THE UNIVERSITY OF ZAMBIA

SCHOOL OF NATURAL SCIENCES

PRODUCT RECOMMENDER SYSTEM FOR TELECOMMUNICATION INDUSTRIES: A CASE OF ZAMBIA TELECOMUNICATIONS COMPANIES BY

MULIZWA SOFT

A Dissertation submitted to the University of Zambia in partial Fulfilment of the requirements for the award of the Master degree of Science In Computer Science.

February 2019

DECLARATION

I, the undersigned, declare that this has not previously been submitted in candidature for any degree. The dissertation is the result of my own work and investigations, except where otherwise stated. Other sources are acknowledged by given explicit references. A complete list of references is appended.

Signature:……………………………………………….

Date:…………………………………………………..

CERTIFICATE OF APPROVAL

This document by MULIZWA SOFT is approved as partial fulfilment of the requirements for the award of the degree of Master of Science in Computer Science of the University of Zambia

Examiner’s Name: ………………………………………………

Signature…………………………… Date: …………………….

Examiner’s Name: ………………………………………………

Signature: …………………………… Date: …………………….

Examiner’s Name: ………………………………………………….

Examiner’s Signature: ……………………… Date: ………………

ACKNOWLEDGEMENT

I give praise and thanks to our God the creator of this universe through our Lord Jesus Christ for giving me good health, understanding and many other blessings during the course of my study. I am extremely grateful to my parent’s, Mrs. Aness Phiri and Mr Evans Mulizwa for providing the moral support I needed to finish my Master’s degree; I am extremely grateful to Mr. Zulu who helped enormously and guided me with his knowledge and experience throughout my research to completion. I would also like to thank the Telecommunication companies for providing part of the data needed to complete this research. Many thanks to my friends, brothers and sisters for the encouragement and support they gave me throughout my studies.

DEDICATION

To my father and family. Thank you so much for all your love, patience, encouragement and support.

List of Acronyms

The acronyms used in this thesis are presented below:

CDR	Call Detail Reports
GSMA	Groupe Speciale Mobile Association
CAGR	Compound, Annual Growth Rate
CB	Content Based
CF	Collaborative Filtering
V.A.S	Value Added Services
KNN	K-nearest neighbor
SMS	Short Message Service
SVM	Support Vector Machine
USSD	Unstructured Supplementary Service Data
SIM	Subscriber Identification Module
OTT	Over-the-Top (OTT)
ICT	Information and Communication Technology
ML	Machine Learning
CAF	Consumer Application Form
ITU	International Telecommunication Union
GSM	Global System for Mobile Communications
OECD	Organization for Economic Co-operation and Development
ARPU	Average Revenue Per User
FTTH	Fiber-to-the-Home
VoIP	Voice over Internet Protocol
IP	Internet Protocol

LIST OF KEYWORDS

Recommender Systems

Telecom products/services

Machine learning algorithms

Big data

Business Intelligence

Call Detail Reports

Collaborative Filtering

Content Based Filtering

Revenue Leakage

Churn

Telecommunication operators

Voice

Customer

Segmentation

Analytics

ABSTRACT

Recommender systems have become increasingly popular in recent years, and are utilized in a variety of areas including movies, music, news, products, research articles, search queries, social tags, and products in general they are designed to automatically generate personalized suggestions of products/services to customers. With the competitiveness that is growing in the telecommunication industry, telecommunication operators seek ways to attract and keep the subscribers on their network, Its notable that telecommunication operators lack the ability to manage their customer retention rate because they do not have a personalized way of recommending products and services to their subscribers, as a result subscribers tend to migrate to new providers. This trend of subscribers migrating to new providers proves to be a severe problem for Telecommunication providers as they experience subscriber base and revenue shrinkage. This dissertation describes a Recommender System for Telecommunication companies using call detail reports (CDR's), machine learning algorithms and big data concepts. Experimental results demonstrate the effectiveness of the proposed approach and the initial application shows that recommender systems can effectively help customers to select the most suitable mobile products or services.

Table Of Contents

CHAPTER ONE
INTRODUCTION TO THE RESEARCH

1.1 Introduction

In this chapter, we introduce the research study. We look at the motivation and significance of the research. The scope, problem statement and aim are given. This is then followed by the objectives, research questions and the research contributions. Finally, we present the organization of the thesis and a summary of the chapter.

1.2 Introduction to the Research Study

Telecom operators are facing increasing challenges in the digital era. Communication tools based on the Internet, such as FaceBook, WhatsApp and Twitter, have dramatically reduced the traditional profits of telecom operators for SMS and voice calls, and they are trying hard to avoid becoming just simple data channels in the digital era. Telecom operators control the last mile for all mobile devices to access the Internet, and therefore will share the future profit from the mobile internet market. Currently, telecom operators are advised to enhance customer loyalty and increase the migration cost for changing the mobile numbers and switching service providers. A large user base is the key to winning market share in the mobile internet arena, and telecom operators are able to secure a huge number of low-end users through subsidizing low-cost android-based devices. Mobile Internet and big data will create tremendous opportunities for telecom operators [1]. Big data is the process of examining large data sets containing a variety of data types, it is used to uncover hidden patterns, unknown correlations, market trends, customer preferences and other useful business information. The analytical findings can lead to more effective marketing, new revenue opportunities, better customer service, improved operational efficiency, competitive advantages over rival organizations and other business benefits. Communications service providers that want to be innovative and maximize their revenue potential must have the right solution in place so that they can harness the volume, variety and velocity of data coming into their organization and leverage actionable insight from that data [2], with the competitiveness that is growing in the telecom industry, telecommunication operators seek for ways to attract and keep the subscribers in their network. Giving away attractive promotions to their subscribers is a powerful and commonly used approach in that context. In order to gain a competitive advantage, the operators need new technologies and methodologies to support more attractive newer patterns of promotions [3]. Telecom businesses today offer hundreds of different mobile products and services such as handsets, mobile plans, prepaid mobiles, and broadband to customers and are constantly exploring new service models that will support customers in their selection and purchase of products and services. Telecom products are always linked with services, referred to hereafter as 'products/services', and have very complex structures and a huge number of choices, for example, a telecom company may have more than 500 telecom products/services in several categories for different groups of customers (individual consumers, small businesses, medium businesses and large businesses). With such a vast number of choices, it is becoming increasingly difficult for customers to find their favourite products quickly and accurately. Only experienced salespeople in a telecom company can make suitable personalized recommendations to customers, which is costly and inefficient [4]. A CDR refers to the information about a charging event collected in a certain format. The information includes call start time, call duration and transferred data amount. Telecommunications companies use CDRs for purposes of billing, extracting business intelligence, fraud detection, etc. However, they face a big data challenge as many telecommunication companies get billions of CDRs per day [5]. Research from GSMA (Groupe Speciale Mobile Association) intelligence found that the mobile industry in Sub-Saharan Africa continues to scale rapidly, reaching 367 million subscribers in mid-2015. However, subscriber growth rates are set to slow sharply over the coming years, with growth in the second half of this decade set to be around 6% compared to 13% in the first half. Mobile operators' revenue growth is slowing across Sub-Saharan Africa, reflecting slowing subscriber growth but also the impact of external factors such as growing competitive pressures and regulatory action. From a compound, annual growth rate (CAGR) of almost 7% for 2010–2015, growth is set to slow to 5% out to 2020 [6].

1.3 Motivation and Significance of the Thesis

A lot of telecommunication industries are coming up in Zambia but their major challenge has been maintaining their customers because most of the service providers do not have attractive products or service while others have poor customer satisfaction and experience. As a result, many subscribers tend to own at least three sim cards from different telecommunication providers, some move from provider to another in order to enjoy cheap services. The study can help telecommunication companies set the right targets and customer segments in planning and decision making and help telecommunication operators to manage their customer retention rate because they do not have a personalized way of recommending products and services to their subscribers.

1.4 Scope

The study will concentrate on the existing machine learning algorithms that are used by Facebook, Amazon or YouTube, apply and integrate them on the subscriber data that is generated by telecommunication companies in order to predict products which are most likely to be used by subscribers for a particular telecommunication company.

1.5 Problem Statement

Telecommunication operators lack the ability to manage their customer retention rate because they do not have a personalized way of recommending products and services to their subscribers as a result subscribers tend to migrate to new providers. This trend of subscribers migrating to new providers proves to be a severe problem for Telecom providers as they experience subscriber base and revenue shrinkage.

1.6 Aim

The main aim is to implement a product recommender system for Telecommunications Companies using machine learning algorithms.

1.7 Objectives

This research was guided by the following objectives:

a) Study and examine the current existing recommender systems used by Netflix, Amazon, and eBay in relation to the products they recommend to their users.

b) Establish challenges telecommunication companies in Zambia face in terms of low revenue, churn, fraud etc.

c) Design and implement a product recommender system which will recommend products that a subscriber is more likely to use.

1.8 Research Questions

This research was guided by the following research questions:

a) How do we analyze the relationship between telecommunication subscribers and telecommunication products?

b) What are the challenges telecommunication companies facing that lead to high revenue loss, churn and bad customer experience?

c) How best will a product recommender system for mobile technology be utilized in order to assist solving the problem of low revenue, churn and fraud?

1.9 Research Contributions

The business processes will enable telecommunication industries to maintain their subscribers base and offer them personalized attractive value-added services products in real time. Implementation of the recommender system machine learning algorithms and python programming language for the Telecommunication Companies was done. Some of this work has been will be published in the International American Journal of Economics.

1.10 Organization of the Thesis

The work done in this thesis is organized into five chapters. Chapter 1 is the Introduction to the Research. In this chapter, we give a brief overview of the work in this thesis. We also give the problem statement, aims and motivation of this thesis. This chapter concludes by the giving an outline of the thesis. Chapter 2 looks at the background theory and related works. In this chapter, we begin by providing a comprehensive review and the background theory of Recommender Systems in all aspects. The research methodology is given in Chapter 3. In this chapter, we look at the methods used to conduct the baseline study and implement the system. In Chapter 4, we present the research findings of the baseline study and the system implementation. Finally, in Chapter 5 the discussion and conclusion are given.

1.11 Summary

In this chapter, we looked at the basic introduction of the work in this thesis. We begin by looking at the challenges that telecommunications companies are facing in Sub-Saharan Africa and narrow it down to Zambia. The motivation, significance and scope of the work in this study are then outlined. Finally, we give the problem statement, outlined the aims, the research contributions and we close this chapter with the outline of the thesis.

CHAPTER TWO
LITERATURE REVIEW

2.1 Introduction

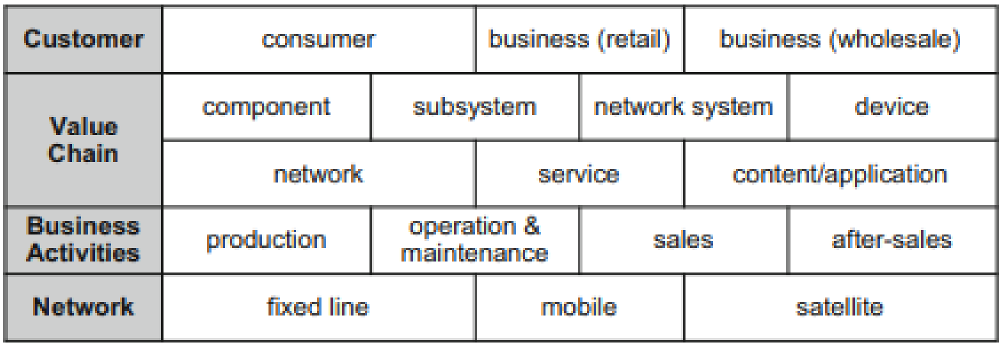
In this chapter, we review the literature to give the background theory and the works related to our study. We begin by extensively looking at Recommender systems and how it can be used improve customer experience and personalization of telecom value added services. This is followed by a brief review of recommender systems and mobile markets. We also look at the software development approaches in the following section. We close this chapter by looking at related works in recommender systems.

2.2 Introduction to Mobile Markets

Sub-Saharan Africa is a predominantly a prepaid market for most of the telecommunications subscribers, the prepaid nature of the region has contributed to the relatively volatile loyalty of its mobile subscribers. It has been noticed that operator subscription numbers are constantly fluctuating due to the non-commitment of subscribers and competition exists not only between operators, but also as a result of over-the-top players, i.e. social media providers such as WhatsApp. These over-the top players use the telecommunications infrastructure to offer data services, which puts an added strain on the already pressured network. The low cost of using these over-the-top services means subscribers are less likely to use operator services such as SMS, USSD and voice [7]. As mobile operators continue to add subscribers to their network they typically reach out to harder to reach areas or segments and either poorer subscribers or multi-SIM subscribers often spend much less than more affluent early adopters of mobile services - reaching to specific niche segments or to remote areas, where competition may be less strong can be costly, diluting margins [8]. Despite all this telecommunication industry faces an increasing churn rate compared with other industries. The types of churn are voluntary and involuntary. If the customer begins the first movement, this is called voluntary churn. There are many reasons for losing the customers such as the phone service changes or the competition. If the company begins to terminate its services to the customer, this results to involuntary churn. In this case customers are churned for reasons like fraud and nonpayment. Increasing churn rate causes a loss of future incomes. Therefore, it is profitable for telecom operators to invest in those customers that already have an experience with the service by renewing their trust rather than trying to attract new customers [9]. The Indian telecom sector for example is witnessing an unprecedented subscriber growth for the past few years. With an increase in the number of service providers which has aggravated the competition and customer acquisition costs in the telecommunication industry but under these circumstances, retaining the existing customers by enhancing customer loyalty and customer value has become a core marketing strategy of most service providers [10]. They have to identify the factors that contribute to loyalty, so as to achieve a competitive edge, customer loyalty has become an important objective for all the telecommunication companies that seeks long term profitability and sustain ability. In any industry, as the competition is becoming intense, the need for retaining the existing customers has become a top priority. Research shows that it costs five times more to acquire a new customer than to retain an old one, hence the success of any business depends on the extent to which it possesses a loyal customer base, as customer loyalty reduces marketing efforts and cost of operations to a greater extent [10]. Customer loyalty is conceptualized into behavioral loyalty and attitudinal loyalty. The behavioral approach focus suggests that loyalty is reflected from the repeated purchasing of brand overtime by a consumer. Attitudinal loyalty is the level of consumer’s psychological attachments and attitudinal advocacy towards the supplier. This means that true loyal acts as an advocate for the company’s products and recommends the products to others, loyalty is beneficial for both the customers and the business. For business, it acts as a competitive advantage, as the customers will not be price–sensitive, shows less switching behavior and make high order purchases. For the customers, it reduces the transaction costs and act as an indication of trust with the service provider [10].

2.3 Telecommunications Market

The telecommunications industry is currently going through a major transformation which creates both opportunities and challenges for fixed operators, mobile operators as well as Internet service providers [11]. New and innovative players are entering the telecommunications market, and this has led to a restructuring of the whole telecommunications industry [12]. Through the fast technological development, increasing market dynamics and deregulation in many countries, the complexity in the telecommunications industry is constantly increasing .Those changes and challenges of the telecommunications industry are the topic of various publications and studies with different focus, including overall market research, value creation and market players regulation and competition, standardization , structures and processes as well as various functional or technical specifics. The first challenge of today’s telecommunications industry is to understand the various players. In the past, the technical realization of communication via mobile or fixed-line networks was the major objective of telecommunications operators.The convergence of voice, video, and data has led to mergers, acquisitions, and partnerships. Increasingly, application and content offers are intermixed with telecommunication services. Entertainment services such as TV offers are linked to traditional communication services, resulting in new competition between TV cable operators and communication network operators. The convergence of telecommunications, media, and hardware industries is an already observed implication. Plunkett (2014, pp. 7–8) points out that the exact composition of the telecommunications industry varies when it comes to including or excluding certain business sectors—e.g., communication equipment or related consulting services. Arlandis and Ciriani (2010, pp. 121–124) relate the telecommunications industry to the information and communication technology (ICT) sector, which they define as an ecosystem consisting of technologies providers, network operators, platform operators, and content providers. Grover and Saed (2003, p. 120) propose a categorization of the telecommunications industry into network providers, tool providers, transaction/service providers, and internet/content providers. When it comes to concrete enterprises offering telecommunication products and services, there is a huge range of different business models, including branded resellers, mobile virtual network operators, or mobile virtual network enablers. There is a variety of characteristics to differentiate those business models—e.g., functional coverage of the value chain or level of control of the communication network. A clear understanding of the market positioning and business scope of a



<table>
<tr><td><b>Customer</b></td><td colspan="2">consumer</td><td colspan="2">business (retail)</td><td colspan="2">business (wholesale)</td></tr>
<tr><td rowspan="2"><b>Value Chain</b></td><td>component</td><td colspan="2">subsystem</td><td colspan="2">network system</td><td>device</td></tr>
<tr><td colspan="2">network</td><td colspan="2">service</td><td colspan="2">content/application</td></tr>
<tr><td><b>Business Activities</b></td><td>production</td><td colspan="2">operation &amp; maintenance</td><td colspan="2">sales</td><td>after-sales</td></tr>
<tr><td><b>Network</b></td><td colspan="2">fixed line</td><td colspan="2">mobile</td><td colspan="2">satellite</td></tr>
</table>

Fig. 2.3.1 Framework for categorizing telecommunications operators[13]

telecommunications operator is an essential prerequisite to support its transformational needs. Therefore, in this section, a categorization along the dimensions customer, value chain, business activities, and network is proposed in fig 2.3.1 . The different dimensions and characteristics are based on a review of existing categorization criteria related to the telecommunications industry. The dimension customer specifies the intended end customer(s) of the telecommunications operator. It is differentiated into consumer, business (retail), and business (wholesale). The value chain starts with the technical hardware and software prerequisites of communication networks (component, subsystem, network system, and device). The network covers all technical aspects required to realize services which might be related to content or applications. The business activities are divided into production, operations and maintenance, sales, and after-sales. The network can be specified by fixed line, mobile, and satellite. The scope of a concrete telecommunications operator might be a complex mixture of the above characteristics. Telecommunications operators are confronted with various challenges that influence their transformational needs. Those challenges are summarized along the dimensions market, products/services, and value chain as shown in (Fig. 2.3.2). The market conditions have changed due to convergence that leads to increased competition. Those changes of the market structures and ecosystem result in new market potentials combined with increased cost and price pressure. Furthermore, these changes

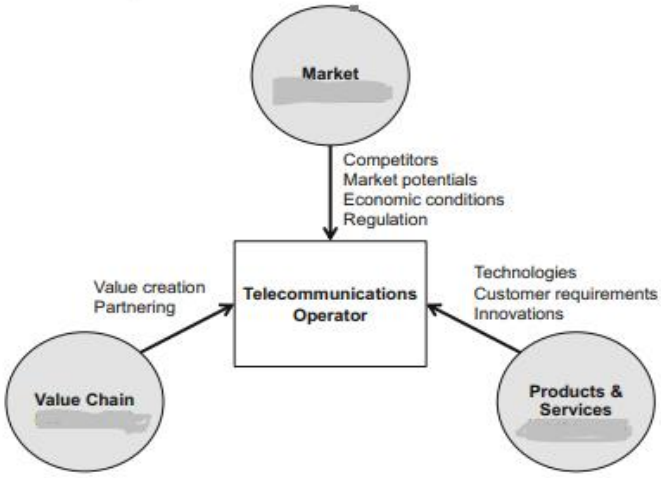


Fig. 2.3.2 Challenges of telecommunications operators

lead to new requirements and challenges for regulators. The value chain reacts to the changed market conditions through increased fragmentation of the value creation and new partnering [13]. In the dimension products and services, telecommunications operators are confronted with the complexity of production systems as well as changed customer demands and usage behavior. Both are related to the requirement of continuous innovations and shorter product development cycles those challenges are an important factor for the transformation of telecommunications operators.The telecommunications market has changed tremendously. Convergence leads to increased competition through Over-the-Top (OTT) providers that offer content and application services on top of existing Telecommunications Operator Market Products & Services Value Chain Competitors Market Potentials Economic Conditions Regulation Value Creation Partnering Technologies Customer Requirements Innovations. In summary, the changed market conditions lead to the disappearance of former revenue sources.

2.3.1 Price Decrease and Cost Pressure

From an economic perspective, the telecommunications industry is an important part of the ICT sector. Global revenue figs are provided by various analysts and research companies. They depend on the exact definition of the industry being applied for their calculation. Plunkett (2014, p. 8) uses a broad definition and estimated a global revenue of 5.4 trillion USD for 2014. The Telecommunications Industry Association (2015) publishes a global revenue of 5.6 trillion USD. Bloomberg3 defines Telecom Carriers as an own industry with a total revenue of 2.1 trillion USD. When it comes to the future trend, these analysts and research companies forecast a slight revenue growth for the next years. However, this revenue growth is decreasing. From a global perspective, the telecommunications industry is a stagnating market.

For a differentiated understanding of the telecommunications industry, the following figs should be considered:

- The worldwide number of fixed-telephone subscriptions has been declining by 12.7% since 2006 in 2014 [14].
- The worldwide number of mobile-cellular telephone subscriptions has more than doubled since 2006 by approximately 152.7% in 2014. However, the growth rate is decreasing [14].
- The worldwide number of broadband subscriptions (fixed and mobile) is increasing. Mobile-broadband subscriptions have especially demonstrated a tremendous growth by approximately 888.9% from 2007 to 2014 [14].
- The market penetration for communication services is constantly increasing: the global estimates for 2015 by ITU are 69% of 3G population coverage, 46 % of households with internet access, and 46 % of individuals with mobile-broadband subscriptions. For the member countries of the Organization for Economic Co-

operation and Development (OECD), the penetration is much higher, with an estimated 81 % for mobile-broadband subscriptions.

- For most communication services, a price decrease can be observed which is a result of the increased competition and ongoing deregulation of the market. For example, ITU report shows decreasing prices for fixed-broadband between 2008 and 2011 with a stagnation since then [14].

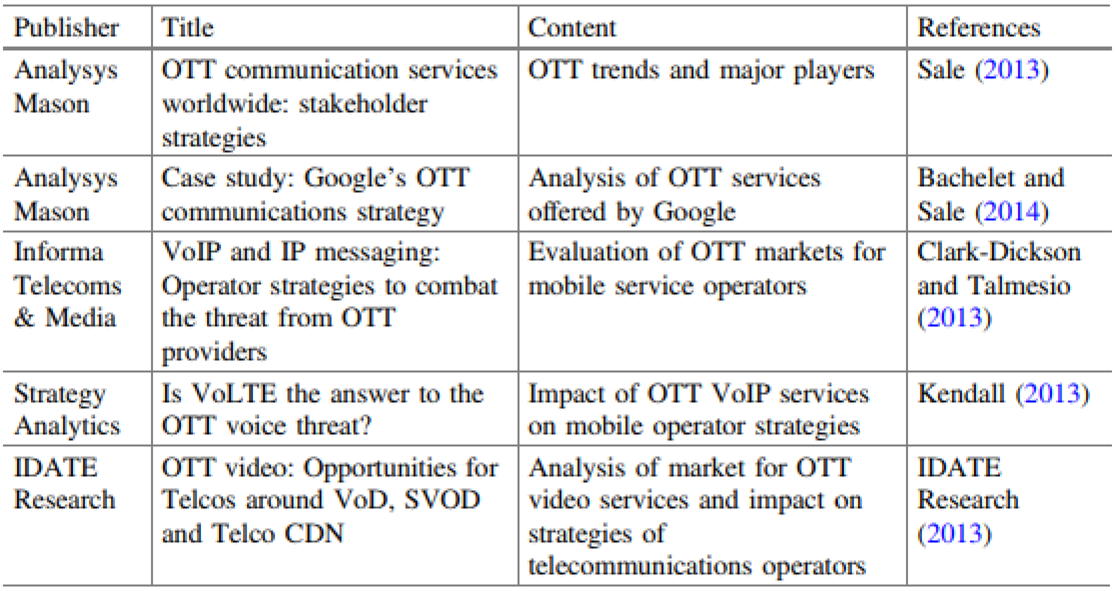
Telecommunications operators are confronted with tremendous changes in the usage behavior in a stagnating market—e.g., compared to a basic mobile phone, using a smartphone generates more than 14 times the data volume [15]. This growth of the data volume has to be handled under the condition of stagnating or even decreasing prices. In the past, traditional communication services—for example, voice telephone—were the major revenue sources of telecommunications operators. Now, the pure transmission is becoming more and more of a commodity for the customer. The increasing demand for high transmission bandwidths still requires extensive investments in network infrastructure. However, those same networks are then beneficial for content and application providers such as Google, Faceproducts, and Netflix, that can profit from the resulting revenues without any participation in the infrastructure investment. Furthermore, those content and application providers even compete with traditional telecommunications operators. As a result, telecommunications operators require innovative services to secure their revenues. Hence, the two contrary conditions of a stagnating and innovative market are mixed. For telecommunications operators, this means the combination of cost reduction and efficiency increase in order to realize the financial flexibility for investments in innovative services. This financial situation is further complicated through new competition caused by the convergence of the market. The technical capability for a broadband transmission requires major investments in fixed or mobile network infrastructure. The value proposition recognized by the customer is related to the communication service. And today those communication services can be offered without owning any network infrastructure. For example, the launch of smartphones—which was seen by the telecommunications operators as an opportunity to introduce new services leading to higher Average Revenue Per User (ARPU)—has actually been a facilitator for the introduction of new services by Over-the-Top (OTT) providers. The new services offered by OTT providers have replaced equivalent telecommunication services—e.g., WhatsApp in the messaging market has replaced the traditional Short Messaging Service (SMS).In the voice market, IP-based products such as Skype and other highly complex enterprise applications have resulted in falling revenues for telecommunications operators. In fact, the usage of Voice over Internet

Protocol (VoIP) is massively changing the telecommunications industry [13]. As a consequence, the traditional voice and messaging markets for telecommunications operators are constantly in decline. A significant part of both historic and predicted telephone and messaging market shifts can be attributed to regulation—either directly related to pricing (e.g., changes in maximum termination or roaming fees), or through the introduction of more competition (e.g., new licensees and wholesale rules). For telecommunications operators, the changed market conditions require higher efficiency and flexibility. In most cases, this leads to transformations of operational structures. These transformations are supported by the reference architecture described in this products. From a strategic perspective, telecommunications operators have to combine their technical capabilities with revenue to create new value propositions. For integrated telecommunications operators—i.e., those operating fixed and mobile network infrastructures—a strategic option is the bundling of communication services and enrichment with content. A typical example is a quadruple-play service combining mobile and fixed telephone, broadband internet, and IPTV. In most cases, this requires partnering with content providers [13]. With those product bundles, telecommunications operators enter the television, video, and media markets. The results are new competitors, such as television cable companies and increased complexity of the value creation. Moreover, those services require a high bandwidth. Therefore, increasing the bandwidth of the offered data connection is an additional strategic option. As example, launching Fiber-to-the-Home (FTTH) services is currently an important topic for telecommunications operators.

2.3.2 Emergence of Over-the-Top (OTT) Providers

The widespread adoption of mobile Internet access has lowered the barriers for many companies to enter the communication services market. Meanwhile, major Internet players have identified opportunities and have also entered these markets. In most cases these services are not necessarily expected to be major drivers of revenue growth; however, they are usually expected to complement the core business, similar to device sales or advertising. The most powerful Internet players are increasingly able to leverage their strengths in the value chain by presenting their communication services as the defaults in devices. From a market perspective, OTT providers are the logical consequence of the changed market conditions. The rising emphasis of application services combined with the convergence in the ICT sector have strengthened new competitors [16]. From a technical perspective, the separation of application and communication services from their technical transportation has supported this trend. In practice, the impact of OTT providers on both telecommunications market and traditional telecommunications operators is discussed in various reports (see Table 2.3.2.1). Telecommunications operators have several strategic options to overcome the challenges arising from OTT providers. Most of the strategies developed and implemented by telecommunications operators to deal with the pressure coming from OTT providers are defensive. The telecommunications operators are aware that OTT communication services are eroding their revenues and, therefore, they need to have a strategy in place to counteract this trend. Blocking VoIP services is a strategy that many telecommunications operators use.

Table 2.3.2.1 Selected reports about OTT market and strategies



Publisher	Title	Content	References
Analysys Mason	OTT communication services worldwide: stakeholder strategies	OTT trends and major players	Sale (2013)
Analysys Mason	Case study: Google's OTT communications strategy	Analysis of OTT services offered by Google	Bachelet and Sale (2014)
Informa Telecoms & Media	VoIP and IP messaging: Operator strategies to combat the threat from OTT providers	Evaluation of OTT markets for mobile service operators	Clark-Dickson and Talmesio (2013)
Strategy Analytics	Is VoLTE the answer to the OTT voice threat?	Impact of OTT VoIP services on mobile operator strategies	Kendall (2013)
IDATE Research	OTT video: Opportunities for Telcos around VoD, SVOD and Telco CDN	Analysis of market for OTT video services and impact on strategies of telecommunications operators	IDATE Research (2013)

Instead of blocking VoIP services, there are some mobile operators that are partnering with OTT providers, and also some mobile operators that are developing their own OTT-like services in their digital business divisions. So far, these two approaches represent the minority of cases. In particular, the attempt to develop own OTT-like services is a strategy which is still in its early stages and which will require a higher maturity level in the digital business areas. On the other hand, the current developments in the OTT market are increasing the pressure on telecommunications operators, giving them only a small window of opportunity to conceive an effective response strategy.

The strategic response alternatives for traditional telecommunications operators to OTT providers can be summarized as follows:

- Accept OTT services: Several telecommunications operators have chosen a hands-off approach to any service that can increase the usage of data, including OTT services. These telecommunications operators believe that the non-occasional nature of communication services such as IP voice and messaging can lead to a strong incentive for customers to purchase a data plan upgrade.
- Attack or absorb OTT services: Many telecommunications operators have decided to attack OTT-based services directly by preventing subscribers from using IP services. This is realized by combining economic and technical aspects that prevent the use of IP services. Another approach is to absorb OTT services by making them ineffective from a customer's perspective. Customers use IP voice and messaging services with the objective to save money. In response, operators are, for example, introducing large voice and messaging bundles with the result that customers do not need to use OTT services in order to save money. In addition, offering services that are similar to the ones offered by OTT providers is a possible strategy. Launching proprietary OTT services is, so far, the least developed option. In the past decade, there have been some attempts by telecommunications operators to deploy instant messaging clients.
- Partner with OTT providers: In some cases, telecommunications operators decide to partner with OTT providers with the objective of benefiting from them. On the one hand, telecommunications operators are afraid that their core services could be marginalized by OTT providers; on the other hand, they are aware that these services can be popular amongst customers. Telecommunications operators that decide to partner with OTT providers might benefit from both the OTT services as well as the OTT brand.

The strategic options listed above are not necessarily mutually exclusive, and many telecommunications operators are active in several of these areas. Price will continue to be the major driver in the voice market. Therefore, telecommunications operators use pricing levers to ensure their voice services are relevant to most smartphone users.

Google is an example of a successful OTT provider. In some areas it is a strong competitor of established telecommunications operators. Google has established comprehensive product and service categories for devices, operating systems, applications and services, content and advertisement so as to service their customers from one source. This provides Google with a competitive advantage in comparison to telecommunications operators specialized in selected categories only. Offering the existing application service via own mobile network capacities (e.g., realized as a Mobile Virtual Network Operator) could be a strategic option that would fit to the ongoing convergence of the whole ICT sector. For traditional telecommunications operators, however, the demand for communication services is directly linked to the existence of attractive content and applications: for example, the growing demand for mobile data services is based on the ever-increasing range of mobile content and applications by, e.g., Google.

This one example highlights the complex interrelation between OTT services and telecommunications operators. The extensive communication services portfolio of OTT providers, their level of control and also the ability to monetize their services present a growing challenge for most telecommunications operators. There are still some operators that have not yet recognized the severe risk of their services being eroded by OTT-based communication services. However, the majority of telecommunications operators have clearly seen the urgent need for developing a strategy for OTT communications.

OTT’s business models develop rapidly and change the traditional revenue models as follows:

- Advertisement is one of the main revenue sources of OTTs;
- Paid subscriptions start to work for OTTs with a larger customer base;
- “Freemium” apps have proved to be an innovative monetization strategy;
- Cloud storage as an add-on service has increased profitability;
- Business intelligence is a powerful tool for content distributors.

In Fig. 2.3.3. a phased approach is outlined to assess the impact of OTTs on the business and thus develop an effective, feasible response strategy tailored to the specific needs. Several telecommunications operators are investing in the development of products and services for vertical markets like energy, automotive, healthcare, and education in order to generate additional revenue streams besides the traditional telecommunications business [17].

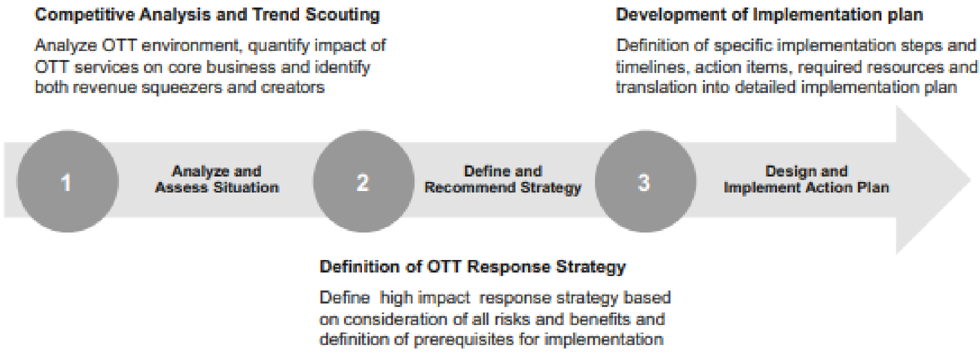

Competitive Analysis and Trend Scouting
Analyze OTT environment, quantify impact of OTT services on core business and identify both revenue squeezers and creators
Development of Implementation plan
Definition of specific implementation steps and timelines, action items, required resources and translation into detailed implementation plan
1
Analyze and Assess Situation
2
Define and Recommend Strategy
3
Design and Implement Action Plan
Definition of OTT Response Strategy
Define high impact response strategy based on consideration of all risks and benefits and definition of prerequisites for implementation

Fig. 2.3.3 OTT response strategy development approach

2.4 Recommender Systems

People find articulating what they want hard, but they are very good at recognizing it when they see it. This insight has led to the utilization of relevance feedback, where people rate web pages as interesting or not interesting and the system tries to find pages that match the “interesting”, positive examples and do not match the “not interesting”, negative examples. With sufficient positive and negative examples, modern machine-learning techniques can classify new pages with impressive accuracy in some cases, text classification accuracy exceeding human capability has been demonstrated. Capturing user preferences is a problematic task. Simply asking the users what they want is too intrusive and prone to error, yet monitoring behavior unobtrusively and then finding meaningful patterns is difficult and computationally time consuming. Capturing accurate user preferences is however, an essential task if the information systems of tomorrow are to respond dynamically to the changing needs of their users [18]. Total information overload becomes increasingly severe in the modern times of omnipresent mass-media and global communication facilities, exceeding the human perception’s ability to dissect relevant information from irrelevant. Consequently, since more than 64 years significant research efforts have been striving to conceive automated filtering systems that provide humans with desirable and relevant information only. During the last two-decades, recommender systems have been gaining momentum as another efficient means of reducing complexity when searching for relevant information. Recommenders intend to provide people with suggestions of products they will appreciate, based upon their past preferences, history of purchase, or demographic information or other types of information [19]. A recommender system consists of three elements as shown in fig (2.4.1). Many recommendation contents which are presented to users have to be made. Then, users’ preferences or behavioral data on these contents must be gathered. Finally, it needs to choose type of recommendation technique about how to analysis these user data and select the optimal content to each user [20].

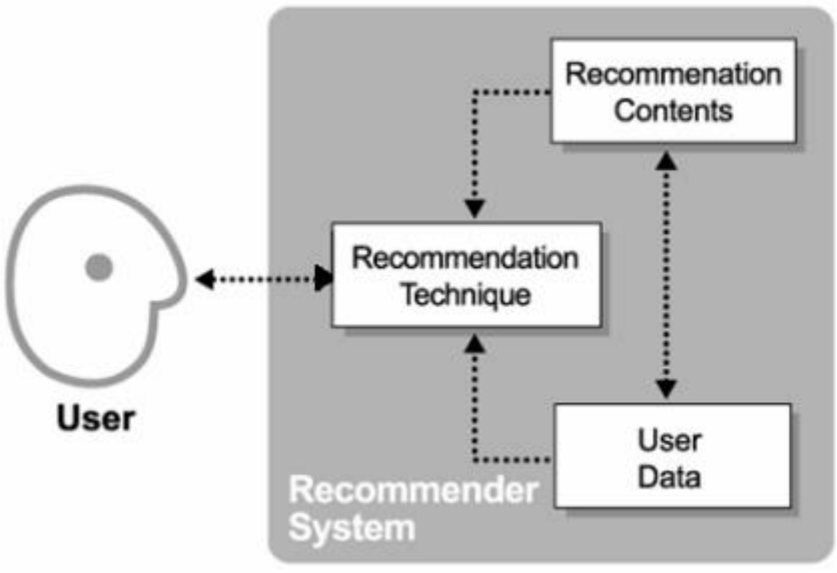


Fig (2.4.1). Constitution of Recommender System [20].

Recommendation systems are used everywhere in the Internet today. From e-commerce sites to restaurants, hotels, tickets, events, and so on are recommended to us everywhere. Have we ever asked ourselves how do they know what will be the best for us? How do they come up with this calculation of showing the items we might like? It is a known fact that Recommender systems are the most successful implementation of web personalization and can be defined as personalized information filtering technology that is used to automatically predict and identify a set of interesting items on behalf of users according to their personal preferences. Recommender systems use the concept of rating to measure users' preferences and a range of filtering techniques, and can be classified in multiple ways according to the nature of the input information. The content-based (CB) methods and collaborative filtering (CF) methods are the most popular techniques adopted in recommender systems. The CB methods recommend products by comparing the content or profile of the unknown products to those products that are preferred by the target user. However, these methods tend to rely heavily on textual descriptions of items, leading to several unsolved problems such as limited information retrieval, new user problems, and overspecialization. Unlike CB methods, CF methods do not involve user profiles and item features when making recommendations. CF methods help people make their choices based on the opinions of other people who share similar interests. There are several kinds of CF methods, among which the most popular approaches are user-based CF and item-based CF. A user-based CF method uses the ratings of users that are most similar to the target user (recommendation seeker) for predicting the ratings of unrated items. More specifically, when making a recommendation, the user-based CF recommender system will first calculate the similarities of all users to the target user by analysing the previous ratings of all users. The system will then select a certain number of most similar users as references, following which it will use the ratings of the selected users on the target item (the unrated item of the target user) to predict the rating of this item for the target user. By contrast, the item-based CF method uses the similarities of items to predict ratings. The major limitations of CF methods are the cold start problem for new users and new items, the sparsity problem, and the long tail problem. These problems have attracted much attention from researchers. A kernel-mapping recommender was proposed in, and the recommendation algorithm performs well in handling these problems. Park et al. used a clustering method to solve the long tail problem. A third approach is the knowledge-based (KB) recommendation approach. This generates recommendations based on business knowledge (business rules) and inferences about a user’s needs and preferences, and because it has functional knowledge about how a particular item meets a particular user need, it is able to reason about the relationship between a need and a potential recommendation. Some KB systems employ case-based reasoning techniques for recommendation. These types of recommenders solve a new problem by looking for a similar past solved problem [21]. Content-based recommendation engine works with existing profiles of users. A profile has information about a user and their taste. Taste is based on user rating for different items, Fig 2.4.2 [22]. Shows an example of content based recommendation.

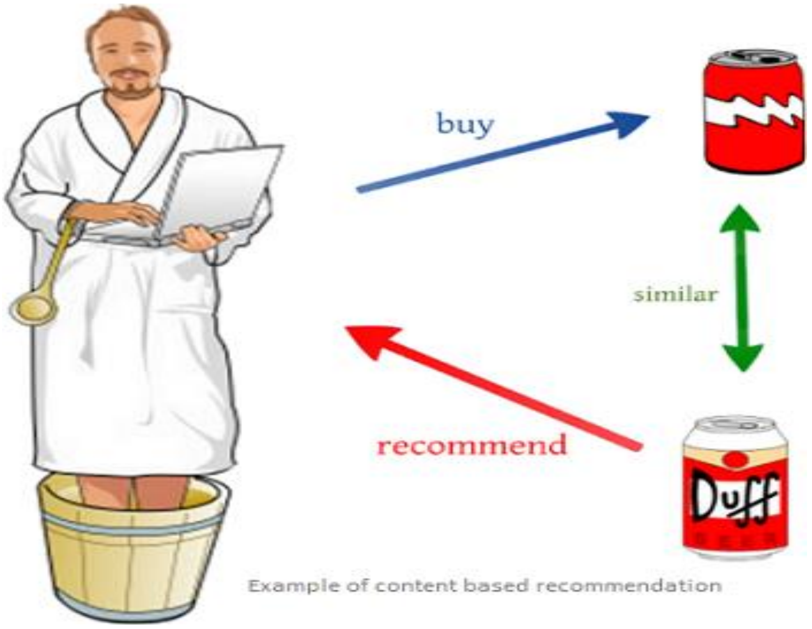


Fig 2.4.2: Content based recommendation [23].

The idea of collaborative filtering is finding users in a community that share appreciations. If two users have same or almost same rated items in common, then they have similar taste, Fig 2.4.3 [23] shows the collaborative filtering example.

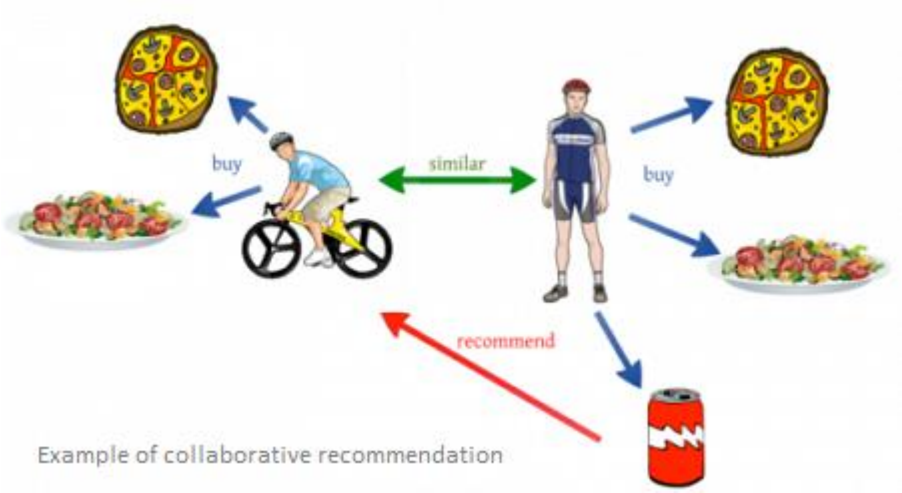


Fig 2.4.3: collaborative filtering [23].

Matching consumers with the most appropriate products is the key to enhancing user satisfaction and loyalty. Therefore, more retailers have become interested in recommender systems, which analyze patterns of user interest in products to provide personalized recommendations that suit a user's taste. Because good personalized recommendations can add another dimension to the user experience, e-commerce leaders like Amazon.com and Netflix have made recommender systems a salient part of their websites in recent years. Such systems are particularly useful for entertainment products such as movies, music, and TV shows [24]. Recommender systems are used by e-commerce sites to suggest products to their customers, the products can be recommended based on the top overall sellers on a site, based on the demographics of the customer, or based on an analysis of the past buying behavior of the customer as a prediction for future buying behavior. Broadly, these techniques are part of personalization on a site, because they help the site adapt itself to each customer, Recommender systems enhance E-commerce sales in three ways:

- Browsers into buyers: Visitors to a Web site often look over the site without ever purchasing anything. Recommender systems can help customers find products they wish to purchase.
- Cross-sell: Recommender systems improve cross-sell by suggesting additional products for the customer to purchase. If the recommendations are good, the average order size should increase. For instance, a site might recommend additional products in the checkout process, based on those products already in the shopping cart.
- Enhance Loyalty: In a world where a site's competitors are only a click or two away, gaining customer loyalties is an essential business strategy, Recommender systems improve loyalty by creating a value-added relationship between the site and the customer. Customers repay these sites by returning to the ones that best match their needs, the more a customer uses the recommendation system – teaching it what they want – the more loyal they are to the site. Even if a competitor company were to build the exact same capabilities, a customer would have to spend an inordinate amount of time and energy teaching the competitor what the company already knows [25].

Recommender systems struggle to catch user needs, and companies have implemented different approaches to tackle this issue. Amazon.com, for instance, immediately recognizes the user's identity and recommends a products, without asking for any user input [26].

The objective of collecting user information is to build a profile that describes a user interests, role in an organization, entitlements, and purchases or other information. The most common techniques are explicit profiling and implicit profiling [24]:

- Explicit profiling: asks each visitor to fill out information or questionnaires by a specific form. This technique has the advantage of letting users tell the recommender system directly what they want.

- Implicit profiling: tracks the visitor's behavior. This technique is generally transparent to the user. The browsing is usually tracked by saving specific user identification and behavior information in log file which keeps the user hits or by a cookie files that is kept at the browser and updated at each visit. For example, Amazon.com logs each customer's buying history and, based on that history, recommends specific purchases.

In Short, we can conclude by saying Recommender systems apply data analysis techniques to the problem of helping users to find the items they would like to purchase at E-Commerce sites by producing a predicted likeliness score or a list of top–N recommended items for a given user (active user who the recommendations made to him). Items recommendation can be made using different methods. Recommendations can be based on demographics of the users, overall top selling items, or past buying habit of users as a predictor of future items these techniques are Collaborative based recommender system, Content-based recommender system and the modern recommender systems intend to mix the two ways this new technique called Hybrid recommender system [24].

2.5. Phases of recommendation process

2.5.1. Information collection phase

This collects relevant information of users to generate a user profile or model for the prediction tasks including user’s attribute, behaviors or content of the resources the user accesses. A recommendation agent cannot function accurately until the user profile/model has been well constructed. The system needs to know as much as possible from the user in order to provide reasonable recommendation right from the onset. Recommender systems rely on different types of input such as the most convenient high quality explicit feedback, which includes explicit input by users regarding their interest in item or implicit feedback by inferring user preferences indirectly through observing user behavior [27]. Hybrid feedback can also be obtained through the combination of both explicit and implicit feedback. In E-learning platform, a user profile is a collection of personal information associated with a specific user. This information includes cognitive skills, intellectual abilities, learning styles, interest, preferences and interaction with the system. The user profile is normally used to retrieve the needed information to build up a model of the user. Thus, a user profile describes a simple user model. The success of any recommendation system depends largely on its ability to represent user’s current interests. Accurate models are indispensable for obtaining relevant and accurate recommendations from any prediction techniques.

2.5.1.1. Explicit feedback

The system normally prompts the user through the system interface to provide ratings for items in order to construct and improve his model. The accuracy of recommendation depends on the quantity of ratings provided by the user. The only shortcoming of this method is, it requires effort from the users and also, users are not always ready to supply enough information. Despite the fact that explicit feedback requires more effort from user, it is still seen as providing more reliable data, since it does not involve extracting preferences from actions, and it also provides transparency into the recommendation process that results in a slightly higher perceived recommendation quality and more confidence in the recommendations [28].

2.5.1.2. Implicit feedback

The system automatically infers the user’s preferences by monitoring the different actions of users such as the history of purchases, navigation history, and time spent on some web pages, links followed by the user, content of e-mail and button clicks among others. Implicit feedback reduces the burden on users by inferring their user’s preferences from their behavior with the system. The method though does not require effort from the user, but it is less accurate. Also, it has also been argued that implicit preference data might in actuality be more objective, as there is no bias arising from users responding in a socially desirable way [28] and there are no self-image issues or any need for maintaining an image for others [29].

2.5.1.3. Hybrid feedback

The strengths of both implicit and explicit feedback can be combined in a hybrid system in order to minimize their weaknesses and get a best performing system. This can be achieved by using an implicit data as a check on explicit rating or allowing user to give explicit feedback only when he chooses to express explicit interest.

2.5.1.4. Learning phase

It applies a learning algorithm to filter and exploit the user’s features from the feedback gathered in information collection phase.

2.5.1.5. Prediction/recommendation phase

It recommends or predicts what kind of items the user may prefer. This can be made either directly based on the dataset collected in information collection phase which could be memory based or model based or through the system's observed activities of the user. Fig. 2.5.1 highlights the recommendation phases.

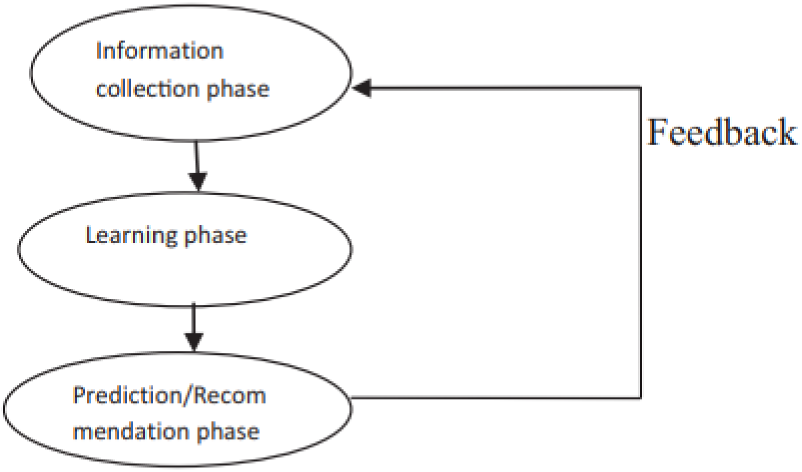


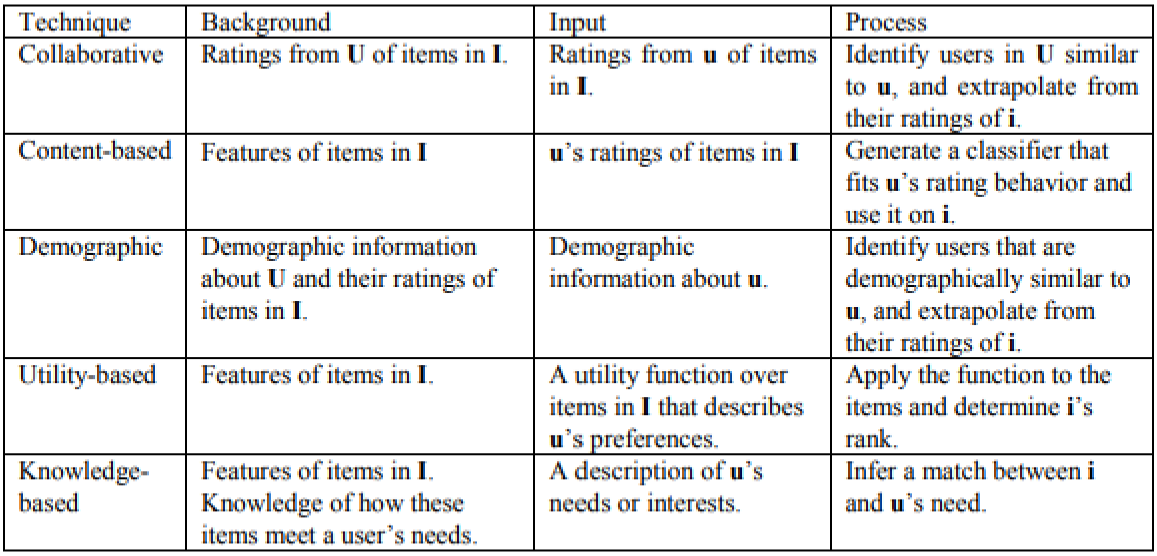
Fig 2.5.1 Recommendation phases.

2.6 Recommendation Techniques

Recommendation techniques (i.e. in Table 2.6.1) have a number of possible classifications [30]. Of interest in this discussion is not the type of interface or the properties of the user's interaction with the recommender, but rather the sources of data on which recommendation is based and the use to which that data is put. Specifically, recommender systems have (i) background data, the information that the system has before the recommendation process begins, (ii) input data, the information that user must communicate to the system in order to generate a recommendation, and (iii) an algorithm that combines background and input data to arrive at its suggestions. On this basis, we can distinguish five different recommendation techniques as shown in Table I. Assume that I is the set of items over which recommendations might be made, U is the set of users whose preferences are known, u is the user for whom recommendations need to be generated, and i is some item for which we would like to predict u's preference. Collaborative recommendation is probably the most familiar, most widely implemented and most mature of the technologies. Collaborative recommender systems aggregate ratings or recommendations of objects, recognize commonalities between users on the basis of their ratings, and generate new recommendations based on inter-user comparisons. A typical user profile in a collaborative system consists of a vector of items and their ratings, continuously augmented as the user interacts with the system over time. Some systems used time-based discounting of ratings to account for drift in user interests [31]. In some cases, ratings may be binary (like/dislike) or real-valued indicating degree of preference. Some of the most important systems using this technique are GroupLens or NetPerceptions, Tapestry and Recommender. These systems can be either memory based, comparing users against each other directly using correlation or other measures, or model-based, in which a model is derived from the historical rating data and used to make predictions. Model-based recommenders have used a variety of learning techniques including neural networks, latent semantic indexing, and Bayesian networks [32].

The greatest strength of collaborative techniques is that they are completely independent of any machine-readable representation of the objects being recommended, and work well for complex objects such as music and movies where variations in taste are responsible for much of the variation in preferences. (Schafer, Konstan & Riedl 1999) call this “people-to-people correlation.” Demographic recommender systems aim to categorize the user based on personal attributes and make

Table 2.6.1 Recommendation techniques



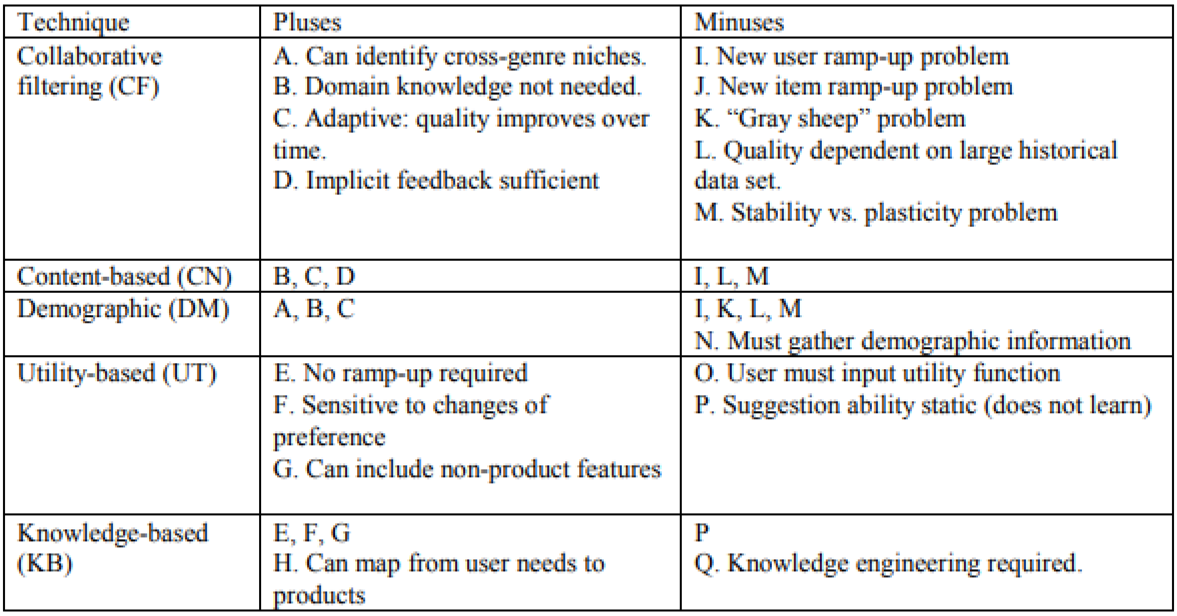
Technique	Background	Input	Process
Collaborative	Ratings from **U** of items in **I**.	Ratings from **u** of items in **I**.	Identify users in **U** similar to **u**, and extrapolate from their ratings of **i**.
Content-based	Features of items in **I**	**u**’s ratings of items in **I**	Generate a classifier that fits **u**’s rating behavior and use it on **i**.
Demographic	Demographic information about **U** and their ratings of items in **I**.	Demographic information about **u**.	Identify users that are demographically similar to **u**, and extrapolate from their ratings of **i**.
Utility-based	Features of items in **I**.	A utility function over items in **I** that describes **u**’s preferences.	Apply the function to the items and determine **i**’s rank.
Knowledge-based	Features of items in **I**. Knowledge of how these items meet a user’s needs.	A description of **u**’s needs or interests.	Infer a match between **i** and **u**’s need.

recommendations based on demographic classes. An early example of this kind of system was Grundy that recommended products based on personal information gathered through an interactive dialogue. The user’s responses were matched against a library of manually assembled user stereotypes. Some more recent recommender systems have also taken this approach. Krulwich (1997), for example, uses demographic groups from marketing research to suggest a range of products and services. A short survey is used to gather the data for user categorization. In other systems, machine learning is used to arrive at a classifier based on demographic data. The representation of demographic information in a user model can vary greatly. Rich's system used hand-crafted attributes with numeric confidence values. Pazzani's model uses Winnow to extract features from users' home pages that are predictive of liking certain restaurants. Demographic techniques form "people-to-people" correlations like collaborative ones, but use different data. The benefit of a demographic approach is that it may not require a history of user ratings of the type needed by collaborative and content-based techniques. Content-based recommendation is an outgrowth and continuation of information filtering research [33]. In a content-based system, the objects of interest are defined by their associated features. For example, text recommendation systems like the newsgroup filtering system NewsWeeder uses the words of their texts as features. A content-based recommender learns a profile of the user's interests based on the features present in objects the user has rated. (Schafer, Konstan & Riedl) call this "item-to-item correlation." The type of user profile derived by a content-based recommender depends on the learning method employed. Decision trees, neural nets, and vector-based representations have all been used. As in the collaborative case, content-based user profiles are long-term models and updated as more evidence about user preferences is observed. Utility-based and knowledge-based recommenders do not attempt to build long-term generalizations about their users, but rather base their advice on an evaluation of the match between a user's need and the set of options available. Utility-based recommenders make suggestions based on a computation of the utility of each object for the user. Of course, the central problem is how to create a utility function for each user. Tête-à-Tête and the e-commerce site PersonaLogic2 each have different techniques for arriving at a user-specific utility function and applying it to the objects under consideration. The user profile therefore is the utility function that the system has derived for the user, and the system employs constraint satisfaction techniques to locate the best match. The benefit of utility-based recommendation is that it can factor non-product attributes, such as vendor reliability and product availability, into the utility computation, making it possible for example to trade off price against delivery schedule for a user who has an immediate need. Knowledge-based recommendation attempts to suggest objects based on inferences about a user's needs and preferences. In some sense, all recommendation techniques could be described as doing some kind of inference. Knowledge-based approaches are distinguished in that they have functional knowledge: they have knowledge about how a particular item meets a particular user need, and can therefore reason about the relationship between a need and a possible recommendation. The user profile can be any knowledge structure that supports this inference. In the simplest case, as in Google, it may simply be the query that the user has formulated. In others, it may be a more detailed representation of the user's needs. The Entree system and several other recent systems (for example, [34]) employ techniques from case-based reasoning for knowledge-based recommendation. (Schafer, Konstan & Riedl) call knowledge-based recommendation the “Editor’s choice” method. The knowledge used by a knowledge-based recommender can also take many forms. Google uses information about the links between web pages to infer popularity and authoritative value. Entree uses knowledge of cuisines to infer similarity between restaurants. Utility-based approaches calculate a utility value for objects to be recommended, and in principle, such calculations could be based on functional knowledge. However, existing systems do not use such inference, requiring users to do their own mapping between their needs and the features of products, either in the form of preference functions for each feature.

2.6.1 Comparing Recommendation Techniques

All recommendation techniques have strengths and weaknesses discussed below and summarized in Table 2.6.2. Perhaps the best known is the “ramp-up” problem [35]. This term actually refers to two distinct but related problems New User: Because recommendations follow from a comparison between the target user and other users based solely on the accumulation of ratings, a user with few ratings becomes difficult to categorize. New Item: Similarly, a new item that has not had many ratings also cannot be easily recommended: the “new item” problem. This problem shows up in domains such as news articles where there is a constant stream of new items and each user only rates a few. It is also known as the “early rater” problem, since the first person to rate an item gets little benefit from doing so: such early ratings do not improve a user’s ability to match against others. This makes it necessary for recommender systems to provide other incentives to encourage users to provide ratings. Collaborative recommender systems depend on overlap in ratings across users and have difficulty when the space of ratings is sparse: few users have rated the same items. The sparsity problem is somewhat reduced in model-based approaches, such as singular value decomposition which can reduce the dimensionality of the space in which comparison takes place. Still sparsity is a significant problem in domains such as news filtering, since there are many items available and, unless the user base is very large, the odds that another user will share a large number of rated items is small. These three problems suggest that pure collaborative techniques are best suited to problems where the density of user interest is relatively high across a small and static universe of items. If the set of items changes too rapidly, old ratings will be of little value to new users who will not be able to have their ratings compared to those of the existing users. If the set of items is large and user interest thinly spread, then the probability of overlap with other users will be small. Collaborative recommenders work best for a user who fits into a niche with many neighbors of similar taste. The technique does not work well for so-called “gray sheep” [36], who fall on a border between existing cliques of users. This is also a problem for demographic systems that attempt to categorize users on personal characteristics. On the other hand, demographic recommenders do not have the “new user” problem, because they do not require a list of ratings from the user. Instead they have the problem of gathering the requisite demographic information. With sensitivity to on-line privacy increasing, especially in electronic commerce contexts, demographic recommenders are likely to remain rare: the data most predictive of user preference is likely to be information that users are reluctant to disclose. Content-based techniques also have a start-up problem in that they must accumulate enough ratings to build a reliable classifier. Relative to collaborative filtering, content-based techniques also have the problem that they are limited by the features that are explicitly associated with the objects that they recommend. For example, content based movie recommendation can only be based on written materials about a movie: actors’ names, plot summaries, etc. because the movie itself is opaque to the system. This puts these techniques at the mercy of the descriptive data available. Collaborative systems rely only on user ratings and can be used to recommend items without any descriptive data. Even in the presence of descriptive data, some experiments have found that collaborative recommender systems can be more accurate than content-based ones. The great power of the collaborative approach relative to content-based ones is its cross-genre or “outside the box” recommendation ability. It may be that listeners who enjoy free jazz also enjoy avant-garde classical music, but a content-based recommender trained on the preferences of a free jazz aficionado would not be able to suggest items in the classical realm since none of the features (performers, instruments, repertoire) associated with items in the different categories would be shared. Only by looking outside the preferences of the individual can such suggestions be made. Both content-based and collaborative techniques suffer from the “portfolio effect.” An ideal recommender would not suggest a stock that the user already owns or a movie she has already seen. The problem becomes quite tricky in domains such as news filtering, since stories that look quite similar to those already read may in fact present some new facts or new perspectives that would be valuable to the user. At the same time, many different presentations of the same wire-service story from different newspapers would not be useful. The DailyLearner system (Billsus & Pazzani, 2000) uses an upper bound of similarity in its content-based recommender to filter out news items too similar to those already seen by the user. Utility-based and knowledge-based recommenders do not have ramp-up or sparsity problems, since they do not base their recommendations on accumulated statistical evidence. Utility-based techniques require that the system build a complete utility function across all features of the objects under consideration. One benefit of this approach is that it can incorporate many different factors that contribute to the value of a product, such as delivery schedule, warranty terms or conceivably the user's existing portfolio, rather than just product-specific features. In addition, these non-product features may have extremely idiosyncratic utility: how soon something can be delivered may matter very much to a user facing a deadline. A utility-based framework thereby lets the user express all of the 5 considerations that need to go into a recommendation. For this reason, Guttman (1999) describes Tête-à-Tête as "product and merchant brokering" system rather than a recommender system. However, under the definition given above, Tête-à-Tête does fit since its main output is a recommendation (a top-ranked item) that is generated on a personalized basis. The flexibility of utility-based systems is also to some degree a failing. The user must construct a complete preference function, and must therefore weigh the significance of each possible feature. Often this creates a significant burden of interaction. Tête-à-Tête uses a small number of "stereotype" preference functions to get the user started, but ultimately the user needs to look at, weigh, and select a preference function for each feature that describes an item of interest. This might be feasible for items with only a few characteristics, such as price, quality and delivery date, but not for more complex and subjective domains like movies or news articles. PersonaLogic does not require the user to input a utility function, but instead derives the function through an interactive questionnaire. While the complete explicit utility function might be a boon to some users, for example, technical users with specific purchasing requirements, it is likely to overwhelm a more casual user with a less-detailed knowledge. Large moves in the product space, for example, from "sports cars" to "family cars" require a complete re-tooling of the preference function, including everything from interior space to fuel economy. This makes a utility-based system less appropriate for the casual browser. Knowledge-based recommender systems are prone to the drawback of all knowledge-based systems: the need for knowledge acquisition. There are three types of knowledge that are involved in such a system: Catalog knowledge: Knowledge about the objects being recommended and their features. For example, the Entree recommender should know that "Thai" cuisine is a kind of "Asian" cuisine. Functional knowledge: The system must be able to map between the user's needs and the object that might satisfy those needs. For example, Entree knows that a need for a romantic dinner spot could be met by a restaurant that is "quiet with an ocean view." User knowledge: To provide good recommendations, the system must have some knowledge about the user. This might take the form of general demographic information or specific information about the need for which a recommendation is sought. Of these knowledge types, the last is the most challenging, as it is, in the worst case, an instance of the general user-modeling problem. Despite this drawback, knowledge-based recommendation has some beneficial characteristics. It is appropriate for casual exploration, because it demands less of the user than utility-based recommendation. It does not involve a startup period during which its suggestions are low quality. A knowledge-based recommender cannot “discover” user niches, the way collaborative systems can. On the other hand, it can make recommendations as wide-ranging as its knowledge base allows. Table 2.6.2 summarizes the five recommendation techniques that we have discussed here, pointing out the pros and cons of each. Collaborative and demographic techniques have the unique capacity to identify cross-genre niches and can entice users to jump outside of the familiar. Knowledge-based techniques can do the same but only if such associations have been identified ahead of time by the knowledge engineer. All of the learning-based techniques (collaborative, content-based and demographic) suffer from the ramp-up problem in one form or another. The converse of this problem is the stability vs. plasticity problem for such learners. Once a user’s profile has been established in the system, it is difficult to change one’s preferences. A steak-eater who becomes a vegetarian will continue to get steakhouse recommendations from a content-based or collaborative recommender for some time, until newer ratings have the chance to tip the scales. Many adaptive systems include some sort of temporal discount to cause older ratings to have less influence, but they do so at the risk of losing information about interests that are long-term but sporadically exercised. For example, a user might like to read about major earthquakes when they happen, but such occurrences are sufficiently rare that the ratings associated with last year’s earthquake are gone by the time the next big one hits. Knowledge and utility-based recommenders respond to the user’s immediate need and do not need any kind of retraining when preferences change. The ramp-up problem has the side-effect of excluding casual users from receiving the full benefits of collaborative and content-based recommendation. It is possible to do simple market-basket recommendation with minimal user input: Amazon.com’s “people who bought X also bought Y” but this mechanism has few of the advantages commonly associated with the collaborative filtering concept. The learning-based technologies work best for dedicated users who are willing to invest some time making their preferences known to the system. Utility- and knowledge-based systems have fewer problems in this regard because they do not rely on having historical data about a user’s preferences. Utility-based systems may present difficulties for casual users who might be unwilling to tailor a utility function simply to browse a catalog.

Table 2.6.2 Tradeoffs between Recommendation Techniques.



Technique	Pluses	Minuses
Collaborative filtering (CF)	A. Can identify cross-genre niches. B. Domain knowledge not needed. C. Adaptive: quality improves over time. D. Implicit feedback sufficient	I. New user ramp-up problem J. New item ramp-up problem K. “Gray sheep” problem L. Quality dependent on large historical data set. M. Stability vs. plasticity problem
Content-based (CN)	B, C, D	I, L, M
Demographic (DM)	A, B, C	I, K, L, M N. Must gather demographic information
Utility-based (UT)	E. No ramp-up required F. Sensitive to changes of preference G. Can include non-product features	O. User must input utility function P. Suggestion ability static (does not learn)
Knowledge-based (KB)	E, F, G H. Can map from user needs to products	P Q. Knowledge engineering required.

2.6.1.1 Hybrid Recommender Systems

Hybrid recommender systems combine two or more recommendation techniques to gain better performance with fewer of the drawbacks of any individual one. Most commonly, collaborative filtering is combined with some other technique in an attempt to avoid the ramp-up problem. Table 2.6.3 shows some of the combination methods that have been employed.

2.6.1.2 Weighted

A weighted hybrid recommender is one in which the score of a recommended item is computed from the results of all of the available recommendation techniques present in the system. For example, the simplest combined hybrid would be a linear combination of recommendation scores. The P-Tango system [36], uses such a hybrid. It initially gives collaborative and content-based recommenders equal weight, but gradually adjusts the weighting as predictions about user ratings are confirmed or disconfirmed. Pazzani’s combination hybrid does not use numeric scores, but rather treats the output of each recommender (collaborative, content-based and demographic) as a set of votes, which are then combined in a consensus scheme. The benefit of a weighted hybrid is that all of the system’s capabilities are brought to bear on the recommendation process in a straightforward way and it is easy to perform post-hoc credit assignment and adjust the hybrid accordingly. However, the implicit assumption in this technique is that the relative value of the different techniques is more or less uniform across the space of possible items. From the discussion above, we know that this is not always so: a collaborative recommender will be weaker for those items with a small number of raters.

2.6.1.3 Switching

A switching hybrid builds in item-level sensitivity to the hybridization strategy, the system uses some criterion to switch between recommendation techniques. The DailyLearner system uses a content/collaborative hybrid in which a content-based recommendation method is employed first. If the content-based system cannot make a required – in PTV, content-based recommendation takes precedence over collaborative responses. Other implementations of the mixed hybrid, ProfBuilder and PickAFlick [37], present multiple recommendation sources side-by-side. Usually, recommendation requires ranking of items or selection of a single best recommendation, at which point some kind of combination technique must be employed.

2.6.1.4 Feature Combination

Another way to achieve the content/collaborative merger is to treat collaborative information as simply additional feature data associated with each example and use content-based techniques over this augmented data set. For example, (Basu, Hirsh & Cohen 1998) report on experiments in which the inductive rule learner Ripper was applied to the task of recommending movies using both user ratings and content features, and achieved significant improvements in precision over a purely collaborative approach. However, this benefit was only achieved by hand filtering content features. The authors found that employing all of the available content features improved recall but not precision. The feature combination hybrid lets the system consider collaborative data without relying on it exclusively, so it reduces the sensitivity of the system to the number of users who have rated an item. Conversely, it lets the system have information about the inherent similarity of items that are otherwise opaque to a collaborative system.

2.6.1.4 Cascade

Unlike the previous hybridization methods, the cascade hybrid involves a staged process. In this technique, one recommendation technique is employed first to produce a coarse ranking of candidates and a second technique refines the recommendation from among the candidate set. The restaurant recommender EntreeC is a cascaded knowledge-based and collaborative recommender. Like Entree, it uses its knowledge of restaurants to make recommendations based on the user's stated interests. The recommendations are placed in buckets of equal preference, and the collaborative technique is employed to break ties, further ranking the suggestions in each bucket. Cascading allows the system to avoid employing the second, lower-priority, technique on items that are already well-differentiated by the first or that are sufficiently poorly-rated that they will never be recommended. Because the cascade's second step focuses only on those items for which additional discrimination is needed, it is more efficient than, for example, a weighted hybrid that applies all of its techniques to all items. In addition, the cascade is by its nature tolerant of noise in the operation of a low-priority technique, since ratings given by the high-priority recommender can only be refined, not overturned [38].

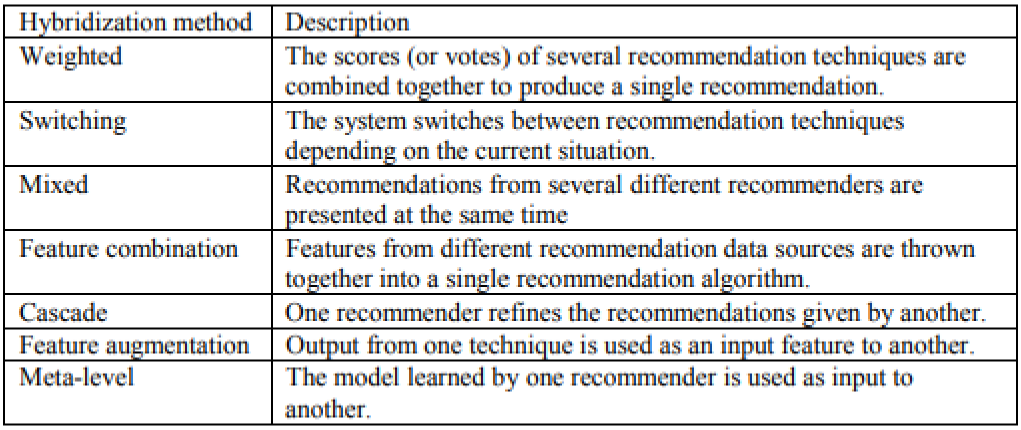
2.6.1.5 Feature Augmentation

One technique is employed to produce a rating or classification of an item and that information is then incorporated into the processing of the next recommendation technique. For example, the Libra system makes content-based recommendations of products based on data found in Amazon.com, using a naive Bayes text classifier. In the text data used by the system is included "related authors" and "related titles" information that Amazon generates using its internal collaborative systems. These features were found to make a significant contribution to the quality of recommendations. The GroupLens research team working with Usenet news filtering also employed feature augmentation. They implemented a set of knowledge-based "filterbots" using specific criteria, such as the number of spelling errors and the size of included messages. These bots contributed ratings to the database of ratings used by the collaborative part of the system, acting as artificial users. With fairly simple agent implementations, they were able to improve email filtering. Augmentation is attractive because it offers a way to improve the performance of a core system, like the NetPerceptions' GroupLens Recommendation Engine or a naive Bayes text classifier, without modifying it. Additional functionality is added by intermediaries who can use other techniques to augment the data itself. Note that this is different from feature combination in which raw data from different sources is combined. While both the cascade and augmentation techniques sequence two recommenders, with the first recommender having an influence over the second, they are fundamentally quite different. In an augmentation hybrid, the features used by the second recommender include the output of the first one, such as the ratings contributed by GroupLens' filterbots. In a cascaded hybrid, the second recommender does not use any output from the first recommender in producing its rankings, but the results of the two recommenders are combined in a prioritized manner [38].

2.6.1.6 Meta-level

Another way that two recommendation techniques can be combined is by using the model generated by one as the input for another. This differs from feature augmentation: in an augmentation hybrid, we use a learned model to generate features for input to a second algorithm; in a meta-level hybrid, the entire model becomes the input. The first meta-level hybrid was the web filtering system Fab. In Fab, user-specific selection agents perform content-based filtering using Rocchio's method to maintain a term vector model that describes the user's area of interest. Collection agents, which garner new pages from the web, use the models from all users in their gathering operations. So, documents are first collected on the basis of their interest to the community as a whole and then distributed to particular users. In addition to the way that user models were shared, Fab was also performing a cascade of collaborative collection and content-based recommendation, although the collaborative step only created a pool of documents and its ranking information was not used by the selection component. A meta-level hybrid that focuses exclusively on recommendation is described by Pazzani (1999) as "collaboration via content". A content-based model is built by Winnow (Littlestone & Warmuth 1994) for each user describing the features that predict restaurants the user likes. These models, essentially vectors of terms and weights, can then be compared across users to make predictions. More recently, (Condliff et al. 1999) have used a two-stage Bayesian mixed-effects scheme: a content-based naive Bayes classifier is built for each user and then the parameters of the classifiers are linked across different users using regression. LaboUr (Schwab, et al. 2001) uses instance-based learning to create content-based user profiles which are then compared in a collaborative manner. The benefit of the meta-level method, especially for the content/collaborative hybrid is that the learned model is a compressed representation of a user's interest, and a collaborative mechanism that follows can operate on this information-dense representation more easily than on raw rating data [38].

Table 2.6.3: Hybridization Methods



Hybridization method	Description
Weighted	The scores (or votes) of several recommendation techniques are combined together to produce a single recommendation.
Switching	The system switches between recommendation techniques depending on the current situation.
Mixed	Recommendations from several different recommenders are presented at the same time
Feature combination	Features from different recommendation data sources are thrown together into a single recommendation algorithm.
Cascade	One recommender refines the recommendations given by another.
Feature augmentation	Output from one technique is used as an input feature to another.
Meta-level	The model learned by one recommender is used as input to another.

2.7 Machine learning

Machine Learning (ML) uses computers to simulate human learning and allows computers to identify and acquire knowledge from the real world, and improve performance on some tasks based on this new knowledge. More formally, [39] defines ML as follows: “A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E”. Although the first concepts of ML originated in the 1950s, ML was studied as a separate field in the 1990s [39]. Today, ML algorithms are used in several areas besides computer science, including business [40], advertising [41] and medicine [42]. Learning is the process of knowledge acquisition. Humans naturally learn from experience because of their ability to reason. In contrast, computers do not learn by reasoning, but learn with algorithms. Today, there are a large number of ML algorithms proposed in the literature. They can be classified based on the approach used for the learning process. There are four main classifications: supervised, unsupervised, semi supervised, and reinforcement learning. Supervised learning happens when algorithms are provided with training data and correct answers. The task of the ML algorithm is to learn based on the training data, and to apply the knowledge that was gained in real data. As an example, consider an ML learning algorithm being used in products classification in a product store. A training set (training data + answers) can be a table relating information about each products to a correct classification. Here, information about each products may be title, author, or even every word a products contains. The ML algorithm learns with the training set. When a new products arrives at the product store, the algorithm can classify it based on the knowledge about products classification it has acquired. In unsupervised learning, ML algorithms do not have a training set. They are presented with some data about the real world and have to learn from that data on their own. Unsupervised learning algorithms are mostly focused on finding hidden patterns in data. For example, suppose that an ML algorithm has access to user profile information in a social network. By using an unsupervised learning approach, the algorithm can separate users into personality categories, such as outgoing and reserved, allowing the social network company to target advertising more directly at specific groups of users. ML algorithms can also be classified as semi-supervised. Semi-supervised learning occurs when algorithms work with a training set with missing information, and still need to learn from it. An example is when an ML algorithm is provided with movie ratings. Not every user rated every movie and so there is some missing information. Semi supervised learning algorithms are able to learn and draw conclusions even with incomplete data. Lastly, ML algorithms might have a reinforcement learning approach. Reinforcement learning occurs when algorithms learn based on external feedback given either by a thinking entity, or the environment. This approach is analogous to teaching dogs to sit or jump. When the dog performs the action correctly, the dog receives a small cookie (positive feedback). It does not receive any cookie (negative feedback) if it performs the wrong action. As an example, in the computer science field, consider an ML algorithm that plays games against an opponent. Moves that lead to victories (positive feedback) in the game should be learned and repeated, whereas moves that lead to losses (negative feedback) are avoided. ML has become quite popular recently with the increase in processor speed and memory size. As a consequence, the field now has a large number of algorithms that use mathematical or statistical analysis to learn, draw conclusions or infer data. This number continues to increase as evidenced by the number of scientific publications that propose variations or combinations of ML algorithms. For that reason, ML algorithms have been categorized based on the purpose for which they are designed. Some examples of classification can be found in [43], and [44], although the field still does not have any standards.

2.8 Related Works

2.8.1 Recommendation in e-commerce.

Amazon.com: we focus on recommender systems in the products section of Amazon.com. Customers who bought: Like many E-commerce sites, Amazon.com (www.amazon.com) is structured with an information page for each products, giving details of the text and purchase information. The Customers Who Bought feature is found on the information page for each products in their catalog. It is in fact two separate recommendation lists. The first recommends products frequently purchased by customers who purchased the selected products. The second recommends authors whose products are frequently purchased by customers who purchased works by the author of the selected products. Eyes: The Eyes feature allows customers to be notified via email of new items added to the Amazon.com catalog. Customers enter requests based upon author, title, subject, ISBN, or publication date information. Customers can use both simple and more complex Boolean-based criteria (AND/OR) for notification queries. Requests can also be directly entered from any search results screen, creating a persistent request based on the search.Amazon.com delivers: Amazon.com Delivers is a variation on the Eyes feature. Customers select checkboxes to choose from a list of specific categories/genres (Oprah products, biographies, cooking). Periodically the editors at Amazon.com send email announcements to notify subscribers of their latest recommendations in the subscribed categories. Products Matcher: The Products Matcher feature allows customers to give direct feedback about products they have read. Customers rate products they have read on a 5-point scale from “hated it” to “loved it.” After rating a sample of products, customers may request recommendations for products they might like. At that point, a half dozens of non-rated texts are presented which correlate with the user’s indicated tastes. Feedback to these recommendations is provided by a “rate these products” feature where customers can indicate a rating for one or more of the recommended products. Customer Comments: The Customer Comments feature allows customers to receive text recommendations based on the opinions of other customers. Located on the information page for each products is a list of 1-5-star ratings and written comments provided by customers who have read the products in question and submitted a review. Customers have the option of incorporating these recommendations into their purchase decision. The use of recommender systems in an e-commerce environment can impact financial performance as well as the intensity of the dialogue with customers. More specifically, recommender systems can enhance e-commerce dialogues in three ways:

- Conversion: Turning Browsers into Buyers increasing the proportion of visitors to a Web-site that make a purchase. Recommender systems help consumers find items that best fit their interests and inclinations; these may include unplanned purchases driven by serendipity from the recommendations made.
- By increasing Cross-sell: Recommender systems improve cross selling by suggesting additional products or services to customers. If the recommendations are good, the average order size increases. For instance, a site might recommend

additional products in the checkout process, based on those products already in the shopping cart.

By building loyalty in a world where competitors are only a click away, building customer-loyalty becomes an essential aspect of business strategy. Recommender systems can improve loyalty by creating a value-added relationship between the site and the customer. Each time a customer visits a website, the system “learns” more about that customer’s preferences and interests and is increasingly able to operationalize this information to e.g. personalize what is offered. By providing each customer with an increasingly relevant experience, a corresponding improvement in the likelihood of that customer returning is achieved. Ultimately, the depth of insight gained into a customer’s preferences and interests can be so great that even if a competitor were to launch an identical, or even superior system, the customer would need to spend an inordinate amount of time and energy “teaching” the competitor to offer a similarly attractive experience [57]. Fig 2.8.1 shows the system architecture of amazon recommender system [58], while Fig 2.8.2 shows the real-world application of amazon recommender system [59].

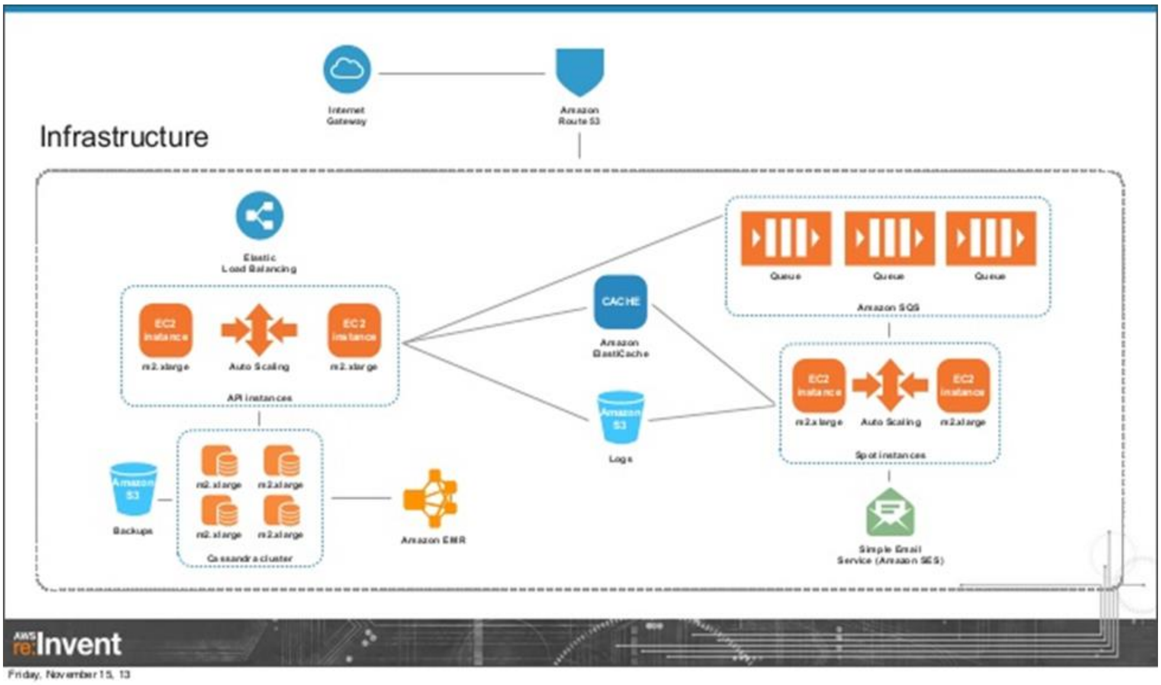
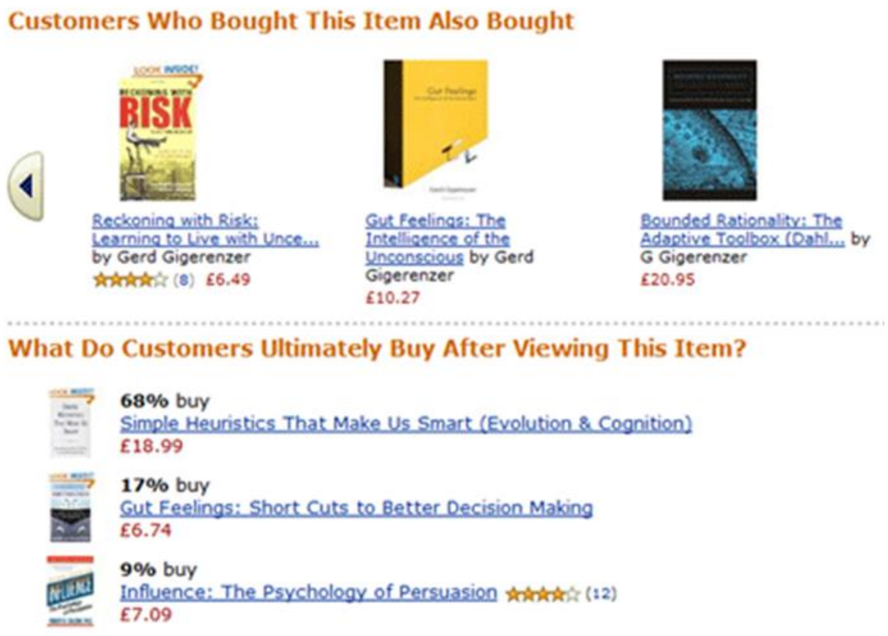


Fig 2.8.1 amazon recommender system architecture.



Customers Who Bought This Item Also Bought

Reckoning with Risk: Learning to Live with Unce... by Gerd Gigerenzer ★★★★☆ (8) £6.49

Gut Feelings: The Intelligence of the Unconscious by Gerd Gigerenzer £10.27

Bounded Rationality: The Adaptive Toolbox (Dahl... by G Gigerenzer £20.95

What Do Customers Ultimately Buy After Viewing This Item?

68% buy Simple Heuristics That Make Us Smart (Evolution & Cognition) £18.99

17% buy Gut Feelings: Short Cuts to Better Decision Making £6.74

9% buy Influence: The Psychology of Persuasion ★★★★☆ (12) £7.09

Fig 2.8.2 amazon recommender system example.

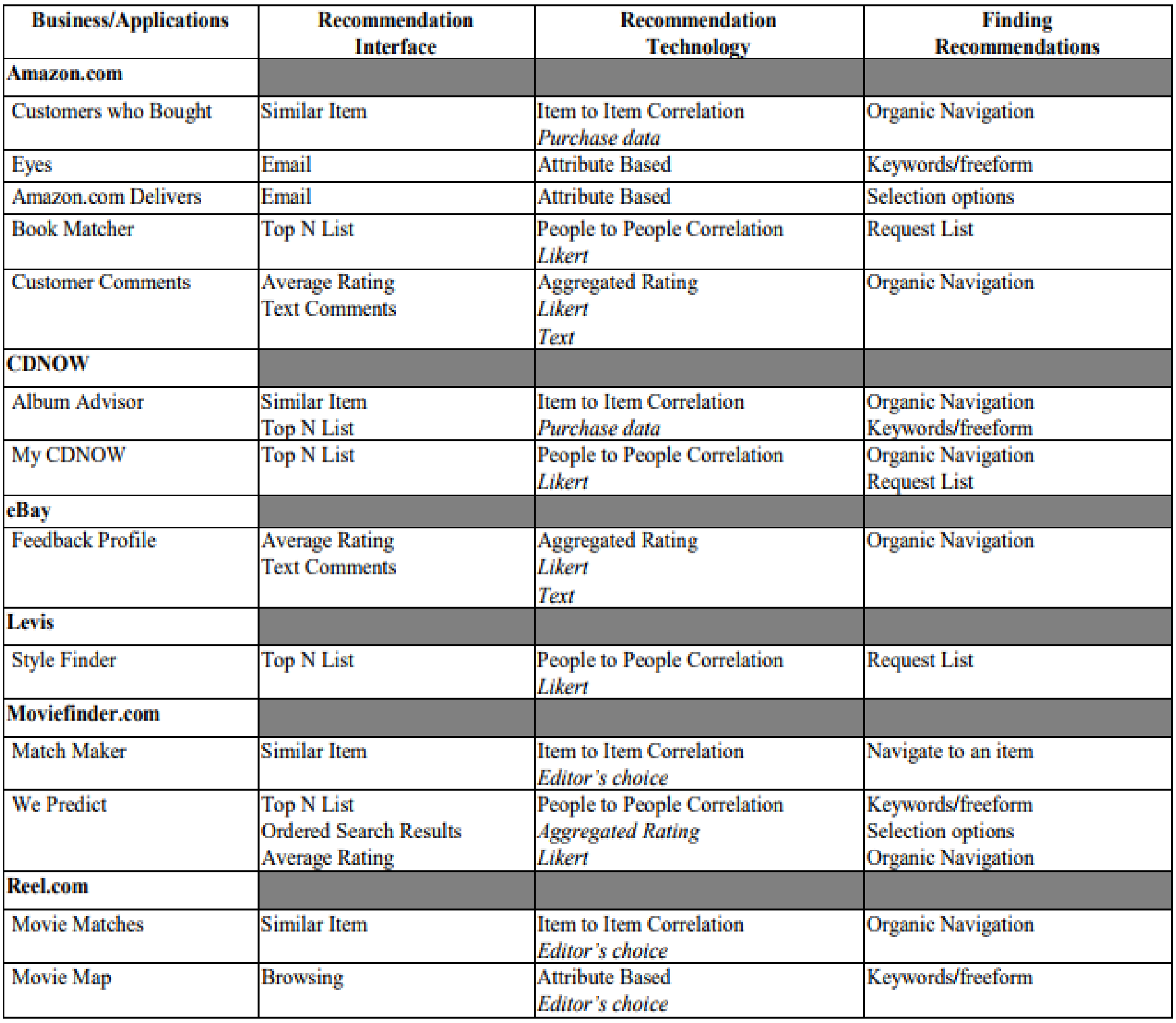
2.8.2 Moviefinder.com Match Maker

Moviefinder.com's Match Maker (www.moviefinder.com) allows customers to locate movies with a similar "mood, theme, genre or cast" to a given movie. From the information page of the movie in question, customers click on the Match Maker icon and are provided with the list of recommended movies, as well as links to other films by the original film's director and key actors. We Predict: We Predict recommends movies to customers based on their previously indicated interests. Customers enter a rating on a 5-point scale -- from A to F – for movies they have viewed. These ratings are used in two different ways. Most simply, as they continue, the information page for non-rated movies contains a personalized textual prediction (go see it – forget it). In a variation of this, customers can use Power find to search for top picks based on syntactic criteria such as Genre, directors, or actors and choose to have these sorted by their personalized prediction or by the all customer average. [59].

2.8.3 CDNOW Album Advisor

The Album Advisor feature of CDNOW (www.cdnow.com) works in two different modes. In the single album mode customers locate the information page for a given album. The system recommends 10 other albums related to the album in question. In the multiple artist mode customers enter up to three artists. In turn, the system recommends 10 albums related to the artists in question. My CDNOW: My CDNOW enables customers to set up their own music store, based on albums and artists they like. Customers indicate which albums they own, and which artists are their favorites. Purchases from CDNOW are entered automatically into the "own it" list. Although "own it" ratings are initially treated as an indication of positive likes, customers can go back and distinguish between "own it and like it" and "own it but dislike it." When customers request recommendations, the system will predict 6 albums the customer might like based on what is already owned. A feedback option is available by customers providing a "own it," "move to wish list" or "not for me" comment for any of the albums in this prediction list. The albums recommended change based on the feedback [56].In Table 2.8.1 we have summarized the applications, interfaces, recommendation technology, and how users find recommendations for all of the example applications. The first column just names each application, under the E-commerce site that houses it. The second column describes the interface that is used for delivering the recommendations. The third column describes the recommendation technology used by the site, and the inputs required by that technology. The fourth column describes how users find recommendations using the application.

Table 2.8.1: Recommender System Examples



Business/Applications	**Recommendation Interface**	**Recommendation Technology**	**Finding Recommendations**
Amazon.com			
Customers who Bought	Similar Item	Item to Item Correlation *Purchase data*	Organic Navigation
Eyes	Email	Attribute Based	Keywords/freeform
Amazon.com Delivers	Email	Attribute Based	Selection options
Book Matcher	Top N List	People to People Correlation *Likert*	Request List
Customer Comments	Average Rating Text Comments	Aggregated Rating *Likert* *Text*	Organic Navigation
CDNOW			
Album Advisor	Similar Item Top N List	Item to Item Correlation *Purchase data*	Organic Navigation Keywords/freeform
My CDNOW	Top N List	People to People Correlation *Likert*	Organic Navigation Request List
eBay			
Feedback Profile	Average Rating Text Comments	Aggregated Rating *Likert* *Text*	Organic Navigation
Levis			
Style Finder	Top N List	People to People Correlation *Likert*	Request List
Moviefinder.com			
Match Maker	Similar Item	Item to Item Correlation *Editor's choice*	Navigate to an item
We Predict	Top N List Ordered Search Results Average Rating	People to People Correlation *Aggregated Rating* *Likert*	Keywords/freeform Selection options Organic Navigation
Reel.com			
Movie Matches	Similar Item	Item to Item Correlation *Editor's choice*	Organic Navigation
Movie Map	Browsing	Attribute Based *Editor's choice*	Keywords/freeform

2.9 Summary

In this chapter, we gave a comprehensive overview of the background theory and some examples of the related works to online recommender systems and telecommunications marketing, that are applied in various fields such as E-commerce and Entertainment. Recommender systems are used in numerous application domains, such as retail, music, content, Web search, querying, and computational advertisements. Some of these domains require specialized methods for adapting recommender systems. All these application domains are Web centric in nature. An important aspect of recommender systems is that they assume the existence of strong user-identification mechanisms in order to track and identify long-term user interests. In many Web domains, mechanisms for strong user identification may not be available. In such cases, direct user of recommendation technology may not be feasible. Furthermore, since new items (advertisements) continually enter and leave the system, certain types of methods such as multi-armed bandits are particularly suitable [60].

CHAPTER THREE
METHODOLOGY

3.1 Introduction

In this chapter, we look at the materials and methods used to conduct this study. We begin by looking at the methodology that was used to conduct the baseline study. This is followed by the methodology that was used to design the models and implement the prototype, the first step was to understand the problems we are trying to solve each marketing problem required a specific algorithm not all algorithms will solve the same problem, we get a data set of CDR's from the telecom company, split it into training set and test set, apply the machine learning algorithms on the data based of the marketing problem that the particular company is facing, get that data and put it in a database for end users to use via the GUI interface so that they will be able to see the products that are recommended to them.

Research Question	**Technical Objective**	**Methodology**
How do we analyse the relationship between telecommunication subscribers and telecommunication products?	Examine the current existing methods used by recommender systems such as Netflix, Amazon, and Facebook Products.	Analyse available technologies used in the telecommunication industries in relation to the recommender systems that currently exist, select the most appropriate machine learning models
What are the challenges telecommunication companies face that lead to high revenue loss, churn and bad customer experience	Establish challenges telecommunication companies in Zambia face in terms of low revenue, churn and fraud.	Review the problem that comes with traditional analysis done by telecom industries like SQL queries and excel
How best will Product recommender systems for mobile technology be utilized in order to assist solving the problem of low revenue, churn and fraud?	Design and implement a product recommender system which will recommend products that a subscriber is more likely to use.	Recommender system will be able to recommend the products that a customer is more likely to use based of their preference of past historical purchases, this will in turn help making good analytical decisions and target the right customers which will increase revenue and reduce churn

3.2 Baseline Study

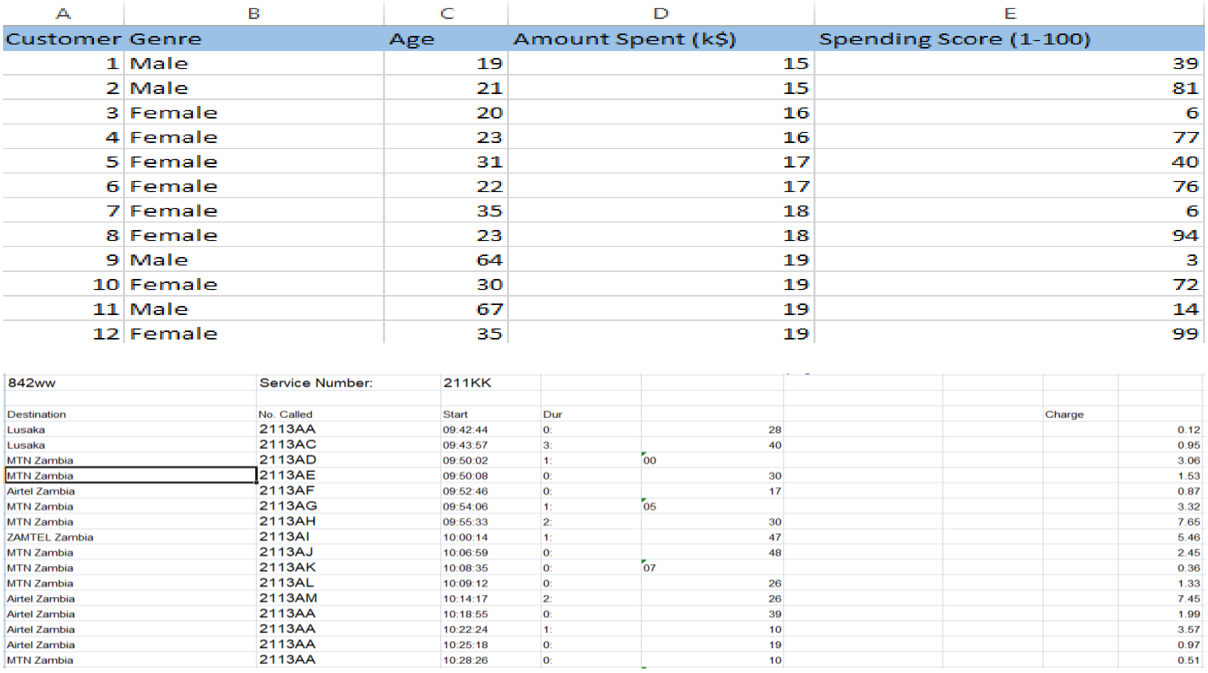
The purpose of the baseline study was to establish the challenges faced by the telecommunication companies regarding revenue leakage and customer experience. Our Research Methodology comprises of qualitative research type [61]. Qualitative case study methodology provides tools for researchers to study complex phenomena within their contexts. When the approach is applied correctly, it becomes a valuable method for researchers to develop theory, evaluate programs, and develop interventions [61]. A descriptive research design was utilized. A descriptive research design involves observing and describing the behavior of a subject without influencing it in any way [62].

3.2.1 Study Setting

The study was conducted in Lusaka using the four-existing mobile telecommunication service providers in Zambia. These companies were selected because they are the current telecommunications companies in Zambia and they are all facing the same problem of failing to maintain their subscriber bases as their numbers fluctuate.

3.2.2 Data Collection

Data collection was carried out over a period of 12 weeks i.e. three months historical data. The historical data dump or raw data was collected from the CDR's for a certain telecommunication company that was at a later stage analyzed to arrive at conclusions from them. SQL procedures were used to generate the SQL dump after the raw CDR's were processed using big data techniques. This data was split into **training set**, **test set** and **validation set,** below is a sample of CDR data extracted as shown in fig 3.1



A	B	C	D	E
Customer	Genre	Age	Amount Spent (k$)	Spending Score (1-100)
1	Male	19	15	39
2	Male	21	15	81
3	Female	20	16	6
4	Female	23	16	77
5	Female	31	17	40
6	Female	22	17	76
7	Female	35	18	6
8	Female	23	18	94
9	Male	64	19	3
10	Female	30	19	72
11	Male	67	19	14
12	Female	35	19	99

842ww	Service Number:	211KK						
Destination	No. Called	Start	Dur				Charge	
Lusaka	2113AA	09:42:44	0:	28				0.12
Lusaka	2113AC	09:43:57	3:	40				0.95
MTN Zambia	2113AD	09:50:02	1:	00				3.06
MTN Zambia	2113AE	09:50:08	0:	30				1.53
Airtel Zambia	2113AF	09:52:46	0:	17				0.87
MTN Zambia	2113AG	09:54:06	1:	05				3.32
MTN Zambia	2113AH	09:55:33	2:	30				7.65
ZAMTEL Zambia	2113AI	10:00:14	1:	47				5.46
MTN Zambia	2113AJ	10:06:59	0:	48				2.45
MTN Zambia	2113AK	10:08:35	0:	07				0.36
MTN Zambia	2113AL	10:09:12	0:	26				1.33
Airtel Zambia	2113AM	10:14:17	2:	26				7.45
Airtel Zambia	2113AA	10:18:55	0:	39				1.99
Airtel Zambia	2113AA	10:22:24	1:	10				3.57
Airtel Zambia	2113AA	10:25:18	0:	19				0.97
MTN Zambia	2113AA	10:28:26	0:	10				0.51

Figure 3.1. Unprocessed call detailed report (objective 4)

3.2.3 Data Processing and Analysis

Data collected was analyzed using statistical or analytical software packages such as DATO, SPSS, Excel, and Tableau. This is because we wanted to do a manual test on whether the data we are going to feed the system will come out as we want it when we use machine learning algorithms, Excel was used to put the important CDR’s fields that we were interested in, Dato was used to for data mining tasks of call detail reports, while Tableau & SPSS was used see and understand their data by connecting the database in order to create visualizations, and share with a click.

3.2.4 Ethical Consideration

Recommender systems are based heavily on feedback from the users, which might be implicit or explicit. This feedback contains significant information about the interests of the user, and it might reveal information about their political opinions, sexual orientations, and personal preferences. In many cases, such information can be highly sensitive, which leads to privacy concerns. Such privacy concerns are significant in that they impede the release of data necessary for the advancement of recommendation algorithm, the data collected will be truncated and hidden from non-researchers or analysts this is because most of the information is private, a call detail report is like a receipt that shows the events of a subscriber.

3.2.5 Limitations of the Baseline Study

The ideal situation was to make this application using three version a website version for smartphone users and backend interface for business intelligence analysts and a USSD application for feature phone users but due to financial constraints I was unable to buy a USSD short code from ZICTA which costs approximately 35 thousand kwacha, a new laptop was required as the old one crushed because it was too small to run the applications.

3.2.6 PRESENTATION OF FINDINGS

The data was summarized and presented in form of tables and figs such as pie charts and bar charts to facilitate understanding.

3.2.6.1 Machine Learning algorithms used

Supervised Learning: This algorithm consists of a target / outcome variable (or dependent variable) which is to be predicted from a given set of predictors (independent variables). Using these set of variables, we generate a function that map inputs to desired outputs. The training process continues until the model achieves a desired level of accuracy on the training data. Examples of Supervised Learning: Regression, Decision Tree, Random Forest, KNN, Logistic Regression [45].

Unsupervised Learning: In this algorithm, we do not have any target or outcome variable to predict / estimate. It is used for clustering population in different groups, which is widely used for segmenting customers in different groups for specific intervention. Examples of Unsupervised Learning: Apriori algorithm, K-means [45].

Reinforcement Learning: Using this algorithm, the machine is trained to make specific decisions. It works this way: the machine is exposed to an environment where it trains itself continually using trial and error. This machine learns from past experience and tries to capture the best possible knowledge to make accurate business decisions. Example of Reinforcement Learning: Markov Decision Process, below are some of the most common machine learning algorithms and some of their potential use cases [45].

3.2.6.2 K-nearest neighbor (k-NN)

The k-nearest neighbor algorithm (k-NN) as shown in Fig 2.7.1 [46] is used to test the degree of similarity between documents and k training data and to store a certain amount of classification data, thereby determining the category of test documents. This method is an instant-based learning algorithm that categorized objects based on closest feature space in the training set [47]. The training sets are mapped into multi-dimensional feature space. The feature space is partitioned into regions based on the category of the training set. A point in the feature space is assigned to a particular category if it is the most frequent category among the k nearest training data. Usually Euclidean Distance is typically used in computing the distance between the vectors. The key element of this method is the availability of a similarity measure for identifying neighbors of a particular document [47]. The training phase consists only of storing the feature vectors and categories of the training set. In the classification phase, distances from the new vector, representing an input document, to all stored vectors are computed and k closest samples are selected. The annotated category of a document is predicted based on the nearest point which has been assigned to a particular category. The goal of cluster analysis is to group or cluster observations into subsets based on the similarity of responses on multiple variables. Observations that have similar response patterns are grouped together to form clusters. Cluster analysis is an unsupervised machine learning method, which means there is no specific response variable included in the analysis. Cluster analysis is often used in marketing, to develop targeted advertising campaigns. This is often referred to as market segmentation. Likewise, health researchers might use cluster analysis to identify individuals at greatest risk for health problems, and to develop targeted health messages based on patterns of health behavior, Cluster analysis can also be used as a data reduction technique that allows us to take many variables and reduce them down to a single categorical variable that has as many categories as the number of clusters identified in the dataset. This categorical variable can then be used in other analysis to predict some response variable of interest, Cluster analysis measures the distance between points in the p-dimensional space, and groups together those observations that are close to each other. The most commonly used distance measuring, K-means cluster analysis, is call Euclidean distance. The Euclidian distance measure determines how close observations are to each other by drawing a straight line between pairs of observations, and calculating the distance between them based on the length of this line.

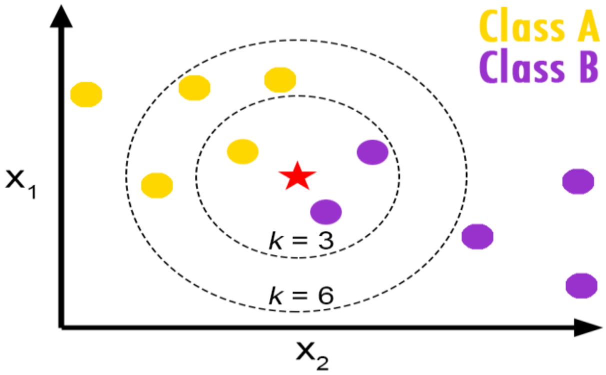


Fig 2.7.1 k-nearest neighbor.

The major drawback of this method is it uses all features in distance computation, and causes the method computationally intensive, especially when the size of training set grows. Besides, the accuracy of k-nearest neighbor classification is severely degraded by the presence of noisy or irrelevant feature.

3.2.6.3 K-Means Cluster Analysis Model

The goal of cluster analysis is to group or cluster observations into subsets based on the similarity of responses on multiple variables. Observations that have similar response patterns are grouped together to form clusters. Cluster analysis is an unsupervised machine learning method, which means there is no specific response variable included in the analysis. Cluster analysis is often used in marketing, to develop targeted advertising campaigns. This is often referred to as market segmentation. Likewise, health researchers might use cluster analysis to identify individuals at greatest risk for health problems, and to develop targeted health messages based on patterns of health behavior, Cluster analysis can also be used as a data reduction technique that allows us to take many variables and reduce them down to a single categorical variable that has as many categories as the number of clusters identified in the dataset. This categorical variable can then be used in other analysis to predict some response variable of interest [48], Cluster analysis measures the distance between points in the p-dimensional space, and groups together those observations that are close to each other. The most commonly used distance measuring, K-means cluster analysis, is call Euclidean distance. The Euclidian distance measure determines how close observations are to each other by drawing a straight line between pairs of observations, and calculating the distance between them based on the length of this line. To demonstrate how cluster analysis works, let’s look at our machine learning model we developed, these models predicts the subscriber segments that are likely to buy a certain product or services from any telecom provider 0 ‘red’ shows not likely while 1 ‘green’ shows most likely as shown in Fig 2.7.2

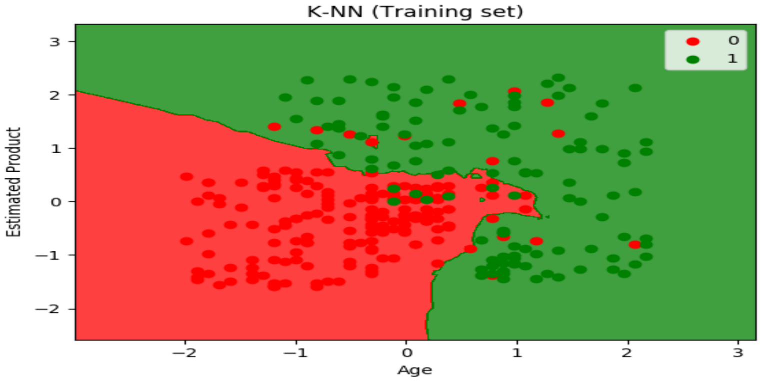


Fig 2.7.2 K nearest neighbor model prediction

3.2.6.4 Naïve Bayes Algorithm

Naïve Bayes classifier is a simple probabilistic classifier based on applying Bayes' Theorem with strong independence assumptions. A more descriptive term for the underlying probability model would be independent feature model. These independence assumptions of features make the features order is irrelevant and consequently that the present of one feature does not affect other features in classification tasks [49]. These assumptions make the computation of Bayesian classification approach more efficient, but this assumption severely limits its applicability. Depending on the precise nature of the probability model, the naïve Bayes classifiers can be trained very efficiently by requiring a relatively small amount of training data to estimate the parameters necessary for classification. Because independent variables are assumed, only the variances of the variables for each class need to be determined and not the entire covariance matrix. Due to its apparently over-simplified assumptions, the naïve Bayes classifiers often work much better in many complex real-world situations than one might expect. The naïve Bayes classifiers has been reported to perform surprisingly well for many real-world classification applications under some specific conditions [50] [51] [52] [53] [54]. An advantage of the naïve Bayes classifier is that it requires a small amount of training data to estimate the parameters necessary for classification. Bayesian classification approach arrives at the correct classification as long as the correct category is more probable than the others. Category's probabilities do not have to be estimated very well. In other words, the overall classifier is robust enough to ignore serious deficiencies in its underlying naïve probability model. The main disadvantage of the naïve Bayes classification approach is its relatively low classification performance compare to other discriminative algorithms, such as the SVM with its outperformed classification effectiveness. Therefore, many active researches have been carried out to clarify the reasons that the naïve Bayes classifier fails in classification tasks and enhance the traditional approaches by implementing some effective and efficient techniques [30] [51] [52] [53] [54]. The Naive Bayes algorithm is a simple probabilistic classifier that calculates a set of probabilities by counting the frequency and combinations of values in a given data set. The algorithm uses Bayes theorem and assumes all attributes to be independent given the value of the class variable. This conditional independence assumption rarely holds true in real world applications, hence the characterization as Naive yet the algorithm tends to perform well and learn rapidly in various supervised classification problems, Naïve Bayesian classifier is based on Bayes' theorem and the theorem of total probability [55]. Bayes theorem named after Rev. Thomas Bayes. It works on conditional probability. Conditional probability is the probability that something will happen, given that something else has already occurred. Using the conditional probability, we can calculate the probability of an event using its prior knowledge. Below is the formula for calculating the conditional probability.

$$P(H|E) = \frac{P(E|H) * P(H)}{P(E)} \qquad (1)$$

Where:

P(H) is the probability of hypothesis H being true. This is known as the prior probability.

P(E) is the probability of the evidence (regardless of the hypothesis).

P(E|H) is the probability of the evidence given that hypothesis is true.

P(H|E) is the probability of the hypothesis given that the evidence is there.

Naïve Bayes has been one of the popular machine learning methods for many years. Its simplicity makes the framework attractive in various tasks and reasonable performances are obtained in the tasks although this learning is based on an unrealistic independence assumption. For this reason, there also have been many interesting works of investigating naive Bayes. Recently the [56] shows very good results by selecting Naïve Bayes with SVM for text classification also the authors in [57] prove that Naive Bayes with SOM give very good results in clustering the documents. The authors in [58] propose a Poisson Naive Bayes text classification model with weight enhancing method, and shows that the new model assumes that a document is generated by a multivariate Poisson model. They suggest per-document term frequency normalization to estimate the Poisson parameter, while the traditional multinomial classifier estimates its parameters by considering all the training documents as a unique huge training document. The [59] presented that naive Bayes can perform surprisingly well in the classification tasks where the probability itself calculated by the naive Bayes is not important. The authors in a review [60] described that researcher shows great interest in naïve Bayes classifier for spam filtering. So, this technique is most widely used in email, web contents, and spam categorization. Naive Bayes work well on numeric and textual data, easy to implement and computation comparing with other algorithms, however conditional independence assumption is violated by real-world data and perform very poorly when features are highly correlated and does not consider frequency of word occurrences.

3.2.6.5 Bayesian Inference Model

Bayesian inference is an extremely powerful set of tools for modeling any random variable, such as the value of a regression parameter, a demographic statistic, a business KPI, or the part of speech of a word. We provide our understanding of a problem and some data, and in return get a quantitative measure of how certain we are of a particular fact [61]. To demonstrate how Naïve Bayes works (i.e. check code in appendix C), let's look at our machine learning model we developed, these models predicts the subscriber segments that are likely to buy a certain product or services from any telecom provider 0 'red' shows not likely while 1 'green' shows most likely as shown in Fig 2.7.3

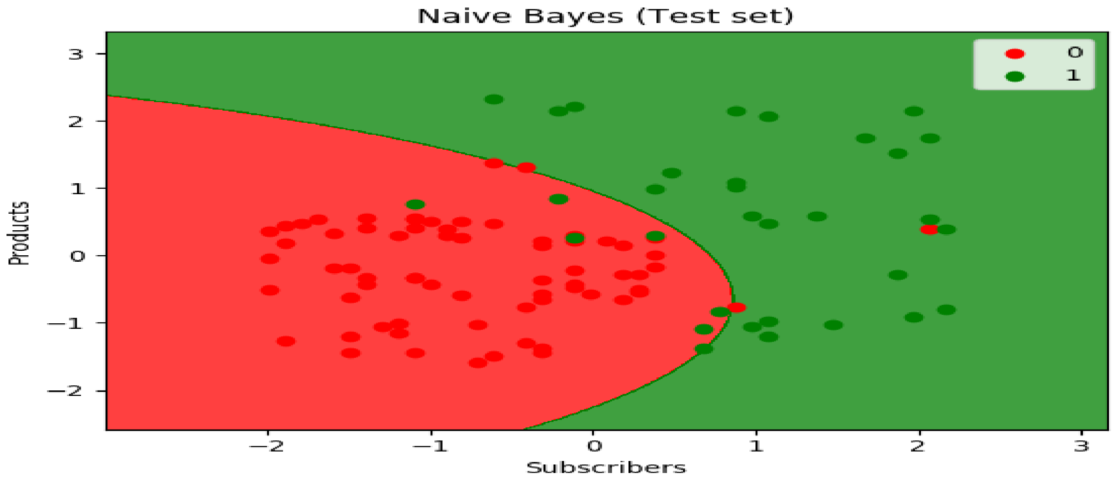


Fig 2.7.3 Naïve bayes

3.2.6.6 Random Forest

Random Forest is a machine learning method. This data mining algorithm is based on decision trees, but proceeds by growing many trees. While decision trees are not very reproducible on future data and are proceed by searching for a split on every variable in every node, Random Forests searches for a split on only one variable in a node. The variable that has the largest association with the target among candidate explanatory variables but only among those explanatory variables that have been randomly selected to be tested for that node, how does it work? First, a small subset of explanatory variables is selected at random. Next, the node is split with the best variable among the small number of randomly selected variables. Then, a new list of eligible explanatory variables is selected on random to split on the next node. This continues until the tree is fully grown, and ideally there is one observation in each terminal mode. The eligible variables set will be quite different from node to node. However, important variables will eventually make it into the tree. Their relative success in predicting the target variable will begin to get them larger and larger numbers of "votes" in their favor. The growing of each tree in a random forest is not only based on subsets of explanatory variables at each node, but also based on a random subset of the sample for each tree in the forest. This process of selecting a random sample of observations is known as Bagging. Importantly, each tree is growing on a different randomly selected sample of Bagged data with the remaining Out of Bag data available to test the accuracy of each tree. For each tree, the Bagging Process selects about 60% of the original sample, while the resulting tree is tested against the remaining 40% of the sample. Thus, the randomly selected bag data and out of bag data, will be a different 60% and 40% of observations for each tree. The most important thing to know when interpreting results of random forests is that the trees generated are not themselves interpreted. Instead, they are used to collectively rank the importance of variables in predicting our target of interest [62]. Figs 2.7.4 [63] and 2.7.5 [64] show random forest tree illustration and random forest flow chart respectively.

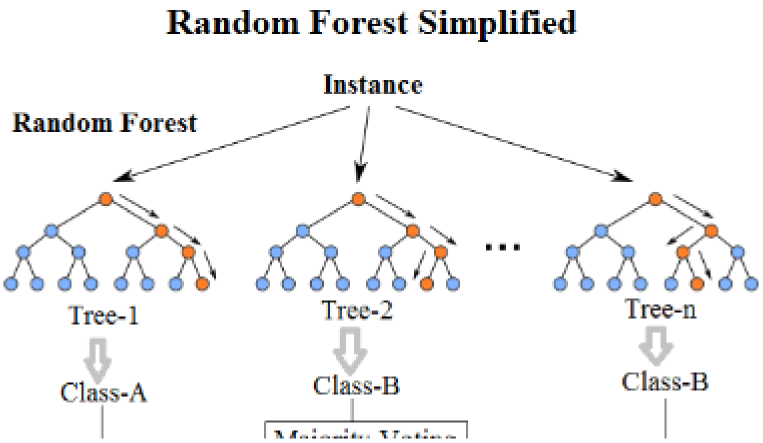


Fig 2.7.4 random forest tree illustration

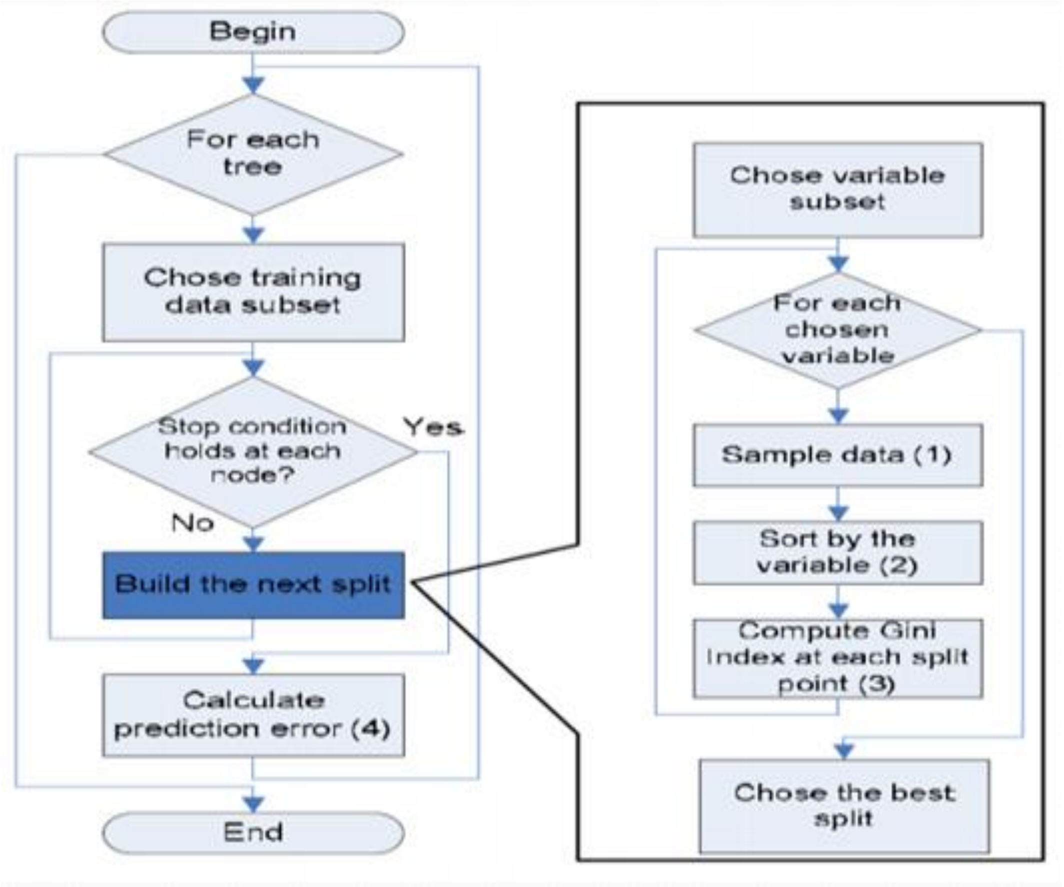


Fig 2.7.5 Random forest flow chat example.

To demonstrate how Random forest works (i.e. check code in appendix C), let’s look at our machine learning model we developed, these models predicts the subscriber segments that are likely to buy a certain product or services from any telecom provider 0 ‘red’ shows not likely while 1 ‘green’ shows most likely as shown in Fig 2.7.6

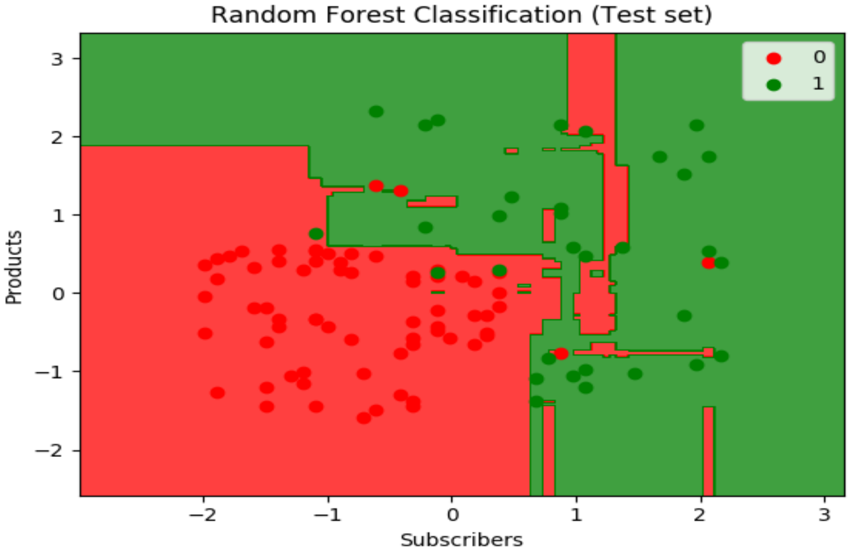
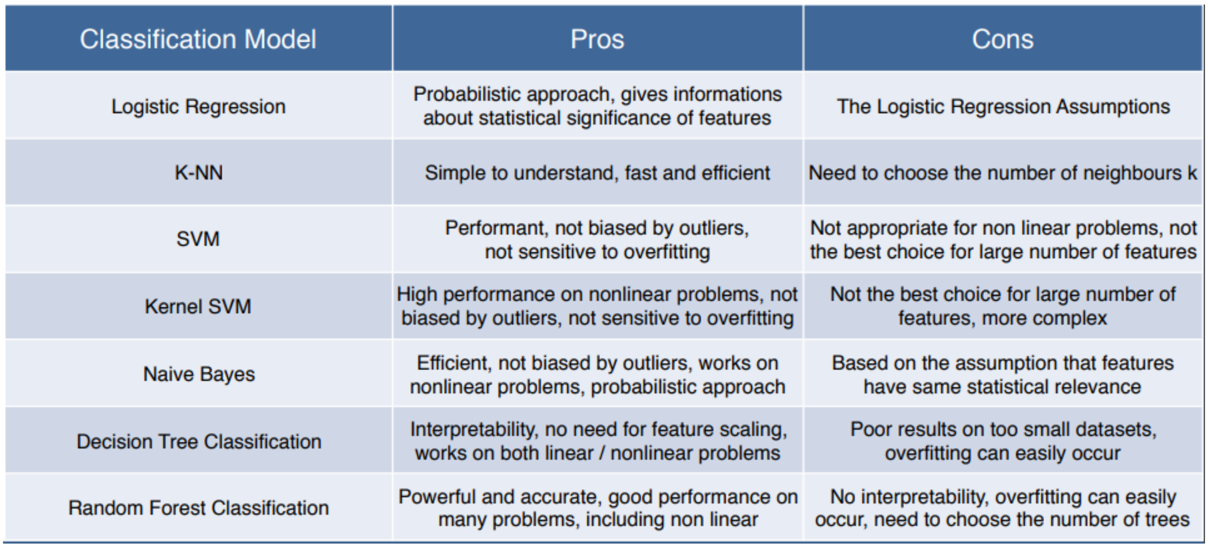


Fig 2.7.6 Random Forest model classifier

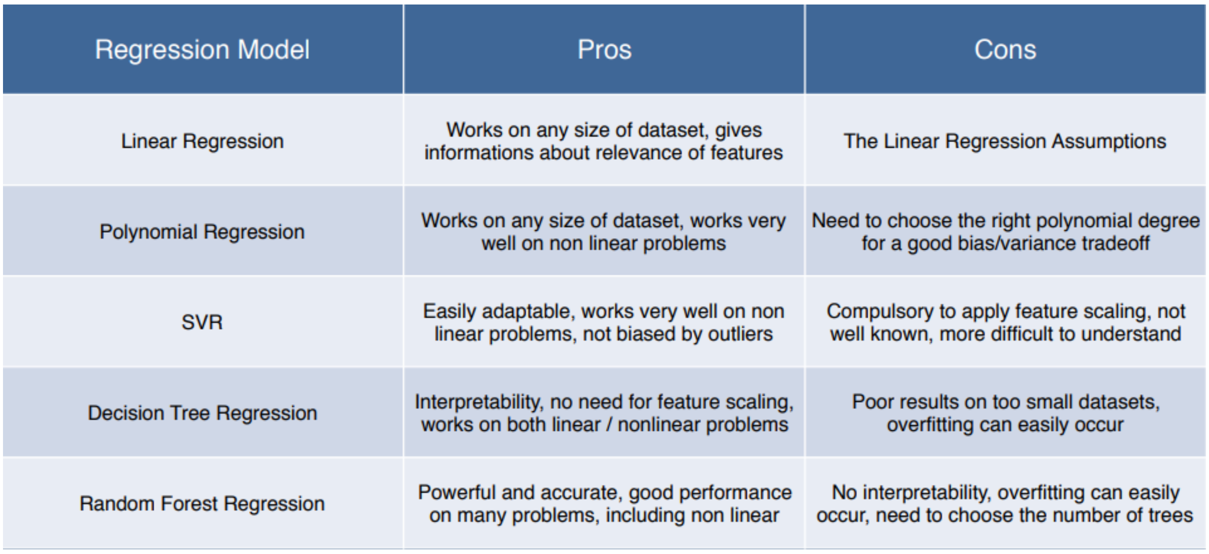
Classification algorithms are used when the desired output is a discrete label. In other words, they're helpful when the answer to your question about your business falls under a finite set of possible outcomes. Multi-label classification captures everything else, and is useful for customer segmentation, audio and image categorization, and text analysis for mining customer sentiment. If these are the questions you're hoping to answer with machine learning in your business, consider algorithms like naive Bayes, decision trees, logistic regression, kernel approximation, and K-nearest neighbors. On the other hand, regression is useful for predicting outputs that are continuous. That means the answer to your question is represented by a quantity that can be flexibly determined based on the inputs of the model rather than being confined to a set of possible labels. Regression problems with time-ordered inputs are called time-series forecasting problems, like ARIMA forecasting, which allows data scientists to explain seasonal patterns in sales, evaluate the impact of new marketing campaigns, and more. Choosing an algorithm is a critical step in the machine learning process, so it's important that it truly fits the use case of the problem at hand. Make sure data scientists and business users align early on at your organization to avoid common pitfalls of building predictive models, tables 2.7.1 and 2.7.2 are showing the pros and cons for each model.

Table 2.7.1 shows you all the pros and the cons of each classification model.



Classification Model	Pros	Cons
Logistic Regression	Probabilistic approach, gives informations about statistical significance of features	The Logistic Regression Assumptions
K-NN	Simple to understand, fast and efficient	Need to choose the number of neighbours k
SVM	Performant, not biased by outliers, not sensitive to overfitting	Not appropriate for non linear problems, not the best choice for large number of features
Kernel SVM	High performance on nonlinear problems, not biased by outliers, not sensitive to overfitting	Not the best choice for large number of features, more complex
Naive Bayes	Efficient, not biased by outliers, works on nonlinear problems, probabilistic approach	Based on the assumption that features have same statistical relevance
Decision Tree Classification	Interpretability, no need for feature scaling, works on both linear / nonlinear problems	Poor results on too small datasets, overfitting can easily occur
Random Forest Classification	Powerful and accurate, good performance on many problems, including non linear	No interpretability, overfitting can easily occur, need to choose the number of trees

Table 2.7.2 shows you all the pros and the cons of each regression model.



Regression Model	Pros	Cons
Linear Regression	Works on any size of dataset, gives informations about relevance of features	The Linear Regression Assumptions
Polynomial Regression	Works on any size of dataset, works very well on non linear problems	Need to choose the right polynomial degree for a good bias/variance tradeoff
SVR	Easily adaptable, works very well on non linear problems, not biased by outliers	Compulsory to apply feature scaling, not well known, more difficult to understand
Decision Tree Regression	Interpretability, no need for feature scaling, works on both linear / nonlinear problems	Poor results on too small datasets, overfitting can easily occur
Random Forest Regression	Powerful and accurate, good performance on many problems, including non linear	No interpretability, overfitting can easily occur, need to choose the number of trees

3.3 SYSTEM ARCHITECTURE AND DEVELOPMENT

This section describes the development of a Telecom Product Recommender System.

How do I know which model to choose for my problem?

you first need to fig out whether your problem is linear or nonlinear, if your problem is linear, you should go for Logistic Regression or SVM and If your problem is nonlinear, you should go for K-NN, Naive Bayes, Decision Tree or Random Forest But from a business point of view, you would rather use:

- Logistic Regression or Naive Bayes when you want to rank your predictions by their probability. For example, if you want to rank your customers from the highest probability that they buy a certain product, to the lowest probability. Eventually that allows you to target your marketing campaigns. And of course, for this type of business problem, you should use Logistic Regression if your problem is linear, and Naive Bayes if your problem is nonlinear.
- SVM when you want to predict to which segment your customers belong to. Segments can be any kind of segments, for example some market segments you identified earlier with clustering

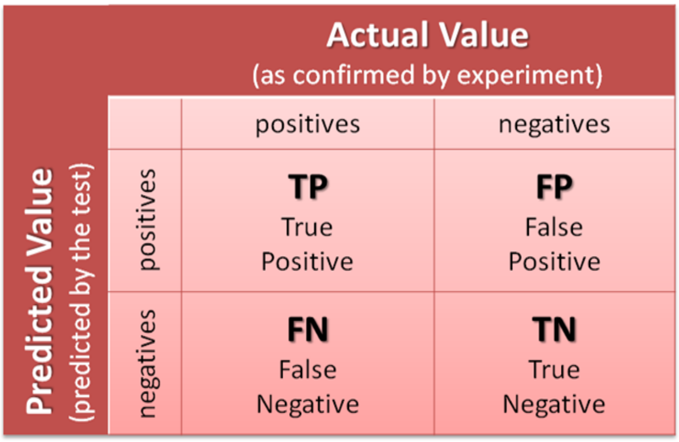
3.3.1 Evaluating Classification Models Performance

3.3.1.1 Confusion Matrix

A confusion matrix illustrates (i.e. fig 3.1) the accuracy of the solution to a classification problem. Given n classes a confusion matrix is a m x n matrix, where Ci,j indicates the number of tuples from D that were assign to class Ci,j but where the correct class is Ci . Obviously, the best solution will have only zero values outside the diagonal a confusion matrix contains information about actual and predicted classifications done by a classification system. Performance of such systems is commonly evaluated using the data in the matrix. Some standards and terms:

- True positive (TP): If the outcome from a prediction is p and the actual value is also p, then it is called a true positive.
- False positive (FP): However, if the actual value is n then it is said to be a false positive.

- Precision and recall: Precision is the fraction of retrieved instances that are relevant, while recall is the fraction of relevant instances that are retrieved. Both precision and recall are therefore based on an understanding and measure of relevance. Precision can be seen as a measure of exactness or quality, whereas recall is a measure of completeness or quantity. Recall is nothing but the true positive rate for the class [66]. Fig 3.1 illustrates a confusion matrix



Predicted Value (predicted by the test)		Actual Value (as confirmed by experiment)	
		positives	negatives
	positives	TP True Positive	FP False Positive
	negatives	FN False Negative	TN True Negative

Fig 3.1 illustrates a confusion matrix

3.3.1.2 Train/Test Split and Cross Validation

In statistics and machine learning we usually split our data into to subsets: training data and testing data (and sometimes to three: train, validate and test), and fit our model on the train data, in order to make predictions on the test data. When we do that, one of two things might happen: we overfit our model or we underfit our model. We don’t want any of these things to happen, because they affect the predictability of our model—we might be using a model that has lower accuracy and/or is ungeneralized (meaning you can’t generalize your predictions on other data).

- **Overfitting:** Overfitting means that model we trained has trained “too well” and is now, well, fit too closely to the training dataset. This usually happens when the model is too complex (i.e. too many features/variables compared to the number of observations). This model will be very accurate on the training data but will probably be very not accurate on untrained or new data. It is because this model is not generalized (or not AS generalized), meaning you can generalize the results and can’t make any inferences on other data, which is, ultimately, what you are trying to do. Basically, when this happens, the model learns or describes the “noise” in the training data instead of the actual relationships between variables in the data. This noise, obviously, isn't part in of any new dataset, and cannot be applied to it.

- **Underfitting:** In contrast to overfitting, when a model is underfitted, it means that the model does not fit the training data and therefore misses the trends in the data. It also means the model cannot be generalized to new data. As you probably guessed (or figd out!), this is usually the result of a very simple model (not enough predictors/independent variables). It could also happen when, for example, we fit a linear model (like linear regression) to data that is not linear. It almost goes without saying that this model will have poor predictive ability (on training data and can't be generalized to other data).

- **Train/Test Split:** As I said before, the data we use is usually split into training data and test data. The training set contains a known output and the model learns on this data in order to be generalized to other data later on. We have the test dataset (or subset) in order to test our model's prediction on this subset. I've loaded in the data, split it into a training and testing sets, fitted a regression model to the training data, made predictions based on this data and tested the predictions on the test data. Seems good, right? But train/test split does have its dangers—what if the split we make isn't random? What if one subset of our data has only people from a certain state, employees with a certain income level but no other income levels, only women or only people at a certain age? (Imagine a file ordered by one of these). This will result in overfitting, even though we're trying to avoid it! This is where cross validation comes in.
- **Cross Validation:** In the previous paragraph, I mentioned the caveats in the train/test split method. In order to avoid this, we can perform something called cross validation. It's very similar to train/test split, but it's applied to more subsets. Meaning, we split our data into k subsets, and train on k-1 one of those subsets. What we do is to hold the last subset for test. We're able to do it for each of the subsets.
- **K-Folds Cross Validation:** In K-Folds Cross Validation we split our data into k different subsets (or folds). We use k-1 subsets to train our data and leave the last subset (or the last fold) as test data. We then average the model against each of the folds and then finalize our model. After that we test it against the test set [67].

3.4 System Architecture

Client

Client is the user interface presented on a web browser. When a customer visits the website of the telecom company, the client browser will send requests to the web server every time the user performs an action, such as login or visiting a new page. When the web server receives the requests, it retrieves the requested resources and sends them back to the client browser.

Web server

Websites are hosted in web servers. A web server consists of two dimensions: the logical server, which is the software that serves the web requests, and the physical server, which is the computer running the logical server and storing all the resources. Based on the web server, the recommendation system web site can be divided into three layers: the presentation layer, business logic layer and data access layer.

- Presentation layer: This layer is responsible for generating the requested web pages and handling the UI logics and events. When a user requests to view a new page, the presentation layer will invoke corresponding methods in the business logic layer, extract the request data, transform the data into HTML page and send it back to the client.
- Business Logic Layer: This layer defines the business rules and processes of the application, and serves as a mediator between the presentation layer and the data access layer. In the recommender system, the business logic layer contains two parts: one part implements the recommender system website business processes and the other part implements the hybrid machine learning-based telecom product recommendation approach.
- Data Access Layer: This layer deals with the data operations of the database and transfers data with the business logic layer. In recommender system, the data access layer is implemented using Entity Framework.

Database Server

The database server is the computer server that runs the database applications. In the recommender system, we use POSTGRESQL Server PGADMIN version 4 as the database application because it is the most compatible with all the Microsoft/Linux technologies we use. The database server can be either the same computer as the web server or a separate server running the database application.

3.5 Recommendation System Development Steps

This recommender system has been developed by the following steps:

- Classification and clustering of existing customers through retrieving and analysing the existing customer profile database. The existing customer profile database has rich profile information about existing customers, such as customer name, customer account(s), and current products/services, re-contract time, and customer usage information.
- This study sets up a set of business rules with the telecom company for existing customers. This study designed and applied five types of business rules: 1) the bundle rules, 2) the fleet rules, 3) the discount rules, 4) the product rules and 5) the special offers. Three examples of the business rules are: Some fixed line products cannot be purchased standalone. They have to be bundled with a fixed broadband product. A customer can receive additional discounts for some products, if they are purchased together. For a period of time, some products may be on special or the business may be promoting those products.
- Establish a customer view from the current customer database. This step involves database information retrieval and incorporation, and the customer view (database) structure design, as well as the physical storage of data in the view (database).
- Design a set of online data collection pages to obtain existing customers' requirements and web-based interface as well as outputs.
- Implement the developed recommendation approach.
- Interface design, including customer data collection, recommendation list generation and related explanations.

- System testing and revision. Test cases are conducted to test and evaluate the performance of the developed intelligent recommender system, using telecom customer data.

3.6 SYSTEM APPLICATION

The main process of recommendation is described as follows:

- To collect customer information. In this step, the rating data of customers are collected in the mobile product/service and handset detail web pages on which a customer can rate a mobile product/service and a handset. The rating value, as well as the customer ID and mobile product/service ID or handset ID, will then be stored in the database.
- To gather data from similar existing customers, including purchase records, usage, website visit history and personal profiles;
- To collect related product data and build a product database and determine main features;
- To analyse the collected data (customer and product), business rules, and predict the ratings of unrated products using machine learning techniques;
- To select the top-K products with the highest predicted ratings as recommendations for customers.

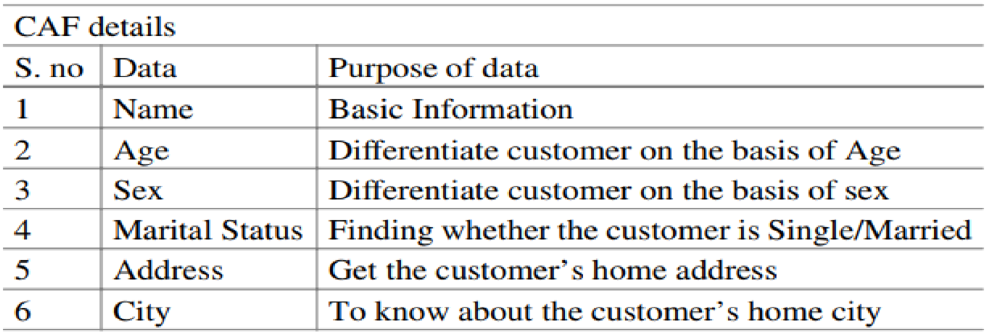
There are two types of recommendations:

- Mobile products/services and handset recommendations: After a customer logs into the homepage, recommender system is able to generate recommendations to the customer. The system will firstly read the approach settings from the configuration file which include parameters such as the number of neighbors and the number of items to be recommended. The system will then load the rating records of users and use the hybrid method to make recommendations. Finally, the system will return a list of recommended handsets.
- Package recommendation: For a postpaid customer whose contract will expire in four weeks' time, the recommender system will automatically recommend a package which includes handsets, plans and extra telecom services.

3.6.1 Types of Data Needed

To generate recommendations, recommendation engine needs certain sets of data. Depending on what type of recommendation needs to be generated, recommendation engine will use specific set. In this dissertation, we will be discussing about the recommendation engines which will be used for recommending Ringtones and Value-added Services (VAS). Consumer application form details are the details that consumer fills while registering for a particular service or set of services provided by the telecom company. It contains details about the consumer's basic information. The set of CAF details which the company has is mentioned in below Table 3.1

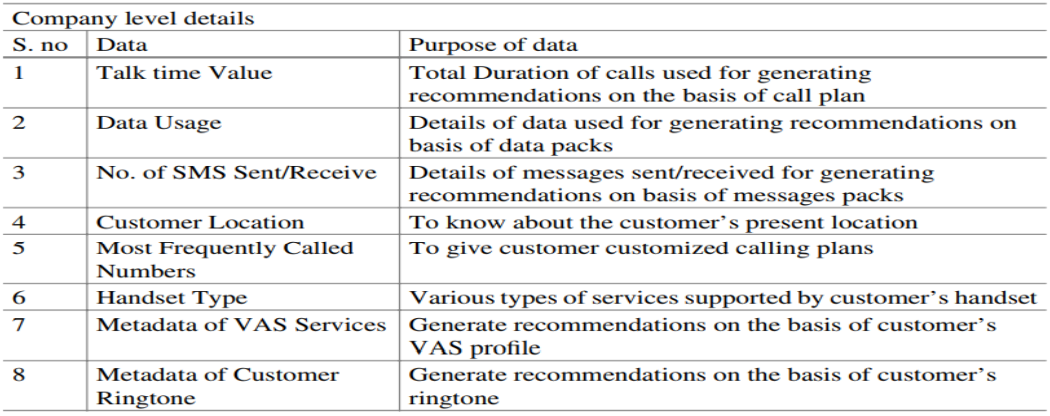
Table 3.1. Customer application form (CAF) details



CAF details		
S. no	Data	Purpose of data
1	Name	Basic Information
2	Age	Differentiate customer on the basis of Age
3	Sex	Differentiate customer on the basis of sex
4	Marital Status	Finding whether the customer is Single/Married
5	Address	Get the customer's home address
6	City	To know about the customer's home city

Company level details are the details of consumer transactions which the company stores in their databases (Table 3.2). Based on these details and metadata of ringtone, recommendation engine generates the ringtone recommendations for the customer. In case of other VAS (Value Added Services), recommendations generated are based on the segment to which consumer profile belongs rather than on the basis of metadata as metadata details are not associated for these kinds of services. The set of company level details which the company maintains in their databases is given below in the table.

Table 3.2 Company level details



Company level details		
S. no	Data	Purpose of data
1	Talk time Value	Total Duration of calls used for generating recommendations on the basis of call plan
2	Data Usage	Details of data used for generating recommendations on basis of data packs
3	No. of SMS Sent/Receive	Details of messages sent/received for generating recommendations on basis of messages packs
4	Customer Location	To know about the customer's present location
5	Most Frequently Called Numbers	To give customer customized calling plans
6	Handset Type	Various types of services supported by customer's handset
7	Metadata of VAS Services	Generate recommendations on the basis of customer's VAS profile
8	Metadata of Customer Ringtone	Generate recommendations on the basis of customer's ringtone

Metadata details of ringtone as shown in Table 3.3, is used in case of recommending caller tones to the customers. In this case, metadata of customer's past and present caller tones are used to know to which segment the customer belongs, and on the basis of that segment, recommendation is generated for the customer.

Table 3.3 Metadata details of ringtone

Metadata details		
S. no	Data	Purpose of data
1	Type	Type of Ringtone e.g. pop, jazz etc.
2	Film/Album	Film/Album to which Ringtone belongs
3	Composer	To generate recommendations on basis of composer
4	Singer	To give recommendations on basis of singer
5	Director	To generate recommendations on basis of music director
6	Language	Language of the ringtone

3.6.2 Proposed System Model

Fig 3.2 represents the architecture of the recommendation engine which works on the basis of customer segmentation and metadata comparison.

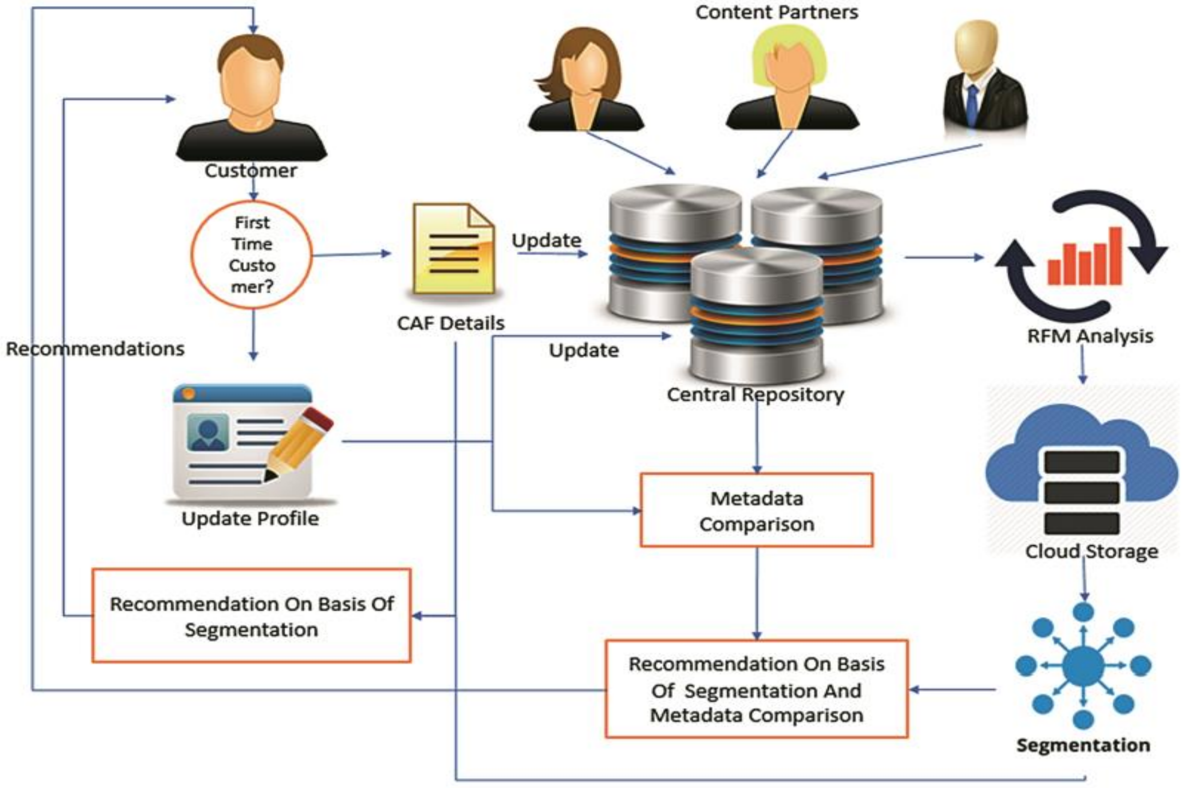


Fig 3.2 proposed system model

3.6.3 Customer Profile

A new customer profile gets created when the customer registers himself/herself for the services of the telecom company. Customer profile contains basic details about the customer like name, age, gender etc. and these are the details which the customer gives himself by filling consumer application form (CAF). Then, it is checked whether the customer is first time customer or he/she wants to change the caller tone of smartphone. In case of former, recommendation engine generates the recommendations on the basis of customer segmentation i.e. the segment to which the customer's profile belongs. But in case of lateral, recommendations are generated on the basis of metadata and customer profile details.

3.6.4 Recommendation on Basis of Customer Segment

When the customer register himself for the first time, then there are no metadata details present for that customer, then the recommendations are generated on the basis of matching the customer profile parameters to the profiles of already existing customers in the repository. The profile which matches maximally is used to know the segment to which the customer profile belongs and on basis of that segment, the recommendation is generated.

3.6.5 Updating Profile and Metadata

If the already registered customer wants to change the caller tone of his/her smartphone, then firstly the metadata details of the customer's record is changed accordingly. So, customer's profile gets updated whenever customer wants to change the caller tone.

3.6.6 Metadata Comparison

In this step, recommendation engine will compare the metadata details of customer profile with the metadata of the other profiles i.e. we will be calculating the similarity index of the customer metadata details with other customer's metadata details and the profile corresponding to the maximum similarity index will be then looked up to see to which segment it belongs on basis of which recommendations will be made. The similarity index between two customer's profiles is represented by a number between −1.0 and 1.0. The possibility of customer liking/selecting particular ringtone will be between −1.0 and 1.0. Similarly, in case of not liking/selecting the number will be between −1.0 and 1.0. For finding similarity index, we will have two sets corresponding to each customer. One corresponding to the customer liking/selecting the caller tune and other for not selecting/liking the ringtone. According to Jaccard's formula [75], the similarity index is calculated as follows

$$A\ (B, C) = |B \cap C| \div |B \cup C| \qquad (1)$$

The calculation involves the division of the total number of common elements in both sets by the total number of the elements in both sets (only counted once). The Jaccard index of two similar sets will always be 1, while for two sets with no common elements will always yield 0. Jaccard index for two profiles on the basis of liking of each parameter is,

$$A\ (B, C) = |S1 \cap S2\ | \div |S1 \cup S2| \quad (2)$$

Now as two customers selecting same ringtone is similar, then two customers not selecting the same ringtone are also similar. So, by changing above equation we get,

$$A\ (B, C) = (|S1 \cap S2\ | + |NS1 \cap NS2|) \div (|S1 \cup S2 \cup NS1 \cup NS2|) \ \ (3)$$

I.e. instead of considering same selection we also have taken into the account the deselection. In denominator, we have taken the total number of selection/deselection that customer has made. Here, we have considered the customer selection or deselection in independent sort of way. But what if customer likes ringtone but other customer doesn't and vice-versa. To take this thing into account we again have to modify the equation as,

$$A\ (B, C) = (|S1 \cap S2\ | + |NS1 \cap NS2\ | - |S1 \cap NS2\ | - |NS1 \cap S2\ |) \div (|S1 \cup S2 \cup NS1 \cup NS2|) \quad (4)$$

Now this equation will give 1.0 if two customer's profiles have same selection/liking for caller tones and −1.0 if two customers have deselection/disliking for caller tones.

3.6.7 Recommendation on Basis of Customer Segmentation and Metadata Comparison

RFM analysis of customer transaction information plus the metadata of profile that matched with the customer's profile gives us the segment to which the customer profile belongs. RFM is method for analysis of market which is used to examine which customers are best ones by examining how recent the customer has made purchase, how often he/she purchases and how much the customer spends on the purchase. It is based on the fact that "80% of business comes from 20% of the customers". Customers are then given ratings on the basis of these three input parameters by the telecom organization. The RFM score and metadata details matched profile gives the segment to which the customer profile belongs. On the basis of which the recommendations are generated for the customers i.e. Hybrid recommendation systems [76].

3.6.8 Content Provider

Content providers belonging to different partners are responsible for uploading the content to the telecom company's central repository on which RFM analysis is performed and output is stored on the central cloud storage, so that results can be accessible anytime and from anywhere. Based on these results, customer segmentation is done.

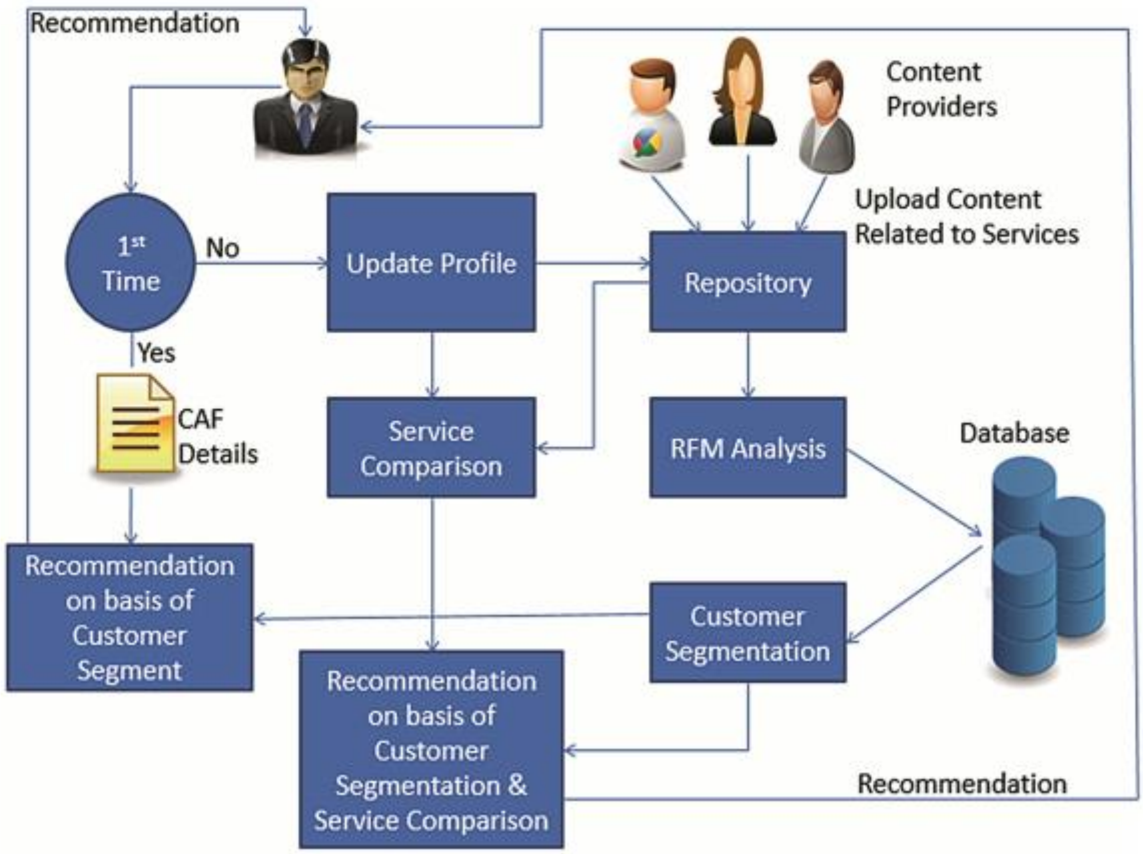


Fig. 3.3. Architecture of model based on customer segmentation and service comparison

The similar type of model can be used in case of Games i.e. based on customer profile, recommendations can be given whether one likes action, racing, puzzles, sports games etc. While new customer is given recommendations based on similar CAF details of other customers who have already subscribed to games, old customers get recommendations based on metadata of games that they have subscribed along with the segment of customer profile to which they belong. Fig 3.3 represents the architecture of the recommendation engine which works on the basis of customer segmentation and Service comparison. In case of other VAS related services, if a new customer wants to subscribe to any service like Astrology, Cricket, Jokes etc., he/she is recommended based on customer profile segmentation of other customers with similar CAF details as we have no idea about the interests of the new customer. Also, there is no meta-data related to these types of services. On the other hand, old customers are given recommendations based on their updated profile and RFM analysis of transaction information of the customer which gives the segment to which the customer profile belongs.

3.7 System Requirements Specification

In the system requirements specification phase of the research study, Object-Oriented Analysis (OOA) was used. System Requirements are descriptions of what the system should do, the services provided by that system and the constraints on its operation. Requirements reflect user needs for a system that serve a certain purpose. A requirement may also be described as a high-level abstract statement of a service that a system should provide or a constraint on the system [68]. Software system requirements can be categorized into functional and non-functional requirements. Sommerville [68], describes functional requirements as statements of services the system should provide and non-functional requirements as constraints on the services or function offered by the system. The System Requirements Specification section, therefore, provides a complete description of all the functionalities and specifications for the telecommunication recommender system.

3.7.1 Specific requirements

3.7.1.1 Interface Requirements

The user needs to click the link to the website. Then he/she needs to register to the system by providing a password and an email, otherwise he/she won’t be able to use the Recommender System. Then, to benefit from the Recommender System he/she needs to be active on the website by purchasing different products, if they have not purchased anything a popular product will be recommended.

3.7.1.2 Functional Requirements

3.7.1.3 General System Requirements

- **GR-001 Registration:** A guest should be encouraged to register in order to take advantage of benefits only open to members, such as receiving personalised recommendations from the system. The registration process will capture personal details, login information and a knowledge base about the user's preferences and interests.
- **GR-002 Login/Logout:** In order to login to the system, a guest must be a member of the web site and hold a valid username and password.
- **GR-003 User Control Panel (UCP):** Upon a successful login the user will be presented with their own User Control Panel (UCP). A UCP will encapsulate the following functionalities: the user's favourite products, user's shopping cart, and statistical data about the user's preferences, user's profile (which is responsible for capturing the user's interest through its knowledge based form) and a list of recommendations for the user
- **GR-004 Browse:** All users will be able to browse products in a certain category or a subcategory within a category.
- **GR-005 Search:** All users will be able to search for a products. There will be two types of search: simple and advanced.
- **GR-006 Search by Company's Name:** The automatic creation of hyperlinked company's names will feature on a product's details. This will enable one products to be related to another.
- **GR-007 View products details:** A products information details will be shown whenever a user request for it.
- **GR-008 Ratings:** The member will be able to rate a products in a numeric scale. The ratings could be updated or deleted by the user who generated them.
- **GR-009 Add to Favourites:** The members will be able to add any products to their favourite products list.
- **GR-010 Add to Shopping Cart:** The members should be able to add any products to their shopping cart.
- **GR-011 Add to Owned Products List:** The members will be encouraged to add any products that they own to their owned products list in order to improve the recommendations.

- **GR-012 Build User’s Profile:** The system will capture the interests of the user automatically when a user logs on to a website. The profile should adapt to changes in user interests over time.
- **GR-013 Shows Personalised Recommendation:** When a member asks for recommendations, the system will display four lists of recommended products based on different algorithms.
- **GR-014 Shows Non-Personalised Recommendation:** Any guest can receive non-personalised recommendations from the system.
- **GR-015 Improve Recommendation Area:** The user may want to exclude a products that had been purchased, added to favourites, browsed or rated from being used in making recommendations. The improve recommendations area will encapsulate the following functionality: favourite products, products the user owns, rated products and products appearing recently in the user’s browsing history. This will allow the user to delete any products that is no longer of interest.

3.7.1.4 Recommendation Module Requirements

3.7.1.4.1 Non- Personalised Recommendations

- RR-001: When the detailed information for a products is browsed, a recommended products list will be provided, based on what the customers have purchased in the past with the browsed products, as well as on what they have browsed
- RR-002: When a user searches for or browses for a products, the system should show the average products rating along with the number of users who have ranked the products using the Mean algorithm.
- RR-003: Any user will be able to view statistical data in the products details page that is based on all viewers’ rankings of products (on a scale from 1 to 5). Any user will also be able to see, for each ranking of a products, the number of users who gave it this ranking.
- RR-004: Non-personalised recommendations will be provided on the website homepage, which will be updated every time the page is browsed, based on the month’s bestselling products and the most recently viewed products for that month.

3.7.1.4.2 Personalised Recommendations

- RR-005: the system will provide products recommendations based on collaborative filtering. The system will implement the User-User Nearest Algorithm using two different similarity measures (Pearson Correlation coefficient and Mean Squared Difference) to evaluate the best similarity measure for the application domain.

- RR-006: the system should provide a list of products recommendations to the user based on content based filtering. This should be done by using the user profile, which builds as a result of the user's interaction with the system.
- RR-007: the system should provide products recommendations based on the combination of content based and collaborative filtering

3.7.1.5 Non-Functional Requirements

The non-functional requirements describe the actual operation of the system and define the qualities of the resulting system.

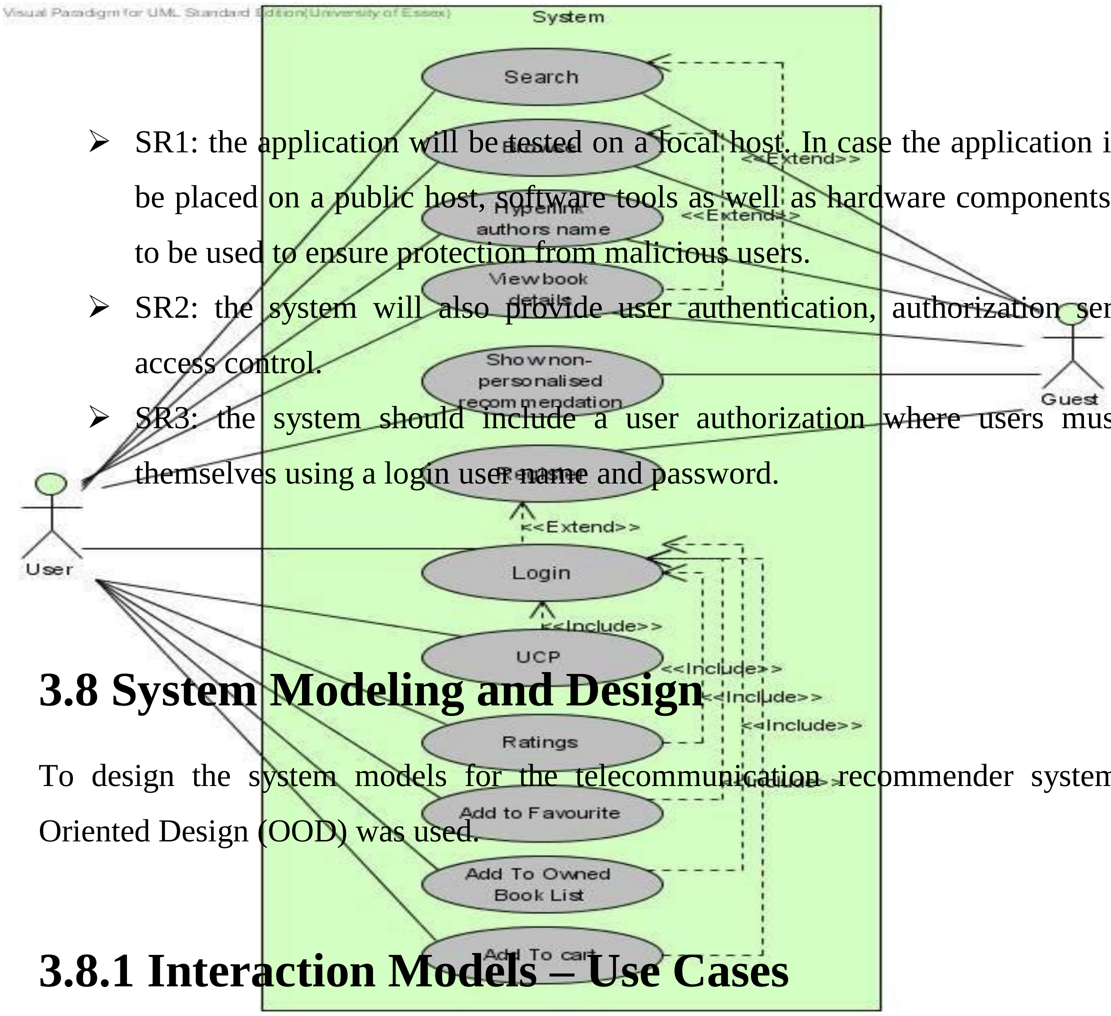
3.7.1.5.1 Interface Requirements

- IR1: the interface will be defined using XHTML for the content and CSS for specifying the layout and style. The CSS will be defined in an external style sheet.
- IR2: the interfaces will have a uniform design throughout the site, and user help should be provided whenever requested.
- IR3: the application developed is expected to be used by non-computer specialists, therefore it must be simple and easy to navigate.

3.7.1.5.2 Performance Requirements

- PR1: the system shall ensure accuracy and consistency of the required services. This will be achieved by extensive testing before each deliverable is 'signed off' as completed, coupled with the usage of exception handling techniques.
- PR2: the system should produce accurate recommendations that match the user's preferences.
- PR3: the database should be capable of handling a potentially large number of users.

3.7.1.6 Information Security and Privacy Requirement



- SR1: the application will be tested on a local host. In case the application is going to be placed on a public host, software tools as well as hardware components will have to be used to ensure protection from malicious users.
- SR2: the system will also provide user authentication, authorization services and access control.
- SR3: the system should include a user authorization where users must identify themselves using a login user name and password.

3.8 System Modeling and Design

To design the system models for the telecommunication recommender system, Object-Oriented Design (OOD) was used.

3.8.1 Interaction Models – Use Cases

From the functional requirements described in section 3, use cases are derived and relationships between them are defined. The following use case diagram (Fig 3.4) shows the main activities a user could perform in the system.

Fig 3.4: Main Use Case Diagram

It assumed that the user can receive different kinds of non-personalised and personalised recommendations based on different algorithms. The following use case diagram (Fig 3.5) shows the activities that users can perform in their control panel

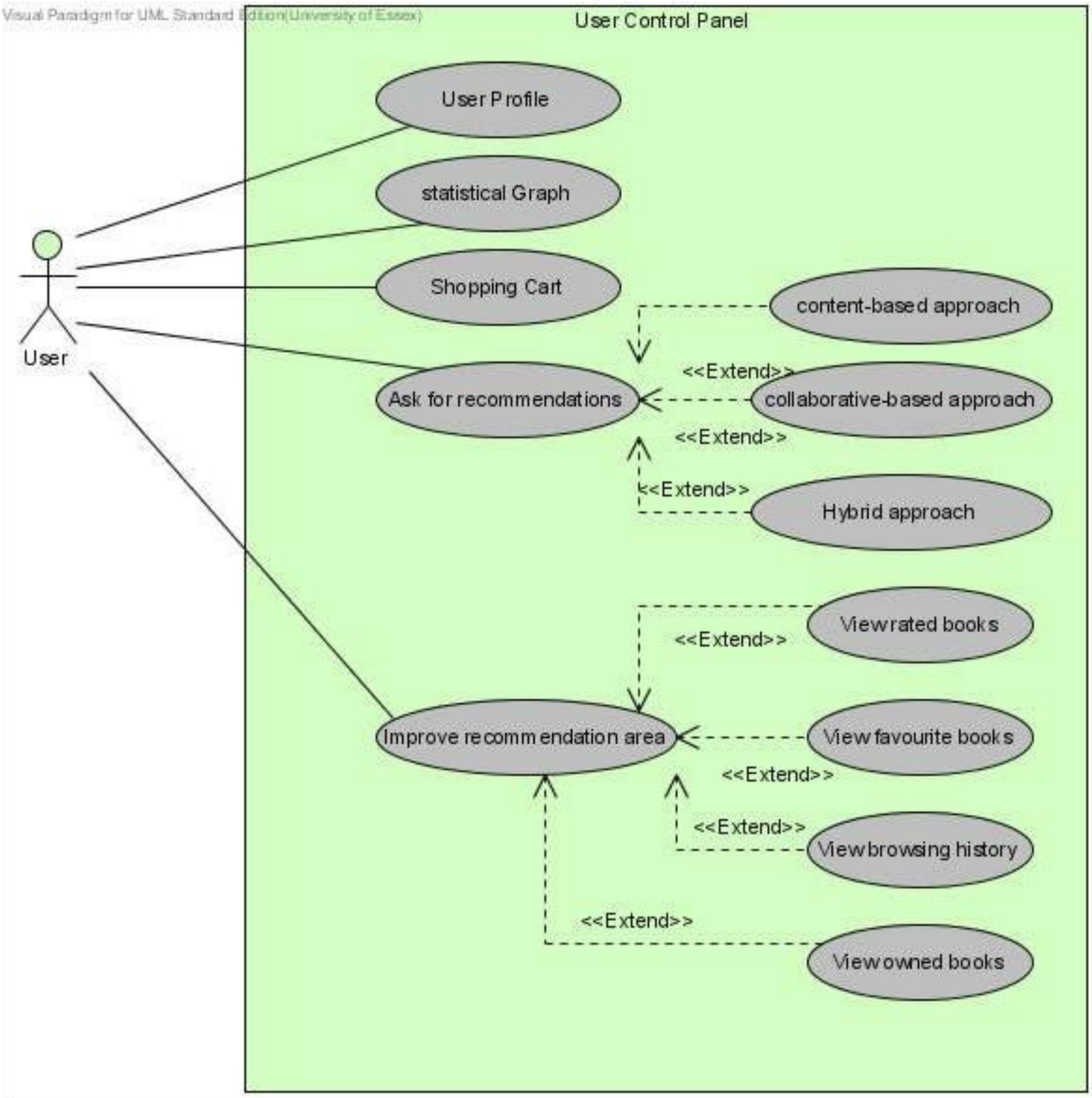


Fig 3.5: User Control Panel Use Case

3.8.2 Use Cases

The use cases have been assigned packages to give a clear structure, as follows:

1. Accounts
    - Register
    - Login
    - Edit profile
    - Statistical
    - View shopping cart

2. Browsing
    - Browse a category
    - Search
    - View products details
    - Hyperlink service provider

3. Learning module

- Add to cart
- Add to owned products
- Add to browsing history
- Add to favorites
- Rate

4. Recommendation module

- Personalized recommendations
- Non-personalized recommendations

5. Improve user recommendation area

- View rated products
- View favorite products
- View browsing history
- View owned products

Due to lack of space, the detailed use cases are provided in Appendix D, and the network activity diagrams are provided in Appendix E.

3.9 Design constraints

In the implementation process of this system, Python, PHP Programming Languages will be the main development languages. Since Python is selected to be the main development language, Python Programming Language is chosen as a standard for the development process of the system. In the process of the documentation of the system, IEEE standards will be used and UML standard will be used while designing the diagrams. Since this system will be a part of much larger system, it must be portable to this larger system. That’s why portability is one of the most important attributes of this system. Since the larger system is a website that has the potential of increasing its number of users, user traffic and number of songs, this system needs to be scale up with the website in the correct order. Therefore, scalability must be the number one attribute that system will have.

3.10 Summary

In this chapter, the materials and methods that were used in the baseline study and the system prototype was developed. A quantitively Methodology was used in this research study. Object-Oriented System Development Methodology that is Use Case driven was used in the system design and implementation. The proposed business process models for telecommunication recommender system were presented after an analysis of the current processes. System machine learning models including interaction models, structural models and data models were presented to provide the means by which the recommender system may be implemented.

CHAPTER FOUR
RESULTS

4.1 Introduction

In this chapter, we present the results that were derived from the baseline study. We also present results for the implementation of the system prototype using screenshots of the system application and hardware.

4.2 Baseline Line Study

In this section, the results from the baseline study derived from analysis of each variable through descriptive statistics are presented. The presentation of the results is in form of tables, bar charts and pie charts.

4.2.1 Demographic Data

Quantitative data was collected from a certain telecommunication located in Lusaka. The demographic study results for customers who purchased certain products and services are shown the in Fig 4.1.

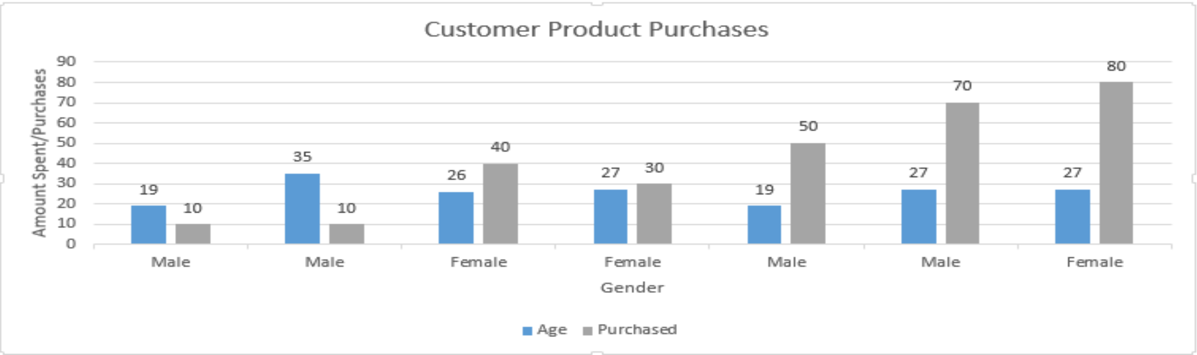


Fig 4.1: Demographic data chart

4.2.2 DATA DUMP SQL DATA EXTRACTED FROM CDR'S

Table 4.1 shows that all the data after extraction from CDRs the data dump or raw data will be collected from the CDR's for a certain telecom company, SQL procedures will be used to collect the SQL dump or raw CDR's will be processed using big data techniques, data collected will be truncated and hidden from non-researchers this is because most of the information is private

Table 4.1: Sql data dump

User ID	Gender	Age	Estimated price	Purchased	product bought
15624510	Male	19	19000	10	Xtratime
15810944	Male	35	20000	10	Xtratime
15668575	Female	26	43000	40	siliza
15603246	Female	27	57000	30	Soche
15804002	Male	19	76000	50	Spaka
15728773	Male	27	58000	70	Chizela
15598044	Female	27	84000	80	Mobile Money

Fig 4.2 shows the average age between men and women who are likely to buy a telecom product after processing raw CDR’s.

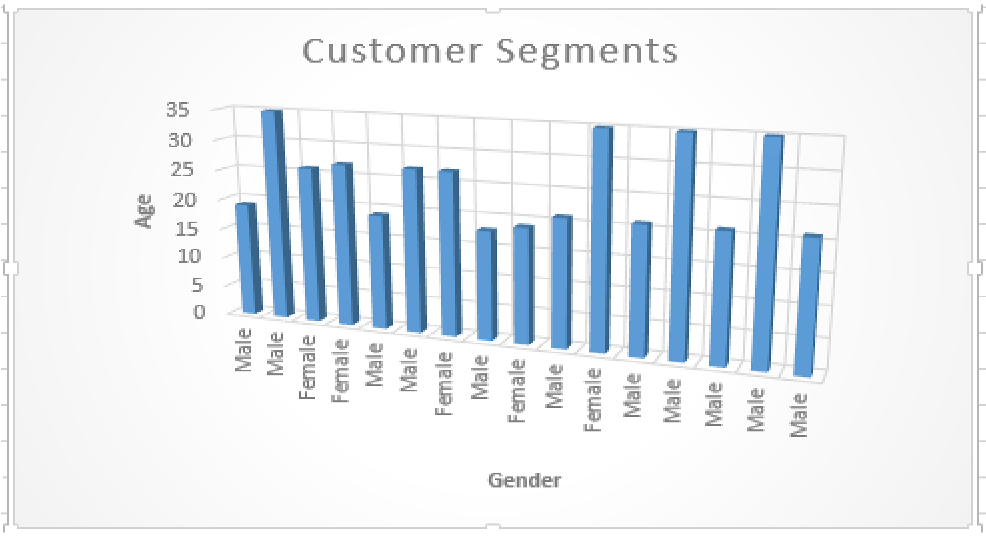
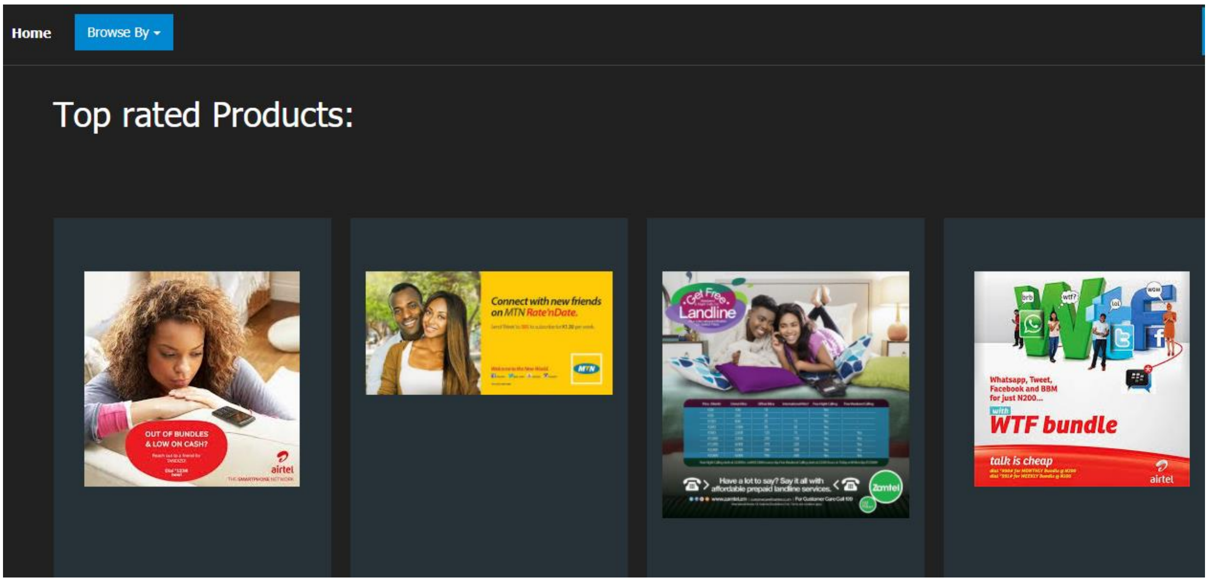


Fig 4.2 average age between men and women

4.3 SYSTEM APPLICATION SCREENSHOTS

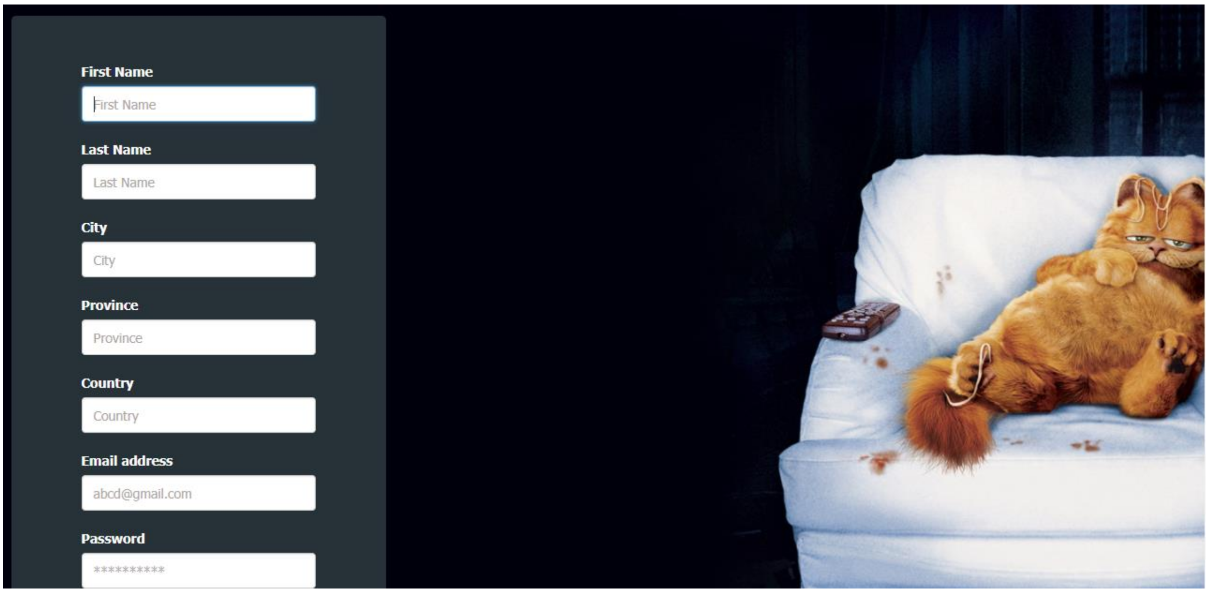
Fig 4.3 shows the most top-rated products before a subscriber logs in, these are based on the most top rated and popular product, system has managed to do a recommendation based on similarity of new user compared to old user historical purchase information.


Home
Browse By
OUT OF BUNDLES & LOW ON CASH?
airtel
Connect with new friends on MTN Rate'nDate.
MTN
Get Free Landline
Have a lot to say? Say it all with affordable prepaid landline services.
Zamtel
Whatsapp, Tweet, Facebook and BBM for just N200...
with WTF bundle
talk is cheap
airtel

Top rated Products:

Fig 4.3 top-rated products

Fig 4.4 shows the signup page for the website if a user is not registered, new users will be given an option to sign up so that they system will be able to recommend them products based on their attributes like location, nearest neighbor etc.



First Name: First Name

Last Name: Last Name

City: City

Province: Province

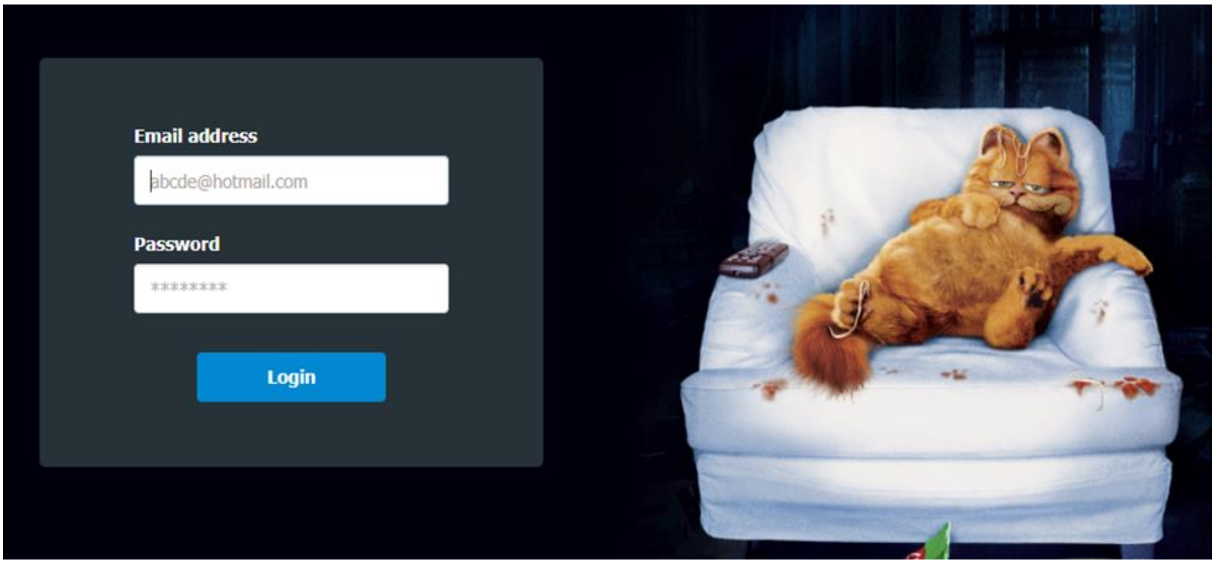
Country: Country

Email address: abcd@gmail.com

Password: **********

Fig 4.4 signup page.

Fig 4.5 shows the login page after user has successfully signed up.

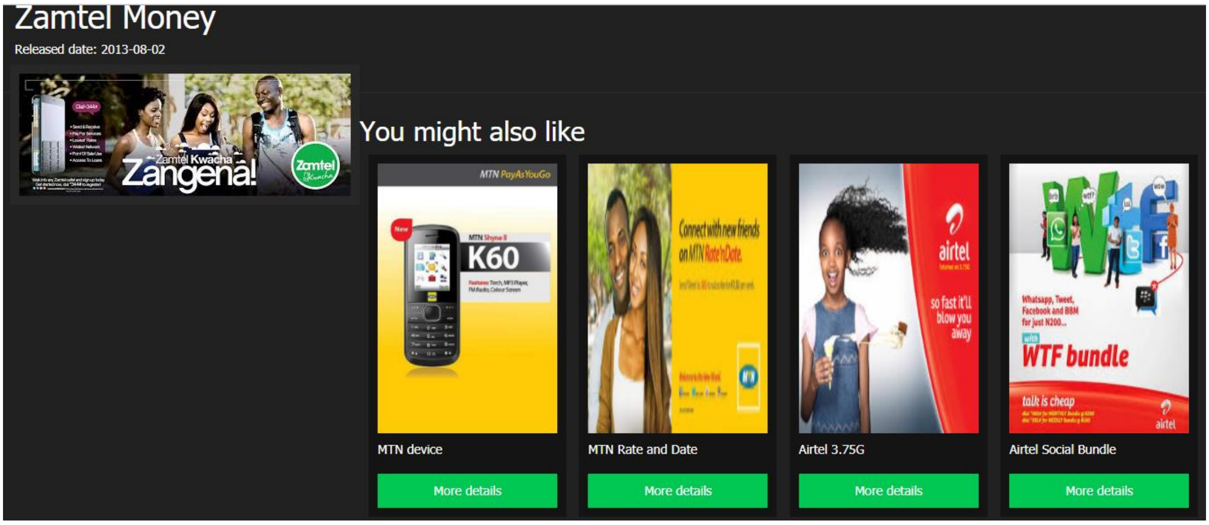


Email address: abcde@hotmail.com

Password: ********

Fig 4.5 login page.

Fig 4.6 below, shows product recommendations to a subscriber, after a successful login a customer can click on a product they like, for example customer bought Zamtel money and the system showed the customer the other products they might like to buy like MTN device, Airtel Soche.


Zamtel Money

Released date: 2013-08-02

You might also like

MTN device

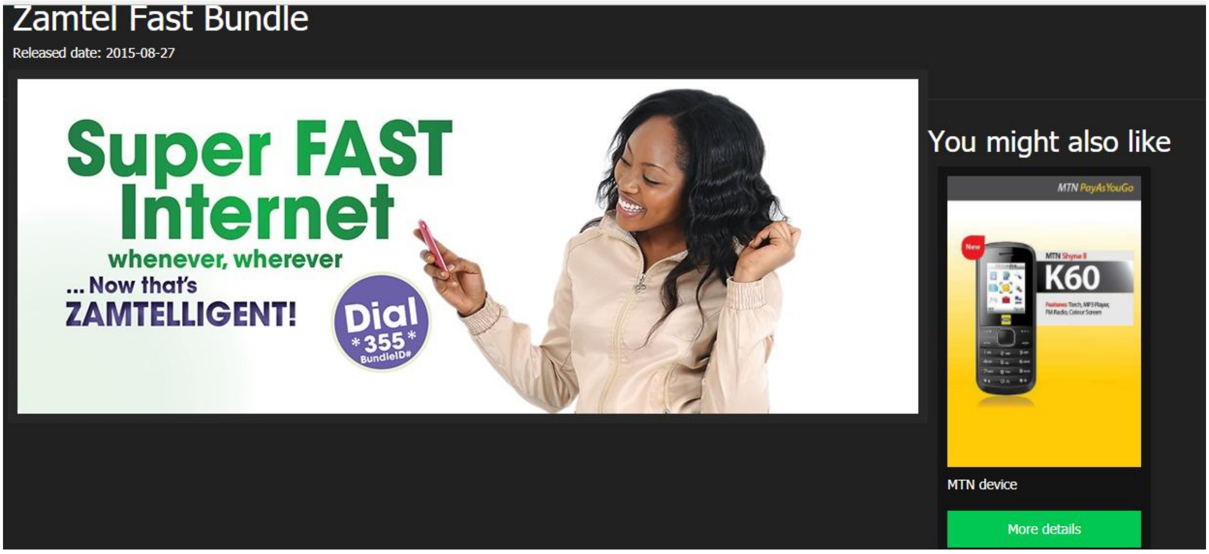
MTN Rate and Date

Airtel 3.75G

Airtel Social Bundle

Fig 4.6 product recommendations

Fig 4.7 shows product being recommended to a subscriber.


Zamtel Fast Bundle

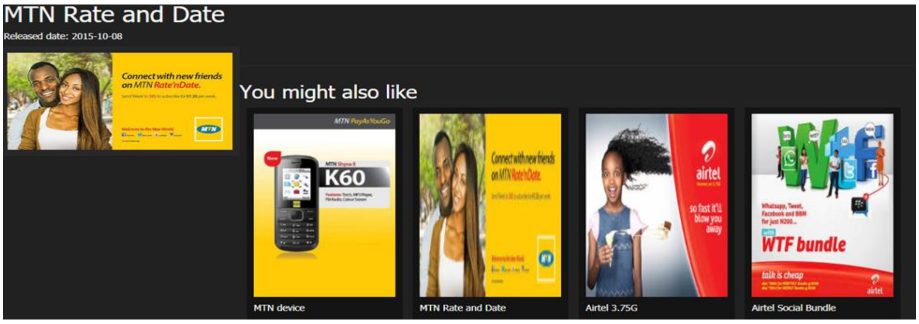
Released date: 2015-08-27

You might also like

MTN device

Fig 4.7 product recommendations

Fig 4.8 shows product being recommend to a subscriber based on what they are likely to use.


MTN Rate and Date

Released date: 2015-10-08

You might also like

MTN device

MTN Rate and Date

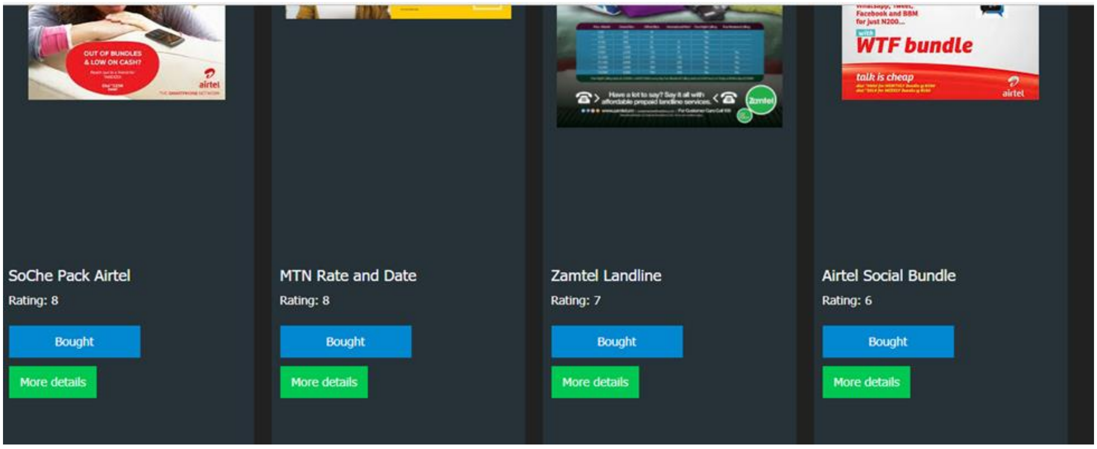
Airtel 3.75G

Airtel Social Bundle

Connect with new friends on MTN Rate'nDate
MTN PayAsYouGo
K60
airtel
so fast it'll blow you away
WTF bundle
talk is cheap

Fig 4.8 product recommendations

Fig 4.9 shows product recommendations to a subscriber.



SoChe Pack Airtel
Rating: 8

MTN Rate and Date
Rating: 8

Zamtel Landline
Rating: 7

Airtel Social Bundle
Rating: 6

Fig 4.9 product recommendations to a subscriber.

4.4 TESTING AND HOSTING

The user study was conducted by asking the candidate several questions related to the system and how they felt about each of these incorporated feature. Some of the questions were:

- **Did you use a similar system before?** This question was initially posted to know if the user had prior experience using any system of similar kind. From this we could find valuable data as they would relate our system with the ones they used before. Our study shows in fig 4.10 that 54 percent of the people among the 18 participants had used a similar kind of system before.

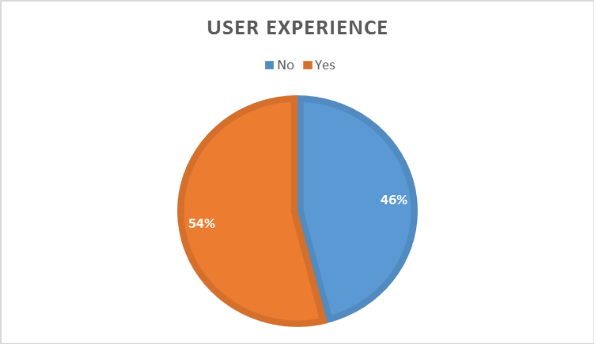


Fig 4.10 product recommendations user experience.

- **How many online recommender system have you used in a year**? We wanted our system to be evaluated by people who had good knowledge with online recommendation system and experience hence we posted this question to know what their level was and interest in online systems. This would particularly guarantee us users who had fair idea about our system and thus the result set and suggestions would be accurate for further development. From the pie chart below in figure 4.11 we observe that many users had taken at least 3 online courses in a year and others took more than 3 too. Many had commented that they did not complete the online shopping fully which may due to several reasons; one being the online products but that was not taken into consideration as we only wanted to know about the basic functionality of the system, recommendation algorithm and other design issues we could further develop.

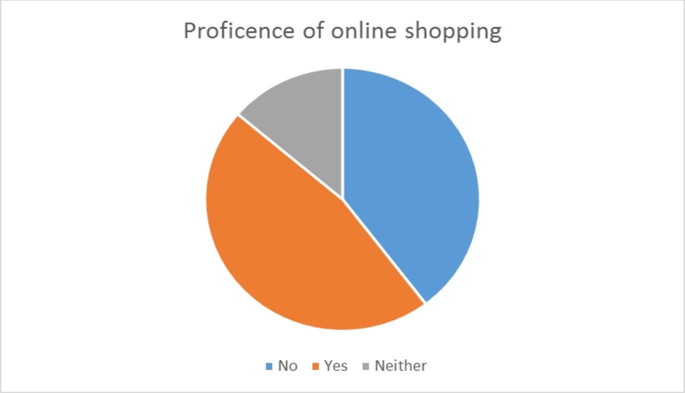


Fig 4.11 proficiency of online shopping

- **Satisfaction with current features?** In order to know how the users felt about the various features and measure the level of satisfaction we posted this in the user study. We observed that among the 16 participants 12 of them liked the idea of a collaborated items with product recommendation. Some of them complained about the basic user interface and thus that remains as a challenge to work and consider the design interface carefully in the future

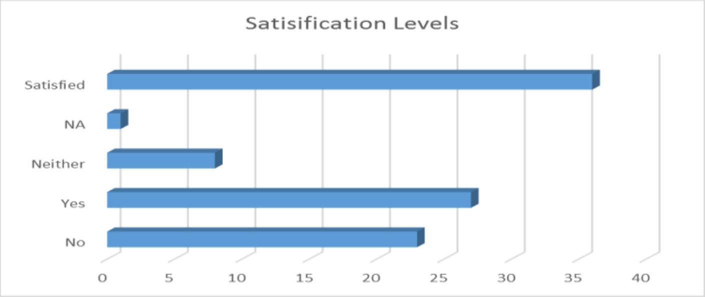


Fig 4.9 proficiency of online shopping

- **Would you like to use our system again?** Finally, we wanted how the users felt about using the system again to gauge how popular the system would be if deployed. To our surprise, 13 amongst the 16 participant said they were interested in our system and would return to a more developed version when deployed. This shows the currently growing interest towards the recommender system and migration from the traditional way of advertising products to customers. There were users who did not want to use this kind of system in the future and were more interested in the traditional way. This potentially maybe due to the flaws in the system or their inclination towards the conventional system of doing things. You can find the questionairs used in **appendix H and I**

4.4 Summary

In this chapter, we successfully analyzed the data that was collected from the extracted Sql data dump used machine learning algorithms and presented the results in form of tables, bar and figs charts. The researcher established that there is an increasing number of mobile operators in the country and they lack a personalized way of recommending products to their subscribers but with the help of machine learning algorithms and data analyses tools, telecommunication companies can predict the services to their customers and from a financial perspective telecommunications operator are confronted with price decrease and cost pressure. Both are related to changed usage behaviors and strong competition in convergent markets. In response, telecommunications operators have to realize new revenue sources through innovative services. Under the condition of globally stagnating telecommunications markets, the challenge is to combine the two contrary objectives of investments in innovations with consistent cost management.

CHAPTER FIVE
DISCUSSION AND CONCLUSION

5.1 Conclusion

Recommender systems made a significant progress over the last decade when numerous content-based, collaborative and hybrid methods were proposed and several "industrial-strength" systems have been developed. However, despite all these advances, the current generation of recommender systems surveyed in this paper still requires further improvements to make recommendation methods more effective in a broader range of applications. In this dissertation, we reviewed various limitations of the current recommendation methods and discussed possible extensions that can provide better recommendation capabilities. These extensions include, among others, the improved modeling of users and items, incorporation of the contextual information into the recommendation process, support for multi-criteria ratings, and provision of a more flexible and less intrusive recommendation process. We hope that the issues presented in this paper would advance the discussion in the recommender systems community about the next generation of recommendation technologies. With increasing number of mobile operators in our country, user is entitled with unlimited freedom to switch from one mobile operator to another if he is not satisfied with service or pricing but with an introduction of a Recommender system, telecommunication companies can use recommender system to suggest products to their customers. The products can be recommended based on the top overall sellers on a site, based on the demographics of the customer, or based on an analysis of the past buying behavior of the customer as a prediction for future buying behavior. Broadly, these techniques are part of personalization on a site, because they help the site adapt itself to each customer

5.2 FINAL DISCUSSION AND FURTHER STUDY

This study proposes a hybrid recommendation approach which combines user-based and item-based collaborative filtering techniques with fuzzy set techniques and knowledge base for mobile product and service recommendation. It particularly implements the approach in a personalized recommender system for telecom products/services. This system has undergone preliminarily testing in a telecom company and achieved excellent performance. As we have mentioned in Section 1, telecom companies have two groups of customers: individual consumers and businesses. This study only focuses on individual consumers. In the future, the recommendation approach and software system will be improved and adapted to develop a mobile product/service recommender system to support business customers. In that situation, a customer (business) may have multiple handsets with different plans, multiple services including fixed-line, SMS, GSM mobiles, access to Facebook, Twitter, and more. The similarity between two customers becomes very difficult and has high uncertainty. A new tree-structure fuzzy measure approach will be developed and used in a new recommendation approach.

5.3 Recommendations

I love sites like Netflix, Pandora, Facebook, bay and Amazon. They all do a fantastic job in finding what I want, and providing me with relevant product recommendations. They manage to dig in deep into their products and “inflate” the goodies that I really like. I am a loyal customer. Loyalty is brought by understanding customers and delivering to them what they want or value. I would recommend telecommunication companies to use such technologies and Integration between them, recommendation engines harvest real-time and historical data from your clients such as frequency of visits to the web pages, which pages they have visited, duration and time spent visiting specific pages, product(s) clicked on, purchases made, etc. Everything is and can be collected. Your customers will now have much more information in their profile without them needing to populate and manually fill in forms, and you can start understanding who they are and how to serve them better.

5.4 Summary

In this chapter, we discussed the results presented in Chapter four. In section 5 we discussed the results obtained from the baseline study. In section 4 we discussed the system implementation. We further presented the conclusion of the study, the recommendations and the future works in the sections that followed.

REFERENCES

[1] Artificial intelligence for the public sector: opportunities and challenges of cross-sector collaboration Slava Jankin Mikhaylov, Marc Esteve, and Averill Campion, Published: 06 August 2018 https://doi.org/10.1098/rsta.2017.0357

[2] Waller, M. A. and Fawcett, S. E. (2013), Data Science, Predictive Analytics, and Big Data: A Revolution That Will Transform Supply Chain Design and Management. J Bus Logist, 34: 77-84. doi:10.1111/jbl.12010.

[3] Ritala, P. (2012), Coopetition Strategy – When is it Successful? Empirical Evidence on Innovation and Market Performance. British Journal of Management, 23: 307-324. doi:10.1111/j.1467-8551.2011.00741.x

[4] Zui Zhang, Hua Lin, Kun Liu, Dianshuang Wu, Guangquan Zhang, Jie Lu,A hybrid fuzzy-based personalized recommender system for telecom products/services,
Information Sciences, Volume 235, 2013, Pages 117-129, ISSN 0020-0255,https://doi.org/10.1016/j.ins.2013.01.025.

[5] Eric Bouillet, Ravi Kothari, Vibhore Kumar, Laurent Mignet, Senthil Nathan, Anand Ranganathan, Deepak S. Turaga, Octavian Udrea, and Olivier Verscheure. 2012. Processing 6 billion CDRs/day: from research to production (experience report). In Proceedings of the 6th ACM International Conference on Distributed Event-Based Systems (DEBS '12). ACM, New York, NY, USA, 264-267. DOI: https://doi.org/10.1145/2335484.2335513

[6] Ogbo, Erezi and Brown, Tim and Sicker, Douglas, Understanding Mobile Service Substitution and the Urban-Rural Digital Divide in Nigeria (March 31, 2017). Available at SSRN: https://ssrn.com/abstract=2944367 or http://dx.doi.org/10.2139/ssrn.2944367.

[7] Bloomfield, Gerald S.and Vedanthan, Rajesh and Vasudevan, Lavanya and Kithei, Anne and Were, Martin and Velazquez, Eric J. Mobile health for non-communicable diseases in Sub-Saharan Africa: a systematic review of the literature and strategic framework for research", journal="Globalization and Health,"2014", doi="10.1186/1744-8603-1.

[8] Porter, G. (2015). Mobile Phones, Mobility Practices, and Transport Organization in Sub-Saharan Africa. Mobility in History 6, 1, 81-88, available from: < https://doi.org/10.3167/mih.2015.060109>.

[9] A Hafez - Int. J. Comput. Electr. Autom. Control Inf. Eng, 2016 - academia.edu:Mining Big Data in Telecommunications Industry:Challenges, Techniques, and Revenue Opportunity.

[10] K.r. rashmi..Determinants of customer loyalty in indian mobile telecom sector-a conceptual analysis.

[11] The telecommunication industry revisited: The changing pattern of partnerships Grover and Saeed 2003.

[12] Value Creation in the MobileMarket - A Reference Model for the Role(s) of the FutureMobile Network Operator: Pousttchi and Hufenbach 2011.

[13] From value chain to value network: Insights for mobile operators: Peppard, Joe; Rylander, Anna.

[14] ITU releases 2015 ICT figs.

[15] Szymon Jaroszewicz: Cross-selling models for telecommunication services.

[16] Peppard, J. and Rylander, A. (2006) From Value Chain to Value Network: Insights for Mobile Operators.

[17] Christian Czarnecki, Christian Dietze:Reference Architecture for the Telecommunications Industry: Transformation of Strategy, Organization, Processes, Data, and Applications

[18] Freund, Y., R. Iyer, R.E. Schapire, and Y. Singer. An efficient boosting algorithm for combining preferences. In Proc. of the 15th Intl. Conference on Machine Learning, 1998.

[19] Jin, R., L. Si, and C. Zhai. Preference-based Graphic Models for Collaborative Filtering. In Proceedings of the 19th Conference on Uncertainty in Artificial Intelligence (UAI 2003), Acapulco, Mexico, August 2003a.

[20] Jin, R., L. Si, C. Zhai, and J. Callan. Collaborative Filtering with Decoupled Models for Preferences and Ratings. In Proc. of the 12th International Conference on Information and Knowledge Management (CIKM 2003), New Orleans, LA, November 2003b

[21]

[18] Freund, Y., R. Iyer, R.E. Schapire, and Y. Singer. An efficient boosting algorithm for combining preferences. In Proc. of the 15th Intl. Conference on Machine Learning, 1998.

[19] Jin, R., L. Si, and C. Zhai. Preference-based Graphic Models for Collaborative Filtering. In Proceedings of the 19th Conference on Uncertainty in Artificial Intelligence (UAI 2003), Acapulco, Mexico, August 2003a.

[20] Jin, R., L. Si, C. Zhai, and J. Callan. Collaborative Filtering with Decoupled Models for Preferences and Ratings. In Proc. of the 12th International Conference on Information and Knowledge Management (CIKM 2003), New Orleans, LA, November 2003b

[21] Towards the Next Generation of Recommender Systems:A Survey of the State-of-the-Art and Possible Extensions Gediminas Adomavicius and Alexander Tuzhilin.

[22] Mahendra, Yohanes Dicky. Sistem rekomendasi objek wisata Yogyakarta dengan pendekatan item-Based Collaborative Filtering. Diss. Sanata Dharma University, 2019.

[23] Mahendra, Yohanes Dicky. Sistem rekomendasi objek wisata Yogyakarta dengan pendekatan item-Based Collaborative Filtering. Diss. Sanata Dharma University, 2019.

[24] Yehuda Koren, Yahoo Research Robert Bell and Chris Volinsky, AT&T Labs—Research....matrix factorization techniques for recommender systems

[25] Recommender Systems in E-Commerce J. Ben Schafer, Joseph Konstan,John Riedl.

[26]

[27] Oard DW, Kim J. Implicit feedback for recommender systems. In: Proceedings of 5th DELOS workshop on filtering and collaborative filtering; 1998. p. 31–6.

[28] J. Buder, C. Schwind Learning with personalized recommender systems: a psychological view Comput Human Behav, 28 (1) (2012), pp. 207-216.

[29] Gadanho SC, Lhuillier N. Addressing uncertainty in implicit preferences. In: Proceedings of the 2007 ACM conference on Recommender Systems (RecSys '07). ACM, New York, NY, USA; 2007. p. 97–104.

[30] Resnick, P. and Varian, H.R., 1997. Recommender systems. Communications of the ACM, 40(3), pp.56-58.

[31] Koji Miyahara, Michael J. Pazzani: Collaborative Filtering with the Simple Bayesian Classifier. PRICAI 2000: 679-689

[32] Adomavicius, G. and Tuzhilin, A., 2004. Recommendation technologies: Survey of current methods and possible extensions.

[33] Information filtering and information retrieval: two sides of the same coin Belkin & Croft 1992.

[34] Schmitt, S. and Bergmann, R., 1999, June. Applying case-based reasoning technology for product selection and customization in electronic commerce environments. In 12th Bled Electronic Commerce Conference (Vol. 273).

[35] Good, N., Schafer, J.B., Konstan, J.A., Borchers, A., Sarwar, B., Herlocker, J. and Riedl, J., 1999, July. Combining collaborative filtering with personal agents for better recommendations. In AAAI/IAAI (pp. 439-446).

[36] Claypool, M., Gokhale, A., Miranda, T., Murnikov, P., Netes, D. and Sartin, M., 1999. Combing content-based and collaborative filters in an online newspaper.

[37] Burke, R., 2002. Hybrid recommender systems: Survey and experiments. User modeling and user-adapted interaction, 12(4), pp.331-370.

[38] Sang-Bum Kim, Kyoung-Soo Han, Hae-Chang Rim, and Sung Hyon Myaeng, "Some Effective Techniques for Naive Bayes Text Classification", IEEE Transactions On

Knowledge and Data Engineering, Vol. 18, No. 11, Pp-1457- 1466, November 2006.

[39] Michalski, R. S., Carbonell, J. G., & Mitchell, T. M. (Eds.). (2013). Machine learning: An artificial intelligence approach. Springer Science & Business Media.

[40] Apte, C. (2010). The role of machine learning in business optimization. In Proceedings of the 27th International Conference on Machine Learning (ICML10) (pp. 1-2).

[41] Cui, Q., Bai, F. S., Gao, B., & Liu, T. Y. (2015). Global Optimization for Advertisement Selection in Sponsored Search. Journal of Computer Science and Technology, 30(2), 295-310.

[42] Kononenko, I. (2001). Machine learning for medical diagnosis: history, state of the art and perspective. Artificial Intelligence in medicine, 23(1), 89-109.

[43] Shalev-Shwartz, S., & Ben-David, S. (2014). Understanding Machine Learning: From Theory to Algorithms. Cambridge University Press.

[44] Kulkarni, S. (Ed.). (2012). Machine Learning Algorithms for Problem Solving in Computational Applications: Intelligent Techniques: Intelligent Techniques. IGI Global.

[45] https://www.analyticsvidhya.com/blog/2017/09/common-machine-learning-algorithms/.

[46] Eui-Hong (Sam) Han, George Karypis, Vipin Kumar;Text Categorization Using Weighted Adjusted k-Nearest Neighbor Classification, Department of Computer Science
And Engineering. Army HPC Research Centre, University of Minnesota, Minneapolis, USA. 1999.

[47] Comparison between Traditional Approach and Object-Oriented Approach in Software Engineering Development, Nabil Mohammed Ali Munassar

[48] Chen, H., Chiang, R.H. and Storey, V.C., 2012. Business intelligence and analytics: from big data to big impact. MIS quarterly, pp.1165-1188.

[49] Heide Brücher, Gerhard Knolmayer, Marc-André Mittermayer; “Document Classification Methods for Organizing Explicit Knowledge”, Research Group Information Engineering,Institute of Information Systems, University of Bern, Engehaldenstrasse 8, CH - 3012 Bern, Switzerland. 2002.

[50] Andrew McCallum, Kamal Nigam; “A Comparison of Event Models for Naïve Bayes Text Classification”, Journal of Machine Learning Research 3, pp. 1265-1287. 2003.

[51] Irina Rish; “An Empirical Study of the Naïve Bayes Classifier”, In Proceedings of the IJCAI-01 Workshop on Empirical Methods in Artificial Intelligence. 2001.

[52] Irina Rish, Joseph Hellerstein, Jayram Thathachar; “An Analysis of Data Characteristics that affect Naïve Bayes Performance”, IBM T.J. Watson Research Center 30.

[53] Pedro Domingos, Michael Pazzani; “On the Optimality of the Simple Bayesian Classifier under Zero-One Loss,Machine Learning”, Vol. 29, No. 2-3, pp.103-130. 1997.

[54] Sang-Bum Kim, Hue-Chang Rim, Dong-Suk Yook,Huei-Seok Lim; "Effective Methods for Improving Naïve Bayes Text Classification", 7th Pacific Rim International Conference on Artificial Intelligence, Vol. 2417. 2002

[55] Performance Analysis of Naive Bayes and J48 Classification Algorithm for Data Classification Tina R. Patil, Mrs. S. S. Sherekar Sant Gadgebaba Amravati University, Amravati.

[56] Dino Isa, Lam Hong lee, V. P Kallimani, R. RajKumar, "Text Documents Preprocessing with the Bahes Formula for Classification using the Support vector machine", IEEE, Traction of Knowledge and Data Engineering, Vol-20, N0-9 pp-1264-1272, 2008.

[57] Dino Isa, V. P Kallimani Lam Hong lee, "Using Self Organizing Map for Clustering of Text Documents",", Elsever, Expert System with Applications-2008.

[58] Sang-Bum Kim, Kyoung-Soo Han, Hae-Chang Rim, and Sung Hyon Myaeng, "Some Effective Techniques for Naive Bayes Text Classification", IEEE Transactions On Knowledge and Data Engineering, Vol. 18, No. 11, Pp-1457- 1466, November 2006.

[59] P. Domingos and M. J. Pazzani, "On the Optimality of the Simple Bayesian Classifier under Zero-One Loss," Machine Learning, vol. 29, nos. 2/3, pp. 103-130, 1997.

[60] Thiago S.Guzella, Walimir M. Caminhas "A Review of machine Learning Approches to Spam Filtering", Elsever, Expert System with Applications-2009.

[61] Box, G.E. and Tiao, G.C., 2011. Bayesian inference in statistical analysis (Vol. 40). John Wiley & Sons.

[62] Schermelleh-Engel, K., Moosbrugger, H. and Müller, H., 2003. Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. Methods of psychological research online, 8(2), pp.23-74.

[63] Dimitriadis, S.I. and Liparas, D., How Random is the Random Forest? Random Forest Algorithm on the Service of Structural Imaging Biomarkers for Alzheimer's.

[64] Liaw, A. and Wiener, M., 2002. Classification and regression by randomForest. R news, 2(3), pp.18-22.

[65] Schafer, J.B., Konstan, J. and Riedl, J., 1999, November. Recommender systems in e-commerce. In Proceedings of the 1st ACM conference on Electronic commerce (pp. 158-166). ACM.

[66] Prasad, B., 2007. A knowledge-based product recommendation system for e-commerce. International Journal of Intelligent Information and Database Systems, 1(1), pp.18-36.

[67] Wang, J., De Vries, A.P. and Reinders, M.J., 2006, April. A user-item relevance model for log-based collaborative filtering. In European conference on information retrieval (pp. 37-48). Springer, Berlin, Heidelberg.

[68] S. Bennett, S. McRobb and R. Farmer, Object-Oriented Systems Analysis and Design using UML, 3rd ed., McGraw Hill, 2006.

[69] Thiago S.Guzella, Walimir M. Caminhas “A Review of machine Learning Approches to Spam Filtering”, Elsever, Expert System with Applications-2009.

[70] Baxter, P. and Jack, S., 2008. Qualitative case study methodology: Study design and implementation for novice researchers. The qualitative report, 13(4), pp.544-559.

[71] C. R. Kothari, Research Methodlogy: Methods and Techniques, 2nd ed., New Delhi: New Age International (P) Ltd, 2004

[72] Lu, J., Shambour, Q., Xu, Y., Lin, Q. and Zhang, G., 2010. BizSeeker: a hybrid semantic recommendation system for personalized government-to-business e-services. Internet Research, 20(3), pp.342-365.

[73] Engle, G., 2011. System and method of collaborative filtering based on attribute profiling. U.S. Patent 8,001,008.

[74] I. Sommerville, Software Engineering, 9th ed., Boston: Addison-Wesley, 2011

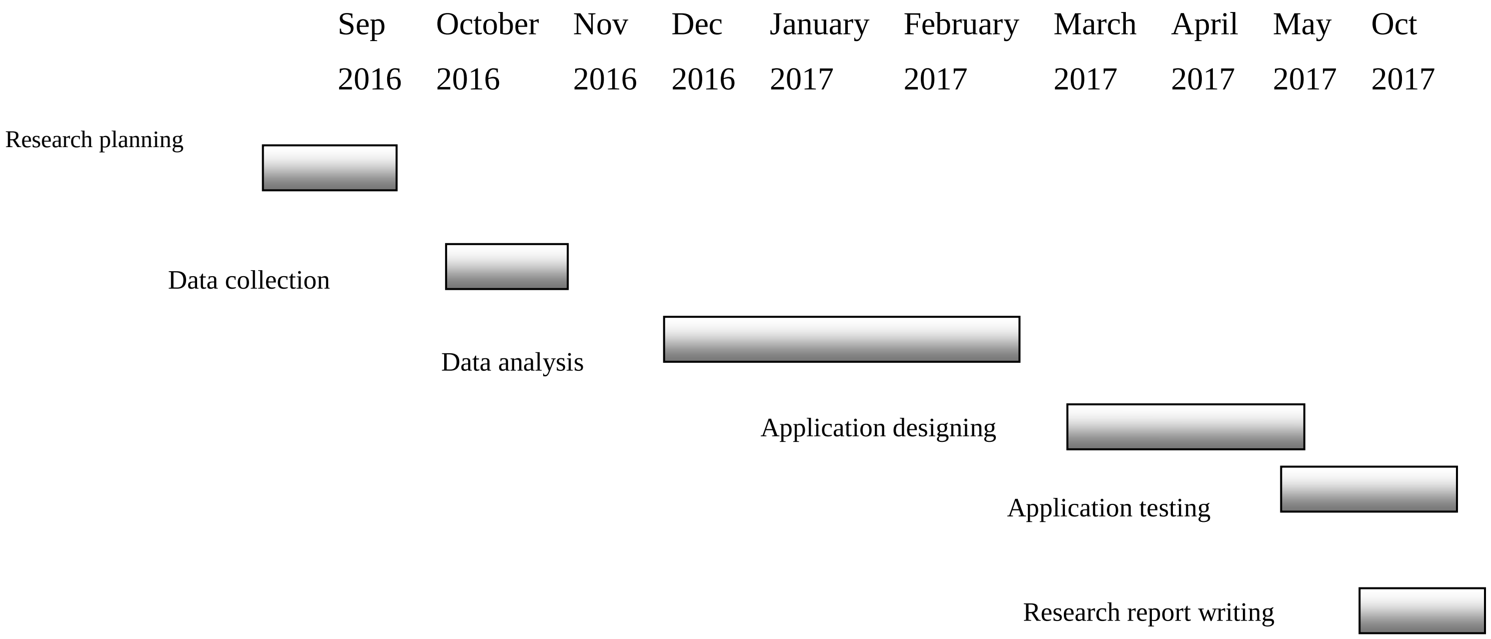
[75] Jaccard, P.: Etude comparative de la distribution florale dans une portion des Alpes ET du Jura.Impr. Corbaz (1901).

[76] Adomavicius, G., Tuzhilin, A.: Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. IEEE Trans. Knowl. Data Eng. 17(6), 734–749(2005).

APPENDIX A

1.0 Timelines Frame for Activities

The research will flow according to the following schedule of tasks:



APPENDIX B

1.0 Budget

The following is my projected budget for the entire research.

Item	Description	Estimated cost (USD)
Research assistant’s stipend	Payment to research assistants	833
Stationery	Paper for questionnaires and other documents if need be	139
Travel	Fuel for traveling	167
Food	Food when travelling	228
Publication	Research publication fee, printing and binding	139
Laptop computer	Purchase of powerful laptop for analysis and development.	800

Total	**2306**

APPENDIX C: Code for the algorithms

K-Nearest Neighbors (K-NN) algorithm code

```
# K-Nearest Neighbors (K-NN)

# Importing the libraries
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd

# Importing the dataset
dataset = pd.read_csv('Social_Network_Ads.csv')
X = dataset.iloc[:, [2, 3]].values
y = dataset.iloc[:, 4].values

# Splitting the dataset into the Training set and Test set
from sklearn.cross_validation import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.25, random_state = 0)

# Feature Scaling
from sklearn.preprocessing import StandardScaler
sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)

# Fitting K-NN to the Training set
from sklearn.neighbors import KNeighborsClassifier
classifier = KNeighborsClassifier(n_neighbors = 5, metric = 'minkowski', p = 2)
classifier.fit(X_train, y_train)

# Predicting the Test set results
y_pred = classifier.predict(X_test)

# Making the Confusion Matrix
from sklearn.metrics import confusion_matrix
cm = confusion_matrix(y_test, y_pred)

# Visualising the Training set results
from matplotlib.colors import ListedColormap
X_set, y_set = X_train, y_train
X1, X2 = np.meshgrid(np.arange(start = X_set[:, 0].min() - 1, stop = X_set[:, 0].max() + 1, step = 0.01),
                     np.arange(start = X_set[:, 1].min() - 1, stop = X_set[:, 1].max() + 1, step = 0.01))
plt.contourf(X1, X2, classifier.predict(np.array([X1.ravel(), X2.ravel()]).T).reshape(X1.shape),
             alpha = 0.75, cmap = ListedColormap(('red', 'green')))
plt.xlim(X1.min(), X1.max())
plt.ylim(X2.min(), X2.max())
for i, j in enumerate(np.unique(y_set)):
    plt.scatter(X_set[y_set == j, 0], X_set[y_set == j, 1],
                c = ListedColormap(('red', 'green'))(i), label = j)
plt.title('K-NN (Training set)')
plt.xlabel('Subscribers')
plt.ylabel('Estimated products')
plt.legend()
plt.show()

# Visualising the Test set results
from matplotlib.colors import ListedColormap
X_set, y_set = X_test, y_test
X1, X2 = np.meshgrid(np.arange(start = X_set[:, 0].min() - 1, stop = X_set[:, 0].max() + 1, step = 0.01),
                     np.arange(start = X_set[:, 1].min() - 1, stop = X_set[:, 1].max() + 1, step = 0.01))
plt.contourf(X1, X2, classifier.predict(np.array([X1.ravel(), X2.ravel()]).T).reshape(X1.shape),
             alpha = 0.75, cmap = ListedColormap(('red', 'green')))
plt.xlim(X1.min(), X1.max())
plt.ylim(X2.min(), X2.max())
for i, j in enumerate(np.unique(y_set)):
    plt.scatter(X_set[y_set == j, 0], X_set[y_set == j, 1],
                c = ListedColormap(('red', 'green'))(i), label = j)
plt.title('K-NN (Test set)')
plt.xlabel('Subscribers')
plt.ylabel('Estimated products')
plt.legend()
plt.show()
```

Support Vector Machine (SVM) algorithm code

```
# Support Vector Machine (SVM)

# Importing the libraries
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd

# Importing the dataset
dataset = pd.read_csv('Social_Network_Ads.csv')
X = dataset.iloc[:, [2, 3]].values
y = dataset.iloc[:, 4].values

# Splitting the dataset into the Training set and Test set
from sklearn.cross_validation import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.25, random_state = 0)

# Feature Scaling
from sklearn.preprocessing import StandardScaler
sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)

# Fitting SVM to the Training set
from sklearn.svm import SVC
classifier = SVC(kernel = 'linear', random_state = 0)
classifier.fit(X_train, y_train)

# Predicting the Test set results
y_pred = classifier.predict(X_test)

# Making the Confusion Matrix
from sklearn.metrics import confusion_matrix
cm = confusion_matrix(y_test, y_pred)

# Visualising the Training set results
```

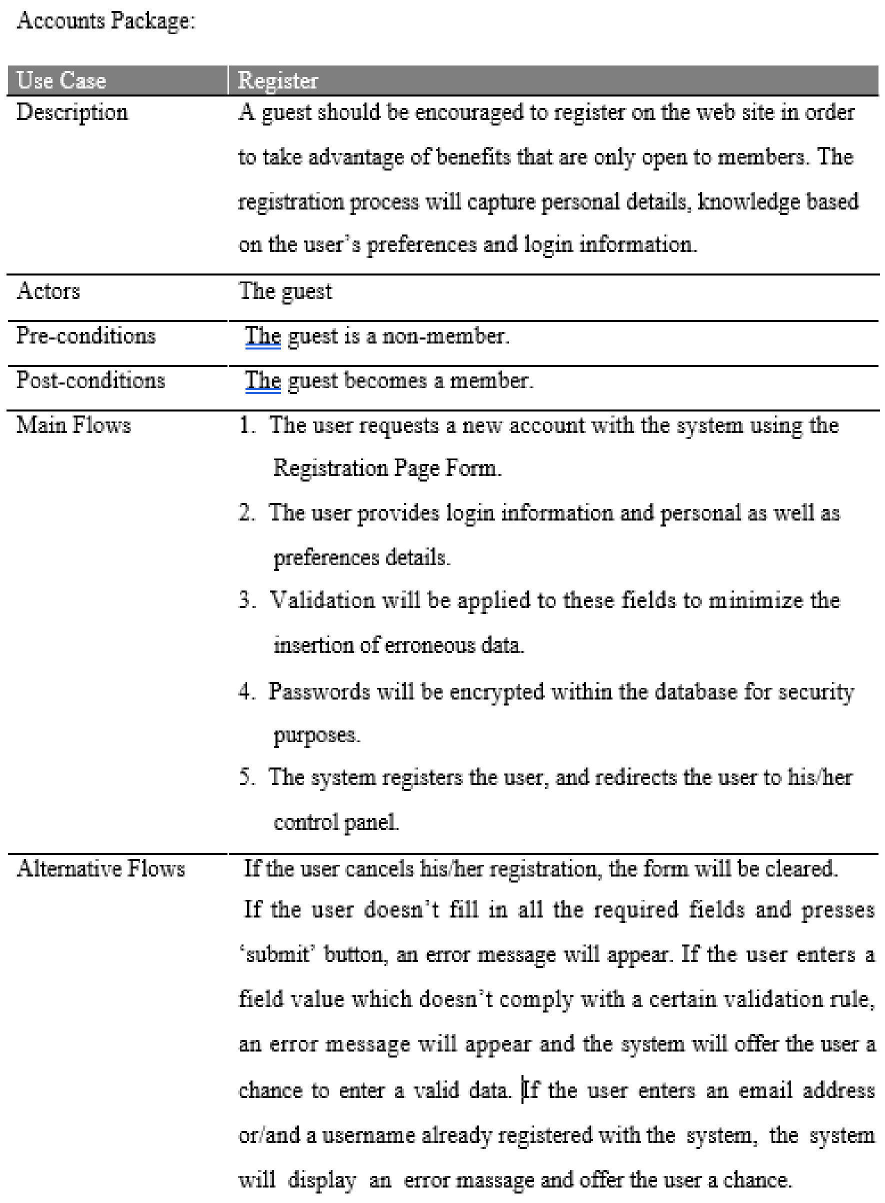
```
from matplotlib.colors import ListedColormap
X_set, y_set = X_train, y_train
X1, X2 = np.meshgrid(np.arange(start = X_set[:, 0].min() - 1, stop = X_set[:, 0].max() + 1, step = 0.01),
                     np.arange(start = X_set[:, 1].min() - 1, stop = X_set[:, 1].max() + 1, step = 0.01))
plt.contourf(X1, X2, classifier.predict(np.array([X1.ravel(), X2.ravel()]).T).reshape(X1.shape),
             alpha = 0.75, cmap = ListedColormap(('red', 'green')))
plt.xlim(X1.min(), X1.max())
plt.ylim(X2.min(), X2.max())
for i, j in enumerate(np.unique(y_set)):
    plt.scatter(X_set[y_set == j, 0], X_set[y_set == j, 1],
                c = ListedColormap(('red', 'green'))(i), label = j)
plt.title('SVM (Training set)')
plt.xlabel('Age')
plt.ylabel('Products')
plt.legend()
plt.show()

# Visualising the Test set results
from matplotlib.colors import ListedColormap
X_set, y_set = X_test, y_test
X1, X2 = np.meshgrid(np.arange(start = X_set[:, 0].min() - 1, stop = X_set[:, 0].max() + 1, step = 0.01),
                     np.arange(start = X_set[:, 1].min() - 1, stop = X_set[:, 1].max() + 1, step = 0.01))
plt.contourf(X1, X2, classifier.predict(np.array([X1.ravel(), X2.ravel()]).T).reshape(X1.shape),
             alpha = 0.75, cmap = ListedColormap(('red', 'green')))
plt.xlim(X1.min(), X1.max())
plt.ylim(X2.min(), X2.max())
for i, j in enumerate(np.unique(y_set)):
    plt.scatter(X_set[y_set == j, 0], X_set[y_set == j, 1],
                c = ListedColormap(('red', 'green'))(i), label = j)
plt.title('SVM (Test set)')
plt.xlabel('Age')
plt.ylabel('Products')
plt.legend()
```

Random Forest Classification algorithm code

```
# Random Forest Classification

# Importing the libraries
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd

# Importing the dataset
dataset = pd.read_csv('Social_Network_Ads.csv')
X = dataset.iloc[:, [2, 3]].values
y = dataset.iloc[:, 4].values

# Splitting the dataset into the Training set and Test set
from sklearn.cross_validation import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.25, random_state = 0)

# Feature Scaling
from sklearn.preprocessing import StandardScaler
sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)
```

```python

# Fitting Random Forest Classification to the Training set
from sklearn.ensemble import RandomForestClassifier
classifier = RandomForestClassifier(n_estimators = 10, criterion = 'entropy', random_state = 0)
classifier.fit(X_train, y_train)

# Predicting the Test set results
y_pred = classifier.predict(X_test)

# Making the Confusion Matrix
from sklearn.metrics import confusion_matrix
cm = confusion_matrix(y_test, y_pred)

# Visualising the Training set results
from matplotlib.colors import ListedColormap
X_set, y_set = X_train, y_train
X1, X2 = np.meshgrid(np.arange(start = X_set[:, 0].min() - 1, stop = X_set[:, 0].max() + 1, step = 0.01),
                     np.arange(start = X_set[:, 1].min() - 1, stop = X_set[:, 1].max() + 1, step = 0.01))
plt.contourf(X1, X2, classifier.predict(np.array([X1.ravel(), X2.ravel()]).T).reshape(X1.shape),
             alpha = 0.75, cmap = ListedColormap(('red', 'green')))
plt.xlim(X1.min(), X1.max())
plt.ylim(X2.min(), X2.max())
for i, j in enumerate(np.unique(y_set)):
    plt.scatter(X_set[y_set == j, 0], X_set[y_set == j, 1],
                c = ListedColormap(('red', 'green'))(i), label = j)
plt.title('Random Forest Classification (Training set)')
plt.xlabel('Subscribers')
plt.ylabel('products')
plt.legend()
plt.show()

# Visualising the Test set results
from matplotlib.colors import ListedColormap
X_set, y_set = X_test, y_test
X1, X2 = np.meshgrid(np.arange(start = X_set[:, 0].min() - 1, stop = X_set[:, 0].max() + 1, step = 0.01),
                     np.arange(start = X_set[:, 1].min() - 1, stop = X_set[:, 1].max() + 1, step = 0.01))
plt.contourf(X1, X2, classifier.predict(np.array([X1.ravel(), X2.ravel()]).T).reshape(X1.shape),
             alpha = 0.75, cmap = ListedColormap(('red', 'green')))
plt.xlim(X1.min(), X1.max())
plt.ylim(X2.min(), X2.max())
for i, j in enumerate(np.unique(y_set)):
    plt.scatter(X_set[y_set == j, 0], X_set[y_set == j, 1],
                c = ListedColormap(('red', 'green'))(i), label = j)
plt.title('Random Forest Classification (Test set)')
plt.xlabel('Subscribers')
plt.ylabel('products')
plt.legend()
plt.show()
```

APPENDIX D: Use Cases



Accounts Package:

Use Case	Register
Description	A guest should be encouraged to register on the web site in order to take advantage of benefits that are only open to members. The registration process will capture personal details, knowledge based on the user's preferences and login information.
Actors	The guest
Pre-conditions	The guest is a non-member.
Post-conditions	The guest becomes a member.
Main Flows	1. The user requests a new account with the system using the Registration Page Form. 2. The user provides login information and personal as well as preferences details. 3. Validation will be applied to these fields to minimize the insertion of erroneous data. 4. Passwords will be encrypted within the database for security purposes. 5. The system registers the user, and redirects the user to his/her control panel.
Alternative Flows	If the user cancels his/her registration, the form will be cleared. If the user doesn't fill in all the required fields and presses 'submit' button, an error message will appear. If the user enters a field value which doesn't comply with a certain validation rule, an error message will appear and the system will offer the user a chance to enter a valid data. If the user enters an email address or/and a username already registered with the system, the system will display an error massage and offer the user a chance.

Use Case	Login
Description	A login page will provide an interface to authenticate a user before using restricted features.
Actors	The user
Pre-conditions	The user is a member.
Post-conditions	The user is logged into the system and will remain logged in until he/she loges out or leaves the website for an extended period of time.
Main Flows	1. The user requests to log in with the system using the Login Page Form. 2. The login form requires the capture of a username and password. 3. The user enters a username and password. 4. The system verifies the username and password. 5. The system signs the user into the website, and redirects the user to his/her Control Panel.
Alternative Flows	If the user is not registered, the system will request the user to register instead of login. If the user enters an incorrect username and/or password, the system will display an error message and offer another chance to enter the correct account details.

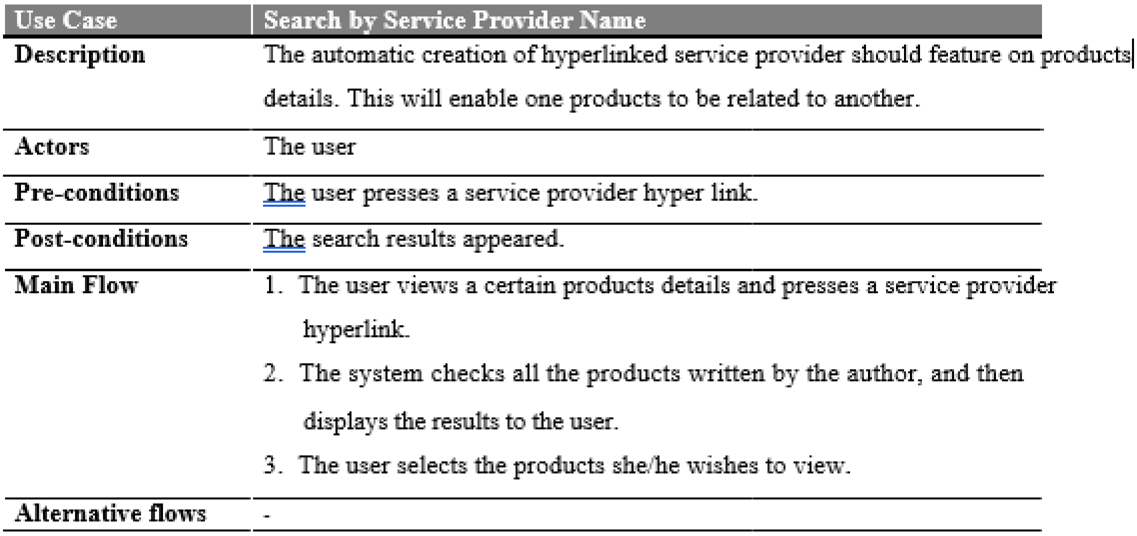
Use Case	Edit Profile
Description	A user may edit his/her profile to change his/her personal or preference information.
Actors	The user
Pre-conditions	The profile page will appear if the user loges in.
Post-conditions	The user's profile has been modified.
Main Flows	1. The user requests to edit their profile. 2. The user enters the new information. 3. The system verifies and saves the changes.
Alternative Flows	If the user enters an incorrect personal data, the system will display an error message to inform the user.

Use Case	Statistical
Description	A statistical chart which represents a user's interest in a certain category.
Actors	The user
Pre-conditions	The statistical page will appear if the user loges in.
Post-conditions	The statistical chart has been showed.
Main Flows	1. The user selects a certain category. 2. The system shows a statistical graph based on the selected category.
Alternative Flows	If the user doesn't select a category, the system will display an error message to inform the user. If The selected category does not contain any data about the user's interest, the system will display a message to inform the user and encourage him/her to evaluate more books.

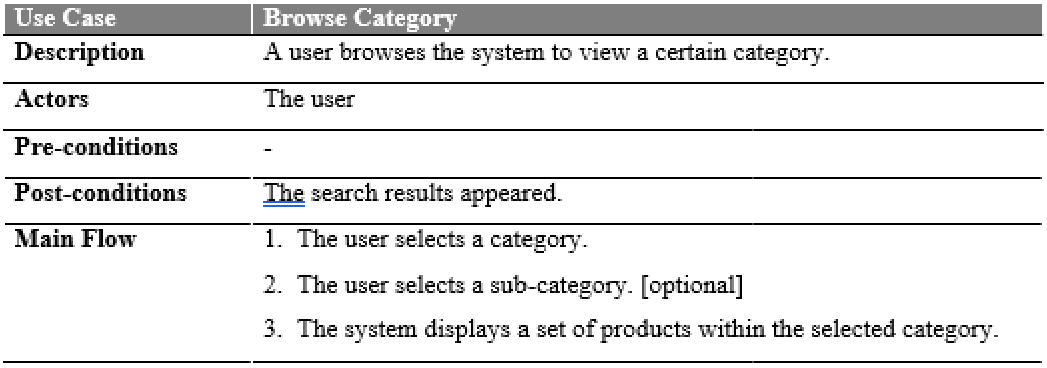
Use Case	View Shopping cart
Description	A user should be able to view his/her shopping cart.
Actors	The user
Pre-conditions	The user must be logged in.
Post-conditions	The shopping cart has been showed.
Main Flows	1. The user requests to view his/her shopping cart. 2. The system shows products in the user's shopping cart.
Alternative Flows	If there are no products in the user's shopping cart, the system will display a message to inform the user.

Browsing Package:

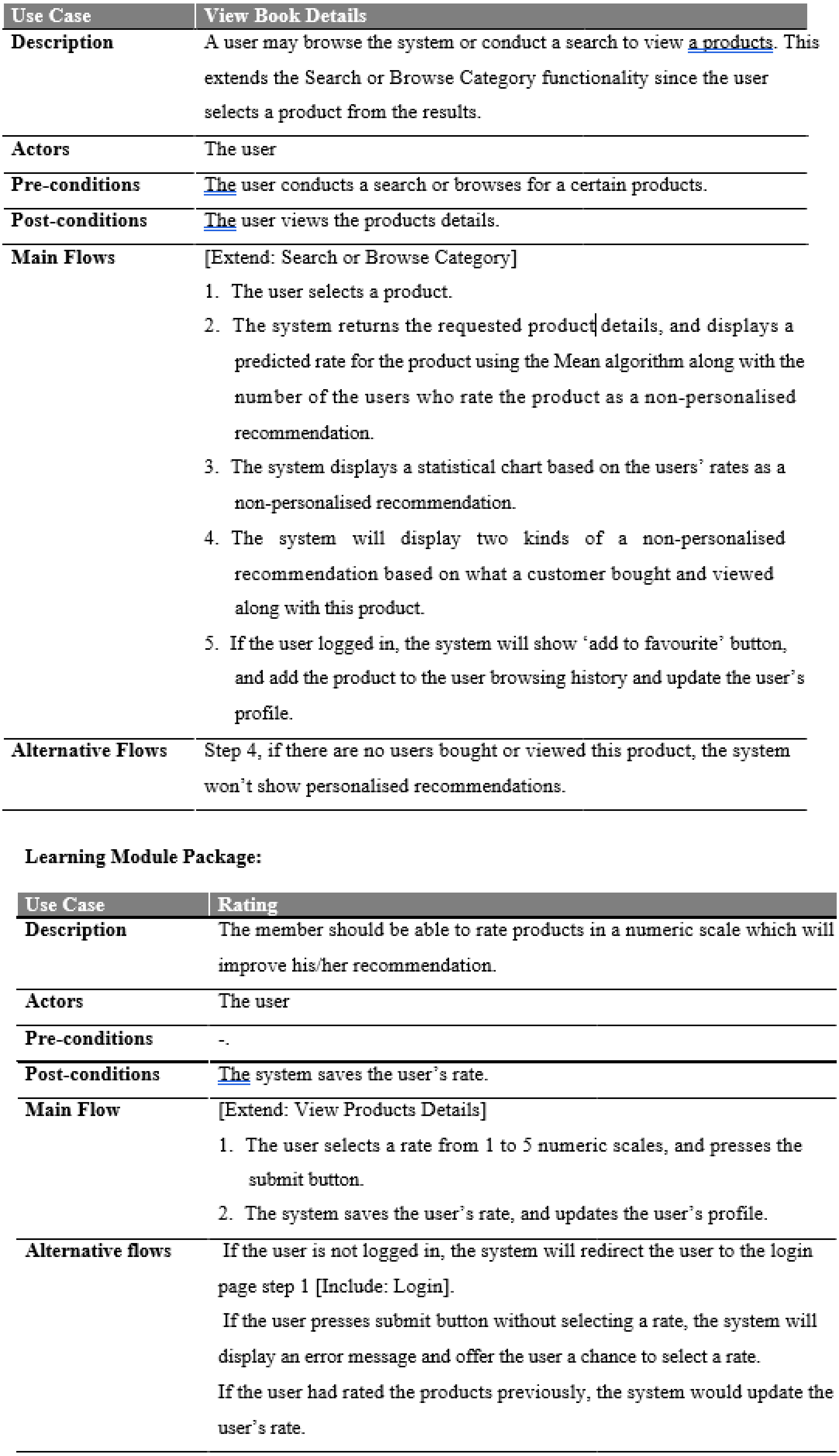
Use Case	Search
Description	There are two types of search: basic and advanced. Any visitor to the site can conduct a search.
Actors	The user
Pre-conditions	The user writes the keyword or fills out advanced search form for books that he/she is looking for.
Post-conditions	The search results appeared.
Main Flows	1. The user enters a keyword before submitting the search form. 2. The system checks for matching entries and displays the results to the user. 3. The user selects the product she/he wishes to view.
Alternative Flows	1. If the user wishes to perform an advanced search, Step 1 will be replaced with a form to the user; the user fills out and submits the advanced search form. 2. If there are no matches to the search term, the system will inform the user.



Use Case	Search by Service Provider Name
Description	The automatic creation of hyperlinked service provider should feature on products details. This will enable one products to be related to another.
Actors	The user
Pre-conditions	The user presses a service provider hyper link.
Post-conditions	The search results appeared.
Main Flow	1. The user views a certain products details and presses a service provider hyperlink. 2. The system checks all the products written by the author, and then displays the results to the user. 3. The user selects the products she/he wishes to view.
Alternative flows	-



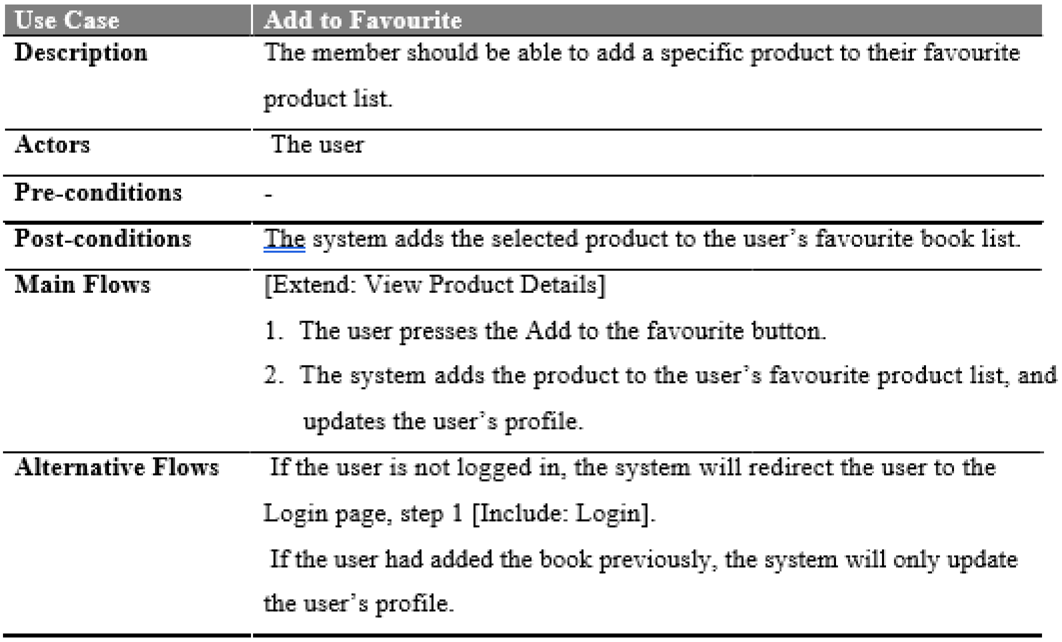
Use Case	Browse Category
Description	A user browses the system to view a certain category.
Actors	The user
Pre-conditions	-
Post-conditions	The search results appeared.
Main Flow	1. The user selects a category. 2. The user selects a sub-category. [optional] 3. The system displays a set of products within the selected category.



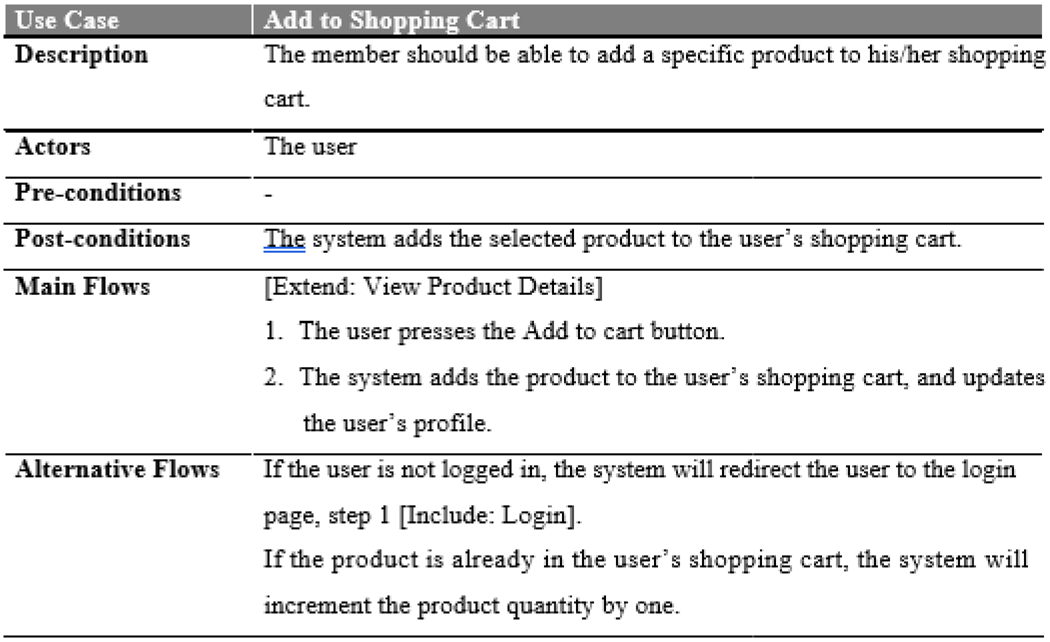
Use Case	View Book Details
Description	A user may browse the system or conduct a search to view a products. This extends the Search or Browse Category functionality since the user selects a product from the results.
Actors	The user
Pre-conditions	The user conducts a search or browses for a certain products.
Post-conditions	The user views the products details.
Main Flows	[Extend: Search or Browse Category] 1. The user selects a product. 2. The system returns the requested product details, and displays a predicted rate for the product using the Mean algorithm along with the number of the users who rate the product as a non-personalised recommendation. 3. The system displays a statistical chart based on the users' rates as a non-personalised recommendation. 4. The system will display two kinds of a non-personalised recommendation based on what a customer bought and viewed along with this product. 5. If the user logged in, the system will show 'add to favourite' button, and add the product to the user browsing history and update the user's profile.
Alternative Flows	Step 4, if there are no users bought or viewed this product, the system won't show personalised recommendations.

Learning Module Package:

Use Case	Rating
Description	The member should be able to rate products in a numeric scale which will improve his/her recommendation.
Actors	The user
Pre-conditions	-.
Post-conditions	The system saves the user's rate.
Main Flow	[Extend: View Products Details] 1. The user selects a rate from 1 to 5 numeric scales, and presses the submit button. 2. The system saves the user's rate, and updates the user's profile.
Alternative flows	If the user is not logged in, the system will redirect the user to the login page step 1 [Include: Login]. If the user presses submit button without selecting a rate, the system will display an error message and offer the user a chance to select a rate. If the user had rated the products previously, the system would update the user's rate.



Use Case	Add to Favourite
Description	The member should be able to add a specific product to their favourite product list.
Actors	The user
Pre-conditions	-
Post-conditions	The system adds the selected product to the user's favourite book list.
Main Flows	[Extend: View Product Details] 1. The user presses the Add to the favourite button. 2. The system adds the product to the user's favourite product list, and updates the user's profile.
Alternative Flows	If the user is not logged in, the system will redirect the user to the Login page, step 1 [Include: Login]. If the user had added the book previously, the system will only update the user's profile.

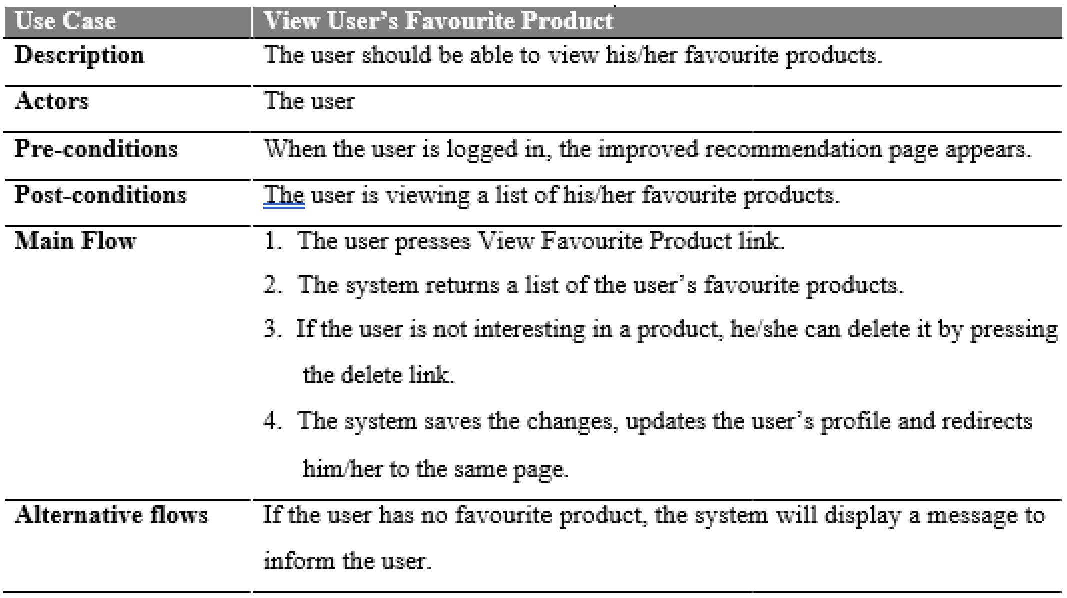


Use Case	Add to Shopping Cart
Description	The member should be able to add a specific product to his/her shopping cart.
Actors	The user
Pre-conditions	-
Post-conditions	The system adds the selected product to the user's shopping cart.
Main Flows	[Extend: View Product Details] 1. The user presses the Add to cart button. 2. The system adds the product to the user's shopping cart, and updates the user's profile.
Alternative Flows	If the user is not logged in, the system will redirect the user to the login page, step 1 [Include: Login]. If the product is already in the user's shopping cart, the system will increment the product quantity by one.

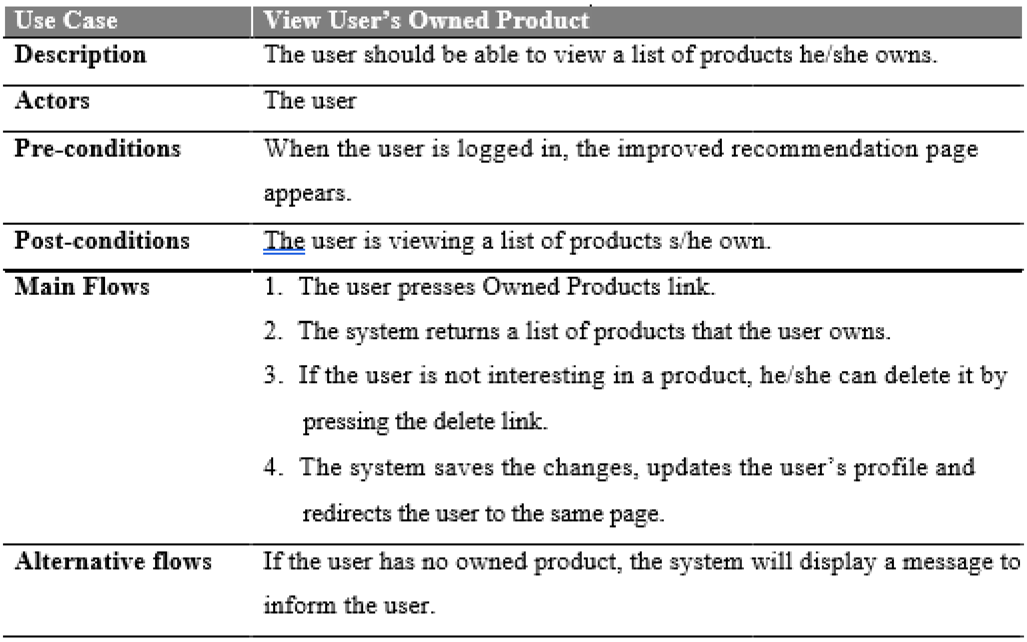
Use Case	Add to Owned Product
Description	The member should encourage marking any product, that she/he owns, in order to improve its recommendations.
Actors	The user
Pre-conditions	-
Post-conditions	The system adds the marked product to the user's owned book list.
Main Flow	[Extend: View Product Details] 1. The user marks *I owned this product* check box. 2. The system adds the product to the user's owned book list, and updates the user's profile.
Alternative flows	If the user is not logged in, the system will redirect the user to the login page, step 1 [Include: Login]. If the product is already in the user's owned product list, the system will only update the user's profile.

Improve Recommendation Package:

Use Case	View Rated Product
Description	The user should be able to view his/her rated products to improve his/her recommendations.
Actors	The user
Pre-conditions	When the user is logged in, the improved recommendation page will appear.
Post-conditions	The user is on his Rated product page viewing his rated books.
Main Flow	1. The user presses View Rated Product link. 2. The system returns a list of rated products selected by the user. 3. [optional] the user selects a product to delete it by pressing the Delete link (no more interesting in this product].[recursive] 4. The system saves the changes, updates the user's profile and redirects the user to the same page.
Alternative flows	If the user has no rated product, the system will display a message to inform the user.



Use Case	View User’s Favourite Product
Description	The user should be able to view his/her favourite products.
Actors	The user
Pre-conditions	When the user is logged in, the improved recommendation page appears.
Post-conditions	The user is viewing a list of his/her favourite products.
Main Flow	1. The user presses View Favourite Product link. 2. The system returns a list of the user’s favourite products. 3. If the user is not interesting in a product, he/she can delete it by pressing the delete link. 4. The system saves the changes, updates the user’s profile and redirects him/her to the same page.
Alternative flows	If the user has no favourite product, the system will display a message to inform the user.

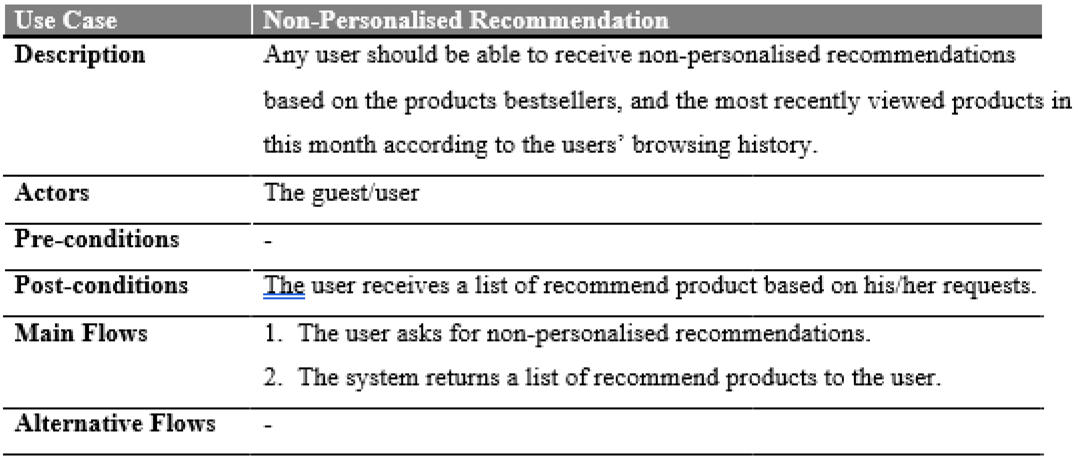


Use Case	View User’s Owned Product
Description	The user should be able to view a list of products he/she owns.
Actors	The user
Pre-conditions	When the user is logged in, the improved recommendation page appears.
Post-conditions	The user is viewing a list of products s/he own.
Main Flows	1. The user presses Owned Products link. 2. The system returns a list of products that the user owns. 3. If the user is not interesting in a product, he/she can delete it by pressing the delete link. 4. The system saves the changes, updates the user’s profile and redirects the user to the same page.
Alternative flows	If the user has no owned product, the system will display a message to inform the user.

Use Case	View User's Browsing History
Description	The user should be able to view his/her browsing history in order to improve his/her recommendation.
Actors	The user
Pre-conditions	When the user is logged in, the improved recommendation page appears.
Post-conditions	The user is viewing a list of his recently browsed books.
Main Flows	1. The user presses Browsing History link. 2. The system returns a list of the recently viewed products by the user. 3. If the user is not interesting in a product, he/she can delete it by pressing the delete link. 4. The system saves the changes, updates the user's profile and redirects him/her to the same page.
Alternative Flows	If the user has no product in their browsing history; the system will display a message to inform the user.

Recommendation Module Package:

Use Case	Personalised Recommendation
Description	The users should be able to ask for personalised recommendations based on their profiles and/or similar users' preferences.These recommendations must be produced using different algorithms in order to evaluate them during the project experiment.
Actors	The user
Pre-conditions	The user is logged in.
Post-conditions	The user receives a list of recommend products based on his/her requests.
Main Flows	1. The user asks for recommendations based on a certain approach (content based, collaborative filtering approach using MSD for similarity measurement, collaborative filtering approach using Pearson correlation for similarity measurement, or the hybrid approach). 2. The system returns a list of recommend products to the user.
Alternative Flows	The user has no recommendations because no similar users are founded and/or the user profile does not contain enough data about the user interest. The system will display a message to inform the user and ask him/her to evaluate more products in order to receive recommendations.



Use Case	Non-Personalised Recommendation
Description	Any user should be able to receive non-personalised recommendations based on the products bestsellers, and the most recently viewed products in this month according to the users' browsing history.
Actors	The guest/user
Pre-conditions	-
Post-conditions	The user receives a list of recommend product based on his/her requests.
Main Flows	1. The user asks for non-personalised recommendations. 2. The system returns a list of recommend products to the user.
Alternative Flows	-

APPENDIX E: Activity Diagrams

Account:

Register

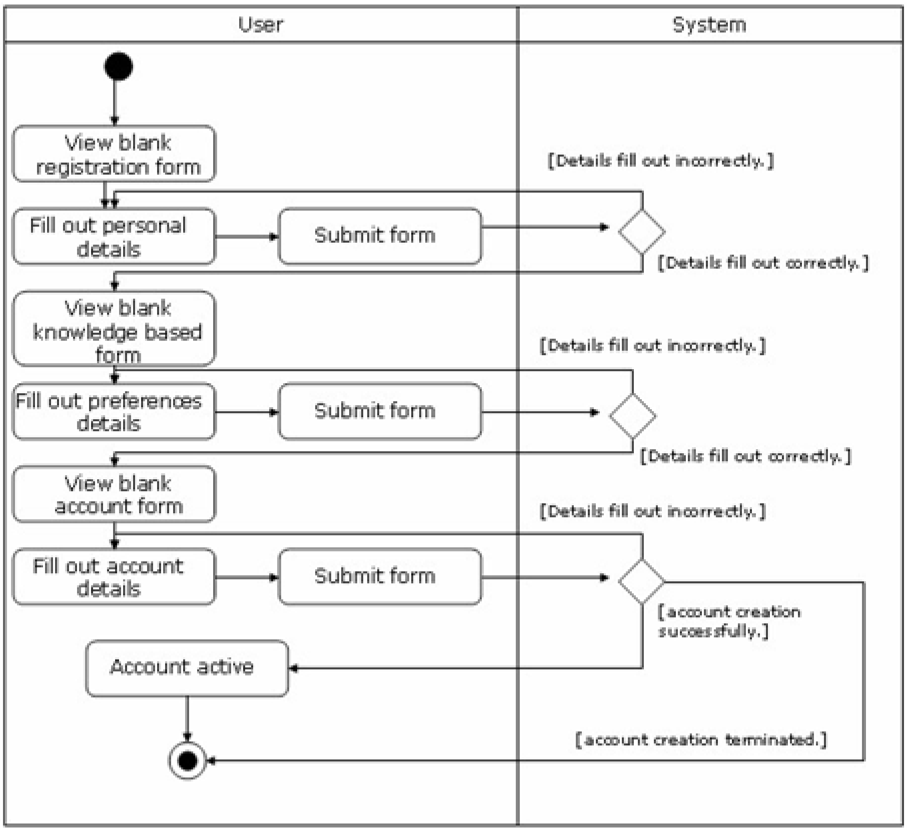

User
System
View blank registration form
[Details fill out incorrectly.]
Fill out personal details
Submit form
[Details fill out correctly.]
View blank knowledge based form
[Details fill out incorrectly.]
Fill out preferences details
Submit form
[Details fill out correctly.]
View blank account form
[Details fill out incorrectly.]
Fill out account details
Submit form
[account creation successfully.]
Account active
[account creation terminated.]

Fig E.1: Activity Diagram to Model the User Registration Process

Login

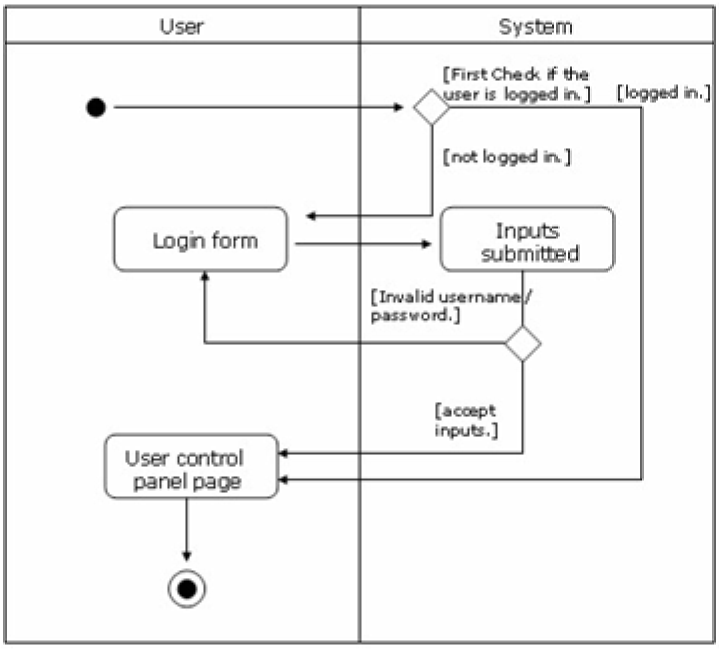


Fig E.2: Activity Diagram to Model the User Registration Process

Search

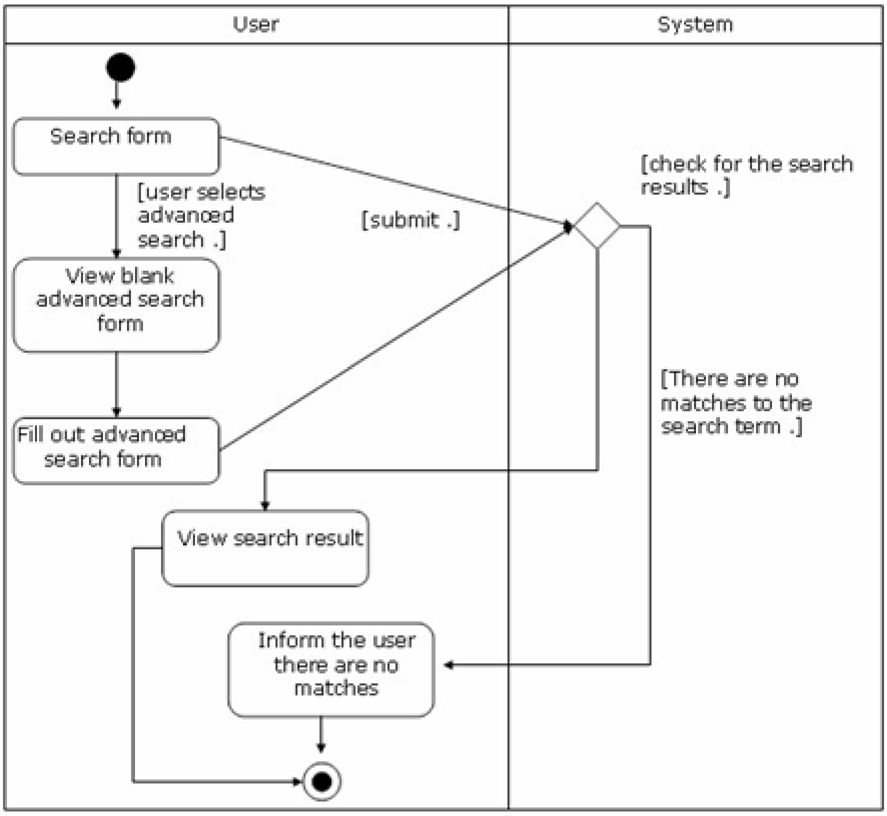
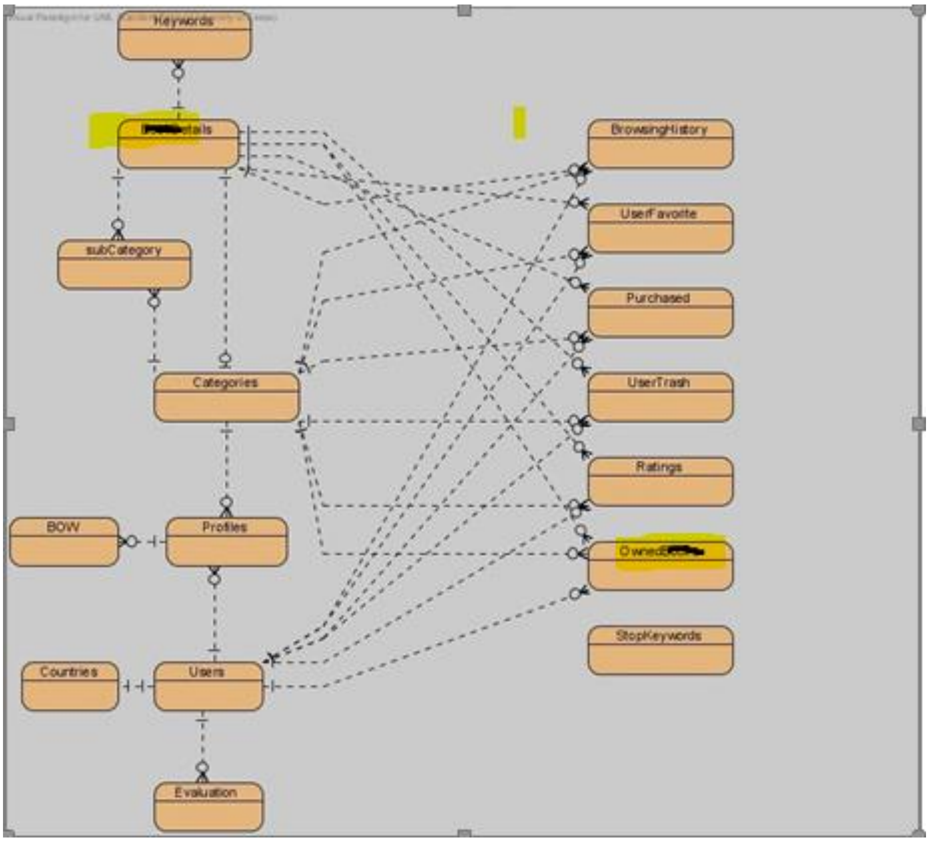

User
System
Search form
[user selects advanced search .]
[submit .]
[check for the search results .]
View blank advanced search form
Fill out advanced search form
[There are no matches to the search term .]
View search result
Inform the user there are no matches

Fig E.3: Activity Diagram – Search

APPENDIX F: Entity-Relationship (ER) Diagram



APPENDIX G: Entity-Relationship (ER) Diagram

Table G.1: Database Table - Users

Table Name: Users	
Attributes	Data Type
UserID (PK)	INT AUTO_INCREMENT
userTitle	VARCHAR(10)
name	VARCHAR(50)
Country	VARCHAR(50)
UserAddress	VARCHAR(50)
Phone	VARCHAR(50)
email	VARCHAR(50)
UserName	VARCHAR(50)

Table G.2: Database Table - ProductDetails

Table Name: ProductDetails	
Attributes	**Data Type**
ProductId (PK)	INT AUTO_INCREMENT
title	VARCHAR(50)
ServiceProvider	VARCHAR(50)
year	INT
ISBN	INT
pages	INT
image	VARCHAR(MAX)
imageLarge	VARCHAR(MAX)
category	VARCHAR(50)
language	VARCHAR(50)
price	MONEY
quantity	INT
description	VARCHAR(MAX)
categoryId(FK)	INT
subCategoryId(FK)	INT
subCategoryId2(FK)	INT
subCategoryId3(FK)	INT

Table G.3: Database Table - Categories

Table Name: Categories	
Attributes	**Data Type**
categoryID (PK)	INT AUTO_INCREMENT
category	VARCHAR(50)
Table Name: SubCategory	
Attributes	**Data Type**
subCategoryId(PK)	INT
subCategory	VARCHAR(50)
categoryId(FK)	INT

Table G.4: Database Table -Ratings

Table Name: Ratings	
Attributes	**Data Type**
userID(PK)	INT
productsID(PK)	INT
categoryID(PK)	INT
Rate	SMALLINT
RateTime	DateTime

Table G.5: Database Table – UserFavourite

Table Name: UserFavourite	
Attributes	**Data Type**
userID (PK)	INT
ProductID (PK)	INT
categoryID(PK)	INT

Table G.6: Database Table - BrowsingHistory

Table Name: BrowsingHistory	
Attributes	**Data Type**
userID (PK)	INT
productID (PK)	INT
categoryId(PK)	INT
visitedDate	DateTime

Table G.7: Database Table – Purchased

Table Name: Purchased	
Attributes	**Data Type**
userID (PK)	INT
ProductID (PK)	INT
categoryId(PK)	INT

Table G.8: Database Table - Profiles

Table Name: Profiles	
Attributes	**Data Type**
userID (PK)	INT
categoryId (PK)	INT
frequent	INT
BOWID(FK)	INT

Table G.9: Database Table - Keywords

Table Name: Keywords	
Attributes	**Data Type**
ProductID (PK)	INT
word (PK)	VARCHAR(50)
frequent	INT

APPENDIX H

Telecom Recommender System Usability Survey

Name : …………………………………………………….

Designation: ………………………………………………

Sex: □**Male** □ **Female**

Company: ………..

Please rate the usability of the system.

- Try to respond to all the items.
- For items that are not applicable, use: **NA**

SA=strongly agree, A=agree N=Neutral, D=disagree, SD=strongly disagree	SA	A	N	D	SD
a. The items recommended to me matched my interests					
b. This recommender system gave me good suggestions					
c. The recommendations I received better fits my interests than what I may receive from a friend.					
d. Some of the recommended items are familiar to me					
e. The items recommended to me are attractive.					
f. This recommender system helped me discover new products					
g. The items recommended to me are similar to each other					
h. I was only provided with general recommendations (e.g., top rated products), which are the same for anyone					
i. This recommender system explains why the products are recommended to me					
j. The information provided for the recommended items is sufficient for me to make a purchase					
k. This recommender system helped me find the ideal item.					
l. This recommender system influenced my selection of items					
m. Finding an item to buy with the help of this recommender system is easy					
n. I understood why the items were recommended to me.					
o. This recommender system made me more confident about my selection/decision					
p. Overall, I am satisfied with this recommender system					
q. I will use this recommender frequently.					
r. I will tell my friends about this recommender.					
s. I would buy the items recommended, given the opportunity.					

APPENDIX I

Questionnaire

Factors affecting customer loyalty in telecom sector in Zambia

Dear participants, this research is intended to identify the factors that affect the customer loyalty. This questionnaire will neither be shared with anyone nor will be used for any commercial purpose, this is only for the purpose of academic research report; your survey responses will be kept confidential.

Profile:

Gender: □ Male □Female

Age: □ Below 18 □18-25 □25-33 □ Above 33

Marital Status: □ Single □Married

Income: □ below 20,000 □ 20,000-40,000 □ 40,000 -60,000 □ Above 60,000

Profession □Student □Salaried □ Self – Employed □ other ____________

1. Which mobile company connection you have subscribed?
   a) MTN
   b) ZAMTEL
   c) AIRTEL
   d) VODAFONE

2. Referring question # 1, did the product purchased from above Company satisfy you?
   a) Yes
   b) No

3. Which services are more helpful to you while using above company services?
   a) Call rates
   b) SMS service
   c) Quality Network
   d) Value Added Services
   e) DATA

4. In total, how long have you been a customer of above Company?
   a) Less than one year

b) One to under three years

c) Three to under five years

d) Five to under ten years

e) Ten years or more

5. Company always provides a proper demonstration on the new products and services?

a) Yes

b) No

6. Have how many times have you switched from your current network to another network operator.

a) once

b) twice

c) More than 3 times

7. I would like to switch from my current network if another operator provides better services.

a) Yes

b) No

Customer Satisfaction

SA=strongly agree, A=agree N=Neutral, D=disagree, SD=strongly disagree	SA	A	N	D	SD
t. I am satisfy with the products and services provided by company.					
u. I would like recommend products of this company to your friends and relatives.					
v. I am properly satisfied after sale service from Company?					
w. Company provides the products and services that best fit with my interest.					

Perceived Price

SA=strongly agree, A=agree N=Neutral, D=disagree, SD=strongly disagree	SA	A	N	D	SD
a. When I buy products, I like to be sure that I am getting my money’s worth.					
b. I generally shop around for lower prices on products, but they still must meet quality requirements before I buy them.					
c. The price are reasonable and affordable from my services provider					
d. I prefer to pay more if quality of product and services worth it.					

Services Quality

SA=strongly agree, A=agree N=Neutral, D=disagree, SD=strongly disagree	SA	A	N	D	SD
a. Quality of services worth more than the price the company charges.					
b. Services provided by company create superiority feelings in me.					
c. Company always keeps improving the quality of services.					
d. I never compromise on the quality of service provided by the operator.					

Trust

SA=strongly agree, A=agree N=Neutral, D=disagree, SD=strongly disagree	SA	A	N	D	SD
a. The operator provides timely information when there are new services.					
b. Company provides true information to the customer.					
c. Company develops an encouraging attitude toward using the products.					
d. The description of products and services is reliable					

Please rate your level of agreement with the following statements (7-1 scale with 7 being completely agree, 4 being neutral, and 1 being completely disagree):

	1	2	3	4	5	6	7
-I believe Company deserves my loyalty	□	□	□	□	□	□	□
-Over the past year, my loyalty to Company has grown stronger	□	□	□	□	□	□	□
-Company values people and relationships ahead of short-term goals.	□	□	□	□	□	□	□
-Customer Care gives valuable information to customer.	□	□	□	□	□	□	□

LIST OF PUBLICATIONS

1. Recommender System for Telecommunication Industries: A Case of Zambia Telecoms.