation task. These factors can increase the user acceptance of services/products as well as user understanding of the recommendation process. In the following subsections, we describe the persuasive and interactive recommender systems.

3.2.1 Persuasive Recommender System

Recommender systems aim to recommend services/products to users. In contrast, persuasive recommender systems aim to persuade users to get some services/products using several strategies based on human-human communication theories. Moreover, try to answer the question of how to integrate persuasiveness into recommender systems? (YOO; GRETZEL; ZANKER, 2012).

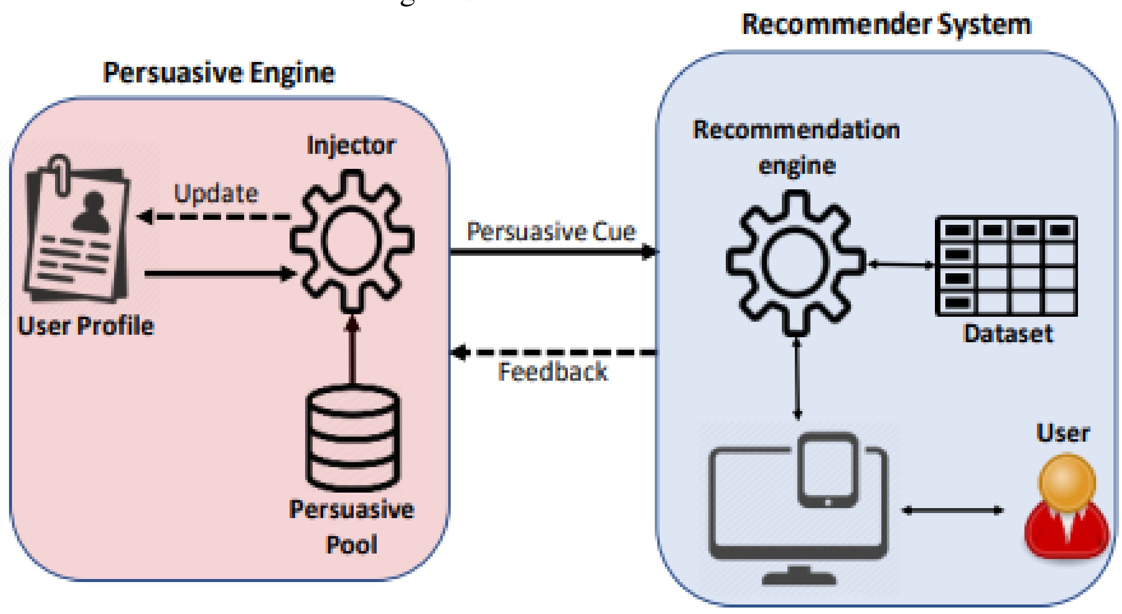
In recent years, several recommender systems have proposed works that integrate persuasion strategies such as informative messages, transparency, presentation of recommendations as well as justification, and explanation (GKIKA; LEKAKOS, 2014; SOFIA et al., 2016).

Heras et al. (HERAS et al., 2017) propose a persuasive educational recommender system based on the argumentation strategy. The argumentation strategy is used to persuade students to use certain learning objects based on their preferences and needs. The learning objects metadata and the user profiles are used in the argumentation strategy.

The persuasive educational recommender system uses the argumentation as an explanation that is transferred to students in different ways. They can be content-based (based on past information of learning objects that the student already rated), knowledge-based (based on the past information that students profile already rated learning objects), and collaborative (based on the similarity of similar students). In this work, they evaluated how an argumentation strategy as an explanation form can persuade students into the acceptance of learning objects.

Alslaity et al. (ALSLAITY; TRAN, 2019) propose PerPer, a personalized persuasive recommender system framework based on the learning automates concepts to choose the best persuasive strategy to apply for a specific user. PerPer framework is composed of two parts: A persuasive engine and a recommendation system. Figure 3.1 shows the PerPer framework. The persuasive engine takes as input the feedback and user profile and based on this, choose the best persuasive strategy to apply in the target user. On the other hand, the recommender system component recommends products using the persuasive strategy chosen by the persuasive component. The strategy can vary from textual cues based on keywords, symbols, visualizations, and human-agents entities.

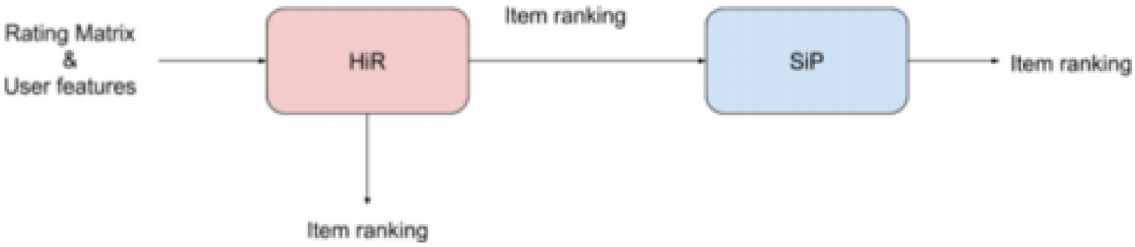
Figure 3.1: PerPer framework



Source: (ALSLAITY; TRAN, 2019)

Sánchez et al. (SÁNCHEZ-CORCUERA et al., 2020) propose a persuasive recommender system to determine the best persuasive strategy for a user profile so that the user uses shared spaces to save electricity. They propose integrating two recommender systems: Hybrid item Ranker(HiR) and Specific impact Predictor (SiP). HiR provides a ranking of items based on the specific user profile and the rating matrix of items. In order to make that, a principal component analysis is used to represent user profiles visually. So, users near in space mean that they may have similar item preferences. A score based on users near in space is multiplying to the ranking of items to determine the ranking of items with the best persuasive strategy. Consequently, SiP receives from HiR as inputs only items that represent the best persuasive strategy for a user profile. Using an active learning approach, SiP predicts the user's rating respect for the items given.

Figure 3.2: Architecture of the persuasive recommender system.



Source: (SÁNCHEZ-CORCUERA et al., 2020)

3.2.2 Interactive Recommender System

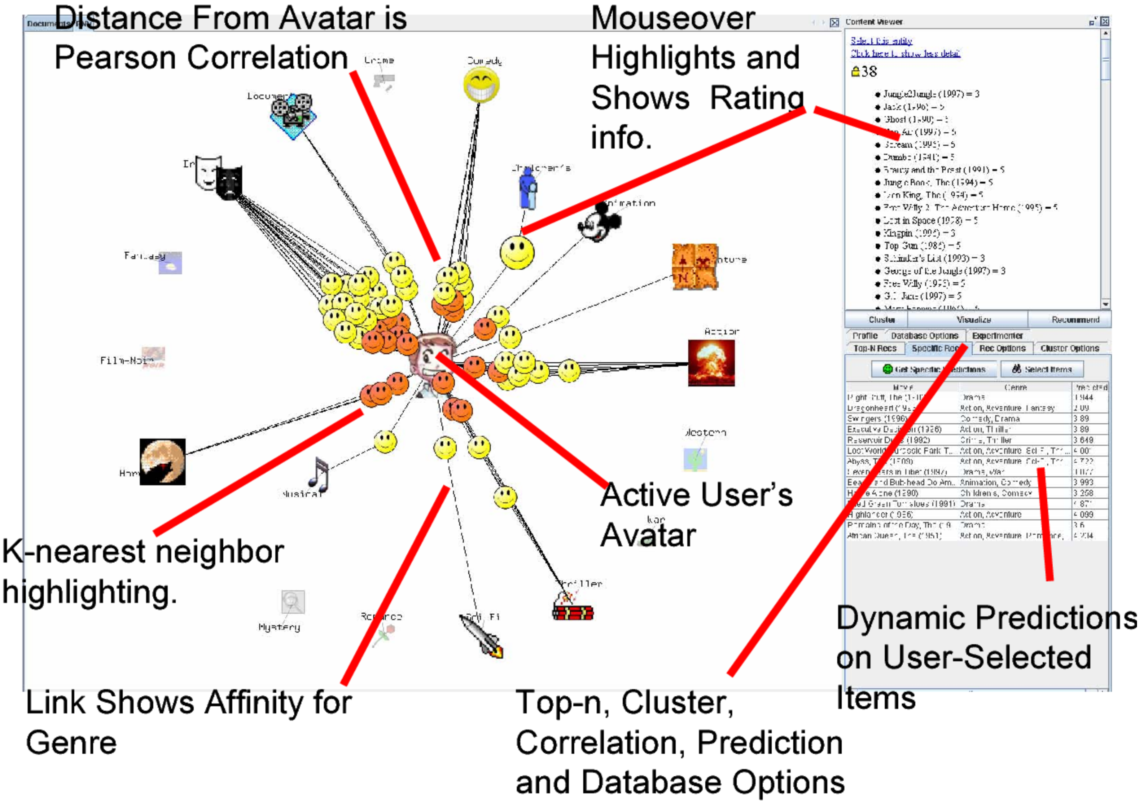
The black-box nature of the recommender systems limits the understanding and acceptance of the recommendations suggested for the users. The interactive recommender systems address the natural questions of the users such as *why reason some recommendation is offered?*, *how the recommendation system obtained the recommendations?* or *is it possible to intervene in the recommendation process?* and so on. In order to address these questions, some studies were performed. He et al. (HE; PARRA; VERBERT, 2016) propose a framework that involves recommendation and visualization components to solve several drawbacks in recommender systems such as transparency (explanation of the internal functionability of the recommender system for users), justification (justify why the recommendation is get by the user), controllability (the user take part in the recommendation process), cold-start problem (alleviating through algorithmics and interactive visualization methods and techniques), and diversity (enhancement of potential products with are not included in the user preferences). On the other hand, Valdez et al. (VALDEZ; ZIEFLE; VERBERT, 2016) analyzed several works that consider human factors and visualization methods that improve user experience and acceptance. Another critical factor is interaction with recommender systems. In this direction, Jugovac et al. (JUGOVAC; JANNACH, 2017) present an overview of the interaction aspects involved in the recommender systems consideration the presentation and interaction in the recommendation process. Interactive recommender systems intersect multiple areas such as the UX/UI design, recommendation system, data visualization, and machine learning to enable users to answer their natural questions, as mentioned above. Some of the most representatives interactive recommender systems that support the present dissertation work are mentioned in the following paragraphs.

O'Donovan et al. (O'DONOVAN et al., 2008) propose an interactive recommender system based on a collaborative filtering approach. This work provides a visual explanation of the collaborative filtering process. Given an active user, it is possible to manipulate the number of user neighbors to recommend items for the active user. The Euclidean distance gives the correlation that exists between the neighbors and the active user that is represented visually. The active user receives recommendations based on the profile of his k-nearest neighbors. This recommender system deals with the cold-start problem because the recommendation is performed based on the k-nearest neighbors, i.e., there are many options to recommend for the active user. The user involvement in the recommendation process helps the user to understand how the recommendations are obtained, and the visualization interface improves the recommendation acceptance. Figure 3.3 shows the visual interface of the PeerChooser recommender system.

Symeonidis et al. propose MoviExplain, a recommender system that integrates explanation. MovieExplain is a hybrid recommender system that combines the content-based filtering and collaborative-based filtering method. Information about the item is used to create an item profile while the user's rating is used to create a user profile. MovieExplain takes as input the user and item profile and obtains recommendations based on the similarity with other items and user profiles. MoviExplain aims to justify the recommendation for increments the recommendation acceptance. Figure 3.4 shows the MoviExplain interface.

Zhao et al. (ZHAO et al., 2010) propose Pharos, a social map-based recommender system. This recommender system aims to help users to gain insights about the contents of communities. In social websites, communities are described by topics and people related to those topics. Pharos allows users to discover interesting communities and topics interactively. Besides, users obtain a content list, and people list recommendations based on their interests. Pharos addresses the cold-start problem because it is based on clusterization as well as the recommendation explanation because it provides an explanation about the community via the visualization that encodes topics and people related to the

Figure 3.3: PeerChooser's Interactive Interface - Visual explanation of the collaborative filtering process.



Source: (O'DONOVAN et al., 2008)

Figure 3.4: MoviExplain Interface - Justification of the recommendation of movies based on the movie and user profiles.

Source: (SYMEONIDIS; NANOPOULOS; MANOLOPOULOS, 2009)

community. The figure 3.5 shows the Pharos interface.

Jin et al. (JIN et al., 2016) propose Paris-Ad, an advertising recommender system. This recommender system allows users to control and understand the process of the

Figure 3.5: Pharos Interface - Communities described by topics and people related to the topics.

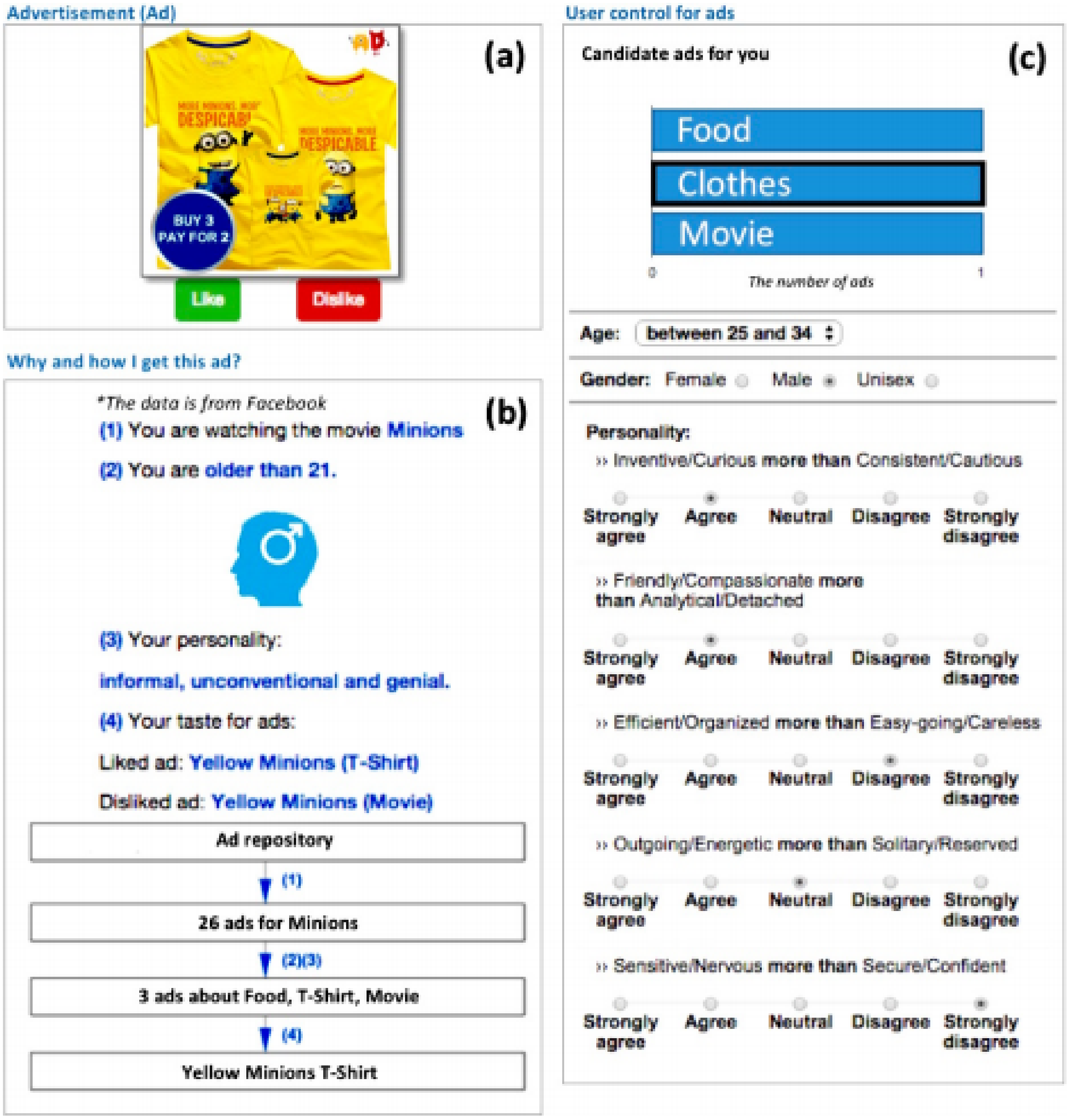
Source: (ZHAO et al., 2010)

recommendation of ads. User demographic information, personality information, and ad preferences from a user is used to create a user profile. When a user watches a trailer movie, Paris-Ad recommends the most relevant ads for the user using his user profile and the trailer movie information. Transparency feature is adopted, showing to the user the information used to obtain the recommendations as well as an explanation about the ads recommended. Besides, Paris-Ad allows users involvement in the recommendation process via the manipulation of user-profiles and ad categories. In Figure 3.6 is shown the Paris-Ad interface.

Aoike et al. (AOIKE et al., 2019) propose an interactive tour planning recommender system using crowd information of tourist spots in order to enhance the experience and satisfaction of tourists. Tourists interact with the recommender systems providing personal specifications such as tour conditions and tour characteristics. In the experiments, they obtained that in 70% of cases, the recommender systems provide alternative tourist plans in spots that are not crowded. Figure 3.7 shows the tour planning interface.

Yu et al. (YU; SHEN; JIN, 2020) propose a hands-free visual dialog interactive recommender system. User comments in a natural language the desired item features.

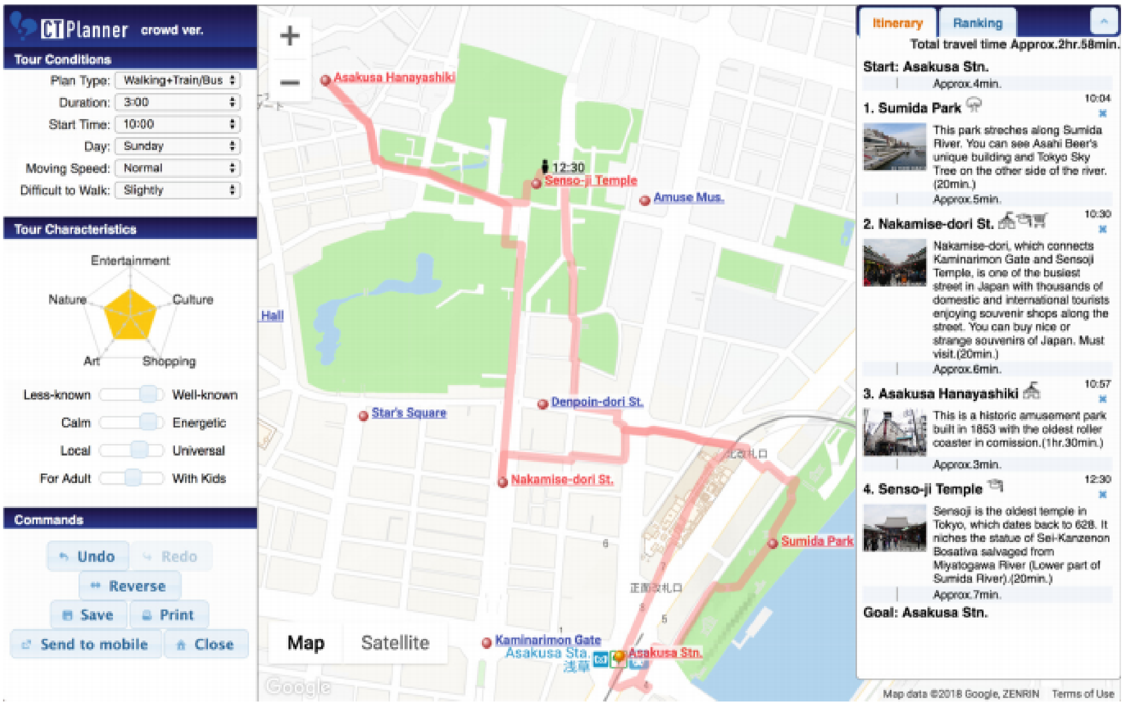
Figure 3.6: Paris-Ad Interface - Explanation of the ad recommendation based on user control and movie information.



Source: (JIN et al., 2016)

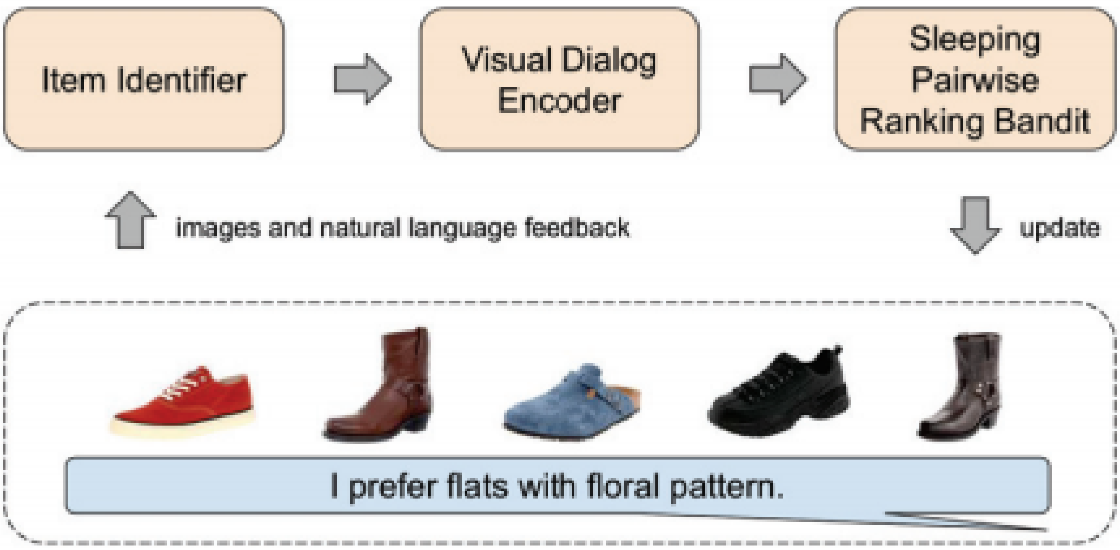
The system provides a list of visual items based on these item features chosen. A neural network model takes as input the user comment that described the desired item features and the list of images, and it identifies the items related to that comments using a bandit algorithm. This process is performed iteratively until the user finds the expected item. Figure 3.8 shows the components and process involved in the hands-free visual dialog interactive recommender system.

Figure 3.7: Tour planning interface.



Source: (AOIKE et al., 2019)

Figure 3.8: Process in the hands-free visual dialog interactive recommender system.



Source: (YU; SHEN; JIN, 2020)

3.3 Association Rules Visualization

The association rule algorithm suffers the overload of rule generation that difficult the rule identification and exploration. In past years, several works were proposed to turn easy the identification and exploration of association rules. Bruzzese et al. (BRUZZESE;

DAVINO, 2008) describe several popular approaches to visualize association rules such as tabular visualization, scatter plot, graph visualization, parallel coordinates, and others. Recently, Hahsler et al. (HAHSLER, 2017) provide aruleviz, aruleviz is an R library that contains the ten most relevant visualization of association rules.

On the other hand, other works propose alternatives to visualize association rules. Hahsler et al. (HAHSLER; KARPIENKO, 2017) propose an interactive visual exploration of association rules based on a grouped matrix representation. Using a clusterization method, the rules are grouped considering an interesting measure such as support, confidence, or lift. The matrix visualization is used to show the grouped rules, and it can be explored in detail. The antecedents and consequents of the rules are displayed in the columns and rows, respectively.

Mukherji et al. (MUKHERJI et al., 2018) propose an interactive rule exploration framework. This framework provides two interactive spaces called parameter and visualization space, respectively. The aims of the first interactive space, called parameter space, provide the visualization of rules in a general manner, this is generated by a dynamic set of parameters. On the other hand, the rules visualization space allows visualizing the rules in detail using a tabular view or a glyph visualization.

In order to obtain a general overview of the visualizations of association rules, a comparative table is built based on representative works of the literature (BRUZZESE; DAVINO, 2008; HAHSLER; CHELLUBOINA, 2011; HAHSLER, 2017; MUKHERJI et al., 2018). The Table 3.2 shows a comparison of the most representative visualization of association rules in terms of visualization technique used, size of rules supported, number of measures supported, interactive features and the level of usability of the visualization.

Table 3.2: Comparison of techniques for association rule visualization

Technique	**Rule set**	**Measures**	**Interactive features**	**Ease of use [--, ++]**
Scatterplot	Large	3	hover, zoom, pan, brush, grid line	++
Two-Key plot	Large	2+order	hover, zoom, pan	++
Matrix-based	Medium	1	hover, zoom, pan, reordering	0
Matrix-based (2 measures)	Medium	2	hover, zoom, pan, reordering	-

Continued on next page

Table 3.2 – *Continued from previous page*

Technique	**Rule set**	**Measures**	**Interactive features**	**Ease of use [--, ++]**
Matrix-based (3D bar)	Small	1	hover, reordering	+
Grouped matrix	Large	1	drill down, inspect	0
Graph-based	Small	2	hover, zoom, pan, brush	++
Graph-based (external)	Large	2	hover, zoom, pan, brush, reordering	+
Parallel coordinates	Small	1	reordering	-
Double-decker	Single rule	2		-
PSpace (Matrix)	Large	2	inspection, zoom, grid line, hover, drill-down	++
RSpace (Glyph)	Large	2	inspection, filtering, sorting	++
Tabular	Medium	1+	hover, filtering, searching, sorting, drill down, pagination	++

4 BROKER-RECSYS - AN INTERACTIVE RECOMMENDER SYSTEM FOR INSURANCE BROKERAGE

In this chapter, we describe Broker-RecSys, an interactive recommender system that supports brokers with the recommendation process of insurance products in their client portfolio. Section 4.1 describes the insurance brokerage domain flow and the functional requirements related to the insurance product recommendation task. Section 4.2 presents an overview of the Broker-RecSys framework and describes its main parts at a high level. Posteriorly, it is detailed each component from each Broker-RecSys part in section 4.3. Finally, section 4.4 presents the Broker-RecSys tool, resources involved in the development process, as well as a running example.

4.1 Recommender System Requirements

In the insurance brokerage domain, there are three main components: Insurer, Broker, and client. Figure 4.1 shows the Insurance brokerage domain flow.

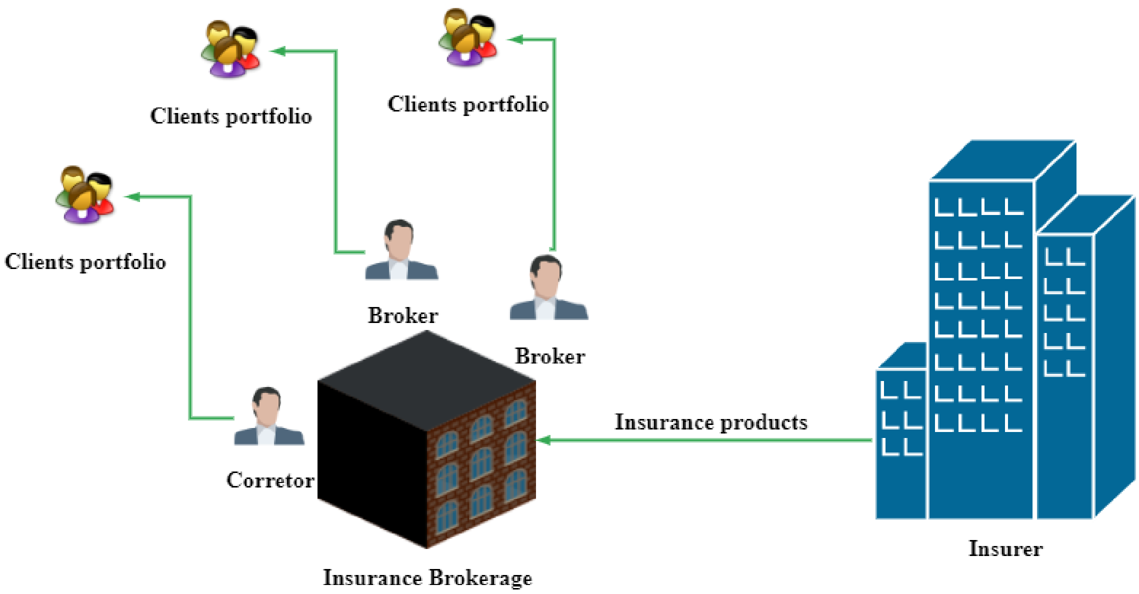
- **Client:** The client is somebody that wants or buys an insurance product to cover risk and be protected from a financial loss.
- **Insurer:** Insurers provide insurance products based on the stratification of people groups. Also, they offer insurance riders that are additional benefits that complement the insurance products.
- **Broker:** An insurance broker is a professional that offers, negotiates, and sells insurance products to clients for compensation.

4.1.1 Requirements into the recommendation process

In collaboration with an insurance broker expert, we determine the functional requirements needed in the recommendation process in the insurance brokerage domain. The insurance broker is an intermediate between the insurer and the clients. They offer insurance products, negotiate, and finally sell the insurance products for their clients. Also, brokers try to retain existing clients offering additional insurance products. To identify clients to offer additional insurance products, they evaluate the number of products that the client already purchased and specific client characteristics as an indicator that those clients are potential clients to offer more insurance products. For example, Gretel, a young girl that lives in the south of Brazil, has a car insurance product purchased, and she is friends of Roberta, a young traveler. They shared similar personal characteristics, such as age, gender, and region. Then, the broker can offer additional products that Roberta has to Gretel, such as insurance travel, and expect to succeed in the recommendation based on their experience i.e., similar people probably buy similar products. Below, we describe the functional requirements involved in the insurance products recommendation task.

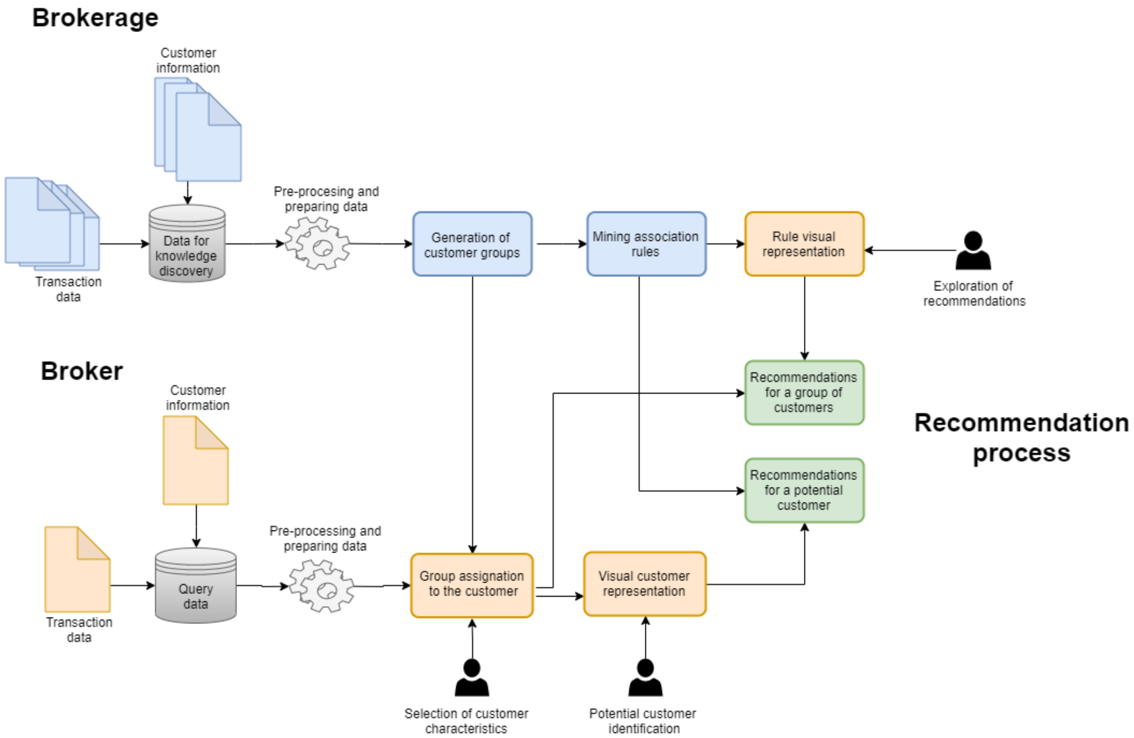
- Cluster clients based on specific client characteristics.
- Filter clients based on specific client characteristics.
- Filter clients based on the number of products purchased.
- Visual customer representation to allow an easy exploratory and identification of clients.
- Recommendation of products for a specific client.
- Recommendation of products for a group of clients i.e., perform marketing campaign.
- Allow the broker to take part in the recommendation process.
- Visual representation of recommendations.

Figure 4.1: Insurance brokerage domain flow in a high level.



4.2 System Overview

Figure 4.2: Broker-RecSys framework



Broker-RecSys is an interactive recommender system that aims to be a support system to help brokers into the recommendation of insurance products in their client portfolio. Using the information of the client portfolio of brokers in a brokerage, Broker-RecSys extracts useful information from the portfolios and help to recommend insurance products for each broker in his client portfolio. Broker-Recsys is a collaborative recommender system because it takes advantage of the isolated client portfolio of brokers, joins all this information in one place, and extracts useful knowledge from the data. After that, the discovered knowledge is consumed and used for all brokers to offers insurance products in their client portfolio.

To help insurance brokers in the recommendation process of insurance products, Broker-RecSys combine data mining methods with user interaction and visualization methods. The recommender system allows insurance brokers to recommend products for two types of target users: specific customer and customer group. Specific customers can receive two types of recommendations based on popularity and purchase behavior. Besides, Broker-RecSys provides a module to perform ultra-segmentation based on interesting characteristics chosen by the broker. In order to help brokers into the identification of potential customers or interesting groups of customers to offer insurance products, the system provides filters and visualizations that turn these tasks easy to perform. Broker-RecSys is composed of three parts: Brokerage, Broker, and Recommendation process. Figure 4.2 shows the propose Broker-RecSys framework. In the following, each part of the Broker-RecSys framework is detailed.

- **Brokerage part:** This part contains three modules: **Integration and preprocessing data**, **generation of customer groups** and **mining association rules**. In the first module **Integration and preprocessing data**, it is integrated the client portfolio of brokers and performed a preparing process of the data for the next module **generation of customer groups**. The module generation of customer groups undertakes to perform the ultra-segmentation of broker-data in their multiple dimensions, i.e., clusterization based on their client characteristics such as gender, age, and region. Finally, inside each customer group generated by the previous component, It is mining the transaction data to obtain the query models used in the recommendation process. The elements of the Brokerage part are colored by light blue color (see Figure 4.2).
- **Broker part:** This part contains four modules: **upload input data**, **group assignment to the customer**, **visual customer representation** and **rule visual representation**. The first module allows the broker to upload his client portfolio to the system, and automatically the portfolio is preparing for the next module **group assignment to the customer**. The **group assignment to the customer** module assigns a group label for each customer in the client portfolio based on the attributes chosen by the broker to group the customers. The third module **visual customer representation** allows the broker to realize the exploration and identification of potential clients. Finally, the **rule visual representation** component allows the broker to visualize the recommendations in differents views for exploring and identify interesting recommendations. The elements of the Broker part are colored with an orange color (see Figure 4.2).
- **Recommendation process part:** After executing the previous components of the brokerage part as well as the broker part, the broker can perform the recommendation process. The recommendation process provides to the broker two levels of recommendations and two types of recommendations. The recommendation level refers to that a broker can recommend insurance products for a specific customer or

a customer group. On the other hand, the types of recommendations indicate that the broker can recommend insurance products based on popularity and purchase behavior. The elements of the Recommendation process are colored by green color (see Figure 4.2).

4.3 Broker-RecSys Components

In this section is detailed each component from the Broker-RecSys framework. As explained above, the Broker-RecSys framework is composed of three parts: Brokerage, Broker, and Recommendation process. Each of them contains a sequence of components that inconjunct builds the interactive recommender system.

4.3.1 Preprocessing and Preparing Data

Data in the real world have a dirty nature. This means that the data can be incomplete (missing attribute values), noisy (it can contain wrong values and outliers values), and inconsistent (there is a disparity in the attribute values). Due that the quality of the data determines the quality of the knowledge results, our first component performs these tasks.

This component is responsible for preprocessing and preparing the data for knowledge discovery (brokerage data) and query data (broker portfolio). We applied the basic techniques for performing the preprocessing and preparing of data such as data cleaning, data integration, data transformation, and data reduction (GARCÍA; LUENGO; HERRERA, 2015).

In the data cleaning step, we fill missing values, remove outliers, and resolve inconsistencies in the attribute values. In the data integration step, we perform the integration of multiple datasets; in this case, client portfolios. This dataset are stored in different formats such as .xlsx, .csv or .txt and with diverse codification (e.g., utf-8, ANSI). The integration of client portfolios consists of combining the multiple client portfolios in a unified client portfolio in the same format file, format data, and codification. In the next step, data transformation uses this unified dataset. In the data transformation step, the data is transformed into a valid format used by data mining algorithms.

Finally, in the data reduction step, we select predefined features determined in concordance with an insurance broker expert based on the importance of insurance products' recommendation. Until this point, The data is pre-processed and prepared and already to feed the next components (see Table 4.1).

Table 4.1: Dataset format of the portfolio of clients used by the Broker-RecSys tool.

Attribute	**Description**
Birthday	Birthday of the customer e.g., 22/02/1850
Region	Region (0,..,9) of the customer extracted from postal code
Gender	Gender of the customer e.g., female, male, etc.
Product	Product purchase by the customer e.g., car, travel, etc
Name	Name of the customer
Id	Identifier of the customer

4.3.2 Customer Groups Generation

Customer segmentation matters. Customer segmentation groups customers into homogeneous sets based on their similar characteristics, and it allows the extraction of useful information. Insurance broker's requirements show that they desire flexibility to recommend products in their client portfolio based on different customer characteristics such as based gender or age and region, and so on. The cluster control requirement enables a more personalized recommendation because it considers only customers with strong similarities, and consequently, the recommendations should be more personalized. Thus, for satisfying this requirement, a module called "customer groups generation" is integrated into Broker-RecSys. This module allows Broker-RecSys to generate all possible groups of clients based on their demographic information combination. Thus, each customer characteristics combination will allow mining the data inside that group to generate recommendations (query models).

More technically, consider n as the number of customer characteristics, then, it generates $2^n - 1$ customer characteristic combinations. Consequently, each customer characteristic combination generates a set of customer groups. In each of these customer groups, the data is mining to create the query models. For example, given the number of features $n = 2$ and the feature vector $F = \{f_1, f_2\}$ where f_1 and f_2 are the customer characteristics. Then the total characteristics combinations in F will be $C = \{\{f_1\}, \{f_2\}, \{f_1, f_2\}\}$. Consequently, it will generate several groups based on each element of the vector C. For example, if we take the first element $\{f_1\}$, it means that the clusterization will be performed based on the feature f_1 or if we take the third element $\{f_1, f_2\}$ that indicate that the clusterization will be performed based on the features f_1 and f_2. Each one of the elements of the feature vector C will generate some groups according to the clusterization algorithm used. For cluster customers, we employed the mean-shift algorithm (COMANICIU; MEER, 2002) due to its capacity to determine clusters through exploiting the density in the data distribution, as explained in subsection 2.2.2. The bandwidth parameter value for the mean-shift algorithm is calculated using a sample point estimator method. The bandwidth value was obtained as the mean of the maximum pairwise distances for each data point in the data distribution with their n neighbors where n represents the 30$\%$ of the data points.

4.3.3 Groups Assignment for Customers

For each customer in the query data (customer portfolio), a label is assigned based on the segments generated by the **customer groups generation** module. For an input customer portfolio vector $CP = \{cp_1, cp_2, .., cp_l\}$ where l is the number of customers and $cp_i = \{fc_1, fc_2, .., fc_p\}$ represent the customer characteristic vector i where $i \in \{1, l\}$ and p represent the number of customer characteristics, This component assign one of the desired segment created by the **customer groups generation** component to each customer cp_i. The assignation occurs considering the customer characteristics selected by the broker in the segmentation of its customer portfolio. For example, considering $p = 2$ then $cp_i = \{fc_1, fc_2\}$ and the customer characteristics selected by the broker was fc_1, then, to assign one segment to the customer i, the **group assigment for client** component recovered the centroids of the segments generated by the customer characteristics combination $\{fc_1\}$ by the "customer groups generation" component. Then, it computes the distance between the cp_i considering $\{fc_1\}$ to the centroid to each segment recovered, and it is assigned to the customer the label of the segment where the distance is minimum. After that, all the customers in the portfolio will obtain a label of segment assigned and can get recommendations based on the customer group to they belong.

4.3.4 Visual Representation of Customers

One goal in this work is to facilitate the exploration and identification of potential customers to offer insurance products in the customer level recommendation. Thus, we consider the number of products purchased by the customer as an indicator to identify potential customers as well as their attributes to allow filtering. The clients that have fewer products bought turn a potential customer to offer additional insurance products; otherwise, the customer does not need much attention.

An expert insurance broker evaluates several visualizations to choose one that best accomplishes the requirements to perform the insurance product recommendation task. Figure 4.3 shows the validation flow of the visual representation of customers. Additionally, Table 4.2 details the visualization type, pros, cons and the expert validation.

Figure 4.3: Validation flow of the Visual Representation of Customers

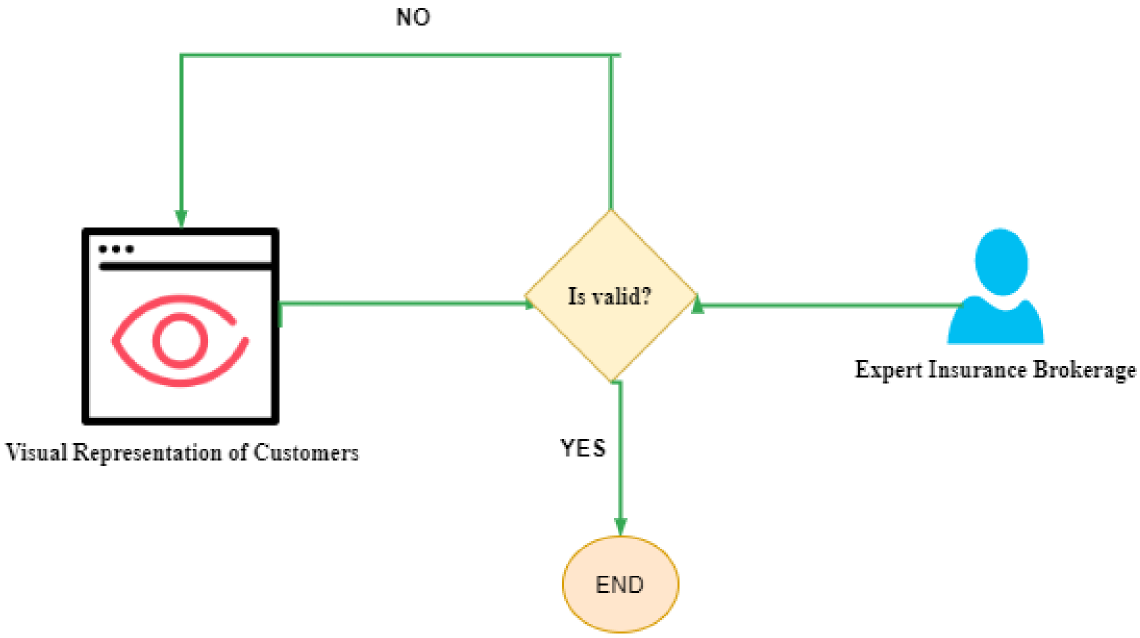
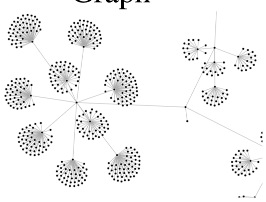
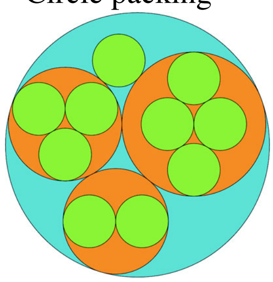


Table 4.2: Validation of the different visualizations proposed for the visual representation of customers.

Visualization Type	**Pros and Cons**	**Expert Validation**
Force-Directed Graph 	Pros: - Arrangement of circles without overlaping. - Cluster objects. - Show/Hide objects. - Color and size encode. - Zoom and pan. Cons: - Unused spaces in cluster formation. - Cluttered in the cluster process until convergence.	No
Circle packing 	Pros: - Arrangement of circles without overlaping. - Cluster objects. - Show/Hide objects. - Color and size encode. - Zoom and pan. Cons:	Yes

We use the circle packing visualization for representing the customers visually as well as the customer groups arranged. Each internal circle in the circle packing represents one customer. The internal circle color is encoded between the divergence of red, yellow, and green colors, which are inspired in the semaphore colors that denote the alert level. More specifically, it means that the internal circle color near to red has few products purchase, and the internal circle color near to green has many products. The internal circles in the circle packing visualization are grouped visually based on the label of the customer group to belong (see Figure 4.7(c)).

On the other hand, customer group visualization helps the broker to know how many customer groups have and understand what customer type the broker has in his portfolio based on customer characteristics. We attach this by encoding customer group information into a visual summarization using visualization such as pie chart and bar chart that allows brokers to explore and identify interesting groups of customers in an easy and faster manner (see Figure 4.10a(b)).

4.3.5 Mining Association Rules and Rules Visualization

Due to the nature of the client portfolio data (past sales) and for obtaining the query models, our recommender system's core is constructed based on the association rule mining method. In the **customer groups generation** component, we got the customer groups based on the customer characteristics combination. For each customer group the transaction data is mining using the apriori algorithm (AGRAWAL; SRIKANT et al., 1994)(for more details see subsection 2.2.1). To generate the major quantity of association rules that allow exploring rarity and variety recommendations, we choose the min-support and min-confidence parameters as 0.01 and 0.01, respectively. To recommend insurance products for a customer is filtered the association rules in accordance with the insurance products purchased by the specific customer selected.

As a natural characteristic of the association rule algorithm is the overhead of rule generation, this increases the effort into the rule exploration and identification process.

The table view and scatterplot visualization are chosen to visually represent the recommendations based on the advantages that they have related to the size of the ruleset, measures, interactive features, and usability in comparison with other visualizations (see Table 3.2).

An expert insurance broker validates the two visualizations i.e., table view and scatter plot visualization embedded in the recommendation system, in order to choose one that best accomplishes the requirements to perform the insurance product recommendation task. Figure 4.4 shows the validation flow of the visual representation of recommendations. Additionally, Table 4.3 details the visualization type, pros, cons and the expert validation.

Figure 4.4: Validation flow of the visual Representation of Recommendations.
d[h]

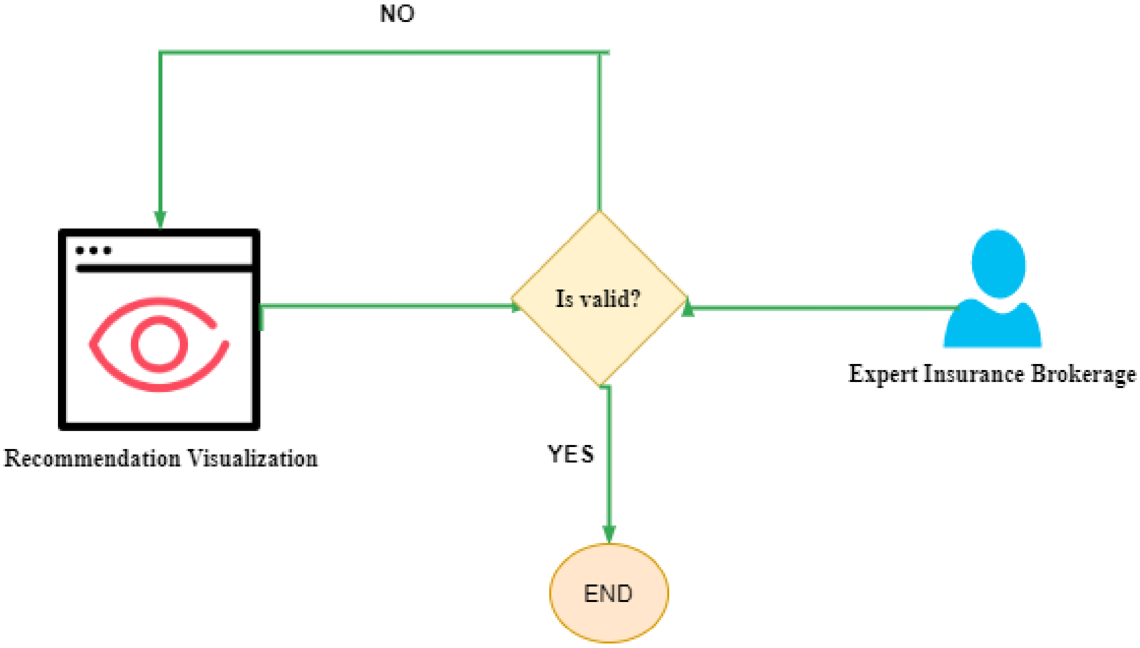
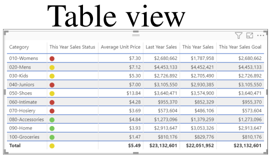
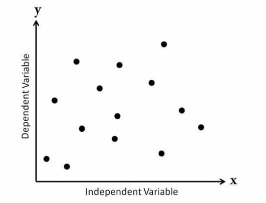


Table 4.3: Validation of the different visualizations proposed for the visual representation of recommendations

Visualization Type	Pros and Cons	Expert Validation
Table view 	Pros: - Works well with a medium rule set. - Encodes measures more than 1. - Interactive features such as hover, filtering, searching, sorting, drill down, pagination. - Ease of use. Cons: - Verbal interaction as well as visual interaction.	Yes

Continued on next page

Table 4.3 – *Continued from previous page*

Visualization Type	Pros and Cons	Expert Validation
Scatter plot 	Pros: - Works well with large rule set. - Ease of use. - Encodes measures up 3. - Show/Hide objects. - Color and size encode. - Zoom, pan, brush and grid line. Cons: - Occlusion in the visualization with large dataset.	Yes

We visually represent the rules using the Table view and Scatterplot visualization to address this problem to turn the process of exploration and identification of rules easy. Additionally, to the support and confidence, we introduce a new measure called customer-affected that denotes how many customers are affected in the client portfolio by the recommendation in the group recommendation level. These two visualizations are chosen considering the advantages to cover the requirements of brokers to perform the recommendation task (see Table 4.3).

4.3.6 Recommendation Explanation

The recommendation explanation plays an important role in the user experience. Provide a recommendation explanation to help users in terms of effectiveness and efficiency (TINTAREV; MASTHOFF, 2007). For this, Broker-RecSys explains the recommendation to turn easy to interpret by final users. We represent the recommendation explanation with a minimalistic textual format that contains the information of measures such as support, confidence, and affected-customers of the recommendation. This explanation tries to answer the human questions related to the recommendation *what product?*

and *why reason?*. For example, in Figure 4.9 an explanation of measure conditional probability (confidence) is shown (**33.33%** of clients that purchased **car and travel insurance** also purchase **business insurance**).

4.3.7 Cold-start Problem

The cold-start problem is a problem that affects recommender systems due to insufficient information that not allows recommending items for the users. For example, this occurs when the recommender systems receive a new user (missing information about the user, e.g., preferences and interests). Similarly, it also can occur when a new item is introduced to the recommender system (missing information about the item, e.g., rating or purchase information). In both cases, the recommender system can not recommend items for the user.

To alleviate the cold-start problem, a problem that affects customers that do not have products purchased (new users), we use the popularity recommendation strategy. This strategy consists of recommending the top-n recommendation. The top-n recommendations are obtained based on the products purchased by a specific customer selected in the group to which the customer belongs. The popular recommendations act as a complement to the recommendation based on purchase behavior. In Figure 4.7 (d) the top-3 popular products for recommend for a customer is shown visually using a bar chart visualization.

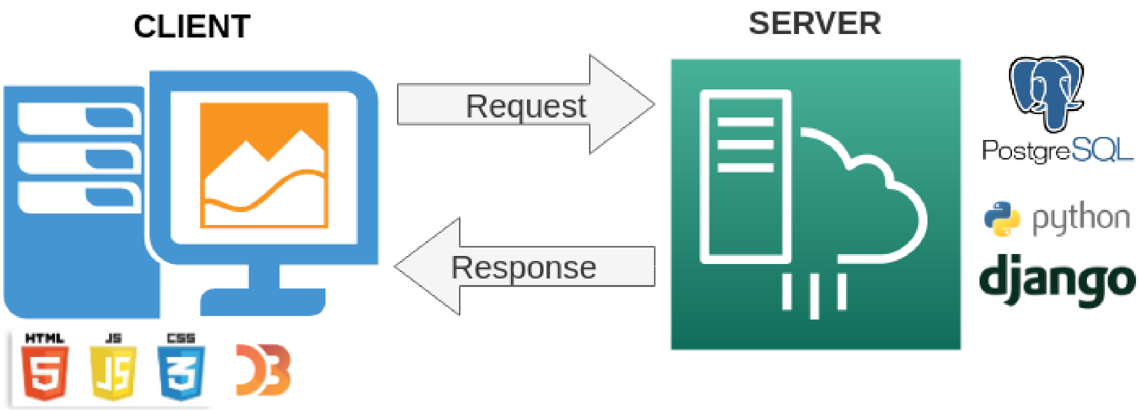
4.4 Broker-RecSys Tool

Broker-RecSys framework is embedded in a web-based system. The resources used to develop the interactive recommender system are as follows: to visualize the data, the d3 javascript library (BOSTOCK; OGIEVETSKY; HEER, 2011) is used. The HTML, JavaScript, and CSS are used as the basic block to build the web system. Python, a popular language for data analysis, was used for building the core of the recommendation system, including popular libraries such as numpy[1] and pandas[2] that are used to implement the several algorithms used in this work. We use PostgreSQL [3] for storage the data and

[1] https://numpy.org/
[2] https://pandas.pydata.org/
[3] https://www.postgresql.org/

Django[4] is used as the web framework based on the MVT architecture to develop the web-based system. The high-level web system architecture is shown in Figure 4.5.

Figure 4.5: Web system architecture

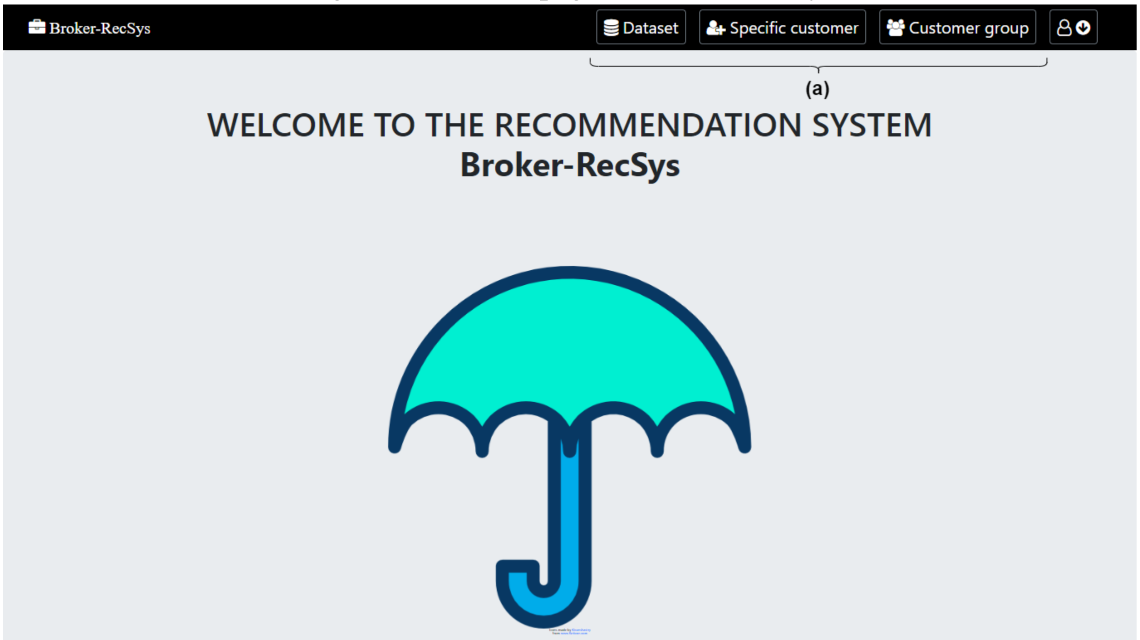


4.4.1 Overview

Broker-RecSys, an interactive recommender system that aims to support insurance brokers into the recommendation process. Figure 4.6 show the home page after a user started a session in the Broker-RecSys system. Figure 4.6(a) shows the three main functionalities that an insurance broker can perform using the Broker-RecSys system.

[4]https://www.djangoproject.com/

Figure 4.6: Home page of Broker-RecSys



Insurance brokers can use the three functionalities in their daily insurance product recommendation activities. The first functionality, "Dataset", allows the insurance broker to upload his clients' portfolios to Broker-RecSys. After that, Broker-RecSys prepare and preprocessing the data (client portfolio) and turn ready to perform the recommendation process at two levels: specific customer and customer group. Then, an Insurance broker can recommend insurance products for a specific customer or a customer group. In the following subsections, a running example of the recommendation process at the two levels are detailed.

4.4.2 Recommendation for a Specific Customer

In the first level, the broker can segment their customers based on interesting characteristics, e.g., gender and region. After this, the broker can filter their customers based on their characteristics (e.g., gender-male, region-0-São Paulo) as well as to explore them visually to identify potential customers to offer recommendations. To identify a potential customer, the broker looks for the color alert level of the customer in the circle packing visualization. After identifying the potential customer, the recommender system provides two types of recommendations for the identified customer: one based on the purchasing behavior and the other based on the most popular products. For example, Pedro is an insurance broker, he uploads his portfolio of clients in the web system and automatically his client portfolio is represented visually (See Figure 4.7(c)). After that, Pedro wants to recommend insurance products for his customers based on age and gender (See Figure 4.7(a)). Then, Broker-RecSys segment the clients using these interesting attributes selected by Pedro and return two groups (See Figure 4.7(c)). Now, Pedro wants to identify potential customers to offer insurance products. For this, Pedro determines that it should be interesting to find male clients with age between 30 and 65 years old that live in regions 5, 6, 8, 9 (See Figure 4.7(b)). Also, Pedro filter clients that have more than one product purchased, obtaining 86 potential customers filtered. Pedro goes to group 2 in the visual customer representation and selects a client. After selection the client, Broker-RecSys provide the information of Alexandre, the client selected as well as provide two types of recommendations to offer for Alexandre: based on popularity (See Figure 4.7(d)) and based on the purchase behavior (See Figure 4.7(e)). At this point, Pedro can offer for Alexandre green card insurance product because more than 40% of clients from the same customer group to which Alexandre belongs bought green card insurance. On the other hand, Pedro wants to know what products he can offer for Alexandre based on his purchase history and why. Pedro interprets the recommendation based on purchase behavior and sees that Alexandre can receive business insurance as a recommendation, and now he wants to ask why? What is the reason?, then, Pedro identify two metrics in the table: the probability of joint purchase (support) and the conditional probability of purchase (confidence). To interpret each of them, Broker-RecSys provides an explanation of each metric for each recommendation (See Figure 4.8 and 4.9). Finally, Pedro decides what recommendations to offer for Alexandre.

Figure 4.7: Visual interface used in the recommendation process for a specific customer. a) User control to cluster customers. b) Customer filters. c) Visual customer representation. d) Popular-based recommendation. e) Purchase behavior recommendation.

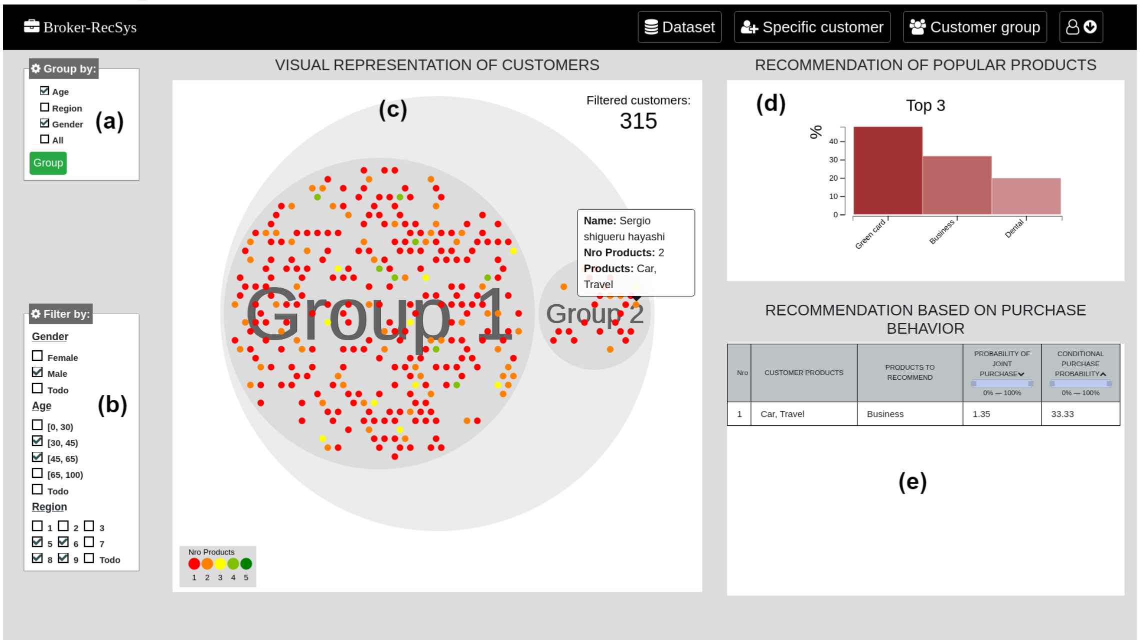


Figure 4.8: Specific customer recommendation - probability of joint purchase (support) explanation

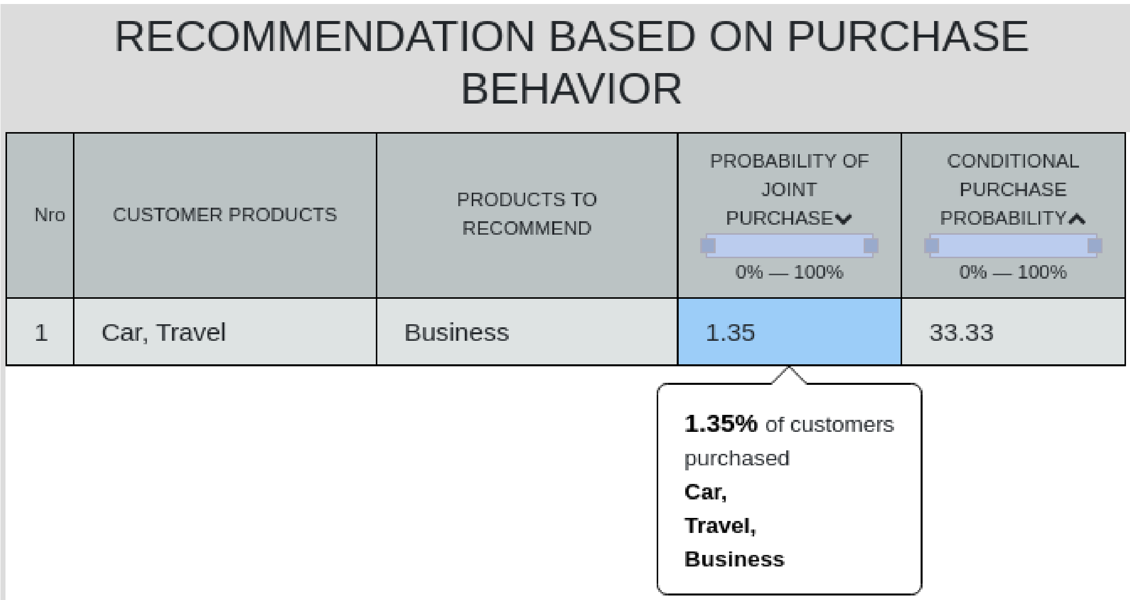
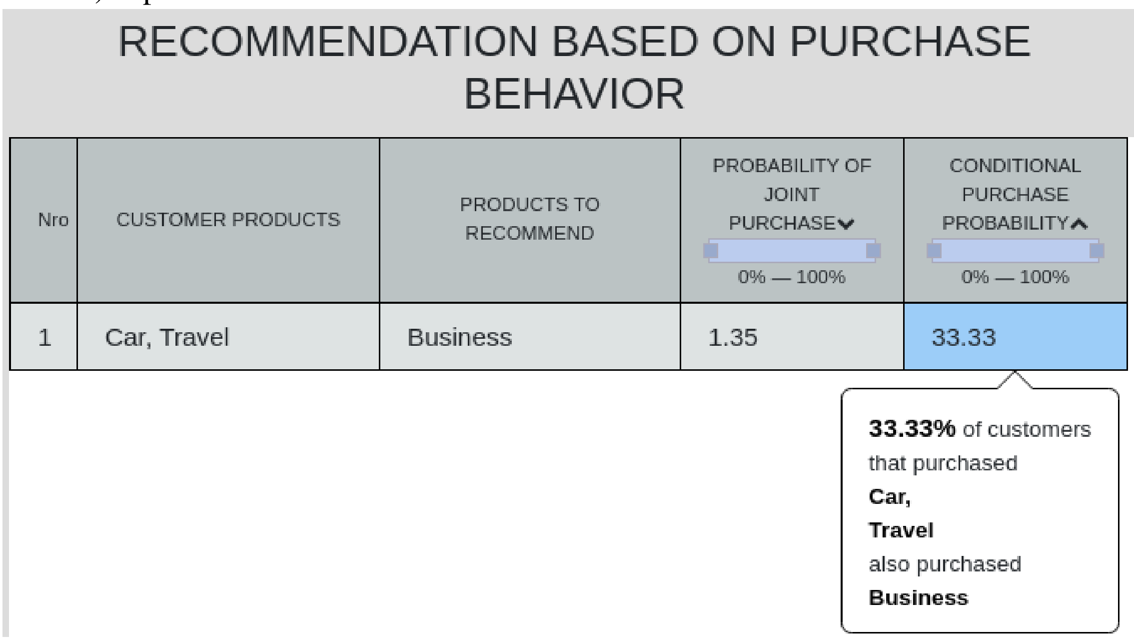


Figure 4.9: Specific customer recommendation - conditional probability of purchase (confidence) explanation

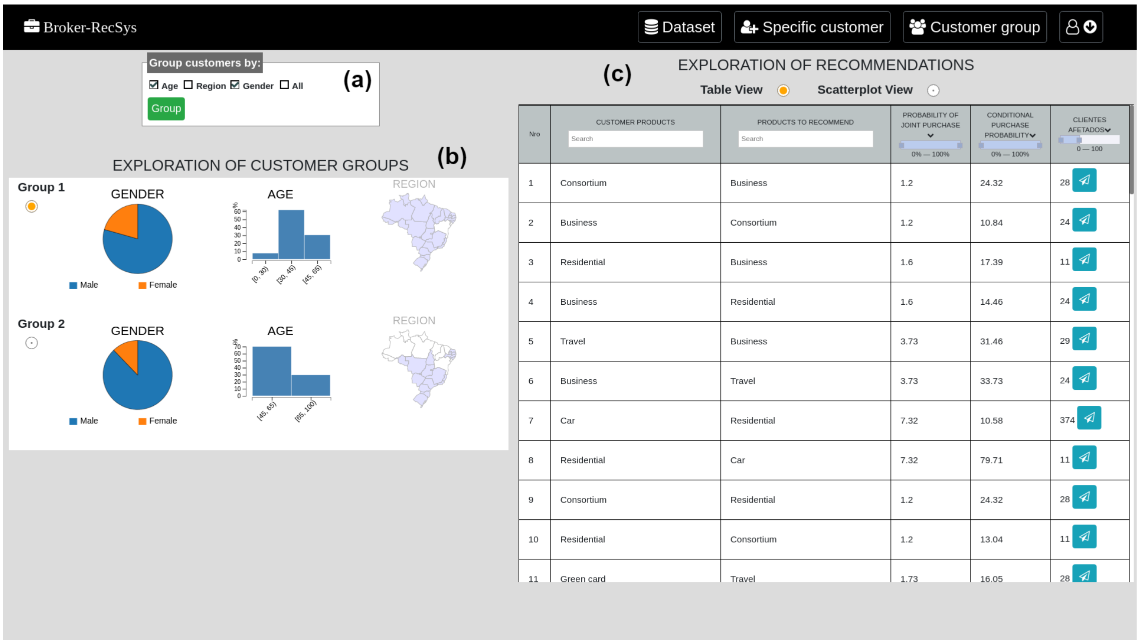


4.4.3 Recommendation for a Customer Group

In the second level, the broker can find interesting recommendations to offer for a group of their customers. The broker visually explores the customer groups. After the broker identified one target group, he can explore the recommendations that can be offered inside the customer group selected. The broker can visualize and explore the recommendations using the Table and Scatterplot visualization (see Fig. 4.10). Maria, an insurance broker, wants to offer insurance products for a customer group, she wants to recommend insurance products based on age and gender. For this, she segments her client portfolio based on these client characteristics (See Figure 4.10 (a)) and obtains two clusters. Next, she explores those customer groups visually (See Figure 4.10(b)) and select the first group that is interesting for her. After she selects the customer group, a set of recommendations to offers in the group selected are showed for Maria (See Figure 4.10(c)). Maria can choose to visualize the recommendation via a table view or a scatterplot view and then explore these recommendations. For each recommendation, she can answer two questions: what products offer? and for why reason?. Maria answers the first question filtering recommendations based on what products Maria wants to offer for her clients, or what products offer for clients that already have specific products purchased. On the other hand, Maria answer the second question using the explanation of each recommendation based on their measures such as probability of joint purchase, probability of conditional purchase and customers affected by the recommendation (See Figure 4.11, 4.12 and 4.13). Thus, Maria can recommend specific products for a customer group. Besides, brokers can plan a marketing campaign for a set of products and send recommendations for their clients via the internet.

Figure 4.10: Recommendation for a customer group. a) User control to cluster customers. b) Visual summarization of customer groups. c) Visualization of product recommendation

(a) Table view



(b) Scatterplot view

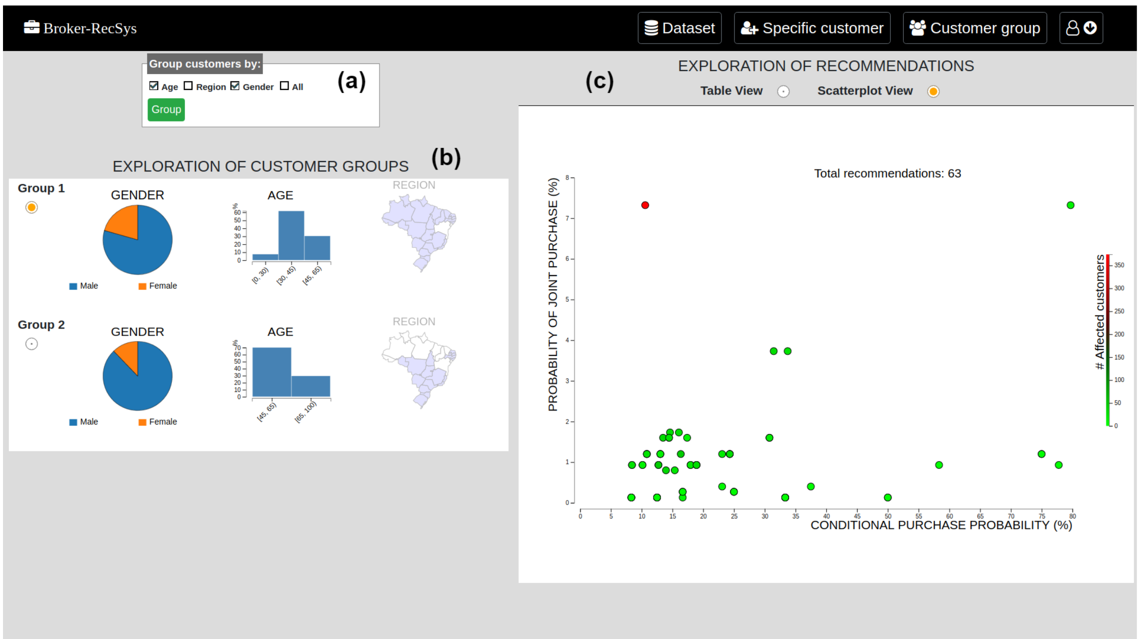


Figure 4.11: Recommendation for a customer group - Probability of joint purchase (support) explanation.

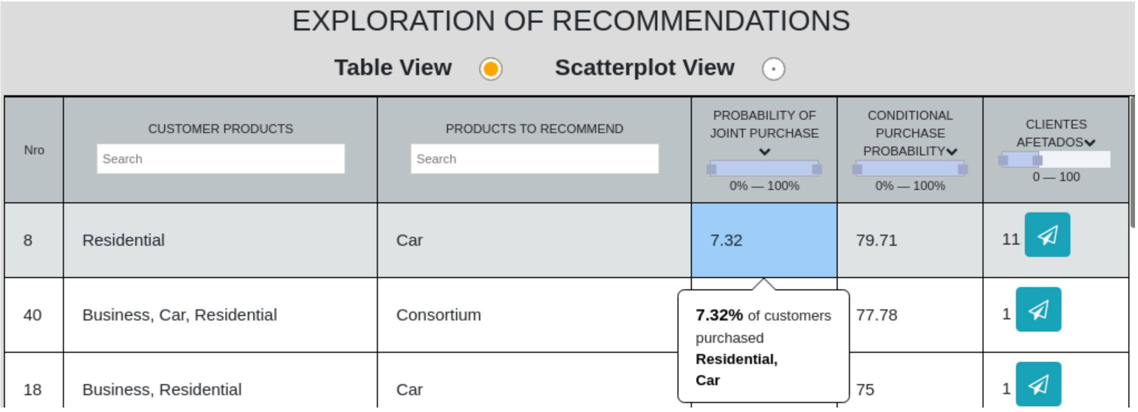


Figure 4.12: Recommendation for a customer group - Conditional probability of purchase (confidence) explanation.

Figure 4.13: Recommendation for a customer group - Affected-customers explanation.

5 EXPERIMENTAL EVALUATION

This chapter presents the experiments that were conducted to evaluate the usability and usefulness dimension of Broker-RecSys. We conducted two user studies: *(i)* local, naive users evaluated the usability dimension based on questionnaires and the eye-tracking analysis, and *(ii)* remote, insurance brokers evaluated the usability and usefulness dimensions based on questionnaires.

5.1 Usability and Usefulness Evaluation

Usability and usefulness are quantitative and qualitative measures obtained from a user in the use of an information system. This usability and usefulness measures are used to evaluate a computational system to increase the user experience.

Usability is defined as “the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use” (ISO, 2018). In this sense, effectiveness measures the satisfactory completion of specifics goals, efficiency measures the resources required to achieve specific goals, and satisfaction is related to the user’s degree of satisfaction in the use of the computational system. Usability focus on user interaction with the different functionalities of the computational system.

On the other hand, usefulness is defined as “the degree to which a user is satisfied with their perceived achievement of pragmatic goals, including the results of use and the consequences of use” (IEC, 2011). In this case, usefulness focuses on user interaction with the information provided by the user’s computational system.

To evaluate Broker-RecSys, we conducted two user studies locally and remotely. Locally, naive users participate in evaluating Broker-RecSys in the usability dimension. In this study case, we combine a widely used evaluation method based on questionnaires and evaluations based on the eye-tracking analysis to make a better analysis related to effectiveness, efficiency, and satisfaction. On the other hand, insurance brokers participate in evaluating Broker-RecSys in the usability and usefulness dimensions. In this study case, the experiment conducted is more focused on user-centered tasks related to the usefulness dimension as well as the evaluation based on questionnaires.

With this two studies, we want to answer our research questions mentioned in Chapter 1:

1. **Broker-RecSys enables naive users to perform insurance products recommendation tasks?**
2. **Broker-RecSys supports insurance brokers in the recommendation process for offer insurance products in their client portfolio?**

A pilot study was conducted previously locally and remotely with ten naive users and three insurance brokers (ATAUCHI; NEDEL; GALANTE, 2019). We use this information to improve the interface design and insurance recommendation tasks for the final test experiment (See Table 5.1).

Table 5.1: User tasks for the test session.

Target of the recommendation task	**Task description**
Specific customer	1. Cluster the clients based on gender. How many groups were formed? 2. Filter the clients by gender female and region 7. How many clients were filtered? 3. Select a client who has 2 products purchased. What products the client purchased? 4. Which best product can be offered to the client selected based on popularity? why? 5. Which best product can be offered to the client selected based on purchase behavior? why?

Continued on next page

Table 5.1 – *Continued from previous page*

Target of the recommendation task	**Task description**
Group of customers	1. Cluster the clients based on gender and age. How many groups are formed? 2. Select the second group. What characteristic has this group? 3. Using a table view, sort and choose the recommendation with the highest conditional probability. What does this mean? 4. With the scatterplot view, what recommendation affect most of customers? What does this mean?

In the following subsections, we detailed a brief concept and formulation of the metrics used in the evaluation of Broker-RecSys.

5.1.1 Usability Questionnaire

We use SUS to evaluate the usability of Broker-RecSys based on the questionnaire, the widely used-based questionnaire that measures perception of usability of a tool (BROOKE et al., 1996). SUS questionnaire is composed of ten items. Each item can receive a rating on the Likert scale between 0 and 5 from strongly disagree to agree strongly. The SUS score varies between 0-100 and determines the level of usability of the tool. The number of participants was considered to calculate SUS score, the position of items i.e., odd or even, and the rating given by each user to each item. SUS score greater than 68 are considered that a tool has good usability. The SUS score is given by Equation 5.2.

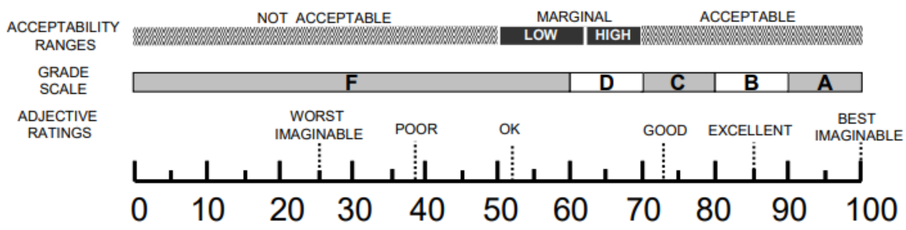
$$SUS\_score\_u = \left(\sum_{i\in\{1,3,5,7,9\}} (R_{u,i} - 1) + \sum_{j\in\{2,4,6,8,10\}} (5 - R_{u,j})\right) * 2.5 \tag{5.1}$$

$$SUS\_score\_total = \frac{\sum_{u \in \{1,..,n\}} SUS\_score\_u}{n} \tag{5.2}$$

where SUS_score_u denotes the score obtained by a user u, $R_{u,i}$ represents the rating given by user u for the item i and SUS_score_total denotes the SUS score obtained from n users.

To interpret the SUS score, Bangor et al. (BANGOR; KORTUM; MILLER, 2009) propose the adjective rating scale that allows obtaining a textual interpretation of the SUS score result. Figure 5.1 shows the SUS interpretation based on the SUS score.

Figure 5.1: SUS score interpretation



Source: (BANGOR; KORTUM; MILLER, 2009)

5.1.2 Eye Tracking Analysis: Effectiveness and Efficiency

The eye-tracking analysis is well adopted in information visualization and psychology research. Information visualization adopts this method to evaluate visualization methods based on the gaze data (KURZHALS et al., 2014; FU; NOY; STOREY, 2017). On the other hand, psychology uses eye-tracking analysis to study the relationship between eye behavior and cognitive processes (MELE; FEDERICI, 2012). The eye-tracking tool is composed of hardware and software technologies that allow us to detect and record gaze data. Although exist several visual analysis based on gaze data (BLASCHECK et al., 2017), the two most used visual analysis based on gaze data is scan path and heatmap visualization (FORSMAN et al., 2013; RASCHKE; BLASCHECK; BURCH, 2014; DRUSCH; BASTIEN; PARIS, 2014). Scan path visualization provides the order of the path taken by the user in exploring the visual stimulus. The scan path method is useful to perform an individual analysis because it is possible to obtain a complete overview of the trajectory followed by the user in the visual stimulus. On the contrary, the heatmap visualization allows determining the locations of concentration of eye fixations of users in the exposition to the visual stimulus. Unlike scan path visualization, heatmap visualization enables us to perform a visual analysis of eye fixation of multiple users while avoiding the occlusion problem that suffers the scan path visualization. Figure 5.2 shows an example of scan path and heatmap visualization.

Figure 5.2: Scanpath (left) and Heatmap (rigth) visualization of the user visual stimulus on a Web page.

Source: (DRUSCH; BASTIEN; PARIS, 2014)

To complement the usability evaluation of Broker-RecSys locally, we perform a visualization analysis from gaze data. Gaze data analysis is used to measure the usability of Broker-RecSys in terms of effectiveness and efficiency. The efficiency is calculated based on the time required to complete the user tasks, and the effectiveness is calculated based on the success or failure in completing the user tasks. Besides, to explore in detail the effectiveness metric, we analyze gaze data using the heatmap visualization. With this, we verify the convergence of regions of the visual stimulus needed to perform and complete the user tasks.

5.1.3 Satisfaction Questionnaire

To measure the user satisfaction in the use of Broker-RecSys, we use a questionnaire strategy after the user finalizes the use of the recommender system. We formulate three questions related to the perception of user satisfaction for the use of Broker-RecSys. These questions are as follow: *What did you find interesting about the recommendation* *system?*, *What you did not find interesting about the recommendation system?* and *What would you like to change in the insurance product recommendation system?*. These questions intend to find the positive, negative, and constructive perception of users for the use of Broker-RecSys. Based on the answers of users for the three questions, we perform a qualitative analysis for determining the satisfaction of users in the use of Broker-RecSys. Besides, this information serves as feedback to improve the recommender system for future works.

5.1.4 Usefulness Questionnaire

The usefulness dimension makes references to the perception of the users for the value of the tool for the tasks that the user needs to do. Thus, to measure the usefulness of Broker-RecSys perceived by the user, we formulate questions related to each functionality of the recommender system. The user, in this case, the insurance broker rate each question with a Likert-scale value between 0 and 5, where 0 denotes strongly disagree and 5 strongly agree. The usefulness score is calculated as the mean of rated values in each question. We determine that an acceptance usefulness score is greater than the middle of the Likert-scale value by each usefulness question i.e., usefulness question score greater than 2.5.

5.2 Participants

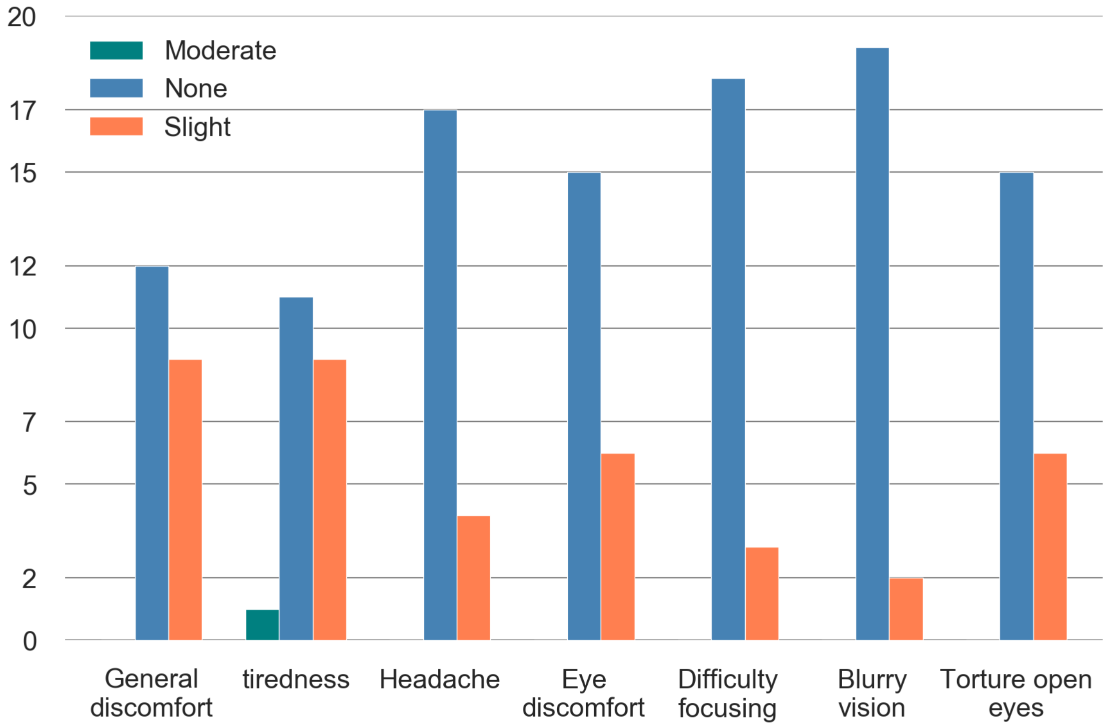
Two types of users participate in the evaluation of Broker-RecSys (see Table 5.2). Each participant type is described in the following paragraphs.

Table 5.2: Type of users that participated in the Broker-RecSys evaluation.

Group	**User type**	**Number of participants**
G1	Professional insurance broker	3
G2	Non-professional insurance broker	21

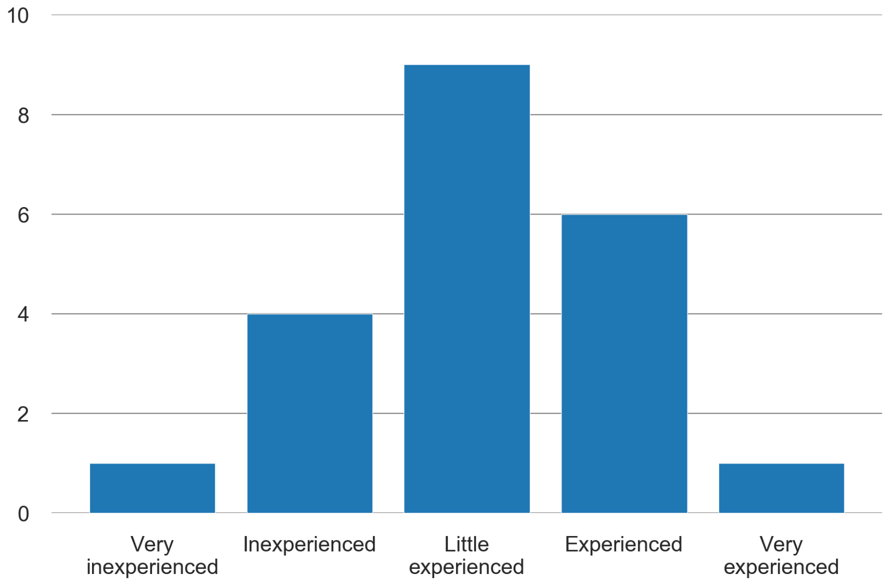
The first type of user (naive users) is inexperienced in the insurance domain, more specific, in the recommendation of insurance products in a portfolio of clients. The naive-users are an excellent option to evaluate the usability of Broker-RecSys. Due to their inexperience, naive users can show if the tool is easy or hard to learn and perform tasks in an unknown domain; in this case, the recommendation of insurance products in a client portfolio. A total of 21 naive users participated in the local experiment. The age of the naive users is composed as follows: 18-25 (10 users), 26-35(9 users), and 36-45(2 users). The gender of the participants is composed of 18 males and 3 females. All naive users are either attending a higher level or they have a higher level. Concerning vision state, 13 naive users have not any vision problem, 6 naive users have myopia and astigmatism, and 2 users have astigmatism. The discomfort state of naive users report that users are in an excellent physical and mental state before the realization of the experiment (see Figure 5.3). Finally, the users are asked about their level of experience in the use of interactive information systems. In Figure 5.4, we can see that, in average, users have an intermediate level of experience in the use of interactive information systems.

Figure 5.3: Physical and mental state of naive users before the test session.



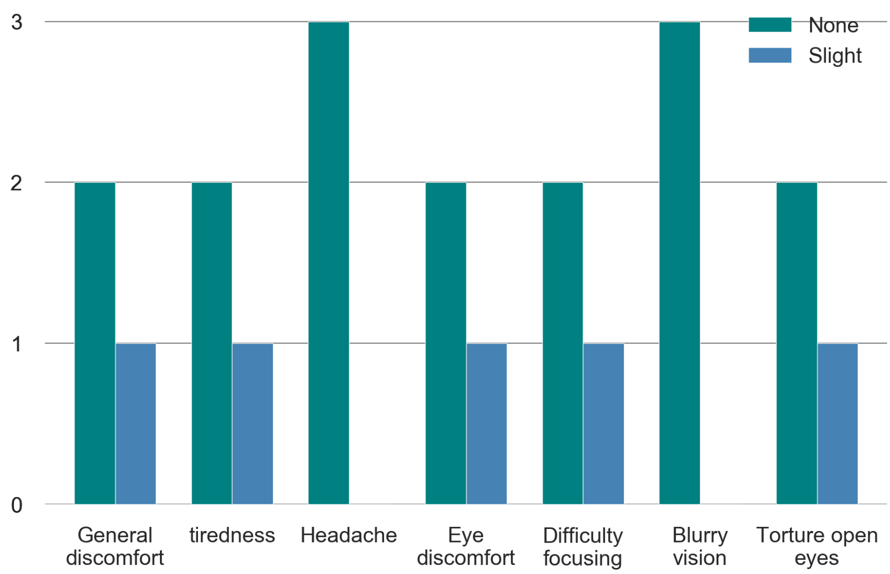
Insurance brokers are professionals in the insurance domain. They sell insurance products and help people to choose the best insurance policy. The objective of Broker-RecSys is to support brokers in the recommendation process of insurance products in their client portfolio. Thus, broker users evaluate Broker-RecSys in the usability and usefulness dimensions. Three insurance brokers participate in the remote experiment. Two insurance brokers are between 26-35 years old, and one insurance broker is older

Figure 5.4: Experience of the naive users in the use of interactive information systems.



than 45 years old. All insurance brokers are males. The level of experience in the use of interactive information systems by insurance brokers is little experienced, experienced, and very experienced, respectively. Figure 5.5 shows the physical and mental state of insurance brokers before the evaluation of Broker-RecSys.

Figure 5.5: Physical and mental state of the broker users before the test session.



5.3 Local Evaluation

A good user interface design can turn a computer system easy to learn and use as well as alleviate exhaustive cognitive processes required by the user to complete specific goals satisfactorily. A well-known practice to design a user interface is to follow the user workflow; this means taking the user as a central point to build the user interface (LEWIS; RIEMAN, 1993). Following this practice, we design Broker-RecSys. The Broker-RecSys interface design is evaluated with naive users based on the assumption that if this type of user has a brief overview of the insurance broker domain, this user can perform tasks related to insurance products recommendation successfully then the user interface design is adequate for its purpose. In this way, it guarantees that the tool will be easy to learn and use and will not require a strong cognitive effort.

5.3.1 Setup Detail and Workflow

We prepared a room with adequate conditions such as a calm environment and without distractions for the development of the experiment (see Figure 5.6). In order to collect the gaze data, an eye tracker is mounted in a notebook. The user is positioned in front of the monitor screen and performs the eye tracker calibration that consists of receiving a sequence of a visual stimulus (red circles). The user does follow up and maintain their fixation in the circles while it enters in a transition of radius between 5 and 50 pixels in 4 seconds approximately. Consequently, the user gaze data is collected. This process is followed until complete the total of visual stimulus and obtain a successful calibration. We consider the eye tracker degree and user distance as $\sim 30^{\circ}$ and ~ 60 cm respectively, as recommended in (TOBII, 2014). More technical details are shown in Table 5.3. After the calibration process, the user performs the recommendation tasks using the interactive recommender system interface while the eye tracker collects the user gaze data. Finally, the user gaze data is analyzed. Figure 5.7 show this process. For collecting gaze data, we implement an eye tracker software using the Tobii Pro SDK [1]. This software is integrated with Broker-RecSys to gather gaze data and record screen activity in real-time while a user uses Broker-RecSys.

[1] http://developer.tobiipro.com/python.html

Figure 5.6: Environment for the test session.

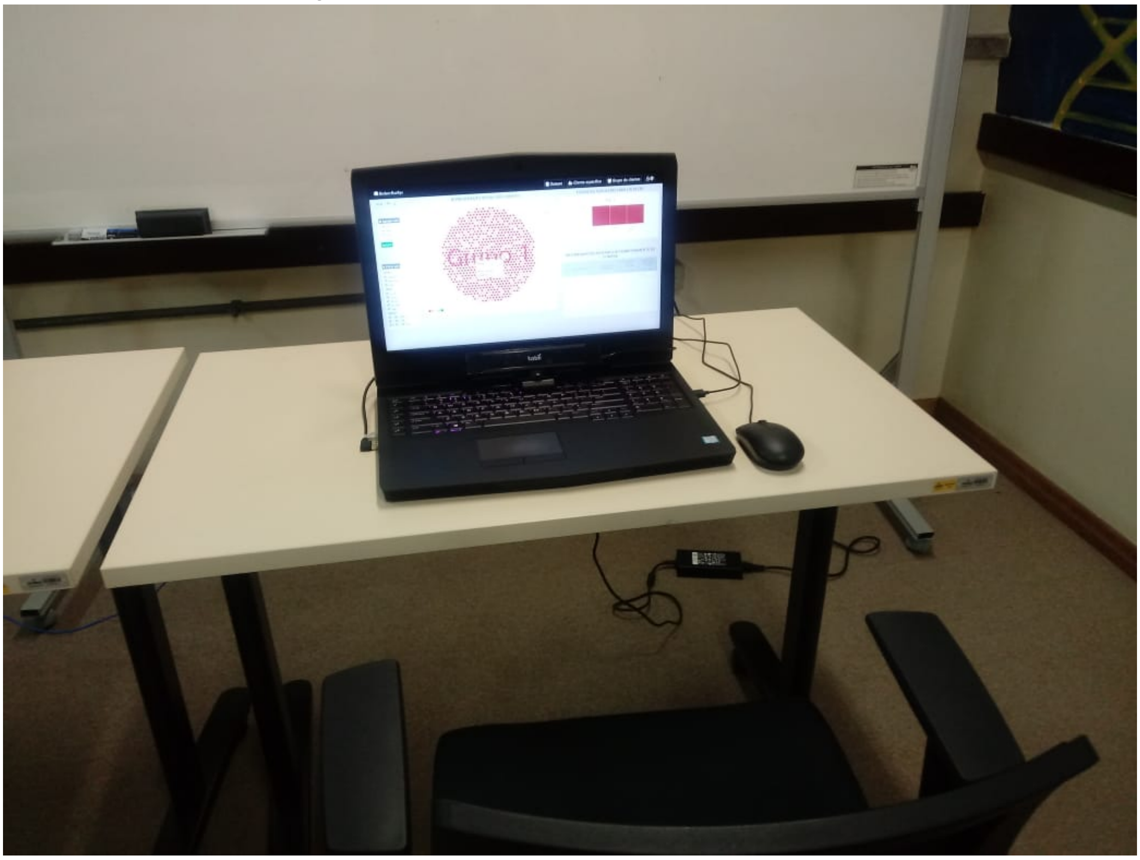


Table 5.3: Setup details of equipment used in the test session.

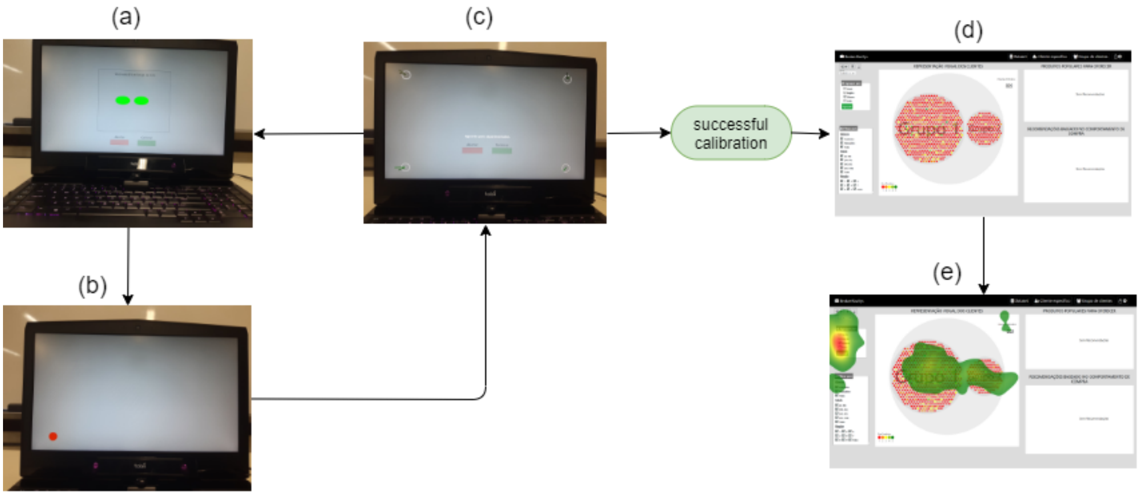
Hardware	Characteristic	Configuration
Notebook	– Processor model Intel(R) Core(TM)2 Quad with 2.83GHz – 8 GB RAM – Screen Resolution 1900x1200	-

Continued on next page

Table 5.3 – *Continued from previous page*

Hardware	Characteristic	Configuration
Eye tracker	– Model: Tobii X2-30 Eye Tracker Compact Edition) – Frequenc: 40hz	– Calibration: 9 stimulus points – User distance: ∼60 cm – Angle inclination: ∼30°

Figure 5.7: Workflow for the eye tracking experiment. (a) User positioning for calibration. (b) Visual stimulus and calibration process. (c) Calibration precision. (d) User tasks and interaction with the user interface while the gaze data is collected. (e) Analysis of the user gaze data.



5.3.2 Procedure

At the beginning of the test session, the moderator introduces the naive user to the study. Here, the moderator gives the user a brief explanation about the setup environment (see Section 5.3.1) as well as basic concepts that will help the naive user in the test session. Following, the user receives a video tutorial that introduces the functionalities of the interactive recommender system to the user. Before starting the test session, the participant's level of discomfort is measured to ensure that the user is in optimal condition for experimenting. The test session consists of two parts: free time exploration and recommendation tasks. In the first part, the user gains a familiarization with the tool as a complement to the video tutorial given previously. In the second part, the user is requested to perform several tasks, and his user gaze data is registered with an eye-tracking tool during the execution of the tasks. These tasks will be used to measure the effectiveness and efficiency of Broker-RecSys. Consequently, the user is evaluated in the dimension usability and satisfaction based on questionnaires. The user filled out a questionnaire form for usability (SUS) as well as the post-test satisfaction questionnaire. A complete explanation of the procedure of the test session is shown in Table 5.4.

Table 5.4: Protocol follow for the naive user in the usability evaluation of Broker-RecSys

Step	Activity	Description	Measurement	Instrument	Analysis
1	Consent form	The user accepts their participation in the study after reading the free and informed consent form.	–	Paper form (see Appendix B)	–
2	Introduction to the study	Presentation of the moderator and their role in the test session, explanation of the setup environment for the test session and the protocol to follow. Additionally it is given a brief explanation about recommender system and insurance domain.	–	Paper form and infographics (see Appendix C)	–

Continued on next page

Table 5.4 – *Continued from previous page*

Step	Activity	Description	Measurement	Instrument	Analysis
3	User characterization (questionnaire):	The user fills their personal information such as name, age, gender as well as their academic background, experience using interactive information systems and their physical and mental level of discomfort.	–	Paper form (see Appendix D)	Informative
4	Introduction to the prototype	The user is provided with a demo (4 minutes) showing the functionality of the interactive recommendation system.	–	Video with audio	–

Continued on next page

Table 5.4 – *Continued from previous page*

Step	Activity	Description	Measurement	Instrument	Analysis
5	Test tasks	**Part 1: Free exploring time** (2 minutes) **Part 2: User tasks (9 tasks).** The user performs 9 tasks related to the insurance products recommendation for a specific customer and for a group of customers respectively (random order), priorizititing the time required and the accuracy to complete the tasks (see Table 5.1).	**M1. Effectiveness and efficiency** Accuracy and time required to complete the tasks. **M2. Visual perception** Distribution of eye fixation over the visual stimulus and screen activity **Part 1:** Familizarization with the prototype **Part 2:** M1, M2	Eye tracking software	– Calculation of number of successful or failed completion of tasks. – Time required to complete the tasks. – Visual attention map: search, attention.

Continued on next page

Table 5.4 – *Continued from previous page*

Step	Activity	Description	Measurement	Instrument	Analysis
6	System Usability Scale (SUS) - questionnaire	10-item Likert scale questionnaire to measure the level of use (ease or lack) of the system.	Satisfaction	Paper form (see Appendix G)	Boxplot visualization (range 0-5) and SUS score
7	Post-test questionnaire	User opinion Questions related to the test session: What are your positive opinion? What are your negative opinion? What would have you like to change to the recommender system?	Satisfaction	Paper form (see Appendix H)	Transcribed and summarized

5.3.3 Results

In this subsection, we describe the results obtained to evaluate the usability of Broker-RecSys in terms of effectiveness, efficiency, and visual stimulus perception and feedback. With these metrics, we aim to answer the first research question (**Broker-RecSys enable naive users to perform insurance products recommendation tasks?**).

5.3.3.1 Effectiveness and Efficiency

To evaluate effectiveness and efficiency of Broker-RecSys, naive users were asked to perform 9 tasks related to the recommendation of insurance products in a client portfolio (see Table 5.1). Effectiveness measures the level to complete tasks i.e., success or failure, while efficiency measures the time required to complete the tasks. Naive users reported 96.19% in completing tasks in the recommendation process for a specific customer with 1.58 minutes in average and 92.86% in completion of tasks in the recommendation process for a group of customers with 1.47 minutes in average (the general and detailed results are described in Table 5.5 and 5.6 respectively). The specific customer and group of customer recommendation process obtained an excellent completion rate near to 100% as well as an average time never higher than 2 minutes to complete the total of tasks in the respective level of recommendation process.

Table 5.5: Task completion rate and average time by type of recommendation process.

Recommendation process	Completion rate	Average time
Specific customer	96.19%	1.58
Group of customer	92.86%	1.47

Table 5.6: User performance by task: Success (S) or failure (F) in completing a task and time spent by task in seconds.

User	Task 01	Task 02	Task 03	Task 04	Task 05	Task 06	Task 07	Task 08	Task 09
Naive user 1	S/5.3	S/22.4	S/12.8	S/23.9	S/30.4	S/7.0	S/34.9	S/15.0	S/19.7
Naive user 2	S/22.1	S/29.1	S/16.7	S/13.9	S/33.1	S/16.1	S/31.3	S/28.0	S/25.0
Naive user 3	S/16.7	S/18.1	S/21.2	S/18.3	S/27.1	S/14.5	S/20.7	S/25.9	F/-
Naive user 4	S/17.3	S/26.5	S/19.6	S/12.0	S/25.7	S/17.3	S/22.1	S/21.7	F/-
Naive user 5	S/20.7	S/27.9	S/21.6	S/22.8	S/23.9	S/8.4	S/25.5	S/32.6	S/18.1
Naive user 6	S/28.3	S/25.1	S/22.5	S/20.3	S/32.0	S/18.9	S/24.3	S/33.2	S/43.7
Naive user 7	S/9.6	F/-	S/11.3	S/8.2	S/23.6	S/5.4	S/26.0	F/-	S/15.8
Naive user 8	S/7.3	F/-	S/15.6	S/10.5	S/20.1	S/7.0	S/24.3	S/39.8	S/24.6
Naive user 9	S/7.6	S/12.6	S/20.4	S/9.7	S/30.4	S/7.3	S/28.1	S/24.6	S/22.1
Naive user 10	S/5.8	S/19.1	S/20.9	S/11.8	S/14.2	S/11.5	S/26.5	S/6.9	S/38.2
Naive user 11	S/7.7	S/20.9	S/21.3	S/24.9	S/25.1	S/9.0	S/33.4	S/29.5	S/27.1
Naive user 12	S/4.7	S/15.4	S/21.9	S/13.1	S/20.4	S/4.1	S/16.3	S/27.3	S/13.4
Naive user 13	S/6.8	S/28.2	S/20.9	S/24.4	S/31.8	S/4.8	S/34.4	F/-	S/44.9
Naive user 14	S/10.5	S/27.3	S/28.9	S/23.1	S/29.3	S/4.8	S/39.3	F/-	S/23.3
Naive user 15	S/15.9	S/23.5	S/15.6	S/10.6	S/21.7	S/5.4	S/15.3	S/19.2	S/17.1
Naive user 16	S/8.7	S/17.7	S/16.9	S/16.6	S/24.2	S/6.3	S/45.3	S/19.8	S/13.4
Naive user 17	S/22.2	S/23.9	S/12.9	S/20.2	S/18.0	S/21.3	S/32.4	S/30.0	S/48.7
Naive user 18	S/14.1	F/-	S/8.8	S/7.8	F/-	S/7.0	S/28.1	S/29.1	S/13.6
Naive user 19	S/19.0	S/25.6	S/27.9	S/25.5	S/38.7	S/12.9	S/25.7	S/27.0	S/27.9
Naive user 20	S/10.7	S/24.7	S/11.4	S/15.7	S/14.8	S/8.8	F/-	S/8.7	S/20.8

Continued on next page

Table 5.6 – *Continued from previous page*

User	Task 01	Task 02	Task 03	Task 04	Task 05	Task 06	Task 07	Task 08	Task 09
Naive user 21	S/9.0	S/18.9	S/20.2	S/6.8	S/9.6	S/10.6	S/27.5	S/33.0	S/24.0
Percentage success	100%	85.72%	100%	100%	95.24%	100%	95.24%	85.72%	90.48%
Average time	12.86 (std=6.72)	22.61 (std=4.77)	18.54 (std=5.25)	16.20 (std=6.32)	24.71 (std=7.24)	9.92 (std=5.06)	28.07 (std=7.28)	25.07 (std=8.59)	25.34 (std=10.91)

5.3.3.2 Visual Attention Map

The visual attention map gives information about the distribution of fixations of users during the performance of the tasks. Figures 5.8, 5.9, 5.10, 5.11, 5.12, 5.13, 5.14, 5.15, 5.16 present the fixations of the users' gaze for the 9 user tasks (see Table 5.1). For each task, the gaze data collected have information about coordinates (x, y) in the notebook screen used as well as the timestamp in milliseconds. Using the coordinates information, we plotting the fixation data points in a representative screenshot related to the task performed. Consequently, we normalized the fixation data points distribution based on the local density using a gaussian function. Analyzing the heatmaps generated by users' gaze data while performing the tasks, we can observe that users' fixation data points converge quite well to the image regions where the answers are. In this way, we can conclude that the Broker-RecSys interface is quite clear and easily understandable.

Figure 5.8: Recommendation for a specific customer: Heatmap-Fixations of the users' gaze for Task 1 (Cluster the clients based on gender. How many groups are formed? - Answer: 2 clusters)

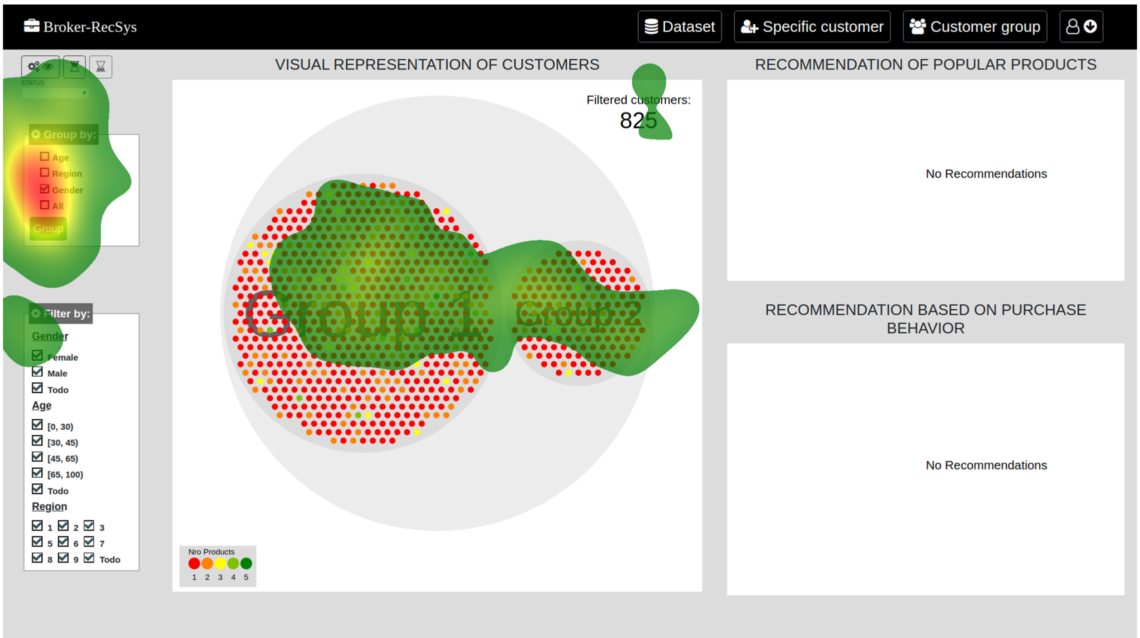


Figure 5.9: Recommendation for a specific customer: Heatmap-Fixations of the users' gaze for Task 2 (Filter the clients by gender (female) and region 7. How many clients are filtered? - Answer: 25 clients)

Figure 5.10: Recommendation for a specific customer: Heatmap-Fixations of the users' gaze for Task 3 (Select a client that have 2 products purchased. Which products the client purchased? - Answer: car and travel insurance)

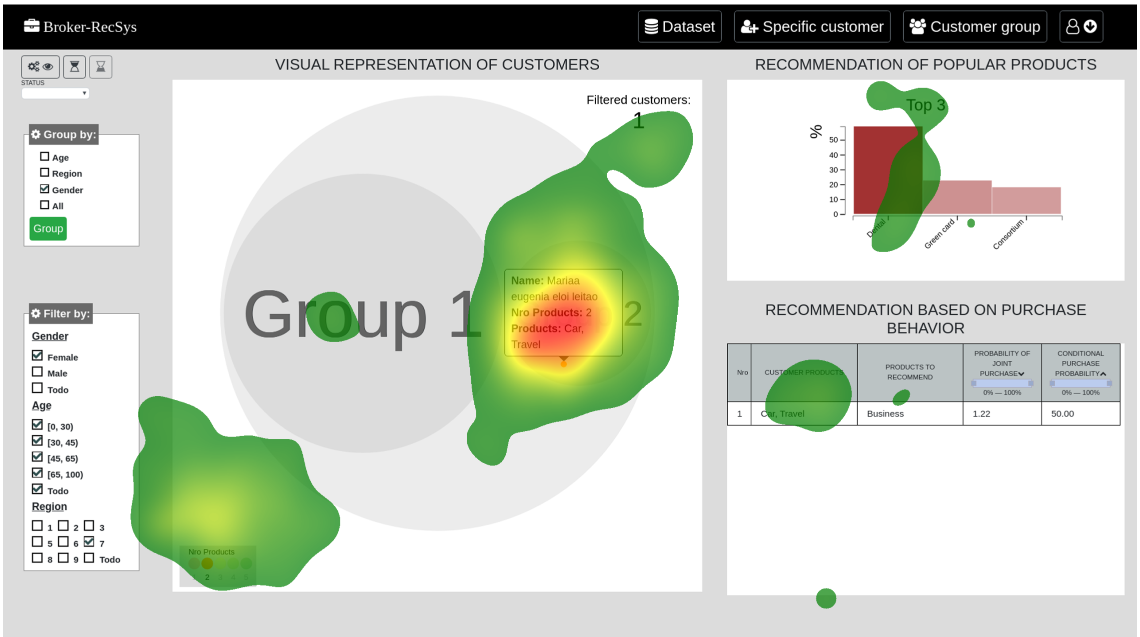


Figure 5.11: Recommendation for a specific customer: Heatmap-Fixations of the users' gaze for Task 4 (What best product can be offered to the client selected based on popularity? Why? - Answer: dental insurance because 59.1% of clients purchased dental insurance)

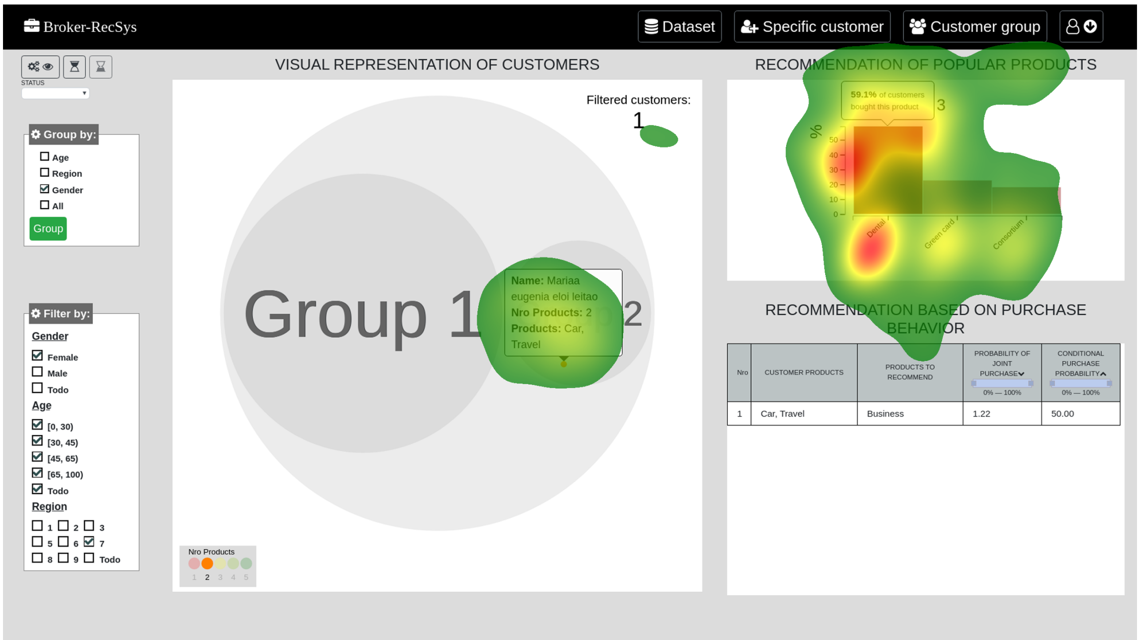


Figure 5.12: Recommendation for a specific customer: Heatmap-Fixations of the users' gaze for Task 5 (What best product can be offered to the client selected based on purchase behavior? Why? - Answer: empresarial insurance because 50% of clients that purchased car and travel insurance also purchased empresarial insurance)

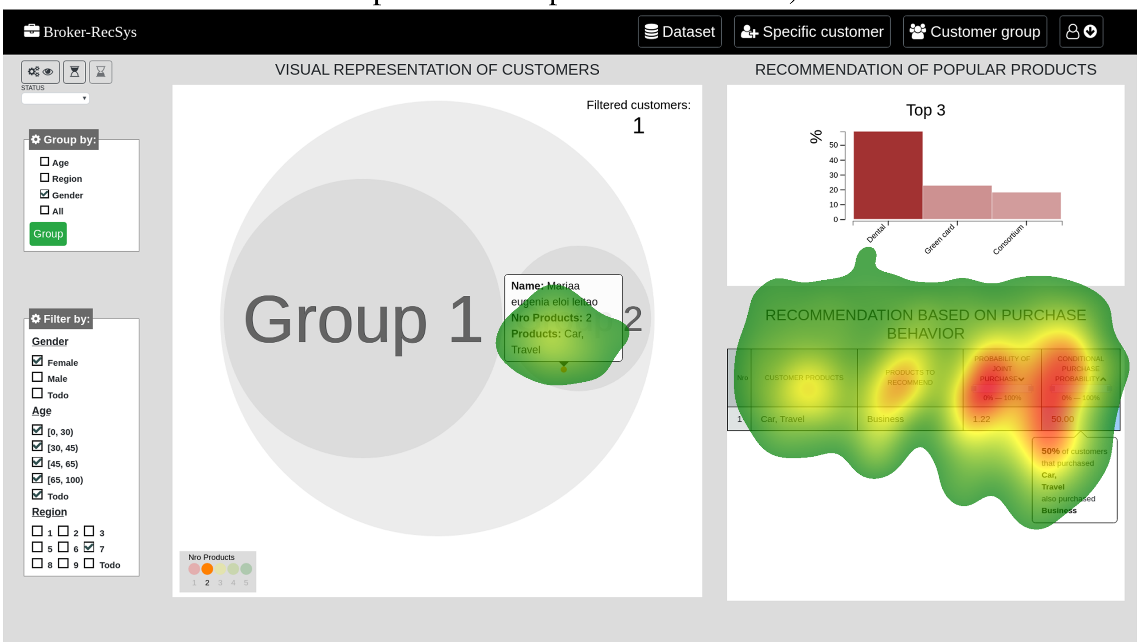


Figure 5.13: Recommendation for a group of customers: Heatmap-Fixations of the users' gaze for Task 1 (Cluster the clients based on gender and age. How many groups are formed? - Answer: 2 groups)

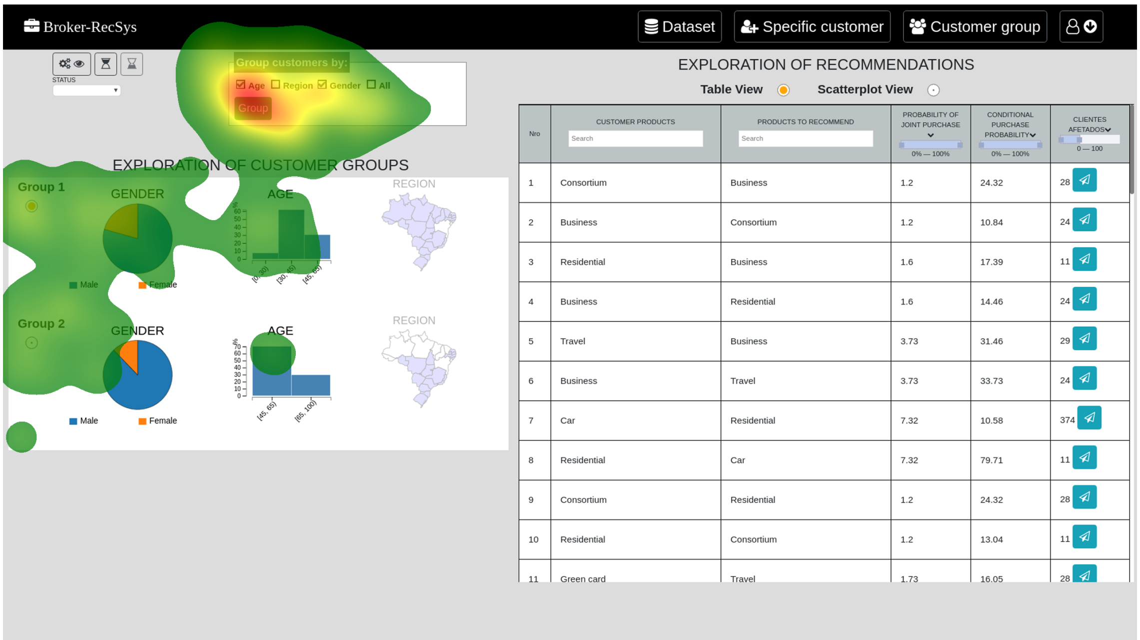


Figure 5.14: Recommendation for a group of customers: Heatmap-Fixations of the users' gaze for Task 2 (Select the second group. Which characteristics this group has? - Answer: mostly male, between 45 and 65 years old, and living in northwestern Brazil)

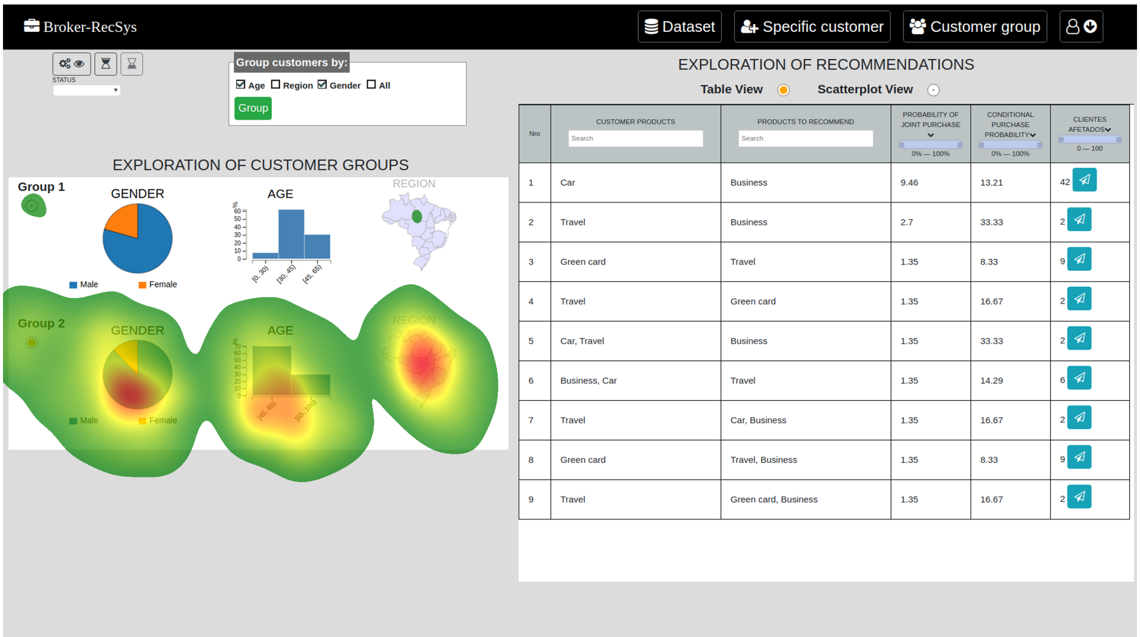


Figure 5.15: Recommendation for a group of customers: Heatmap-Fixations of the users' gaze for Task 3 (With a table view, sort and choose the recommendation with the highest conditional probability. What does this mean? - Answer: 33.33% of the clients that purchased car and travel insurance also purchased empresarial insurance)

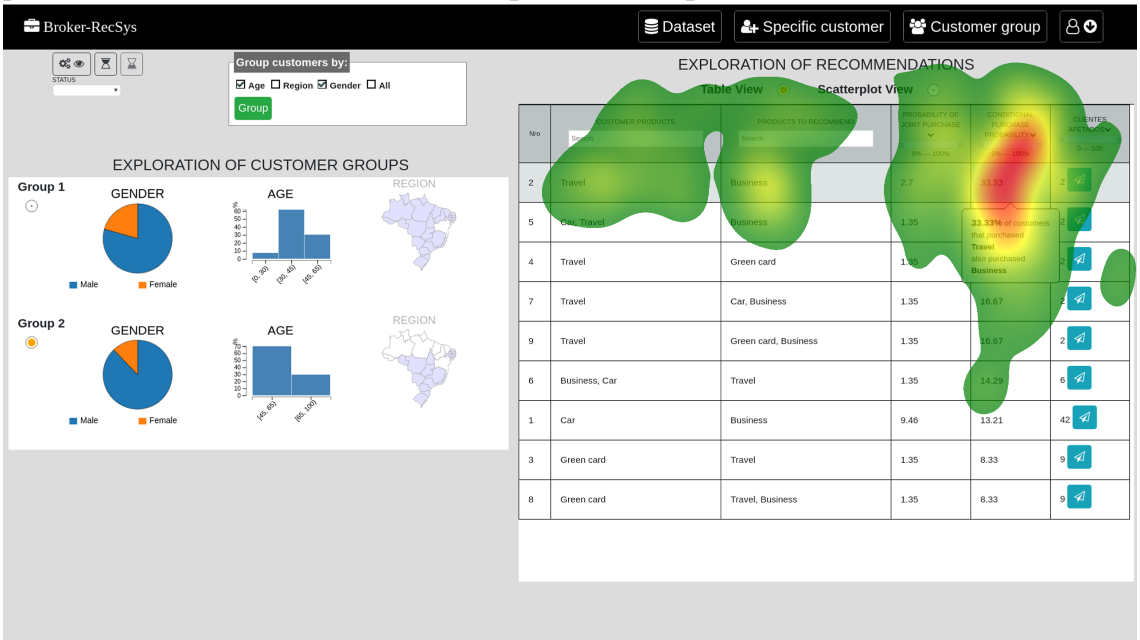
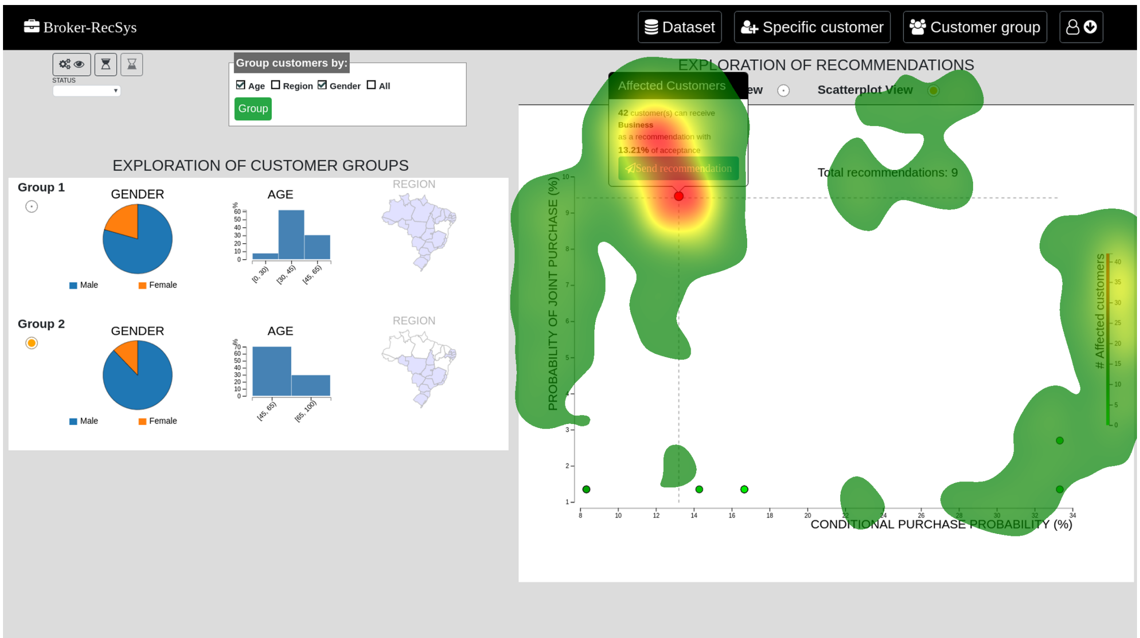


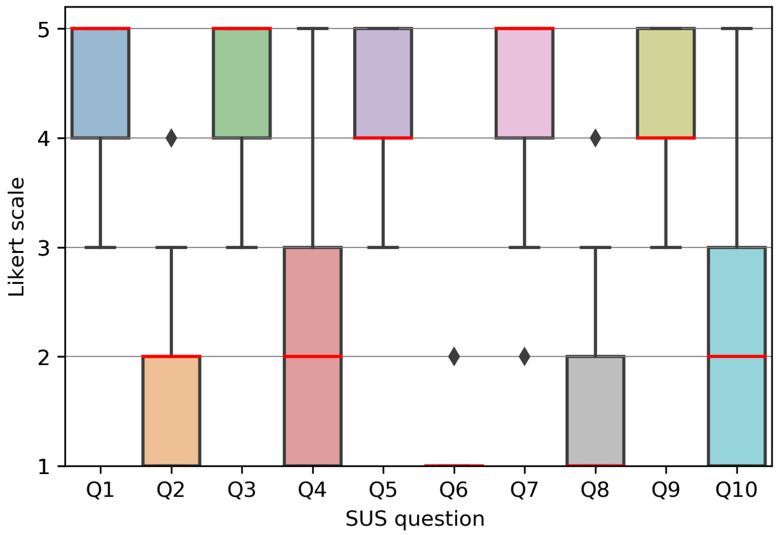
Figure 5.16: Recommendation for a group of customers: Heatmap-Fixations of the users' gaze for Task 4 (With the scatterplot view , which recommendation affects most the customers? What this means? - Answer: empresarial insurance because 38 customers can receive this recommendation)



5.3.3.3 System Usability Questionnaire

The SUS rating distribution from naive users is shown in Figure 5.17. The SUS score obtained from naive user's ratings was 81.9 (std=15.73). This result means that Broker-RecSys has excellent usability perceived by naive users using the subjective scale (see Subsection 5.1.1).

Figure 5.17: SUS-questionnaire rating distribution



5.3.3.4 Post-Questionnaire

Feedback is essential to understanding what users think about Broker-RecSys and how to improve the recommender system based on the naive user's perception. For this, the naive user fills a post-questionnaire where we ask them three questions. The first question (*What did you find the recommendation system interesting?*) is looking to obtain the most interesting features of Broker-RecSys as perceived by naive users.

For this question, some users mentioned that they found it interesting to cluster customers based on similarity and obtain different recommendations (popularity and purchase behavior). Others mentioned that they liked the visualization that makes the recommendation task easy and helps obtain a better understanding of the recommendation process. The second question (*What you did not find interesting about the recommendation system?*) looking for knowing what features of Broker-RecSys do not cause good perception in naive users. Here, some users mentioned that the results in a table were not easy to understand until they focused on the cell of the table and obtained an explanation of the recommendation measure. Others mentioned that the dashboard would be improved in colors e.g., the header of the table. Some users did not like scatterplot visualization. Finally, any users mentioned that the controls for cluster and filter customers were a little confused. The last question (*What would you like to change in the insurance product recommendation system?*) looking for what features in Broker-RecSys that can be improved based on the perception of naive users. Some users mentioned that they would like to obtain a graphic representation of recommendations. Any mentioned that they would like to change the contrast of some part of the dashboard.

5.4 Remote Evaluation

To evaluate the usefulness of Broker-RecSys for insurance brokers in the recommendation process, we asked three experts on the insurance domain to use our system. For this, Broker-RecSys was deployed on the internet using Heroku [2], a cloud platform. Insurance brokers were invited via email to participate in the usability and usefulness evaluation of Broker-RecSys.

5.4.1 Procedure

Firstly, the broker receives a google form that detailed information about the study, such as purpose and objectives, the time required to evaluate the recommender system, and the term of consent. After brokers accept participating in the study, it is collected the personal information of participants to know their profiles. Following, the participant receives infographic information to introduce him/her to the study as well as a demo showing the functionalities of Broker-RecSys. At this point, it is requested that brokers explore Broker-RecSys using their client portfolio. Then, for helping to explore the recommender system and its functionalities, it is requested that brokers perform a set of tasks (See Table 5.1) as same as naive users. After completing to perform the set tasks, the broker is able to evaluate Broker-RecSys in the usefulness dimension. For this, 9 questions are presented to insurance brokers (see Appendix E and F). Following, insurance brokers evaluate Broker-RecSys in the usability dimension based on the SUS questionnaire (See Appendix G). Finally, a post-questionnaire is provided to the broker to obtain the satisfaction of insurance brokers related to the use of Broker-RecSys. The procedure for this

[2]https://www.heroku.com/home

study is detailed in Table 5.7.

Table 5.7: Protocol followed for the broker user in the usability evaluation of the interactive recommender system

Step	Activity	Description	Measurement	Instrument	Analysis
1	Consent Form	The user accepts his participation in the study after reading the free and informed consent form.	–	Paper form (see Appendix B)	–
2	Introduction to study	Presentation of the moderator and his role in the test session, explanation of the protocol to follow. Additionally is given a brief explanation about recommender system and insurance domain.	–	Paper form and infographics (see Appendix C)	–
3	User Characterization (questionnaire):	The user fill his personal information such as name or alias, age, gender as well as his academic background, experience using interactive information systems and his physical and mental level of discomfort.	–	Paper form (see Appendix D)	Informative

Continued on next page

Table 5.7 – *Continued from previous page*

Step	Activity	Description	Measurement	Instrument	Analysis
4	Introduction to the prototype	The user is provided with a demo (4 minutes of duration) showing the functionality of the interactive recommendation system.	–	Video with audio	–
5	Broker-RecSys use	**Part 1: Free exploring time** (2 minutes) **Part 2: Free task execution**	–	–	–
6	Usefulness questionnaire	The user answer question related to the utility of the recommender system	Usefulness	Paper form (see Appendix E and F)	Boxplot Visualization (range 0-5)
7	System Usability Scale (SUS) - questionnaire	10-item Likert scale questionnaire to measure the level of use (ease or lack) of the system.	Satisfaction	Paper form (see Appendix G)	Boxplot visualization (range 0-5) and SUS score

Continued on next page

Table 5.7 – *Continued from previous page*

Step	Activity	Description	Measurement	Instrument	Analysis
8	Post-test questionnaire	User opinion Questions related to the test session: What are your positive opinion? What are your negative opinion? What would have you like to change to the recommender system?	Satisfaction	Paper form (see Appendix H)	Transcribed and summarized

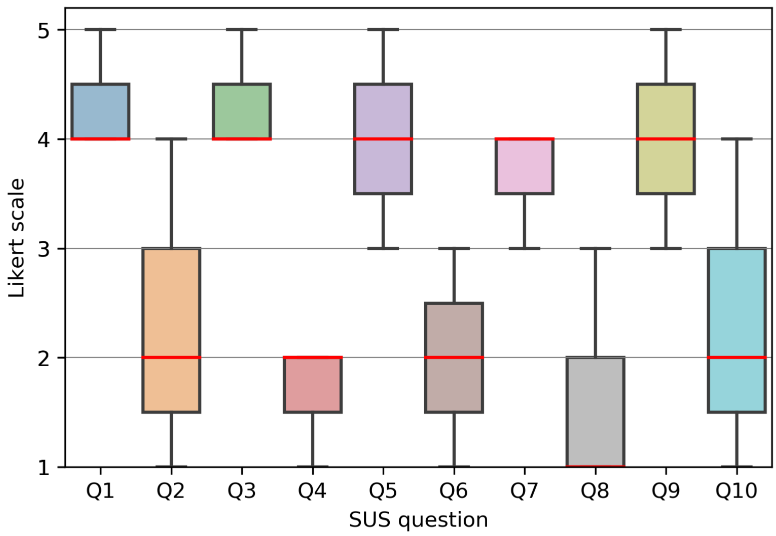
5.4.2 Results

In this subsection, we describe the results obtained in terms of usability and usefulness dimensions perceived by insurance brokers. With these two dimensions, we aim to answer the research question (**Broker-RecSys supports insurance brokers in the recommendation process for offers insurance products in their client portfolio?**).

5.4.2.1 System Usability Questionnaire

The SUS rating distribution from broker users is shown in Figure 5.18. The SUS obtained from insurance broker users was 75.8 (std=18.76), where a score higher than 68 is considered good usability based on the adjective scale (see subsection 5.1.1). In this case, the SUS score is less than the usability perceived by naive users 81.9 but greater than 68. However, in both cases, the SUS scores' adjective scale denotes that Broker-RecSys has good usability.

Figure 5.18: SUS-questionnaire rating distribution

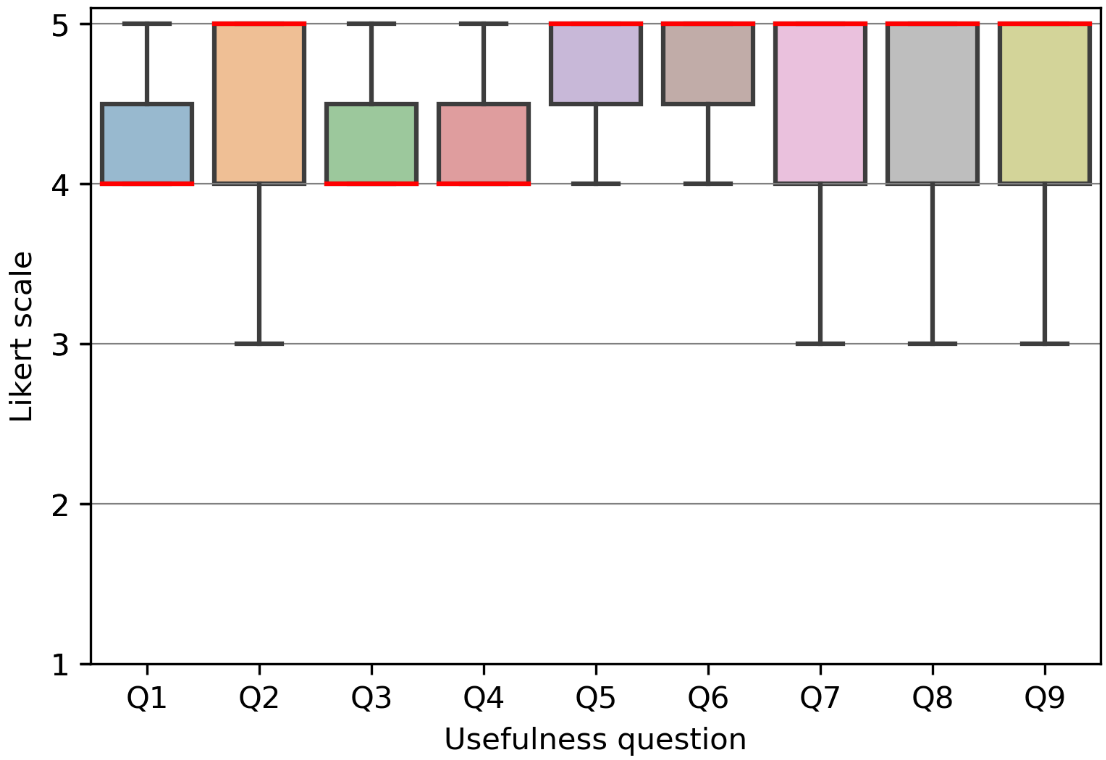


5.4.2.2 System Usefulness Questionnaire

The usefulness questionnaire was designed to measure the level of utility of Broker-RecSys perceived by insurance brokers. Five questions are formulated for a specific customer recommendation and four questions for a group of customer (See Appendix E and F). Figure 5.19 shows the distribution of rating by brokers in each usefulness question. We can see the highest median rating of over 4 points in all usefulness questions greater than

2.5. This result shows that Broker-RecSys is considered useful by an insurance broker in the recommendation process.

Figure 5.19: Usefulness-questionnaire rating distribution



5.4.2.3 Post-Questionnaire

Broker users were invited to answer three questions in order to obtain feedback from them. In the first question (*What did you find the recommendation system interesting?*), one broker answered that he found the diverse visualizations of the data and the approach used for the recommendation interesting. The other one mentioned that he found interesting the recommendation of products based on group tendency (popularity). The last one mentioned that he found interesting the exploration of customer groups visually and the possible recommendations for these customers. In the second question (*What you did not find interesting about the recommendation system?*), insurance brokers mentioned that it is useful the recommendation based on popularity and purchase behavior but only as an entry point instead of a personal assumption. Finally, the third question (*What would you like to change in the insurance product recommendation system?*), insurance brokers did not mention something to change in the system.

5.5 Discussion

Nowadays, recommender systems are vital in several domains. In the insurance domain, recommender systems help agents/brokers recommend the most relevant products. Although the insurance recommender systems achieved high accuracy in the prediction of recommendations, these systems have a lack in the user experience. Users need to know not only what products recommend, moreover, why, and how the recommendation was obtained as well as, if possible, participate in the recommendation process.

In the selling task, activities of insurance brokers consist of exploring and identifying potential customers to offer insurance products and identifying interesting recommendations for a group of customers to perform a marketing campaign. In this sense, Broker-RecSys is designed to help insurance brokers into the recommendation process.

We propose Broker-RecSys, an interactive recommender system framework that integrates user interaction, data visualization, and data mining methods to alleviate several drawbacks related to recommender systems such as cold-start problem, controllability, explanation, and user experience in the recommendation task. The finding obtained in this research work indicates that Broker-RecSys is usable and usefulness for an insurance broker in the recommendation process.

In the following subsections, we discuss the results obtained in the evaluation of Broker-RecSys in the usability and usefulness dimensions.

5.5.1 Usability Dimension

The usability level of Broker-RecSys is determined by naive and broker users. Naive users are users without knowledge about the insurance domain. On the contrary, broker users are users that have strong knowledge about the insurance domain and the tasks involving it.

To determine the usability level of Broker-RecSys, we formulate a research question **Broker-RecSys turn able naive users to perform insurance products recommendation tasks?**. With this formulation, we want to demonstrate that if naive users can use and perform insurance tasks related to the recommendation tasks quickly and without much effort, the tool will be fast and clearly to use. In the local fashion, we evaluate usability of Broker-RecSys in terms of effectiveness, efficiency and satisfaction (See subsection 5.3.3.1 and 5.4.2.3).In the effectiveness metric, how well users get complete the user tasks, we obtained 96.19% and 92.86% in completion the user's tasks in the recommendation process for a specific customer and a group of customers, respectively. On the other hand, we measure the efficiency of the tool based on how many time required the user to perform the recommendation process for a specific customer and a group of customers. In results, we obtained 1.58 and 1.47 minutes for each recommendation process, respectively. Besides, we explore the visual stimulus distribution of eye fixation from users during the performing of tasks. The heatmap visualization shows an accurate area related to the tasks performed by naive users 5.3.3.2. Additionally, we evaluate the satisfaction of naive users in the use of Broker-RecSys. The naive users are asked to answer three questions: what did you find interesting about the recommendation system? What did not you find interesting about the recommendation system? and What would you like to change in the insurance product recommendation system?. The first question, naive users, mentioned that they found interesting the user interaction and visualizations involving in the recommendation process. In the second question, naive users mentioned that they do not like some parts of the dashboard, such as scatterplot, which is more complicated to interpret than the table visualization. In the last question, naive users suggest making some visual changes in the appearance of the dashboard and insurance product representation.

Additionally, we perform an evaluation based on the SUS questionnaire to evaluate the usability perception of users (naive and insurance broker users) in the use of the tool. Results showed that Broker-RecSys obtained a usability score of 81.9 (std=15.73) from naive users and 75.8 (std=18.76) from insurance brokers, respectively. In both cases, the usability is greater than 68 that corresponds to good usability according to the subjective scale form the SUS score results (See Subsubsection 5.3.3.3 and 5.4.2.1). Using the two-sample t-test method with $\alpha = 0.05$ of confidence, we determine p-value = 0.64 where $p > \alpha$; this indicates that there is no significant difference between the two groups of users respect to the SUS score.

5.5.2 Usefulness Dimension

To evaluate the usefulness of Broker-RecSys for the target users i.e., insurance brokers, we formulate a set of usefulness questions (see Appendix E and F) related to the recommendation task. The usefulness questions intend to obtain the perception of utility from insurance brokers using Broker-RecSys in their client portfolio. With the results obtained from the usefulness questions, we aim to answer our second research question, **Broker-RecSys supports insurance brokers in the recommendation process for offers insurance products in their client portfolio?**. Results obtained show that the insurance broker rated with a higher score Broker-RecSys functionalities over or equal to 4. These results showed that insurance brokers consider Broker-RecSys usefulness for their daily activities related to insurance products recommendation.

On the other hand, similar to naive users, we asked to broker to answer three questions to measure the satisfaction degree in the use of Broker-RecSys. In the first questions(what did you find the recommendation system interesting?), brokers answer that found interesting the visualizations because it helps them to understand their client portfolio data easily to perform the recommendation of products. In the second question (What did not you find interesting about the recommendation system?), they mentioned that they did not like that the system does not integrate a multicriteria approach, i.e., enable put additional features of insurance products such as benefits, risks, premiums, terms, and others. In the last question (what would you like to change in the insurance product recommendation system?), they do not mention anything to change in Broker-RecSys.

Based on the usability and usefulness results, we can answer our two research questions with an affirmative response. **Broker-RecSys turn able naive users to perform insurance products recommendation tasks?** Yes, Broker-RecSys turn able naive users to perform tasks related to insurance domain in an easy and fastest manner. **Broker-RecSys supports insurance brokers in the recommendation process for offers insurance products in their client portfolio?** Yes, Broker-RecSys supports insurance brokers in the recommendation process for offers insurance products in their client portfolio.

6 CONCLUSION

In this work, we presented and evaluated Broker-RecSys, an interactive insurance product recommender system framework to support brokers in the recommendation process for offering insurance products in their customer portfolio at two levels: recommendations for a specific customer; and recommendations for a group of customers.

Broker-RecSys integrate user interaction, data visualization, and data mining methods to alleviate several drawbacks such as controllability, user interaction interface, explanation, and cold-start problem. Controllability allowed insurance brokers to obtain variety in recommendations for offers in their client portfolio. The user interaction interface helped insurance brokers in the human visual interpretation during the recommendation process. Explanation supported insurance brokers in interpreting the recommendations, and alleviating cold-start problem allowed insurance brokers to offer recommendations for new users without prior information about the insurance products purchase.

We evaluated Broker-RecSys in terms of usability and usefulness dimensions with two types of participants: 21 naive user participants and 3 expert insurance brokers. The evaluation occurs for naive and broker users locally and remotely, respectively. We combined evaluation based on questionnaires and the evaluation based on the eye-tracking analysis. Results in the usability dimension evaluation showed that Broker-RecSys is easy and fast to use. On the other hand, the usefulness dimension evaluation results showed that Broker-RecSys is useful to support insurance brokers in the recommendation process. Our results suggest that combining user interaction, data visualization ,and data mining can support users in the recommendation process.

For future works, we plan to introduce a knowledge-based insurance recommender approach to Broker-RecSys to enable more sophisticated recommendations based on the restriction of insurance plans. To measure cluster formation accuracy, we plan to use a classifier algorithm and take this information and show to insurance brokers as an indicator of the degree of similarity. To improve the visual understanding of recommendation in Broker-RecSys, we plan to perform a comparative study with alternatives for visualization of recommendations based on product information. To improve user interface, we plan to perform changes in interface design considering the colorblind user problem. Also, we look to answer questions such as: How helpful are the recommendations made? or How many recommendations were converted on sales?.

REFERENCES

ADOMAVICIUS, G.; TUZHILIN, A. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. **IEEE Transactions on Knowledge & Data Engineering**, IEEE, n. 6, p. 734–749, 2005.

AGGARWAL, C. C. Content-based recommender systems. In: **Recommender Systems**. [S.l.]: Springer, 2016. p. 139–166.

AGGARWAL, C. C. Knowledge-based recommender systems. In: ____. **Recommender Systems: The Textbook**. Cham: Springer International Publishing, 2016. p. 167–197. ISBN 978-3-319-29659-3. Disponível em: <https://doi.org/10.1007/978-3-319-29659-3_5>.

AGGARWAL, C. C.; PHILIP, S. Y. Data mining techniques for associations, clustering and classification. In: SPRINGER. **Pacific-Asia Conference on Knowledge Discovery and Data Mining**. [S.l.], 1999. p. 13–23.

AGRAWAL, R.; IMIELIŃSKI, T.; SWAMI, A. Mining association rules between sets of items in large databases. In: ACM. **Acm sigmod record**. [S.l.], 1993. v. 22, n. 2, p. 207–216.

AGRAWAL, R.; SRIKANT, R. et al. Fast algorithms for mining association rules. In: **Proc. 20th int. conf. very large data bases, VLDB**. [S.l.: s.n.], 1994. v. 1215, p. 487–499.

ALSLAITY, A.; TRAN, T. Towards persuasive recommender systems. In: IEEE. **2019 IEEE 2nd International Conference on Information and Computer Technologies (ICICT)**. [S.l.], 2019. p. 143–148.

AMATRIAIN, X. et al. Data mining methods for recommender systems. In: **Recommender systems handbook**. [S.l.]: Springer, 2011. p. 39–71.

AMIN, S. A.; PHILIPS, J.; TABRIZI, N. Current trends in collaborative filtering recommendation systems. In: SPRINGER. **World Congress on Services**. [S.l.], 2019. p. 46–60.

ANDRIENKO, G. et al. Visual analytics of mobility and transportation: State of the art and further research directions. **IEEE Transactions on Intelligent Transportation Systems**, IEEE, v. 18, n. 8, p. 2232–2249, 2017.

AOIKE, T. et al. Utilising crowd information of tourist spots in an interactive tour recommender system. In: **Information and Communication Technologies in Tourism 2019**. [S.l.]: Springer, 2019. p. 27–39.

ARBATLI, E. C.; JOHANSEN, R. M. **A Heatmap for Monitoring Systemic Risk in Norway**. [S.l.], 2017.

ATAUCHI, P. D.; NEDEL, L.; GALANTE, R. Broker-insights: An interactive and visual recommendation system for insurance brokerage. In: SPRINGER. **Computer Graphics International Conference**. [S.l.], 2019. p. 155–166.

BAFNA, P.; PRAMOD, D.; VAIDYA, A. Document clustering: Tf-idf approach. In: IEEE. **2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)**. [S.l.], 2016. p. 61–66.

BANGOR, A.; KORTUM, P.; MILLER, J. Determining what individual sus scores mean: Adding an adjective rating scale. **Journal of usability studies**, Citeseer, v. 4, n. 3, p. 114–123, 2009.

BLASCHECK, T. et al. Visualization of eye tracking data: A taxonomy and survey. In: WILEY ONLINE LIBRARY. **Computer Graphics Forum**. [S.l.], 2017. v. 36, n. 8, p. 260–284.

BOBADILLA, J. et al. Recommender systems survey. **Knowledge-based systems**, Elsevier, v. 46, p. 109–132, 2013.

BOSTOCK, M.; OGIEVETSKY, V.; HEER, J. D^3 data-driven documents. **IEEE transactions on visualization and computer graphics**, IEEE, v. 17, n. 12, p. 2301–2309, 2011.

BRESSAN, M. et al. The limits of popularity-based recommendations, and the role of social ties. In: ACM. **Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining**. [S.l.], 2016. p. 745–754.

BROOKE, J. et al. Sus-a quick and dirty usability scale. **Usability evaluation in industry**, London–, v. 189, n. 194, p. 4–7, 1996.

BRUZZESE, D.; DAVINO, C. Visual mining of association rules. In: **Visual Data Mining**. [S.l.]: Springer, 2008. p. 103–122.

ÇANO, E.; MORISIO, M. Hybrid recommender systems: A systematic literature review. **Intelligent Data Analysis**, IOS Press, v. 21, n. 6, p. 1487–1524, 2017.

CHEN, M.; MAO, S.; LIU, Y. Big data: A survey. **Mobile networks and applications**, Springer, v. 19, n. 2, p. 171–209, 2014.

CHENG, L. C.; WU, C.-C.; CHEN, C.-Y. Behavior analysis of customer churn for a customer relationship system: an empirical case study. **Journal of Global Information Management (JGIM)**, IGI Global, v. 27, n. 1, p. 111–127, 2019.

COMANICIU, D.; MEER, P. Mean shift: A robust approach toward feature space analysis. **IEEE Transactions on Pattern Analysis & Machine Intelligence**, IEEE, n. 5, p. 603–619, 2002.

COMANICIU, D.; RAMESH, V.; MEER, P. The variable bandwidth mean shift and data-driven scale selection. In: IEEE. **Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001**. [S.l.], 2001. v. 1, p. 438–445.

COX, M. A.; COX, T. F. Multidimensional scaling. In: **Handbook of data visualization**. [S.l.]: Springer, 2008. p. 315–347.

DEGELE, J. et al. Identifying e-scooter sharing customer segments using clustering. In: IEEE. **2018 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC)**. [S.l.], 2018. p. 1–8.

DRUSCH, G.; BASTIEN, J.; PARIS, S. Analysing eye-tracking data: From scanpaths and heatmaps to the dynamic visualisation of areas of interest. **Advances in Science, Technology, Higher Education and Society in the Conceptual Age: STHESCA**, AHFE Conference, v. 20, p. 205, 2014.

DU, F. et al. Personalizable and interactive sequence recommender system. In: **Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems**. [S.l.: s.n.], 2018. p. 1–6.

FORSMAN, J. et al. Integrated information visualization to support decision making for use of antibiotics in intensive care: design and usability evaluation. **Informatics for Health and Social Care**, Taylor & Francis, v. 38, n. 4, p. 330–353, 2013.

FU, B.; NOY, N. F.; STOREY, M.-A. Eye tracking the user experience–an evaluation of ontology visualization techniques. **Semantic Web**, IOS Press, v. 8, n. 1, p. 23–41, 2017.

GARCÍA, S.; LUENGO, J.; HERRERA, F. **Data preprocessing in data mining**. [S.l.]: Springer, 2015.

GKIKA, S.; LEKAKOS, G. Investigating the effectiveness of persuasion strategies on recommender systems. In: IEEE. **2014 9th International Workshop on Semantic and Social Media Adaptation and Personalization**. [S.l.], 2014. p. 94–97.

GOMEZ-URIBE, C. A.; HUNT, N. The netflix recommender system: Algorithms, business value, and innovation. **ACM Transactions on Management Information Systems (TMIS)**, ACM, v. 6, n. 4, p. 13, 2016.

GÖRTLER, J. et al. Bubble treemaps for uncertainty visualization. **IEEE transactions on visualization and computer graphics**, IEEE, v. 24, n. 1, p. 719–728, 2017.

HAHSLER, M. arulesviz: Interactive visualization of association rules with r. **R Journal**, v. 9, n. 2, 2017.

HAHSLER, M.; CHELLUBOINA, S. Visualizing association rules: Introduction to the r-extension package arulesviz. **R project module**, p. 223–238, 2011.

HAHSLER, M.; KARPIENKO, R. Visualizing association rules in hierarchical groups. **Journal of Business Economics**, Springer, v. 87, n. 3, p. 317–335, 2017.

HE, C.; PARRA, D.; VERBERT, K. Interactive recommender systems: A survey of the state of the art and future research challenges and opportunities. **Expert Systems with Applications**, Elsevier, v. 56, p. 9–27, 2016.

HERAS, S. et al. Using argumentation to persuade students in an educational recommender system. In: SPRINGER. **International Conference on Persuasive Technology**. [S.l.], 2017. p. 227–239.

HINDUJA, A.; PANDEY, M. Multicriteria recommender system for life insurance plans based on utility theory. **Indian Journal of Science and Technology**, v. 10, n. 14, p. 1–8, 2017.

HWANGBO, H.; KIM, Y. S.; CHA, K. J. Recommendation system development for fashion retail e-commerce. **Electronic Commerce Research and Applications**, Elsevier, v. 28, p. 94–101, 2018.

IEC, I. 25010 (2011). systems and software engineering–systems and software quality requirements and evaluation (square)–system and software quality models. **International Organization for Standardization, Geneva, Switzerland**, 2011.

ISO, I. 9241-11: 2018: Ergonomics of human-system interaction—part 11: usability: definitions and concepts. **International Organization for Standardization, Geneva, Switzerland**, 2018.

JAIN, A. K. Data clustering: 50 years beyond k-means. **Pattern recognition letters**, Elsevier, v. 31, n. 8, p. 651–666, 2010.

JIN, Y. et al. Go with the flow: effects of transparency and user control on targeted advertising using flow charts. In: ACM. **Proceedings of the International Working Conference on Advanced Visual Interfaces**. [S.l.], 2016. p. 68–75.

JOLLIFFE, I. T.; CADIMA, J. Principal component analysis: a review and recent developments. **Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences**, The Royal Society Publishing, v. 374, n. 2065, p. 20150202, 2016.

JUGOVAC, M.; JANNACH, D. Interacting with recommenders—overview and research directions. **ACM Transactions on Interactive Intelligent Systems (TiiS)**, ACM, v. 7, n. 3, p. 10, 2017.

KAEWKIRIYA, T. Framework for prediction of life insurance customers based on multi-algorithms. In: IEEE. **2017 Seventh International Conference on Innovative Computing Technology (INTECH)**. [S.l.], 2017. p. 74–78.

KARIMI, M.; JANNACH, D.; JUGOVAC, M. News recommender systems–survey and roads ahead. **Information Processing & Management**, Elsevier, v. 54, n. 6, p. 1203–1227, 2018.

KLEINSCHMIDT, P.; HANRIEDER, W. The future of sensors, materials science or software engineering? **Sensors and Actuators A: Physical**, Elsevier, v. 33, n. 1-2, p. 5–17, 1992.

KUMAR, P.; SINGH, D. Integrating data mining and ahp for life insurance product recommendation. In: SPRINGER. **International Conference on Computational Intelligence and Information Technology**. [S.l.], 2011. p. 596–602.

KURZHALS, K. et al. Evaluating visual analytics with eye tracking. In: ACM. **Proceedings of the Fifth Workshop on Beyond Time and Errors: Novel Evaluation Methods for Visualization**. [S.l.], 2014. p. 61–69.

KYLVAJA, M.; KUMPULAINEN, P.; KONU, A. Application of data clustering for automated feedback generation about student well-being. In: **Proceedings of the 1st ACM SIGSOFT International Workshop on Education through Advanced Software Engineering and Artificial Intelligence**. [S.l.: s.n.], 2019. p. 21–26.

LECUN, Y. The mnist database of handwritten digits. **http://yann. lecun. com/exdb/mnist/**, 1998.

LESAGE, L. et al. A recommendation system for car insurance. **European Actuarial Journal**, Springer, p. 1–22, 2020.

LEWIS, C.; RIEMAN, J. Task-centered user interface design. **A Practical Introduction**, 1993.

LI, W. et al. Parallel coordinates based visualization for high-dimensional data. In: **Proceedings of the 2019 3rd International Conference on Big Data Research**. [S.l.: s.n.], 2019. p. 161–165.

LIKA, B.; KOLOMVATSOS, K.; HADJIEFTHYMIADES, S. Facing the cold start problem in recommender systems. **Expert Systems with Applications**, Elsevier, v. 41, n. 4, p. 2065–2073, 2014.

LIU, E. et al. Navigating health literacy using interactive data visualisation. In: IEEE. **2016 IEEE 16th International Conference on Bioinformatics and Bioengineering (BIBE)**. [S.l.], 2016. p. 44–50.

LIU, Z. et al. Coreflow: Extracting and visualizing branching patterns from event sequences. In: WILEY ONLINE LIBRARY. **Computer Graphics Forum**. [S.l.], 2017. v. 36, n. 3, p. 527–538.

LOEPP, B.; ZIEGLER, J. Towards interactive recommending in model-based collaborative filtering systems. In: **Proceedings of the 13th ACM Conference on Recommender Systems**. [S.l.: s.n.], 2019. p. 546–547.

LOH, S. et al. Identifying similar users by their scientific publications to reduce cold start in recommender systems. In: **WEBIST**. [S.l.: s.n.], 2009. v. 9, p. 593–600.

LU, J. et al. Recommender system application developments: a survey. **Decision Support Systems**, Elsevier, v. 74, p. 12–32, 2015.

MAATEN, L. v. d.; HINTON, G. Visualizing data using t-sne. **Journal of machine learning research**, v. 9, n. Nov, p. 2579–2605, 2008.

MAO, X. et al. Utilizing multi-source data in popularity prediction for shop-type recommendation. **Knowledge-Based Systems**, Elsevier, v. 165, p. 253–267, 2019.

MCGEE, F. et al. The state of the art in multilayer network visualization. In: WILEY ONLINE LIBRARY. **Computer Graphics Forum**. [S.l.], 2019. v. 38, n. 6, p. 125–149.

MELE, M. L.; FEDERICI, S. Gaze and eye-tracking solutions for psychological research. **Cognitive processing**, Springer, v. 13, n. 1, p. 261–265, 2012.

MIAO, Z. et al. Joint prediction of rating and popularity for cold-start item by sentinel user selection. **IEEE Access**, IEEE, v. 4, p. 8500–8513, 2016.

MONSIVAIS, P. et al. Data visualisation to support obesity policy: case studies of data tools for planning and transport policy in the uk. **International Journal of Obesity**, Nature Publishing Group, v. 42, n. 12, p. 1977–1986, 2018.

MUKHERJI, A. et al. Fire: a two-level interactive visualization for deep exploration of association rules. **International Journal of Data Science and Analytics**, Springer, p. 1–26, 2018.

NAJAFABADI, M. K.; MOHAMED, A. H.; MAHRIN, M. N. A survey on data mining techniques in recommender systems. **Soft Computing**, Springer, v. 23, n. 2, p. 627–654, 2019.

NATARAJAN, S.; MOH, M. Recommending news based on hybrid user profile, popularity, trends, and location. In: IEEE. **2016 International Conference on Collaboration Technologies and Systems (CTS)**. [S.l.], 2016. p. 204–211.

Netflix. **How Netflix's Recommendations System Works**. 2019. <https://help.netflix.com/en/node/100639>, Last accessed on 2019-10-17.

O'DONOVAN, J. et al. Peerchooser: visual interactive recommendation. In: ACM. **Proceedings of the SIGCHI Conference on Human Factors in Computing Systems**. [S.l.], 2008. p. 1085–1088.

OLSON, D. L.; LAUHOFF, G. Association rules. In: **Descriptive Data Mining**. [S.l.]: Springer, 2019. p. 67–76.

PANDEY, A. K.; RAJPOOT, D. S. Resolving cold start problem in recommendation system using demographic approach. In: IEEE. **2016 International Conference on Signal Processing and Communication (ICSC)**. [S.l.], 2016. p. 213–218.

PARYUDI, I.; ASHARI, A.; TJOA, A. M. Personality estimation using demographic data in a personality-based recommender system: A proposal. In: **Proceedings of the 21st International Conference on Information Integration and Web-based Applications & Services**. [S.l.: s.n.], 2019. p. 156–160.

PATEL, K. A.; THAKRAL, P. The best clustering algorithms in data mining. In: IEEE. **2016 International Conference on Communication and Signal Processing (ICCSP)**. [S.l.], 2016. p. 2042–2046.

PAZZANI, M. J. A framework for collaborative, content-based and demographic filtering. **Artificial intelligence review**, Springer, v. 13, n. 5-6, p. 393–408, 1999.

PERDANA, A.; ROBB, A.; ROHDE, F. Interactive data visualisation for accounting information: a three-fit perspective. **Behaviour & Information Technology**, Taylor & Francis, v. 38, n. 1, p. 85–100, 2019.

PIRINGER, H.; KOSARA, R.; HAUSER, H. Interactive focus+ context visualization with linked 2d/3d scatterplots. In: IEEE. **Proceedings. Second International Conference on Coordinated and Multiple Views in Exploratory Visualization, 2004.** [S.l.], 2004. p. 49–60.

QAZI, M. et al. An insurance recommendation system using bayesian networks. In: ACM. **Proceedings of the Eleventh ACM Conference on Recommender Systems**. [S.l.], 2017. p. 274–278.

QAZI, M. et al. Designing and deploying insurance recommender systems using machine learning. **Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery**, Wiley Online Library, p. e1363, 2020.

RAHMAN, S. S. A.; NORMAN, A. A.; SOON, K. Myins: a cbr e-commerce application for insurance policies. **Electronic Commerce Research**, v. 5, n. 1, p. 373–380, 2006.

RAMASAMY, S.; NIRMALA, K. Disease prediction in data mining using association rule mining and keyword based clustering algorithms. **International Journal of Computers and Applications**, Taylor & Francis, v. 42, n. 1, p. 1–8, 2020.

RASCHKE, M.; BLASCHECK, T.; BURCH, M. Visual analysis of eye tracking data. In: **Handbook of human centric visualization**. [S.l.]: Springer, 2014. p. 391–409.

RAZAK, S. F. A.; TAN, S. C.; LIM, W.-S. Developing a rule base for recommending insurance products. In: SPRINGER. **The 8th International Conference on Robotic, Vision, Signal Processing & Power Applications**. [S.l.], 2014. p. 321–326.

REDDY, S. et al. Content-based movie recommendation system using genre correlation. In: **Smart Intelligent Computing and Applications**. [S.l.]: Springer, 2019. p. 391–397.

RICCI, F.; ROKACH, L.; SHAPIRA, B. Recommender systems: introduction and challenges. In: **Recommender systems handbook**. [S.l.]: Springer, 2015. p. 1–34.

ROKACH, L. et al. Recommending insurance riders. In: ACM. **Proceedings of the 28th Annual ACM Symposium on Applied Computing**. [S.l.], 2013. p. 253–260.

SALTON, G. Automatic text processing: The transformation, analysis, and retrieval of. **Reading: Addison-Wesley**, v. 169, 1989.

SÁNCHEZ-CORCUERA, R. et al. Persuasion-based recommender system ensambling matrix factorisation and active learning models. **Personal and Ubiquitous Computing**, Springer, p. 1–11, 2020.

SHINDE, S. K.; KULKARNI, U. Hybrid personalized recommender system using centering-bunching based clustering algorithm. **Expert Systems with Applications**, Elsevier, v. 39, n. 1, p. 1381–1387, 2012.

SOBHANAM, H.; MARIAPPAN, A. Addressing cold start problem in recommender systems using association rules and clustering technique. **2013 International Conference on Computer Communication and Informatics**, p. 1–5, 2013.

SOFIA, G. et al. Investigating the role of personality traits and influence strategies on the persuasive effect of personalized recommendations. In: **4th Workshop on Emotions and Personality in Personalized Systems (EMPIRE)**. [S.l.: s.n.], 2016. p. 9.

SRIDEVI, M.; RAO, R. R. Decors: A simple and efficient demographic collaborative recommender system for movie recommendation. **Advances in Computational Sciences and Technology**, v. 10, n. 7, p. 1969–1979, 2017.

SYMEONIDIS, P.; NANOPOULOS, A.; MANOLOPOULOS, Y. Moviexplain: a recommender system with explanations. **RecSys**, v. 9, p. 317–320, 2009.

TARNOWSKA, K. A.; RAS, Z. W.; JASTREBOFF, P. J. Recommender solutions overview. In: **Decision Support System for Diagnosis and Treatment of Hearing Disorders**. [S.l.]: Springer, 2017. p. 17–33.

TATAR, A. et al. From popularity prediction to ranking online news. **Social Network Analysis and Mining**, Springer, v. 4, n. 1, p. 174, 2014.

TINTAREV, N.; MASTHOFF, J. A survey of explanations in recommender systems. In: IEEE. **2007 IEEE 23rd international conference on data engineering workshop**. [S.l.], 2007. p. 801–810.

TOBII. **User's Manual Tobii X2-30 Eye Tracker**. 2014. <https://www.tobiipro.com/siteassets/tobii-pro/user-manuals/tobii-pro-x2-30-eyetracker-user-manual.pdf>. [Online; accessed 08-February-2019].

TSAI, C.-H.; BRUSILOVSKY, P. Explaining recommendations in an interactive hybrid social recommender. In: **Proceedings of the 24th International Conference on Intelligent User Interfaces**. [S.l.: s.n.], 2019. p. 391–396.

VALDEZ, A. C.; ZIEFLE, M.; VERBERT, K. Hci for recommender systems: the past, the present and the future. In: ACM. **Proceedings of the 10th ACM Conference on Recommender Systems**. [S.l.], 2016. p. 123–126.

VERHOEF, P. C.; DONKERS, B. Predicting customer potential value an application in the insurance industry. **Decision support systems**, Elsevier, v. 32, n. 2, p. 189–199, 2001.

VOZALIS, M.; MARGARITIS, K. G. Collaborative filtering enhanced by demographic correlation. In: **AIAI symposium on professional practice in AI, of the 18th world computer congress**. [S.l.: s.n.], 2004.

WALDNER, M. et al. A comparison of radial and linear charts for visualizing daily patterns. **IEEE transactions on visualization and computer graphics**, IEEE, v. 26, n. 1, p. 1033–1042, 2019.

WANG, C.-H. Association rule mining and cognitive pairwise rating based portfolio analysis for product family design. **Journal of Intelligent Manufacturing**, Springer, v. 30, n. 4, p. 1911–1922, 2019.

WANG, C.-H. et al. Mental disorders and medical comorbidities: Association rule mining approach. **Perspectives in psychiatric care**, Wiley Online Library, v. 55, n. 3, p. 517–526, 2019.

WANG, D. et al. A content-based recommender system for computer science publications. **Knowledge-Based Systems**, Elsevier, v. 157, p. 1–9, 2018.

WANG, X.; ZHANG, Y.; YAMASAKI, T. Earn more social attention: User popularity based tag recommendation system. In: **Companion Proceedings of the Web Conference 2020**. [S.l.: s.n.], 2020. p. 212–216.

WARD, M. O.; GRINSTEIN, G.; KEIM, D. **Interactive data visualization: foundations, techniques, and applications**. [S.l.]: CRC Press, 2010.

WILLIAMSON, B. Digital education governance: data visualization, predictive analytics, and 'real-time'policy instruments. **Journal of Education Policy**, Taylor & Francis, v. 31, n. 2, p. 123–141, 2016.

WITTEN, I. H. et al. **Data Mining: Practical machine learning tools and techniques**. [S.l.]: Morgan Kaufmann, 2016.

WONG, K.-C. A short survey on data clustering algorithms. In: IEEE. **2015 Second International Conference on Soft Computing and Machine Intelligence (ISCMI)**. [S.l.], 2015. p. 64–68.

XU, W. et al. A novel intelligence recommendation model for insurance products with consumer segmentation. **Journal of Systems Science and Information**, De Gruyter, v. 2, n. 1, p. 16–28, 2014.

YAN, Y.; XIE, H. Research on the application of data mining technology in insurance informatization. In: IEEE. **2009 Ninth International Conference on Hybrid Intelligent Systems**. [S.l.], 2009. v. 3, p. 202–205.

YANG, X. et al. Application of visualization techniques in automatic identification system data analysis. In: **Proceedings of the 5th International Conference on Frontiers of Educational Technologies**. [S.l.: s.n.], 2019. p. 63–67.

YOO, K.-H.; GRETZEL, U.; ZANKER, M. **Persuasive recommender systems: conceptual background and implications**. [S.l.]: Springer Science & Business Media, 2012.

YU, T.; SHEN, Y.; JIN, H. Towards hands-free visual dialog interactive recommendation. In: **Proceedings of the AAAI Conference on Artificial Intelligence**. [S.l.: s.n.], 2020. v. 34, n. 01, p. 1137–1144.

ZARRINKALAM, F.; KAHANI, M. A multi-criteria hybrid citation recommendation system based on linked data. In: IEEE. **2012 2nd International eConference on Computer and Knowledge Engineering (ICCKE)**. [S.l.], 2012. p. 283–288.

ZHANG, C.; ZHANG, S. **Association rule mining: models and algorithms**. [S.l.]: Springer-Verlag, 2002.

ZHANG, P. et al. Customer loyalty and customer relationship management. In: SPRINGER. **International Conference on Electronic Commerce, Web Application, and Communication**. [S.l.], 2011. p. 432–436.

ZHAO, S. et al. Who is talking about what: social map-based recommendation for content-centric social websites. In: ACM. **Proceedings of the fourth ACM conference on Recommender systems**. [S.l.], 2010. p. 143–150.

ZHAO, W. X. et al. Exploring demographic information in social media for product recommendation. **Knowledge and Information Systems**, Springer, v. 49, n. 1, p. 61–89, 2016.

ZHIYUAN, H. et al. Application of big data visualization in passenger flow analysis of shanghai metro network. In: IEEE. **2017 2nd IEEE International Conference on Intelligent Transportation Engineering (ICITE)**. [S.l.], 2017. p. 184–188.

ZIBRICZKY, D. Recommender systems meet finance: a literature review. In: **FINREC**. [S.l.: s.n.], 2016.

ZIHAYAT, M. et al. A utility-based news recommendation system. **Decision Support Systems**, Elsevier, v. 117, p. 14–27, 2019.

APPENDIX A — RESUMO EXPANDIDO

Os sistemas de recomendação são ferramentas que auxiliam aos usuários na tomada de decisão na escolha de um produto ou serviço baseado nas suas preferências, necessidades ou interesses. Os sistemas de recomendação são amplamente usados em muitas aplicações como marketing, noticias, comércio eletrônico e assim por diante. Não obstante, muitos desses sistemas atuam como uma caixa preta ao não levar em conta ao usuário no processo de recomendação.

A natureza da caixa preta limita o entendimento de como a recomendação foi obtida e porque essa recomendação é a mais adequada para o usuário. Em consequência, limita a aceitação da recomendação recebida pelo usuário. Por outro lado, os sistemas interativos de recomendação podem solucionar essas limitações, pois combinam métodos de interação com o usuário, visualização de informações e sistema de recomendação. Os sistemas interativos de recomendação permitem responder questões dos usuários como porque o cliente recebeu certa recomendação?, como foi obtida essa recomendação?, é possível que o usuário possa se envolver no processo de recomendação?.

No domínio de corretoras, o corretor de seguros é um intermediário entre a seguradora e o cliente. Os corretores de seguros oferecem, negociam e vendem produtos de seguros na sua carteira de clientes. Atividades relevantes que podem auxiliar ao corretor de seguro são: melhorar sua confiança no cliente a través de uma recomendação acertada de produtos de seguro baseado em estratificação de clientes, o seu lucro ao identificar potenciais clientes para oferecer produtos de seguros, oportunidades para fazer campanha de marketing em seu portfólio de clientes ao identificar potenciais recomendações de produtos para um grupo de clientes.

Este trabalho apresenta o Broker-RecSys, um framework de sistema interativo de recomendação de produtos de seguros para apoiar os corretores no processo de recomendação para oferecer produtos de seguros em seu portfólio. O sistema opera em dois níveis para oferecer recomendações: recomendações para um cliente específico; e recomendações para um grupo de clientes. Procurando oferecer recomendações personalizadas, o Broker-RecSys fornece um módulo para realizar segmentação de clientes com base em características específicas do cliente que são relevantes para o corretor. Dois tipos de recomendações são fornecidos pelo Broker-RecSys: com base na popularidade e no comportamento de compra. O Broker-RecSys integra diversas interações e métodos de visualização de dados para auxiliar ao corretor de seguros na representação visual de seus clientes, na identificação de potenciais clientes para oferecer produtos de seguros, na visualização das recomendações, na exploração de grupo de clientes a través da visualização da sumarização das informações de grupos estratificados, baseado em específicas características dos clientes. Todos eles auxiliam ao corretor no processso de recomendação para que seja fácil e rápido de realizar tarefas de recomendação na sua carteira de clientes. Ademais, Broker-RecSys permite ao corretor de seguro na recomendação de produtos de seguros para novos clientes, clientes que não tem informação de compras no passado.

O Broker-RecSys é avaliado nas dimensões de usabilidade e utilidade. Para a avaliação, foram usados métodos baseados em questionários e métodos baseados na análise de rastreamento ocular. A avaliação foi feita local e remotamente. No experimento local, participaram 21 usuários que não tinham conhecimentos no domínio de seguros. As medidas consideradas na avaliação local foram: eficiência, eficácia, usabilidade subjetiva e satisfação. Por outro lado, no experimento remoto, participaram 3 corretores de seguros. As medidas consideradas na avaliação remota foram: usabilidade subjetiva, satisfação e utilidade.

Os resultados obtidos na dimensão de usabilidade mostraram que o Broker-RecSys obteve uma eficiência em tempo de 1,58 e 1,47 minutos para realizar as tarefas de recomendação para um cliente específico e para um grupo de clientes respectivamente. Na eficácia, Broker-RecSys obteve 96% e 92% de percentagem de completação das tarefas com sucesso na recomendação de produtos para um cliente específico e para um grupo de clientes respectivamente. Além disso, o Mapa de calor gerado a partir da fixação dos olhos dos participantes, mostra o uso eficiente da interface visual do usuário na realização das tarefas de recomendação de produtos de seguros. A avaliação do SUS mostrou uma usabilidade subjetiva excelente e boa por parte de participantes sem conhecimento no domínio de seguros e participantes corretores de seguro respectivamente. Baseado nesses resultados, o Broker-RecSys permite que usuários sem conhecimento no domínio de seguros possam executar tarefas de recomendação de produtos de seguros.

Por outro lado, os resultados na dimensão de utilidade mostraram que o Broker-RecSys obteve 4,6 de score, mostrando a sua utilidade no processo de recomendação de produtos de seguros. Baseado nesse resultado, O Broker-RecSys auxilia corretores de seguros no processo de recomendação de produtos de seguros na sua carteira de clientes.

Os resultados alcançados sugerem que os métodos de mineração de dados, combinados aos métodos de interação e visualização de dados podem auxiliar os usuários no processo de recomendação.

APPENDIX B — INFORMED TERM OF CONSENT

Broker-RecSys: Evaluation of an interactive insurance product recommendation system.

Introduction
This is a study carried out by student Paul Dany Flores Atauchi as part of his master's thesis carried out at the Informatics Institute of the Federal University of Rio Grande do Sul (UFRGS) in the period 2018-2019.

Purpose and objectives
Broker-RecSys is a web system that aims to assist the insurance broker in the process of recommending products in his client portfolio. The objective of the research is to investigate how visualization, interaction and data analysis can help the insurance broker in the insurance product recommendation process.

Experiment duration
The test session needs approximately 15-30 minutes.

*Obligatorio

FREE AND CLARIFIED CONSENT TERM

You are being invited to participate in the experiment called "Evaluation of an interactive insurance product recommendation system". The procedure to be followed is detailed below.

Procedure:

1) Consent form: The user accepts his participation in the study after reading the informed consent form.

2) Introduction to the study: The user receives an explanation of the study and its purpose, information on the configuration of the environment for the test session. Brief explanation about the insurance domain, recommendation system, usability test and explanation of the test session protocol.

3) User characterization: The user fills in basic personal information that was used for statistical purposes.

4) Introduction to the prototype: It is to provide the user with a 4 minute demo (video with audio) showing the functionalities of the prototype.

5) Conducting the test session: The experiment consists of 2 parts.

- Part 1: The user can freely explore the interactive recommendation system for 2 minutes.
- Part 2: The user performs a sequence of tasks related to the recommendation of insurance products.

6) Satisfaction questionnaire (SUS): The user fills out a questionnaire to measure their satisfaction with the use of the interactive recommendation system.

7) Finally, the user is interviewed to assess their preferences regarding the use of the interactive insurance product recommendation system.

Anonymity
The data obtained in carrying out this experiment will be used only in this study and all types of anonymous personal information will be maintained.

Observation
- Non-broker user: The client portfolio used in the experiment is a synthetic dataset for evaluation.
- Broker user: The client portfolio used in the experiment is the insurance broker's own client portfolio.

I read and understood the consent form and the purpose of the study, as well as the importance of this study, the procedures involved in it, its possible benefits and risks. I had the opportunity to ask about the study and all my doubts were answered. I understand that I am free to decide not to participate in this research.

By signing my name, I agree to voluntarily participate in this study.

Sign name *

Tu respuesta

Siguiente

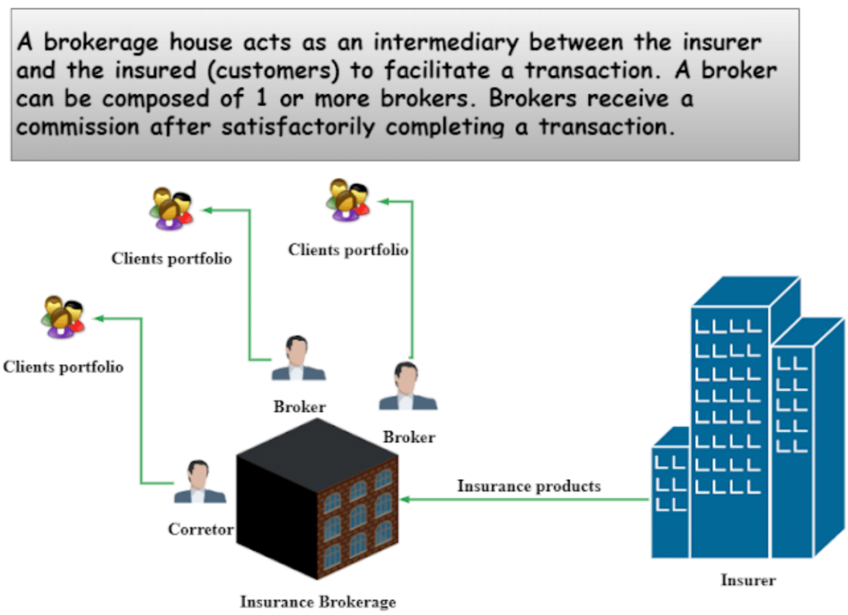
APPENDIX C — INTRODUCTION TO THE STUDY

Broker-RecSys: Evaluation of an interactive insurance product recommendation system.

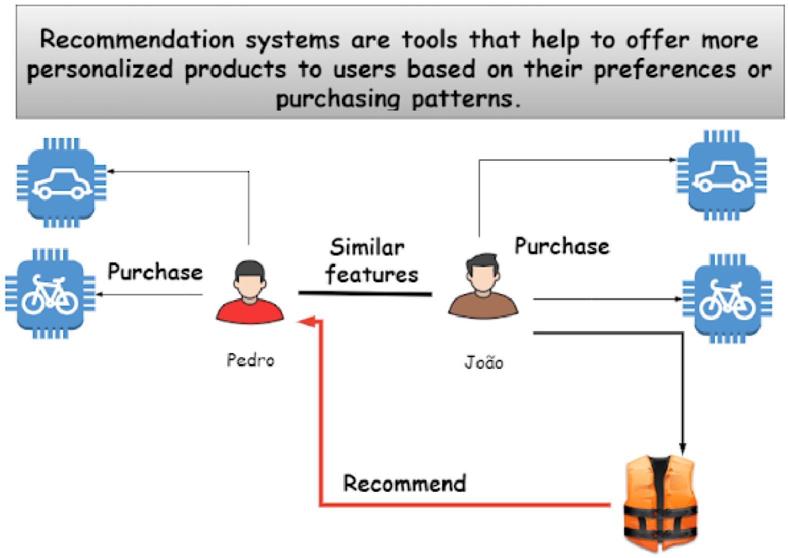
Introduction to the study

Next, infographics about insurance broker, recommendation system and usability test are shown.

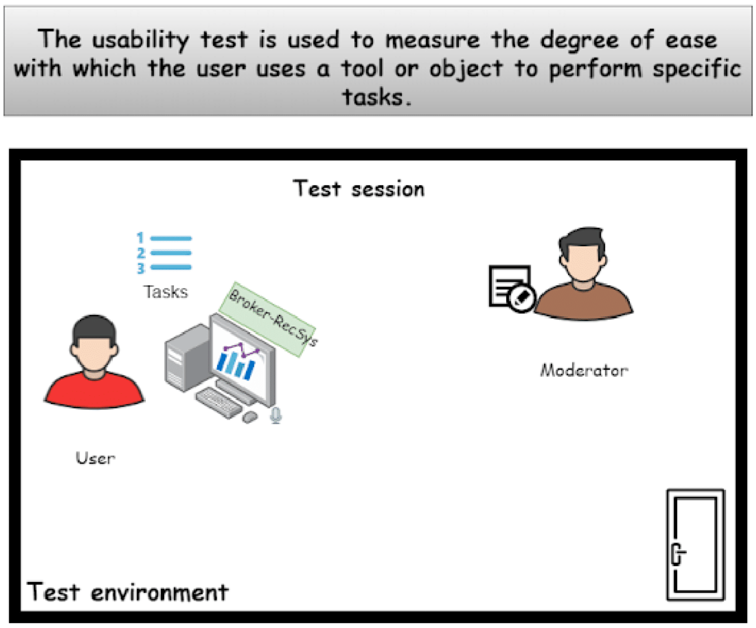
Insurance brokerage



Recommendation System



Usability test



Atrás Siguiente

APPENDIX D — USER CHARACTERIZATION

Broker-RecSys: Evaluation of an interactive insurance product recommendation system.

*Obligatorio

User characterization

We would like to meet you, feel free to answer the following questions.

Are you an insurance broker? *

- ◯ Yes
- ◯ No

Age *

- ◯ 18-25
- ◯ 26-35
- ◯ 36-45
- ◯ >45

Gender *

- ◯ Female
- ◯ Male
- ◯ Prefer not to say
- ◯ Otros:

Formation *

Elegir

How experienced do you consider yourself in the use of interactive information systems? *

	1	2	3	4	5	
Very inexperienced	◯	◯	◯	◯	◯	Very experienced

Do you have any vision problems? *

- ☐ None
- ☐ Myopia
- ☐ Astigmatism
- ☐ Daltonism
- ☐ Hypermetry
- ☐ Otros:

How is your level of discomfort now? *

	None	Light	Moderate	Grave
General discomfort	○	○	○	○
Tiredness	○	○	○	○
Headache	○	○	○	○
Nuisance in the eyes (itching, pain, etc.)	○	○	○	○
Difficulty focusing	○	○	○	○
Blurry vision	○	○	○	○
Torture with eyes open	○	○	○	○

Atrás Siguiente

Nunca envíes contraseñas a través de Formularios de Google.

Google no creó ni aprobó este contenido. Denunciar abuso - Condiciones del Servicio - Política de Privacidad

Google Formularios

APPENDIX E — USEFULNESS QUESTIONNAIRE - RECOMMENDATION FOR A SPECIFIC CUSTOMER

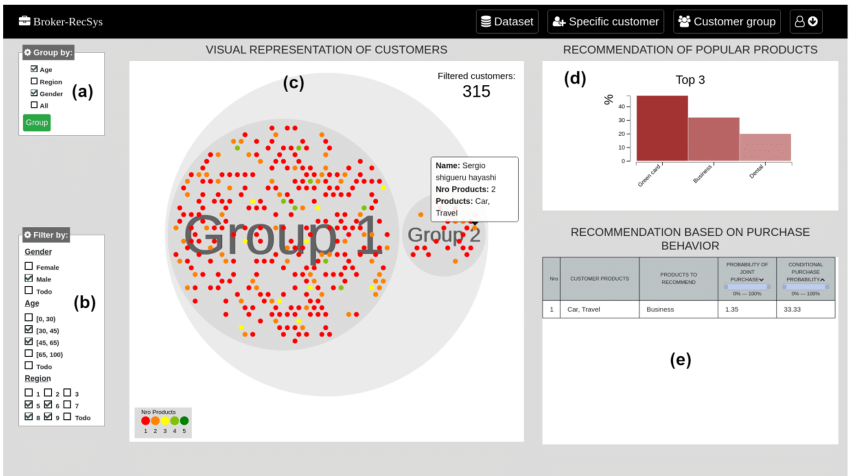
Broker-RecSys: Avaliação de um Sistema interativo de recomendação de produtos de seguros.

*Obligatorio

Recommendation of insurance products for potential customers

This questionnaire aims to assess the utility of the tool in recommending products to potential customers.

Dashboard: product recommendations for potential customers.



I know that by segmenting my customers I can offer a more personalized recommendation. *

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	I totally agree

I have managed with little effort to explore and identify potential customers to recommend insurance products. *

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	I totally agree

I found the popular recommendations useful to offer my clients. *

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	I totally agree

I found the recommendations based on buying behavior useful to offer to my customers. *

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	I totally agree

I got to know why that recommendation, it helped me to better understand the recommendation and the decision making. *

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	I totally agree

Atrás Siguiente

Nunca envíes contraseñas a través de Formularios de Google.

Google no creó ni aprobó este contenido. Denunciar abuso - Condiciones del Servicio - Política de Privacidad

Google Formularios

APPENDIX F — USEFULNESS QUESTIONNAIRE - RECOMMENDATION FOR A CUSTOMER GROUP

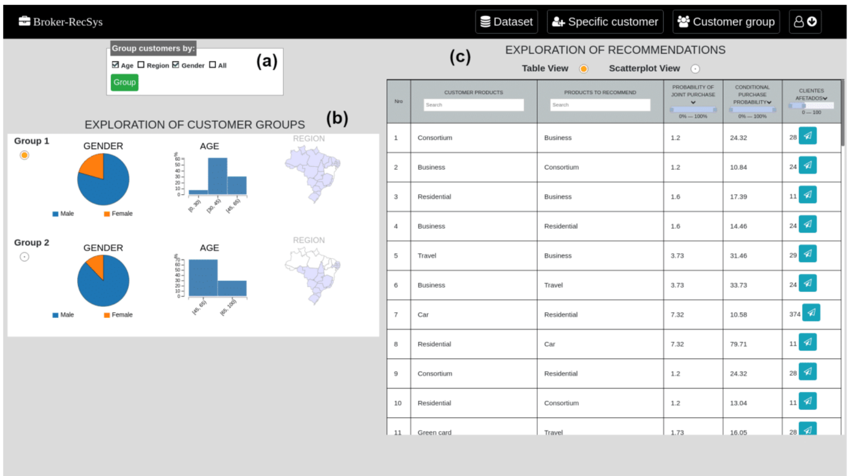
Broker-RecSys: Avaliação de um Sistema interativo de recomendação de produtos de seguros.

*Obligatorio

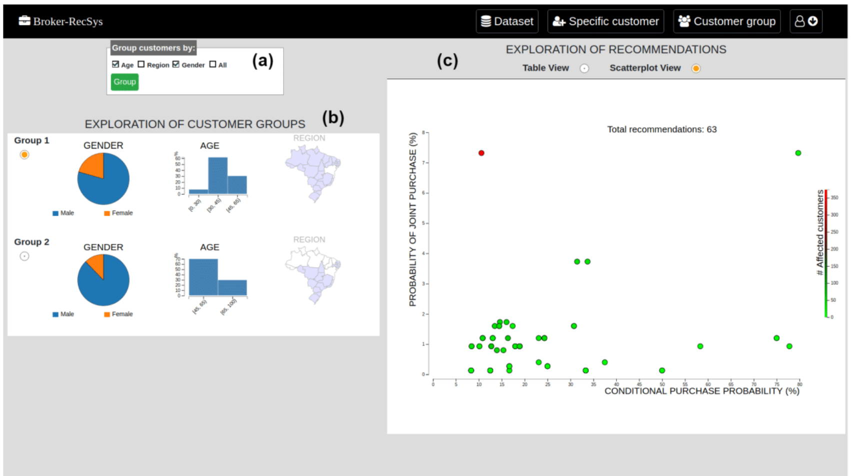
Insurance product recommendation for customer groups

This questionnaire aims to evaluate the use of the tool in recommending products to a group of customers.

Dashboard: recommendations with a table view.



Dashboard: recommendations with a scatter plot view.



I was able to segment customers to find out what types of customer groups are in the customer portfolio. *

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	I totally agree

I was able to identify an interesting group of customers by viewing their information. *

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	I totally agree

I was able to explore and identify interesting recommendations using the table view *

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	I totally agree

I was able to explore and identify interesting recommendations using the scatter plot view *

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	I totally agree

Atrás Siguiente

Nunca envíes contraseñas a través de Formularios de Google.

Google no creó ni aprobó este contenido. Denunciar abuso - Condiciones del Servicio - Política de Privacidad

Google Formularios

APPENDIX G — SUS - SYSTEM USABILITY SCALE QUESTIONNAIRE

Broker-RecSys: Evaluation of an interactive insurance product recommendation system.

*Obligatorio

Usability of the recommendation system

This questionnaire will serve to measure user satisfaction in using the recommendation system.

Evaluate the recommendation system by answering the following questions. Select the assessment that best represents your opinion.

I think I would like to use this recommendation system frequently (To perform similar tasks) *

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	I totally agree

I find the recommendation system unnecessarily complicated. *

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	I totally agree

I found the recommendation system easy to use. *

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	I totally agree

I think I would need help from a person with technical knowledge to use the recommendation system. *

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	I totally agree

I think the various functions of the recommendation system are very well integrated. *

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	I totally agree

I think the recommendation system is very inconsistent.

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	I totally agree

I imagine that most people would learn how to use this recommendation system quickly. *

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	I totally agree

I found the recommendation system very complicated to use. *

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	I totally agree

I felt very confident using the recommendation system. *

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	I totally agree

I had to learn several new things before I was able to use the recommendation system. *

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	I totally agree

Atrás Siguiente

Nunca envíes contraseñas a través de Formularios de Google.

Google no creó ni aprobó este contenido. Denunciar abuso - Condiciones del Servicio - Política de Privacidad

Google Formularios

APPENDIX H — POST-QUESTIONNAIRE

Broker-RecSys: Evaluation of an interactive insurance product recommendation system.

*Obligatorio

Post-questionnaire

To conclude the experiment, can you leave us your criticisms and suggestions about the recommendation system?

What did you find interesting about the recommendation system? *

Tu respuesta

Didn't you find the recommendation system interesting? *

Tu respuesta

What would you like to change in the insurance product recommendation system? *

Tu respuesta

Atrás Siguiente

Nunca envíes contraseñas a través de Formularios de Google.

Google no creó ni aprobó este contenido. Denunciar abuso - Condiciones del Servicio - Política de Privacidad

Google Formularios