UNIVERSITY OF CANBERRA

AUSTRALIA'S CAPITAL UNIVERSITY

A Novel Data Mining and Knowledge Discovery Framework for Digital Library Recommendations System based on User's Feedback and Personalization.

Maram Abdulrahman Almaghrabi

April 2021.

Abstract

The research field of digital libraries involves a combination of different subfields from a wide variety of domains, and the usability and popularity of digital libraries depend on the satisfaction and user experiences of people while engaging with digital library systems. The digital library collections are nothing but large-scale resource search and retrieval systems, that are increasingly playing an important role in our knowledge society. They can be defined as catalogued and curated collections of different types of information resources, that are stored in a digital format in digital places. They can be viewed as online resource hubs, where it is possible to access information anytime and anywhere. It is important that every digital library needs to be equipped with a wide set of tools, assisting the users in their resource search and retrieval activities, and enhance their user-experience with these activities, leading to better usability of the system, and help users in finding relevant information and getting tailored information based on the personalized feedback, feelings and sentiments, ratings, preferences, and metadata build within the system. One of the tools that facilitates such a wide set of tailored and personalisation option, is called the *recommender system*, synonymously called the *recommendation system*, as well. The digital library systems equipped with recommendation system tools aim to provide personalisation features and assist users in finding relevant information based on their preferences, based on their previous choices, maintained in their respective profiles, and other similar users with either similar profiles or usage behaviour and pattern. The success of the recommendation process and tailored retrieval performance of the recommender system depends on the quality and quantity of the information used about the items and users, and their profiles stored from each interaction since the profiles represent the information needs of the users from the user-item interactions. To identify the needs of an individual user, and provide better personalisation, the design of the recommender system

needs to be based on models that maintain detailed item descriptions and accurate users' profiles, their preferences, and details of item-user interactions, and the performance and robustness of the entire recommendation system rely on the modelling accuracy on these factors. Due to massive information available about items and resources, often users are faced with ambiguity and diversity of information, due to the enormous growth of the Internet, with increasingly complex and massive volumes of the information, leading to difficulties in isolating the content that fits their needs. Due to this, users are not always certain of their information needs, nor do they know exactly how to describe what they want. The most challenging task involved in building tailored and personalized recommendations is acquiring information on user needs and their preferences, when there is limited information about users, particularly when the users are new, or do not engage with the system often. These problems make it difficult to profile users accurately and provide quality recommendations. Traditional approaches to address this issue in the research literature has been based on the concept of building taxonomical representation of the item and model the interaction information about each user and the item, and the relevance of each item to users and several other users of similar type. Such item representations or a taxonomical model can assist in determining users' preferences. These models involve a hierarchical structure, comprising the coarse-grained representation to the fine-grained representations, for symbolizing the set of categories or topics, items, and users. However, due to the complex relationship between concepts, items, and users, from massive data sources, these taxonomic representation models are limited in their capabilities, and identification of each user's information need is still a challenging task, particularly, when the user interaction with the items is limited, or the user happens to be new to the system. This thesis attempts to address some of these challenges, and proposes a novel computational framework based on deep machine learning techniques for building recommendation systems for digital libraries (modelled similar to a large-scale resource search

and retrieval system), by including user feedbacks, preferences, explicit ratings, sentiments, and metadata information. In particular, the two deep learning models proposed in this thesis, the Multilayer Convolutional Neural Networks (MCNN) model and Collaborative Filtering based Multistage Deep Neural Network architecture (CFMDNN) model, allow better interaction and collaboration events to be captured between users' feedbacks, explicit ratings, sentiments, and metadata information. The evaluation of the proposed computational framework based on several publicly available datasets shows significant enhancement in the recommendation systems performance, outperforming the related baselines. Further, the extension of these novel deep learning models for a new multilingual context (Arabic /English) digital library systems context, has led to promising results. In summary, this thesis makes significant contributions to recommender systems for digital library systems applications, with models based on multiple components including, explicit ratings, sentiment analysis based on text feedback, and metadata information for both Arabic and English languages contexts.

Acknowledgments

"In the Name of Allah, the Most Beneficent, the Most Merciful. I Thank Allah for granting me the patience, health, guidance and determined to complete my thesis successfully. First of all, I am extremely grateful to Prof Girija Chetty who is my supervisor for her help and guidance. She helps me to make my research contributions more accurate. I can say that she is a very supportive advisor in many aspects. The words cannot give her right as a wonderful supervisor worthy of appreciation. I am also, grateful to my wonderful family, thank you for giving me strength to reach for the stars and chase my dreams, words cannot express how grateful I am, special thanks to my Father who passed away. May Allah have mercy on his soul and grant him Paradise. He was indeed an inspirational man. To my beloved Mother who always support me and pray for me your prayer for me was what sustained me thus far. To the love of my life, my husband thank you for supporting me I cannot thank you enough for encouraging me throughout this journey. To my brothers who always stand up with me and encouraging me thank you all for your continuous support to me. To my baby girl Tamara, this work made for you, I am sure you will be smart girl and we will be proud of you.

Furthermore, I would thank all my friends from faculty of Science & Technology you were amazing and supportive. Also, special thanks to my country Saudi

Arabia, Saudi Arabian Culture Mission, and my University (Umm Al-Qura University) for give me this opportunity and support on my education life.

In addition, my memories of Canberra will be never forgotten. I will always remember these four years I have spent in Australia, especially in Canberra. Finally, I do not deny that this journey was difficult, especially with the Coronavirus Pandemic, but with the grace of Allah and the directions of my supervisor and supporting of my family and my country Saudi Arabia, I was able to complete my degree and achieve my dream.

Publications arising out of this thesis.

M. A. Almaghrabi and G. Chetty, "A Novel Data Mining Testbed for User Centred Modelling and Personalisation of Digital Library Services," 2017 IEEE 13th International Conference on e-Science (e-Science), Auckland, 2017, pp. 434-435.

M. Almaghrabi and G. Chetty, "Impact Rating Scales on Recommender System and Using Deep Learning and Neural Network Models to Improve Rating Prediction," 2018 5th Asia-Pacific World Congress on Computer Science and Engineering (APWC on CSE), Nadi, Fiji, 2018, pp. 207-213.

M. Almaghrabi and G. Chetty, "A Deep Learning Based Collaborative Neural Network Framework for Recommender System," 2018 International Conference on Machine Learning and Data Engineering (iCMLDE), Sydney, Australia, 2018, pp. 121-127.

M. Almaghrabi and G. Chetty, "Investigating Deep Learning Word2vec Model for Sentiment Analysis in Arabic and English languages for User's reviews," *2019 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE)*, Melbourne, Australia, 2019, pp. 1-6, doi: 10.1109/CSDE48274.2019.9162391.

M. Almaghrabi and G. Chetty, " Multilingual Sentiment Recommendation System based on Multilayer Convolutional Neural Networks (MCNN) and Collaborative Filtering based Multistage Deep Neural Network Models (CFMDNN).," *2020 IEEE/ACS 17th International Conference on Computer Systems and Applications (AICCSA),* Antalya, Turkey.

M. Almaghrabi and G. Chetty, " Improving Sentiment Analysis in Arabic and English Languages by Using Multi-Layer Perceptron Model (MLP)," *2020 IEEE International Conference on Data Science and Advanced Analytics (DSAA),* Sydney, Australia.

M. Almaghrabi and G. Chetty, " Deep Machine Learning Digital library recommendation system based on Metadata for Arabic and English Languages" *2020 7th IEEE CSDE 2020, the Asia-Pacific Conference on Computer Science and Data Engineering,* Gold Cost, Australia.

Table of Contents

Table of Contents   xiii

List of Figures.

List of Figures

List of Figures   xix

List of Tables

i. *Abbreviations*

Sentiment Analysis	SA
Digital Libraries	DLs
Resource Description and Access	RDA
Joint Steering Committee	JSC
Anglo American Cataloguing Rule Second Edition	AACR2
Functional Requirements for Authority Data	FRAD
Functional Requirements for Bibliographic Records	FRBR
Open Archive Initiative	OAI
Neural Networks	NNs
Machine Readable Cataloguing Record	MARC
Collaborative Filtering Based Multistage Deep Neural Network architecture	CFMDNN
Bag of Word	BOW
Collaborative Filtering	CF
Natural Language Processing	NLP
Australian National Bibliographic Database	ANBD
Automatic Speech Recognition	ASR
Language Modelling	LM
Convolutional Neural Network	CNN
Deep Belief Network	DBN
Support Vector Machines	SVM
Digital Subscriber Line	DSL

Automatic Discretization	AD
Recurrent Neural Network	RNN
Frequent Pattern Mining Approach	FP-Growth
Online Public Access Catalogue	OPAC
Receiver Operating Characteristic	ROC
Mean Absolut Error	MAE
Generalized Matrix Factorisation	GMF
Collaborative Filtering Deep Recommender architecture	CFDR
Continuous Bag of Words	CBOW
Recursive Neural Networks	ReCNN
Artificial Intelligence	AI
Feed Forward Networks	FFNs
Long Short-Term Memory	LSTM
Memory Network	MemNN
Multi-layer Perceptron	MLP
Multistage Convolutional Neural Network	MCNN
Root Mean Square Error	RMSE
Saudi Digital Library	SDL
Content –Based Recommendation	CBR
Open Archives Initiative Protocol for Metadata Harvesting.	OAI-PMH
Ministry of Higher education	MOHE

Chapter 1 Introduction

Over the last few decades, digital library systems have been continuously evolving. With the rapid advancements in information and communication technologies, the range and quality of services offered by digital libraries have improving significantly, with better user experience, improved personalized recommendations, and sensitivity to user preferences. These improvements are due to efficient data mining and machine learning techniques continuously analysing the data stored in the digital library repositories and warehouses and providing different user-oriented services based on the *search history*, *usage patterns*, and *users' profiles*. Having a computational workflow/pipeline of identifying users based on their search history, usage patterns and user profiles, grouping them into cohorts/categories, and recommending appropriate resources/services lead to better personalisation, and improved user experience and satisfaction.

Digital Libraries (DLs) are usually defined as the information collections that provide similar services to users through several technologies (Callan et al., 2003). Similar information collections can be found in personal data, business data or online data, and could be in text, images, or video form. The digital library users can access these collections by using a computer that is connected to a network. Some of the previous research attempts in improving the DL services involved, scientists exploring different data mining techniques on DL data to get a better understanding of how the users interact with the DL systems (Witten and Frank, 2002). With the implementation of Resource Description and Access (RDA) cataloguing standards by several digital libraries worldwide, it has become relatively easier to implement in-depth analytics and knowledge discovery from DL data stores, in terms of the users, their DL resource usage patterns, their borrowing behaviours, and their profiles. DL resource description based on RDA cataloguing standards provides several advantages, including, better responsiveness, improved user acceptance, costing efficiency, more flexibility, and seamless continuity. Due to their significant benefits, several DL systems with similar architectures got deployed widely in several countries. The Joint Steering Committee (JSC) for the Development of RDA devised Resource Description and Access is a replacement for the Anglo-American Cataloguing Rules Second Edition (AACR2), the library domain's widely used content standard (Oliver, 2010). The RDA's structure is based on the conceptual models of the Functional Requirements for Bibliographic Records (FRBR) and the Functional Requirements for Authority Data (FRAD) (Hart, 2010). The RDA structure is also supported by the Open Archives Initiative Protocol for

Metadata Harvesting (OAI-PMH). OAI is also supported by the Digital Library Federation, the Coalition for networked information and the National Science Foundation. The OAI-PMH protocol from OAI is more convenient to apply in a digital environment and has been adopted widely in the development of Digital Library environments. For example, the first Saudi Digital Library (SDL) based on this protocol standards was established in 2010, with a primary aim to support the educational process and meet the needs and requirements of modern students, researchers, and professionals etc. Currently, it has over 24,000 e-books on various scientific topics and has subscribed to nearly 300 International, local, and regional publishers (MOHE, 2010; Saudi Digital Library, 2011). The development of digital libraries in Australia began in the 2000s (Velasquez and Campbell-Meier, 2015). The emergence of DLs in several nations has instilled great promise in the economic, social, technical, cultural and scholarly spheres, due to their potential to provide substantial advantages in building knowledge communities.

However, though the digital libraries have been around for several years, both in developing and developed nations, their uptake in terms of increased user's engagement with the content and services offered by digital libraries has been minimal. They turned out to have generated more hype, along with poor user experience, rather than facilitate transformation into knowledge communities (Lynch, 2012). This could be due to the issues associated with metadata captured in digital library systems. Similar issues exist in E-commerce application domains, such as online shopping and purchasing, but the user engagement with online e-commerce systems has not reduced much, as compared to digital libraries. This could be due to the perception that personalisation of digital libraries with use of tools such as recommenders comes at the cost of the metadata getting captured and personal private information getting compromised. This might be true, since most of the digital library systems design have been based on RDA cataloguing standards, which allows mining of the metadata, containing user profiles, user identities and browsing behaviours, without any personalisation because users may not be aware of their preferences being stored, or the user might be new with no browsing history. What is ideally needed here, is giving users the right products based on what they need with high accuracy by using data mining and deep learning techniques for a personalized recommendation. The data mining and knowledge discovery framework proposed in this thesis tries to address some of these short comings of current digital library systems, based on advances in data mining and machine learning technologies to build personalized digital library recommender systems.

1.1 Thesis Context and Motivation.

This thesis presents research towards the development of a data mining and knowledge discovery framework for digital library recommendation systems, based on novel approaches involving cutting edge technologies from multiple domains, including Digital Libraries (DLs), recommendation systems, data mining and knowledge discovery, deep learning, sentiment analysis, text processing, resource description and access (RDA) and metadata harvesting methods. Digital Libraries (DLs) which are defined as collections of information a provide better user experience if equipped with new technologies that identify the user preferences, and these associated services can be delivered to user communities, using a variety of technologies (Callan et al., 2003). One of the most significant components of DLs is the circulation system. A circulation system or library lending services is a component which enables users to borrow books or any types of media by requesting items and get approval from the library. In addition, another important tool which helps in enhancing user experiences (following on with other online services such as eCommerce/online shopping), is the recommender component in the digital library systems. The recommender service components are defined as a part of information retrieval pipeline, facilitating additional information for users that can be of equal or more interest (McGinty and Smyth, 2006).

A recommender component in digital library systems is based on comparing users' profile with similar interests by using metadata, user' feedback and ratings, and provide information/displays on additional items of interest. Recommender components provide user personalisation, and can improve customer satisfaction, and hence have become one of the essential components of online web-based services and included in digital library systems as well. They can be defined as a part of the technology component, that allows tailoring the content and presentation of a web-based application for everyone according to his/her preferences and characteristics (Perkowitz and Etzioni, 1999; 2000). The information needed for personalisation is stored in a user model. While the personalized recommendations services can improve user satisfaction, it comes at a cost. Hence it is important to develop automated frameworks for digital library recommendation services that helps users to get what they need. The research proposed in this thesis makes contributions towards developing a recommendations framework based on novel data mining and machine learning techniques targeted for digital libraries use case. More specifically, this thesis aims to study to which extent data mining can extract user 'preferences from metadata, uses' rating and user' feedback and recommend items to the user.

Rest of this Chapter is organised as follows. Next Section states the problem definition. Section 1.3 presents the thesis objectives, followed by Section 1.4 which presents the description of the contributions of this thesis. Subsequently, section 1.5 presents the research questions. Section 1.6 describes the datasets used in this thesis, including the sample dataset developed for the digital library system use-case studied in this thesis, and other publicly available datasets used. Next, Section 1.7 presents the problem definition describing the challenges associated with building digital library recommender systems, and the complexity associated with several aspects for enhancing the recommendations, including sentiment analysis, ratings, personalization and metadata information. Finally, Section 1.8 present the overall structure of the thesis and contributions made in this work.

1.2 Problem Definition.

Data mining term was introduced in the mid-1990s in the United States of America and involved combining the terms from statistics and information technologies, such as databases and artificial intelligence. Data mining is defined as process which finds useful information from large amount of data by extracting some salient features from it (Bharati and Ramageri, 2010). However, Digital libraries are considered more complex systems as compared to traditional libraries. The research field of digital libraries is generally referred to as a union of different subfields from a variety of domains. The domains are combined with new research issues and aim at realizing their full potential (Nürnberg et al., 1995). Chen and Liu (2001) stated that digital libraries present a set of complex and complicated research issues. Further, Chen (2003) found that digital libraries experience the issues of user centred aspects and interface/presentation of information to different types of users. Many users tend to feel the inappropriate presentation of information and poor interfaces in digital library systems, causes a lot of problems related to the information overload (Kim et al., 2003). Therefore, there is a need for better information presentation to different users, and this is possible with appropriate analytics driven information extraction and presentation. Data mining and machine learning techniques can be promising as it can allow analytics to be done on the large data, extract relevant and useful information and provide a customised information to the users in digital libraries (Witten and Frank, 2002; Costabile and other authors, 1999; Han et al., 2012). In contrast with traditional libraries, DLs can make information directly available and easy to find. This is due to the Resource Description and Access (RDA) cataloguing standards used in digital libraries, that tend to tag the metadata and extract names and entities linking the users and resources from previous borrowings records.

Witten and Frank show that the data mining techniques have much to offer to improve the quality and types of services offered with digital libraries to their users (2002). In their study, they have recommended Greenstone open-source toolkit as an efficient tool to build digital library system prototypes. The study of data mining: a competitive tool to the digital library (Lone and Khan, 2014) reports the benefits of using data mining and analytics capabilities in digital libraries towards improving the libraries' services such as analytics related to client information, weblogs, circulation, and information. Also, data mining tools can help manage data and allow better visualisation, in terms of reports generation and explanation/justification, and the potential for better information presentation and visualisation customised for different users (Lone and Khan, 2014). Data mining tools in digital library systems can allow better user segmentation or user cohort segmentation based on the user profiles, their interests, and their borrowing records/behaviours. This can help in the provision of better recommender services in digital libraries, as this will provide personalisation of resources being offered by digital libraries. This is achieved by casting the specific recommendation system problem as a classification or clustering problem, and using different approaches based on data mining and machine learning technologies for solving this problem. By casting it as a classification problem, it is possible to classify or cluster different users, their profiles and their interests at individual levels and cohort levels and enhance the quality of recommendation services being provided by Digital Library systems. The need for better recommender systems in digital libraries has been felt, since there is information overload on the web (Mnih and Salakhutdinov, 2008), and the excess data has complicated retrieval of information for the users or even making choice with available data difficult, particularly for real time retrieval scenarios. The availability of efficient and accurate recommender systems can be beneficial in handling information overload and giving recommendations to the users according to their needs. The systems can therefore, facilitate the search for content that cannot be easily retrieved, and help enhance the user experience, if available in real time. Researchers have, thus, used the recommender systems as a preferable choice for obtaining information (Tang and Wang, 2018). Digital library recommendation systems are based on the retrieval techniques which utilise the algorithms for screening and sorting out the required information.

1.3 Thesis Objectives.

The objectives of this thesis are:

- To investigate the extent of integration of data mining techniques for improving digital library borrowing system services and capture user's preferences by including recommender components in their services.

- To investigate the power of open-source tools such as Greenstone/Librarika for implementing digital library system prototypes, and examine issues that cannot be investigated in live and real digital libraries deployed, including issues corresponding to multilingual contexts (Saudi Arabia/Australia).

- To investigate whether the inclusion of RDA and metadata information can allow improved recommendations to be provided, both in terms of quality and accuracy of recommendations being provided.

- To investigate the potential of advances in machine learning techniques, including deep learning approaches, deep learning techniques, for providing better recommendations.

- To investigate the potential of sentiment analysis techniques, which helps the users to provide their feedback and sentiments expressed as text comments, for enhancing the quality and accuracy of recommendations in multilingual contexts (English and Arabic language for example).

1.4 Thesis Contributions.

This thesis presents an interdisciplinary study and makes contributions to multiple disciplines connected with digital library systems. A novel data mining and knowledge discovery framework is proposed in this thesis, for developing an integrated digital library recommendation system by applying new deep learning and data mining techniques based on including user feedback, involving sentiment analysis, ratings and metadata for Arabic and English languages. The experimental evaluation of the proposed computational framework was done with several publicly available datasets, as well as validated on a prototype use case scenario, developed in-house, comprising of a users' borrowing records dataset collected from Librarika, an open-source library management system. In addition, we have used Saudi digital

library (SDL) dataset by extracting data from their dataset by using data mining techniques to collect resources metadata for both languages English and Arabic from SDL. By applying data mining and deep learning algorithms to find the relationships between items in user's profile, an improved recommendation system was developed based on the user preferences and feedbacks. The system intends to aid in the extraction of complex features from data of high dimension and use them in the building of models relative to the inputs and outputs in the system (Papernot et al., 2016). The capture of user opinions in both languages Arabic and English from user feedbacks by using different machine learning classifiers a multi-lingual recommender system with enhanced performance and robustness.

1.5 Research Questions.

To address the research aims and objectives following research questions have been formulated:

1. What are the traditional approaches to providing recommendations in patterns like digital library systems services?
2. How can we provide better recommendation services by using advances in data mining and machine learning algorithms, including the recent deep learning algorithms, by linking the borrower information and resources information?
3. Are there enough resources, approaches, corpora, sentiment analysis datasets and neural network approaches for providing recommendations in multilingual systems context, particularly in English and Arabic multilingual contexts?

1.6 Thesis Datasets.

In this thesis, we used different datasets from different sub-domains, as large digital library systems consist of resource collections of multiple different types including, the books, audio, music, food, hotel and video collections. All datasets used are available publicly, except the Saudi Digital Library (SDL) datasets we extracted from the Saudi digital library database after their permission. The details of each dataset are presented in experimental work done for investigating each research question in the other Chapter of this thesis. The datasets used are:

1. Librarika dataset.
2. MovieLeans20M dataset.
3. Amazon Book dataset.
4. Book-Crossing dataset.
5. Food Fine Amazon dataset.

6.   Arabic Movie dataset.

7.   Arabic Hotel dataset.

8.   Amazon digital Music dataset.

9.   Booking Hotel dataset.

10.   Saudi Digital Library (SDL) datasets in Arabic and English Languages.

1.7 Challenges in building digital library recommender system based on sentiment analysis, ratings, personalization, and user history.

Though there are several data mining and deep learning-based recommendation system approaches proposed for digital libraries and other similar systems, there are still several challenges and issues in terms of quality and accuracy of recommendations being provided, and this could be due to isolated and disparate strategies used for mining the data and extracting useful and personalized recommendations.

Some of these include:

- Presentation and Visualisation issues, leading to users unable to manage huge amount of data and that makes search experiences in DLs and other similar systems searching more difficult for users.

- Multilingual context management to address the needs of the international community. For instance, none of the previous work we have surveyed, shown the options of multilingual recommender services, such as in the context of digital libraries and other similar systems in Saudi Arabia and Australia.

- The quality and quantity of recommendations provided by the current systems need significant improvement.

- The privacy issues involved in maintaining user records in digital library systems. For example, the user records such as in the University of Canberra library in Australia does not keep user records as that is in American system.

- The need for a large size of high-quality data sets for building better recommender systems based on new data mining and machine learning techniques (such as deep learning techniques). Deep learning techniques require large datasets with rich user-item interaction information to provide good recommendations to the users.

- Most of the Machine-Readable Cataloguing record (MARC) or biographic record of digital libraries everywhere, do not give user/borrower relevant documents.
- Poor user interfaces with lack of interactivity, with no options for enabling users to put their preferences.
- Many libraries have databases from different disciplines, but no one know about it, which cost libraries lot of money, without much return on investment.
- Insufficient datasets available in multilingual contexts, especially in Arabic language context.

1.8 Thesis Structure

The thesis is organised as follows. *Chapter 2* is a comprehensive review of existing research that is relevant in recommender systems, digital libraries, deep learning, and sentiment analysis in both English and Arabic language context. It identifies and justifies the research context and the problems from which the research questions were derived. This chapter also explains the role of sentiment analysis approaches for building better recommendation systems. Finally, it gives the methodologies that have been used to build a recommendation system in digital libraries for both languages.

Next, *Chapter 3* presents the basic proof of concept prototype for the recommendation systems based on traditional data mining algorithms for a digital library use case scenarios. It presents details of digital library test bed developed using Librarika open-source tools for investigating different traditional data mining approaches for user/cohort segmentation.

Chapter 4 presents the novel computation framework developed in this thesis based on deep machine learning and the experimental evaluation of the algorithmic approaches developed in the framework for several publicly available datasets. The algorithms based on Collaborative Filtering (CF) and Collaborative Filtering based Multistage Deep Neural Network architecture (CFMDNN), with different rating scales have shown that CFMDNN algorithms results in good performance.

Chapter 5 investigates sentiment analysis of users' feedback involving text comments, based on different NLP text processing techniques including the Word2Vec model and bag of word (BOW) model.

Chapter 6 presents deep learning approaches based on natural neural networks for Arabic sentiment analysis by using SVM, MLP, CNN and MCNN models and investigates the performance of digital library recommendation system for multilingual contexts, comprising Arabic and English language contexts.

Chapter 7 presents digital library recommendation system approach based on Metadata and hybrid deep learning techniques, and improvements in recommendation system performance achieved.

Chapter 8 summarises the results and compares the improvements achieved with different data mining and deep learning approaches for different types of digital library services considered in this thesis (user segmentation and resource recommendation services). Also, the Chapter summarizes, how the proposed computational framework based on novel data mining and deep leaning techniques, with the inclusion of user preferences in terms of explicit ratings, sentiments with textual comments, and metadata information can enhance the quality and accuracy of recommendations in multilingual digital library system application scenarios. Finally, the Chapter provides directions for future research based on the findings from this thesis and concludes with references and bibliography related to this work.

1.9 Chapter Summary.

This chapter presents an introductory information about the thesis topic. It gives you information and an explanation of topic and road map of the thesis. The chapter ends with research objectives, problem definition, research questions this thesis tries to address, and contributions from this work. The next chapter gives background information and a brief literature review on digital libraries, RDA, data mining algorithms, recommendation system, and the challenges associated with building a recommendation system for digital library application scenarios. It shows the application of the recommendation system and sentiment analysis and gives the methodologies that used to build a recommendation system in digital libraries for both languages Arabic and English. Finally, applies algorithms with metadata to build a digital library recommendation system

Chapter 2 Literature review and Background

2.1 Introduction.

After giving an introduction of the thesis, this chapter describes the background and the basic concepts of data mining technology, and the state of the art in the field for building recommendation systems in the digital library context. The definition of a recommender system has been improved over the past 14 years. In Resnick and Varian's seminal article, the authors describe a recommender system as follows:

"In a typical recommender system people provide recommendations as inputs, which the system then aggregates and directs to appropriate recipients" (Resnick and Varian, 1997).

To be precise, a recommendation system involves modelling of human computer interaction, by embedding user's preferences, finding the relationship between users and items, and understanding their needs. This chapter will also provide a review of the literature relevant to the field of study, the line of investigation and the research methodology used. The aim is to review and analyse the current literature, which is considered as one of the most powerful tools in the hand of researchers, helps researchers to focus on the specific field and sets the background rationale. This chapter describes the digital libraries field, data mining algorithms and RDA on recommendation system. Also, is going to explain sentiment analysis and deep learning and presents some applications. Followed by methodologies used for English and Arabic languages based on sentiment analysis. Finally, shows the deep learning recommender system in DLs based on metadata. Figure 1 shows the literature review structure used for this thesis:
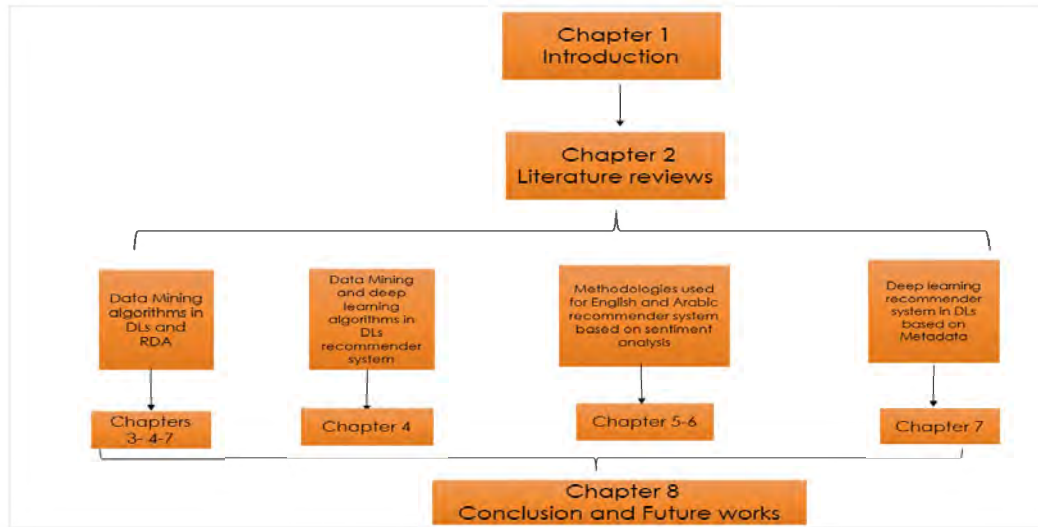
Figure 1: literature reviews structures

Section 2.3 discusses previous studies about data mining and the benefits of RDA in digital libraries and related concepts. There are several data quality issues for information retrieval in digital libraries context, such as missing values. Also, there are differences between end goals of data mining and information retrieval process. The literature review tries to address some of these issues and validates the lack of prior work on studies covering data mining in digital libraries, with a focus on recommendation system, and use of open source data repositories for building a Digital Library (DL) systems.

In this research, a novel computational framework is proposed, based on advances in data mining and machine learning technologies, including deep learning, text processing, sentiment analysis and RDA metadata to build better recommendation systems for Digital Library application scenarios.

Section 2.4 reviews the traditional data mining algorithms used in recommendation system for DLs. Though there are few studies about building a recommendation system for books based on traditional data mining algorithms, there are not many approaches that leverage the benefits of recent advances in data mining, machine learning and associate technologies in text processing, metadata mining and deep learning, for enhancing the performance of recommendation systems. In this thesis, the novel computational framework proposed takes into consideration several of these recent advances. Also, since it is difficult to access the digital library databases, a prototype testbed developed has been developed, based on open source technology tools for assessing the personalization and user experience with different components of digital library services, and impact of apply data mining methodologies on them. Some of the methods include, the association rules, Support degree (Zhou et al., 2014), *Apriori*

algorithm, FP-Growth, classification, Naïve Bayes, Decision Tree and K-NN algorithms. Performance evaluation of each data mining method was done using several performance metrics, including the ROC (Receiver operating characteristic), the Confusion Matrix and the True /False positive rates.

Section 2.5 investigates the literature regarding DLs recommender systems based on deep learning and sentiment analysis. The literature reviews show that all studies used deep learning and sentiment analysis for social media such as Twitter and used it on e- commerce and retail scenarios, such as Amazon and Netflix scenarios. Also, the literature review confirms that most of applications developed using deep learning and sentiment analysis are in the English language and did not include multilingual contexts. This shortcoming has been addressed in this work, by considering Arabic language sentiment analysis.

In this research, the focus of the computational framework proposed for Digital Library recommendations involves the use of explicit ratings, text based feedback for sentiment analysis in multilingual contexts (English and Arabic), for the development of novel deep learning architectures – MCNN and the CFMDNN algorithms, in addition to traditional Word2Vec, RNN, MLP, and SVM algorithms. The experimental evaluation of the proposed computational framework was done with several in-house, private, and publicly available databases, using several performance metrics, including Mean Absolute Error (MAE), RMSE and F-measure.

Sections 2.6, 2.7 and 2.8 review the applications of recommendation system based on sentiment analysis and deep learning techniques. Section 2.9 discusses the methodologies used for English and Arabic sentiment analysis. Section 2.10 discusses the use of sentiment analysis for recommendation systems. Finally, section 2.11 reviews any works on recommender systems for DLs (Digital Libraries) based on metadata. The comprehensive literature review shows a gap in existing approaches, particularly in using a combination of explicit ratings, text-based sentiments, and metadata in multilingual contexts for recommendations in Digital Library application, which this thesis attempts to address. Several data mining techniques to extract metadata from digital library. In this research, the proposed computational framework based on a combination of explicit ratings, text-based sentiments, and metadata in multilingual contexts for recommendations in Digital Library application context, leads to better recommendation system performance, and improved user experience. The next Section reviews the background work on Digital Library Systems.

2.2 Digital Libraries.

The establishment and work of the digital library in Saudi Arabia are discussed in MOHE (Ministry of Higher education). (2010). The digital libraries of Australia have shown increased activity in recent years. Several works are reporting the use of data mining techniques for providing personalized recommendations in digital libraries, including the investigations by Han et al., (2012), Dushay (2002), Kim, Choo, and Chen (2003), Cullen (2005), Kovacevic, Devedzic, and Pocajt (2010), and others.

It is a common saying that "*We are living in the information age*". In reality, we are living in the data age, with Terabytes of data storage in our computer networks every day from information science, medicine and engineering fields. This growth is a result of the computerization of our society. (Han et al., 2012).

Data mining process consists of several steps, including known as knowledge discovery, knowledge mining from data and the knowledge extraction step. The goal of a data mining process is to find the patterns within data, and requires three main steps to discover the knowledge, comprising exploration, pattern identification and deployment (Bharati and Ramageri, 2010), as shown in Figure 2 below:
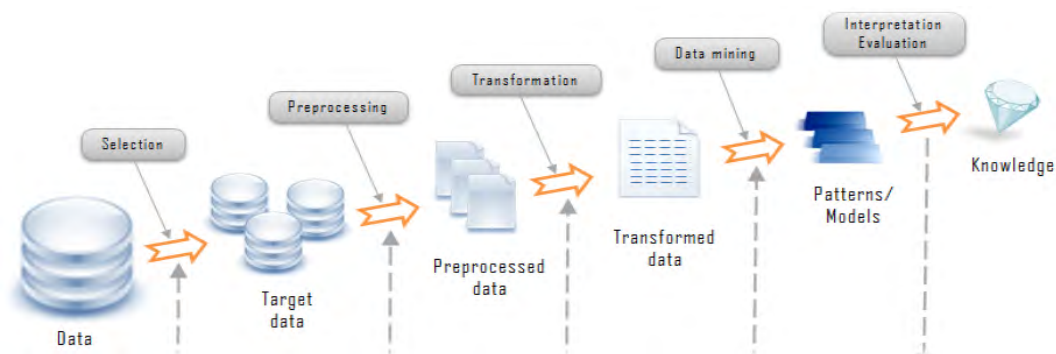
Figure 2: Data Mining Process [Source: Bharati and Ramageri, 2010

2.3 Data mining Algorithms and RDA in digital libraries (DLs):

This Section reviews the relation between data mining techniques and RDA enabled Digital Library systems, and whether they facilitated the creation of personalized recommendations. The first digital library was built in 1975 by Roger Christian, who reported it in the book Toward Paperless Information Systems (Chang and Chen, 2006). Callan et al (2003) conducted a study on Selection Resource in multimedia at digital libraries. This study investigates the issues of resources selection at digital libraries. Moreover, Callan et al (2003) proposed a project

called the MIND that uses traditional data mining algorithms to provide recommendations in digital libraries. Another study by (Mishra and Mishra, 2013) have used data warehousing approach for analysing the data from different perspectives in digital libraries. Further, there are some studies which use similar algorithmic approaches, but their goal is information retrieval, rather than mining the data. According to Bin (2013), the goal of the information retrieval process is to extract the information corresponding to the exact query request, whereas the goal of the data mining process is random knowledge discovery, within the documents. (Mishra and Mishra, 2013). Several studies have done previously show that the data mining and information retrieval can be made to complement each other, and benefit from advances in each field. Yassir and Nayak (2012), reported a study by interviewing 50 participants with the targeted questionnaire. The study showed that there are several issues with information retrieval systems, including missing values, ethical aspects, and attribution of value with meaning, that changes over time. By using tailored and efficient data mining techniques, it is possible to address the issues affecting information retrieval systems. Digital Library systems can be viewed as an information retrieval system, and by applying efficient data mining techniques, it is possible to tailor the digital library systems for personalized recommendations (Mishra and Mishra, 2013). Few other studies have revealed importance of data mining approaches for improving quality of services provided by digital libraries (Guenther, 2000), and provide personalized information services and recommendations, improving information and knowledge transfer process. According to Zhang (2011), data mining processes play an important role in improving the quality of service in digital library systems, as shown in the Figure 3 below:
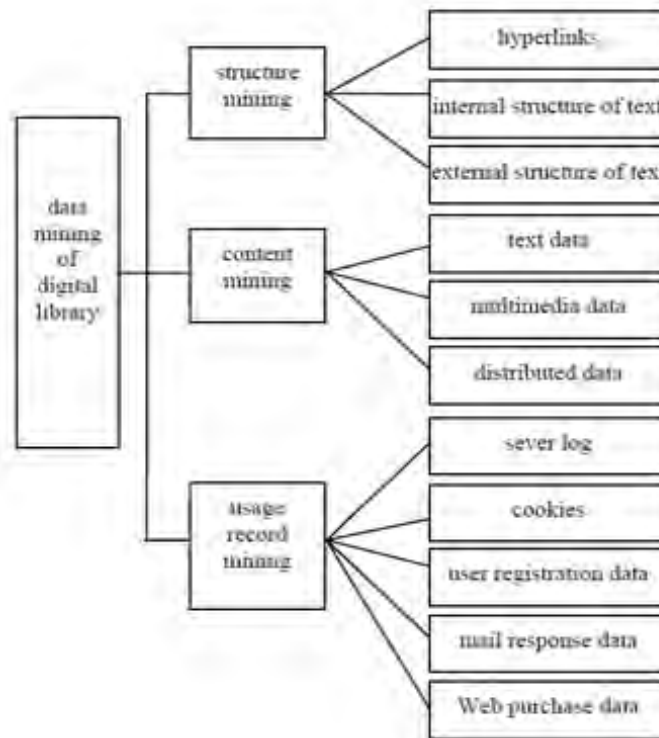
Figure 3: Data Mining in Digital library [Source: Zhang , 2011]

Lone and Khan (2014) also, demonstrated that the data mining process can help digital libraries to improve the libraries' services such as client information access tracking, weblogs, and circulation information tracking. However, most of the previous studies have only focused on the traditional data mining approaches for digital libraries, with single user data records, and did not take into consideration some of the recent advances in both data mining and machine learning fields, as well as information retrieval approaches for providing enhanced services. Some of these advances include using multiple sources of information, or using feedback from explicit ratings, or analysing the sentiments from the text feedback provided by uses or using the metadata information from the modern RDA equipped information retrieval processes in digital library systems. Except a few isolated studies focussing on examining the description of the electronic resources by RDA: Resource Description and Access by (García and Monroy, 2017), discussing electronic resource and entities, attributes and relations. A different type of resource description and cataloguing standard is used by American Libraries, called AACR2 (published under the auspices of AACR fund) (Oliver, 2010). There are many differences between RDA and AACR2 in definition of the resources, particularly for defining the abbreviations, dimensions, recording, inaccuracies, dashes, brackets, sources, edition

statement, publication statement, publication date and copyright, extend and physical description of sound recording. RDA is a new standard for representing all types of resources in the digital environment. Also, another model, the FRBR (Functional Requirements for Bibliographic Records) is a 1998 recommendation of the International Federation of Library Associations and Institutions (IFLA) to restructure catalogue databases to reflect the conceptual structure of information resources, and it describes the grouping of different entities in the bibliographic environment and their arrangement in a hierarchical format. It is a conceptual model of the bibliographic environment which includes the data that the library wants to make available to the users (Strader, 2017). However, RDA is a content-based standard and not a type of metadata model, but more suitable for use with any standard formats to describe information materials such Dublin Core, as it focuses on those parts of the information that the user often needs to know; and on what information should be recorded and how it is to be recorded to help the users to search large databases and indexes and then find, identify, select, and obtain the source that related to the information that they are looking for it (2010), and is shown in Figure 4 below.
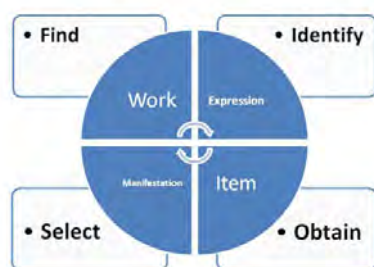
Figure 4: FRBR Model . [Source :Tillett, B., 2005]

The FRBR model meets these needs through the bibliographic network of entities, attributes and characteristics of Work, Expression, Physical Appearance and Item. However, as outlined earlier, there are many benefits when using RDA and data mining in digital libraries.

Some of the prior work on using data mining techniques for digital libraries application contexts focussed on different end goals. Some of them were aimed at solving the classification problems, where the users/readers were categorized into five clusters, and each cluster has its own characteristics. Few other works focussed on using data mining techniques for not just classification, but also regression, rule generation, discovery, summarizing, dependency modelling and sequence analysis aspects. Chang and Chen (2006), used three techniques involving data analysis, building of data warehouse and apply data mining algorithms for a digital library application context, resulting in better service provision, and conforming to the cataloguing standards of the country. Due to different cataloguing standards and protocols used

in each country, the data mining algorithm suite developed differs, and not uniform across the digital library application scenario. For example, Libraries Australia adopted RDA as the preferred cataloguing standard from April 2013 and implemented several changes to support the creation and exchange of RDA data (Velasquez and Campbell, 2015). The OAI-PHM protocol used by Kim et al (2005), for MEDLIN database, uses a metadata harvesting method, to work with RDA, and extract metadata, helping in building better recommendations. system. In general, in many of the previous work surveyed, seven steps appear to be the common practice for the data mining process for library data, and these steps focus on how to select data, pre-process data, transform data, store data, mine data and evaluate mining result (Chang and Chen, 2006).

In a multilingual context (focussing on Arabic language context), Saudi digital library (SDL) services are provided to all Saudi universities under one umbrella and these services were made available for universities users as well. SDL uses two ways to retrieve information, using Summon research engine or direct access to database library. The use of Summon Service research engine increases the value of library, by delivering an unprecedented research experience. This unique service increases resource usage, connects users with librarians, provides the easiest and most efficient way for users to discover library collection.

The Australian digital libraries, on the other hand, use a different approach for providing the services, use a different approach to collect all digital collections' information in one place. There are many tools for querying digital collections, including Trove which is a free search engine created by the National Library of Australia contains many electronic collections in a different field such as Books, Ephemera, Journals, Manuscripts, Maps & Aerial photographs, Music and Pictures (National Library of Australia, 2010). However, in 2015 the National Library encouraged all its digital materials to a new platform and began digitizing content for partner organizations, and the old National Library digital delivery service was stopped in 2016. According to Burrows (1999), the new platform aimed to provide an integrated platform for all Australian digital libraries, allowing cooperation among different organizations and universities for providing resources from their website for their users.

In addition, Libraries Australia provide a resource sharing service for sharing digital documents and is managed by the National Library of Australia for Australian libraries and their users. Their key mission is to support a seamless workflow between Australian libraries and provide data to underpin Trove, by using data from Australian National Bibliographic Database ANBD, which records the location details of over 50 million items held in most Australian academic, research, national, state, public and special libraries. (Libraries Australia, 2015). Thus, by

comparing the architecture of digital library systems across two different contexts, Australia and Saudi Arabia, it can be seen that Australia does not have a specific digital library but has many digital resources that comes from different sources. On the other hand, Saudi Arabia has a specific library which is called Saudi digital Library SDL as shown in Figure 5 below:
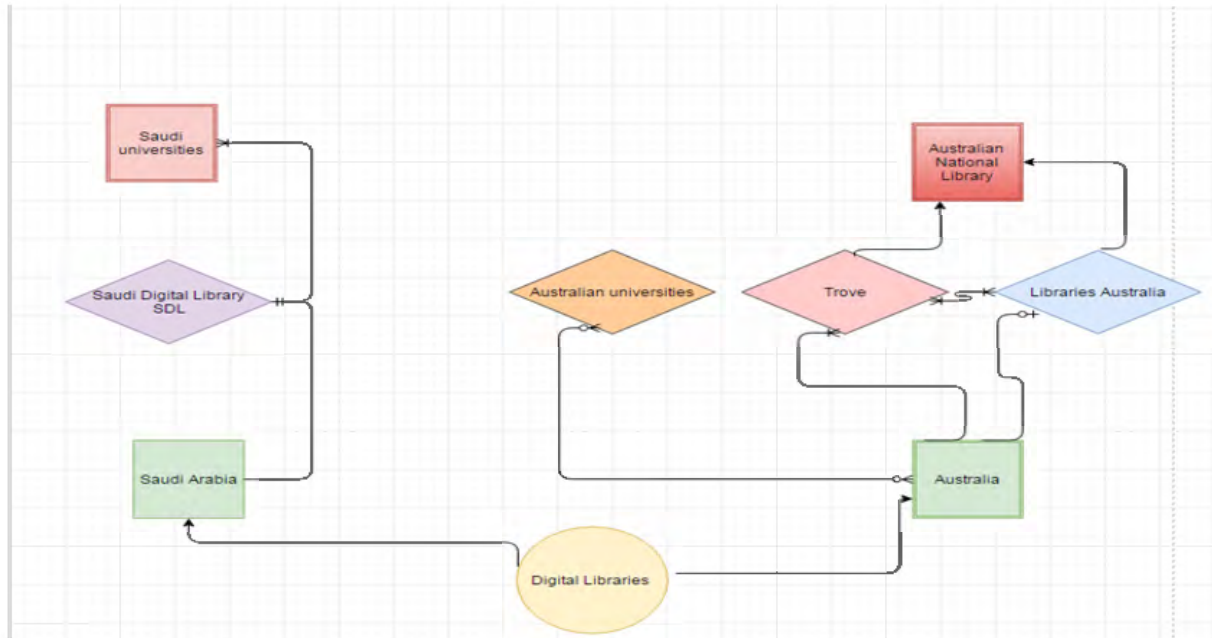
Figure 5: Summaries of digital libraries in Saudi Arabia and Australia

Due to these differences, the data mining approaches used by different digital library systems are not the same and use different algorithmic techniques to provide retrieval services. Yin and Han (2003) proposed a classification system called the predictive association rules algorithm call "CPAR". Antonie and Zaine (2002) used association rules mining to build a classification model. Li et al (2001) proposed a classification model called CMAR which based on multiple association rules to construct a spatial classification system, whereas Chen et al (2001) have been used a classification model based on association rule to classify the data coming from a different application context (underwater acoustic signal).

The architecture used in Kobson digital library in Serbia (Kovacevic et al., 2010) is based on k-means clustering and the naïve Bayes classification together to improve digital library services. This system also uses a combination of different data mining techniques for better retrieval performance and service provision, including association analysis, classification, and clustering. Uppal & Chindwani (2013), also used association analysis, clustering, and sequential pattern mining on library dataset to extract frequent book sequences borrowed by students. Their system uses four steps to extract knowledge from data, namely data cleansing and integration, data selection, data mining and pattern evaluation, as shown in Figure 6 below:
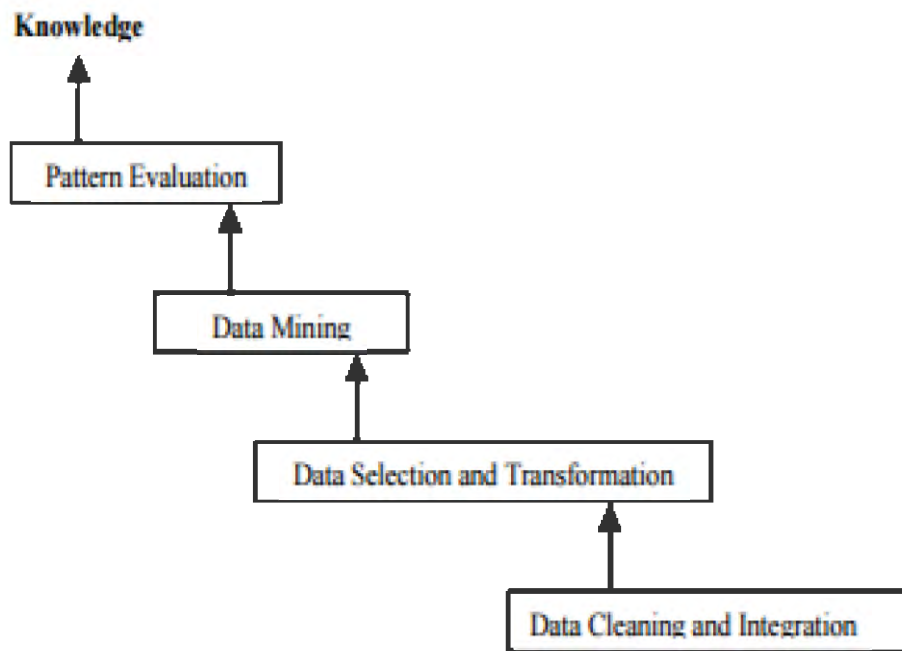
Figure 6: Steps of extracting knowledge from data [Source: Uppal & Chindwani, 2013].

The FP-Growth algorithm, proposed by Song and Wei, (2011) has gathered significant interest in the research community, due to its ability to capture the relationship between books and users, to find needs of readers more clearly. The findings from their work showed that the FP-Growth algorithm can get high accuracy of reader's need by analysing association rules. The approach proposed by Bin (2013) involving four data mining models for retrieval which are classification, association, clustering and regression models. The classification model was used to extract and segment important data into different classes. The Association model is used to make and construct relationship in database based on the entry. The Clustering model involves the grouping of data and features for similar categories, using the historical data to predict future trends. This combination of these different models in this work was one of the earliest personalized retrieval and recommendation technique available and served as a reference model for most of the subsequent works. Prior to this work, Chen and Chen, (2007) developed an approach for recommendation service for digital library context based on a combination of two algorithms, the Association rule mining and Ant colony optimization. This approach allowed discovery of interesting association, and helped find the relationship from a large dataset as shown in Figure 7 below:
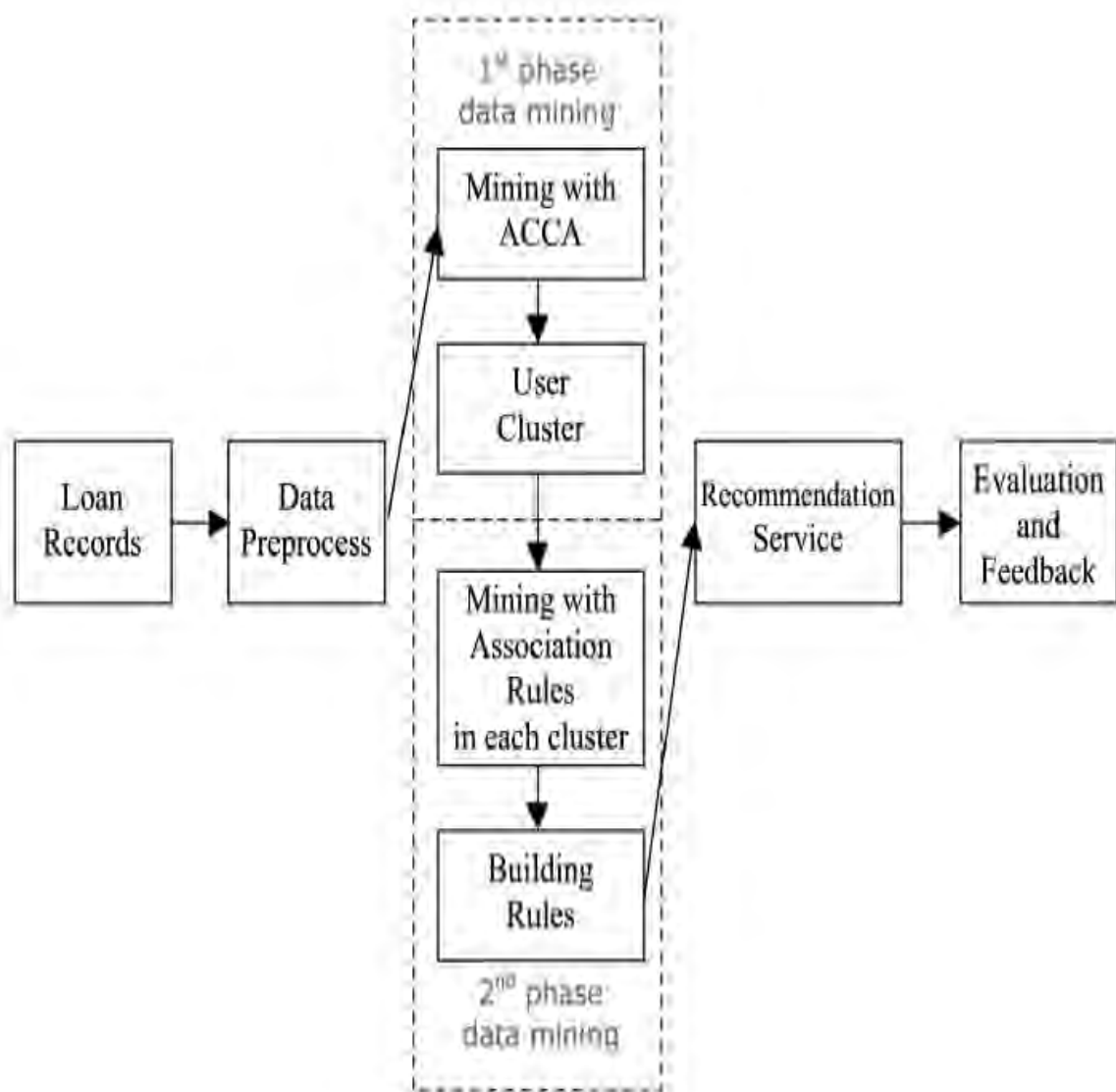
Figure 7: Data mining process in Digital library

Moreover, Ant colony optimization algorithm used in this approach, attempts to find the shorted path for data, analogous to strategy used by real ants, for finding the shortest way to food. This method separated users into several clusters based on records, and then users who have similar interest are collected into one cluster, as showed in Figure 8 below:
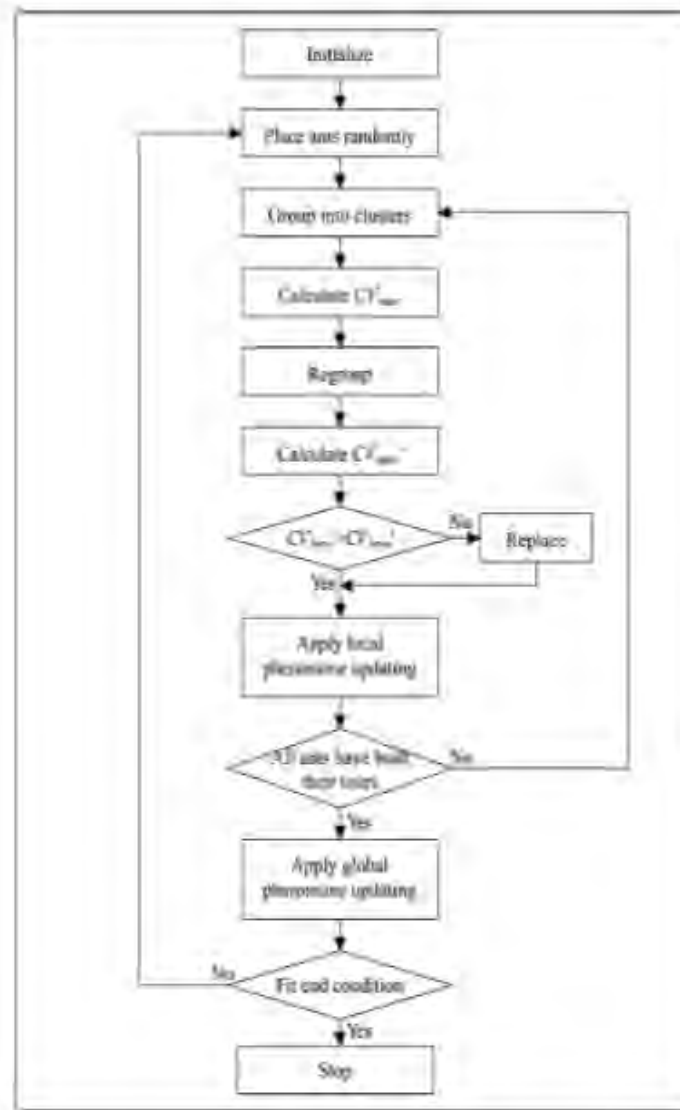
Figure 8: Colony Optimization Algorithm [Source: Chen & Chen, 2007].

The Next Section discusses the subsequent work on different data mining approaches for recommender services for digital libraries context.

2.4 Data mining algorithms used for providing recommendations in Digital libraries.

There are four different steps involved in extracting knowledge from different data sources, repositories, and data stores. These steps include data cleansing and integration, data selection, data mining and pattern evaluation as shown in the Figures 9 and 10 below:
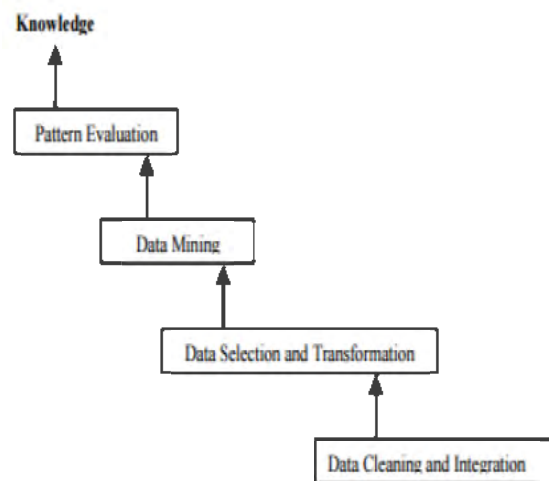
Figure 9: The process of extracting knowledge from data [Source: Uppal & Chindwani,2013].
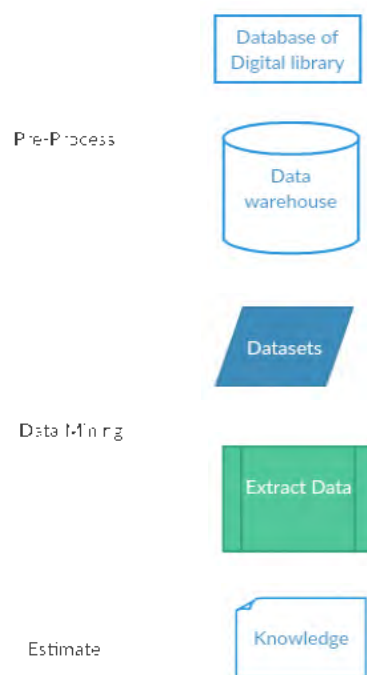
Figure 10: The process of data mining in digital library

Providing recommendations is an additional step involved in this process, and can be the most useful step, for enhancing the quality of services and user experience in digital libraries contexts. This additional step of providing recommendations typically involves producing a list of recommendations to the user and may requires a different set of data mining algorithms. These algorithms help libraries personalise recommendation of books or any resources for users

based on borrower or personal loan records. There have only been a few studies on book recommendations based on library loan records, due to the difficulty in accessing borrower and personal loan records (due to privacy issues). Thabtah et al. (2004), proposed one of the frequently used method based on collaborative filtering. other similar works, the Melvyl recommender project, proposed by Whitney and Schiff (2006), uses a weighted graph model, that is like the association rule method. One of the recent developments, using association rule mining on circulation records for book recommendation was proposed by Rustogi et al (2017), involves the use of collaborative filtering as well as association rules, showing this method has better performance. Further, the authors of "Comparative Study of K-NN, Naïve Bayes and Decision Tree Classification Techniques" have studied book recommender system using frequent pattern algorithm (FP-Growth) and showed that FP-Growth makes effective book recommendation to the users of the library and well applicable in analysing large data in a library in Taiwan. (Jadhav& Channe, 2013). The results showed that naïve Bayes classification can be used together to improve digital library services. Tsuji et al. (2012) in "Use of library loan records for book recommendation" did an empirical study of the application of data mining techniques in library system, and used association analysis, classification, and clustering approaches. This study discusses the application of association analysis, clustering, and sequential pattern mining on library dataset to extract frequent book sequences borrowed by students. According to this study involving the use of library loan records for book recommendations, several methods were used for generating recommendations, namely the collaborative filtering, the association rules and the use of the Amazon system, for mining 1,854,345 loan records from 39,442 users of the university library and recommended books to 33 undergraduate and graduate students based on the collaborative filtering, association rules and the Amazon system (method used not publicly available). The result was that the method used by Amazon (not publicly available) is the best, followed by association rules method, and least performing approach is the collaborative filtering.

Data mining process at its heart uses machine learning algorithms for extracting the data or knowledge from one or more sources. The machine learning models here determines the performance of the data mining process. The different types of machine learning models used for this task are *predictive*, *descriptive* and *correlation* models. Predictive models are designed to predict the value of some properties such as predicting a potential purchase for the customer. Descriptive models are divided into two types: clustering patterns that allow the assembly of individuals, events, products in clusters, and the correlation models allow for the identification of relationships between them. Some of the recent works use different types of models called

deep machine learning models, that are completely data driven and discussed in the next section in detail.

2.5 Deep learning models for recommendations in Digital Library Systems.

Deep learning is known as deep structured learning or hierarchical learning. It is a new subfield of Artificial Intelligence (AI) and Machine Learning (ML) area, and is completely data driven, using data from multiple sources such as images, sound, and text to make sense out of data. It should be mentioned that there are differences between machine learning and data mining. Machine learning provides algorithms to learn from data, extracting relationships between different attributes, and build a model, solve the problem using this model, and predict the future trend, such as identifying a new customer. Data mining on the other hand, is used to explain the data, discover knowledge from a large amount of information in database, for further decision-making based on this data. Therefore, data mining is used to uncover unknown data and knowledge, and machine learning is used to uncover known data and knowledge (Jegan, 2017). Deep learning is a subfield of machine learning field and considered one of the most an important developments in the field, for understanding complex data representations at multiple levels, based on different types of hierarchical learning architecture and algorithms driven by structure and functionality of the human brain, and used in artificial intelligence field. Deep Learning algorithms are useful when dealing with learning from a large amount of unsupervised data, and typically learn from data in a greedy layer-wise fashion (Najafabadi et al., 2015). Including recommender component in a data driven machine learning or AI system can lead to better personalisation, as it can include the end user in the decision making loop, and enhance the user experience and satisfaction, and is increasingly being used recently for online search and retrieval in different fields. There are many companies such as Amazon, Netflix and eBay providing personalisation retrieval, using recommender system components in AI/Machine Learning based search and retrieval platforms, to not only meet users' needs, but also help them in finding what they really need (McGinty and Smyth, 2006). The traditional recommender systems without deep learning based algorithmic models, are based on using different set of criteria for capturing user preferences based on the content, users' profile, and interaction between users and content. These algorithms are called collaborative filtering or content-based filtering methods and use matrix factorisation techniques to model the interaction between the user and the content. Though these methods work well if there is sufficient interaction between the users and content, as there is enough interaction data, but for users who don't interact much with the system or for the new users, there is insufficient interaction data, and leads to poor

performance. The usage of deep learning techniques for providing recommendations led to good results in several other fields involving complex search and retrieval, including computer vision, speech recognition, language processing and so on. Therefore, other similar search and retrieval systems and associated end user experience as in personalized recommendations in digital library systems can see similar benefits with deep learning-based algorithm engines (Shokri and Shmatikov, 2015). However, the use of massive data collections, especially used for building deep learning models, might present critical privacy issues, and lead to compromised safety of the users (Yonetani et al., 2017). The data at the central storage system may also be the subject of surveillance, an implication that the essence of personal privacy is challenging to attain (He et al., 2017). Hence the type of deep learning algorithms chosen or developed have to take into consideration security and privacy aspects and need to be privacy aware. These algorithms depending on the application and the security requirements, need to be customised or designed with due diligence (Willemsen et al., 2016), (Bendersky et al., 2017). Deep learning architectures due to their opaqueness and inherently black box type structures, tend to address these requirements well, as compared to other transparent shallow learning-based algorithms.

2.6 Sentiment analysis for recommendations in Digital Library Systems.

For enhancing the quality of recommendations in any search and retrieval platforms, include the digital library systems, there have been several research efforts to include different types of user feedback. These include not only used of implicit and explicit ratings, but also the use of text-based feedback comments. Use of text-based feedback allows sentiments to be captured and can help understand user needs better. However, including sentiment analysis in decision making loop in providing recommendations, requires use of linguistic NLP (natural language processing) techniques for mining the text. This is more challenging as compared to simple ratings-based recommendations to be generated and the use of NLP techniques requires comprehensive language dictionaries to be used and created in different multilingual contexts and environments, the system is expected to perform. For an English language NLP system there are established language dictionaries tailored for different application scenarios, but this is not the case for other language contexts. Traditional English language processing approaches use of N-gram approaches and TF-IDF based statistical techniques for building *bag-of-words* models, More recently, some form of numeric representation of words, called *word2vec* models are being used, for words to be processed effectively and to ease processing complexity and load.

The bag-of-words model is an approach for extracting and representing specific features from text for use in NLP modelling. It is commonly applied in machine learning for text processing, due to its simplicity and flexibility. The Word2vec models, on the other hand, encompass a group of related models that create word embeddings. Further, the shallow neural networks used for text processing, are aimed to rebuild linguistic contexts of words. These models have featured predominantly in related works and past studies. Some of the related work is discussed next. Both Bag-of-Words and Word2vec Models have been explored widely for sentiment analysis of customer reviews (Chakraborty et al., 2018). Dai et al (2015) have adopted a standard bag of-features model to forecast the opinion class of film reviews based on consumer's inputs. Their outcomes indicated that the bag-of-words models outdid simple decision-making frameworks that implemented hand-picked features for views classification (Bansal and Srivastava, 2018). One of the challenges in building NLP models for sentiment analysis, is time-consuming aspects involved in handling dominant human language to make an algorithm (Gehrmann et al., 2018). Further, many authors showed that such models were inefficient in terms of manual evaluation of lexicon and lend itself to low accuracy (Stojanovic et al.2017; Yuan et al. 2016; Nazir et al.2018). To address this problem, models based on semi-supervised framework for sentiment analysis of consumer reviews were proposed by Walkowiak et al. (2018); Kim et al. (2017); Zhang et al. (2016); and Vinayakumar et al. (2018). Also, used para-graph vectors to determine the syntactic and semantic links of a review text using semi-supervised models were proposed by El-Din, (2016); Dehghani et al. (2017); and Zhang et al. (2016). The concatenation of review embedding, and product embedding gathered from paragraph vectors when applied, outperformed all previous approaches used in sentimental analysis of customer reviews. Several studies compared Bag-of-Words and Word2vec Models on sentiment classification (Campr & Ježek, 2015; Zhang et al.2017; Barry, 2107) while (Joulin et al. 2016; Huang et al. 2019) performed sentiment analysis of clinical datasets. They found that these two models complemented each other. Also, (Vinayakumar et al. 2018; Ragini et al. 2018; Krishna et al.2019) supported the combined use of Word2Vec and bag of words for sentiment analysis of patient's summaries. In terms of accuracy, Stojanovic et al (2017) supported Word2Vec as the most precise and functional word embedding approach. It was highly effective in con-verting customer reviews into substantive vectors (Kaur, 2019). The only problem with the model was that it required huge chunks of reviews for training and yielding the exact vectors.

As can be seen from the several works discussed above, for personalisation or recommendation generations in search and retrieval systems including digital library systems, Sentiment

Analysis (SA) plays an important role. It helps to extract opinions of the people opinions about different products. According to Abbasi et al. (2008) the text extraction and mining could be objective or subjective, as the text contains negatives and positives opinions, and can be inferred using machine learning based models for sentiment analysis at the sentence level. However, most research studies of this type, have been using natural language network in English. Very few works in other multilingual contexts, particularly the Arabic language exist. Some of the studies reported are discussed here. Aly and Atiya used *Large Arabic Book Review* LABR, containing 63,257 book reviews, collected from Goodreads website. They divided reviews into positive and negative. Reviews with rating 4 or 5 labelled as positive and 1 or 2 labelled as negative. However, rated 3 considered as neutral reviews. They used deep learning by applying SVM (support vector machines), MNB (Multinomial Naïve Bayes) and BNB (Bernoulli Naïve Bayes). The result was 90% accuracy for sentiment classification based on NLP text, and 50% for rating classification based on explicit ratings (2013). Mourad and Darwish proposed lexicon-based approach by using two datasets which are MSA (modern standard Arabic) news articles and dialectal Arabic microblogs from Twitter (2013). They used two Lexicon MPQA (multi perspective question answering) for English subjectivity and ArSenti for Arabic. In addition, they clean all words in Tweets by using tokenization and stemming. Mountassir et al. (2012) used three different classification models, which are NB (Naïve Bayes), SVM (Support Vector Machine) and KNN (k- Nearest Neighbourhood) models built with two datasets Movies and Sports datasets. Before model building and evaluation, the authors performed pre-processing. They found that the pre-processing helps to improve classification performance. Further, El-Beltagy and Ali used lexicon-based approach to provide a sentiment classification of Egyptian Arabic text. The result was 83.8% accuracy, accepted as a good result (2013). Sentiment analysis has been viewed as an important aspect in Arabic language context as well, because the number of internet users in Middle-East are increasing, and they are sharing their opinions, reviews and feedbacks. Around 69.2% of internet users use the internet in Middle East. They use Internet for searching about products or services (Ghallab et al.,2020), and well above world average, as can be seen in the Figure 11.
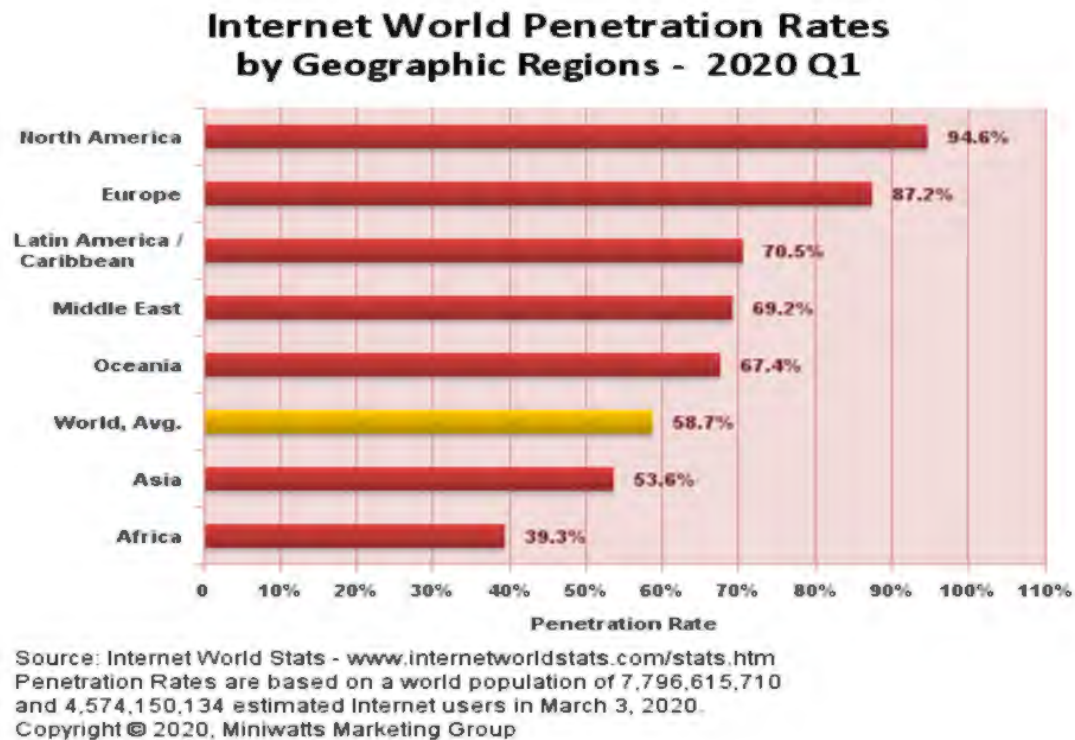
Figure 11: *Internet World Usage.*

Refaee and Rieser built a multilingual sentiment analysis system, based on a hybrid rule-based and supervised learning approach. Their results show that the performance of the supervised method is higher than the unsupervised method for an English dataset, as compared to the accuracy achieved for an Arabic dataset (2016).

According to Boudad et all. (2018) though the research on sentiment analysis has started in 2008, there was limited work on classification at the sentence level and considered a problem in Arabic NLP. Next two Sections describe other related works done on recommendation systems and sentiment analysis for different application areas.

2.7 Applications of Recommendation Systems.

Due to the numerous benefits it offers to its user, recommendation systems have been widely used in various fields. The recommendation system is an application which offers personalisation and customized recommendation of a product or service on the internet (Christakopoulou et al., 2018). The system achieves this by reducing online data overload to improve the level of client satisfaction. To reduce information overload, the application sieves and analyses data and provides the client only with the information they require (Sandoval, 2015). It is, therefore, an application which saves users' time and enhances the satisfaction

level, by analysing large amount of information to bring out an accurate specific result to the user.

They can be used for different applications. For example, the government uses e-government applications to help provide services and information to the citizens. As a result, there has been a surge in information request in government portals and websites (Vateekul et al., 2016). The abundance of information, which at times is not effectively categorised, has rendered people with little ability to make effective decisions (Cambria et al., 2017). The above situation can result in inefficiencies in service delivery since it hinders the retrieval of the required information. Therefore, the recommender systems can be used to help in solving such problems by offering the services or information the users may be interested based on the analysis of their user profiles (Dehghani et al., 2017). Besides, the governments have used the recommender systems to help citizens make decisions on the best candidate to vote for. The system provides information about each candidate during the election campaign and recommends the best one putting into consideration the user's preferences (Dwork et al.,2015).

Moreover, in businesses, recommender system can be used to help managers to maintain current databases through product recommender system. Besides, it is used to help business owners select authentic online sellers. In addition, recommender systems are used to assist in building stable business models by using intelligence, sourced online (Elkahky et al., 2015). Further, the banking sector has equally utilized the system to create effective investment portfolios by matching the client profiles with information from the system. In addition, Internet-based companies, such as Amazon, use the E-commerce recommender systems in their online platforms to assist clients in choosing the products to purchase (Ghosh et al., 2016). These recommender applications define the product to be recommended based on the information about the product total sales, the current trends in buying, the customer demography, and location of the customer (Hongliang and Xiaona, 2015). Using the said information, this system matches the needs of the client and the data from the website to customize a product for the customer.

2.8 Applications of Sentiment Analysis.

Businesses use sentiment analysis to improve their competitive edge by monitoring and managing company reputation (Thai et al., 2016). This helps businesses to implement flexible and insightful programs into the way they present their products. Also, organizations use sentiment analysis to track perception of the brand, find patterns and trends, and monitor the influencers of the industry (Li et al., 2017). As a result of monitoring the company can

implement various ways of counter-competition, such as using automated media monitoring systems and monitoring reviews of the brand on various online platforms.

Sentiment analysis is also used in performing market research and competitor analysis by gathering information (reviews, comments, news articles, and general perception) from different platforms. Information such as competitor contents, perception, and insights can also be gathered by sentiment analysis (Poria et al., 2016). After obtaining such information, the company will be able to effect the necessary changes to remain relevant in the business environment. In doing so, the sentiment analysis tools can help investigate the demands and needs of the market to adjust the business to gain the edge over the competitors (Soni and Sharaff, 2015). Finally, sentiment analysis is used in product analytics. Here a specific item produced by the company is monitored in terms of the clients' comments and remarks regarding the product. Sentiment analysis also assists in improving the performance of a product in each locality or a particular demographic space (Hassan and Mahmood, 2017). The comments used on a product will help to define its niche in the market and ascertain whether any adjustments are required to ensure market traction.

2.9 Progress in Recommender Algorithms for Digital Libraries.

Like other search and retrieval platforms, the algorithms for providing recommendation services in digital libraries are based on the retrieval techniques which utilize the algorithms for screening and sorting out the required information.   Thus, Content-based recommender algorithms receiver the customer requests, and then send distinct recommendations as per the client requirement. It does the above by filtering and sifting the information and offering content or products to the client depending on their profile (Véras et al., 2015). On the other hand, Collaborative filtering-based recommender algorithms coordinates, assesses and then gives information related to what other clients have done before, comparing it with the temperament of the current customer. Finally, a more sophisticated Hybrid recommendation algorithm uses a mixture of the above procedures, and results in higher efficiency, as compared to the said two systems.

The CORE is a recommender service that goes through a large amount of data to find articles that are relevant to the client search query (Covington et al., 2016). The service is based on a plug-in algorithm that is inbuilt in the repository system, specifically for E-prints. When a client looks for a piece of information, the plug-in sends the data on a query to the CORE. The data query may range from the object identifier to the metadata, or when such information is

conceivable. By using the CORE suggestion calculator, the plug-in sends a list of recommended resources to the user. The above procedure eases the search for information, making the tool helpful in digital library systems.

2.10 Use of Sentiment Analysis for Recommendation Systems.

There are many methodologies used for English and Arabic recommendation system that incorporate sentiment analysis. Portugal and Cowan (2018) used a Caption Generation method, which works by associating parts of the image with certain words. In a study conducted to verify this methodology, researchers removed fragments on the image dataset. Then, they employed a Deep Belief Network (DBN) algorithm to model words that were closely related with the image, and those that were close to the fragments. After taking 10,000 pictures from Dataset and Al-Jazeera networks, the researchers generated sentences using the dependency tree relations. The results were positive and showed that both Arabic and English languages represented caption recognition (2018). The above technique, therefore, seems to fit well for including sentiment analysis in both the languages for recommendation systems. Aslanian et al. (2016), used an Automatic Speech Recognition (ASR) methodology based on deep learning and machine learning to recognize spoken Arabic and English digits. To ascertain its efficiency, an in-house dataset was used for building a model in a multi-speaker context with fixed settings. Consequently, the researchers found out that word error rate had significantly reduced, and the cross-entropy had a descent gradient (Aslanian et al. 2016). During the study, the researchers used the LSTMs algorithm to recognize words of the two languages. The switchboards used had an achievement rate of 16.4 WER, an indication of high efficiency using ASR in a recommendation system of English and Arabic based on the sentiment analysis (Abid et al.,2019). Bansal et al. (2016), and Alharbi & de Doncker, (2019) used a Language Modelling (LM) methodology on recommendation systems of English and Arabic by building an LM character that fits both phonemes (Bansal et al., 2016). The LM methodology allows both Convolutional Neural Network (CNN) or Deep Belief Network (DBN) algorithms to be incorporated into language modelling. In a study of English and Arabic, LM methodology gave better results than the baselines, which were tested using morpheme levels. The above results indicate that language modelling is crucial in the recommendation systems. Also, Dialect Detection methodology has been used by Batmaz et al (2018), for its ability to discriminate languages that share a common task as a way of identifying various English and Arabic dialects. The research showed that high-order character-based models such as N grams models and DBN algorithms are the best approaches (Batmaz et al., 2018). Also, their findings were that the

shallow classification methods such as Support-Vector Machines (SVM), and logistic regression methods are not very efficient in carrying out the functionalities of the technique. The reason for the above was based on the analysis that identification of the character level of the dialects of the two languages is challenging in the Digital Subscriber Line (DSL), because its model needs to stay fixed upon every character to enable it to run through several convolutions. As a result, the filter widths seek permission from the convolution algorithms to get the vector representation used for predicting the dialect of each language (Bisio et al., 2017). The result of the above research, when applied against the *F –measure* gave a 43.8% success rate, which indicates that dialect detection is a useful methodology in deep learning and machine learning used for language recommendation system based on sentiment analysis in multilingual contexts (Beel et al., 2016).

Similar methods were used for a text categorisation task as well, and one such method based on deep learning and machine learning technique specific to the categorization of English and Arabic languages was proposed by Dai et al. (2018). This method utilizes three stages, with the first recognizing the placement of punctuation marks, pronouns and verbs. The rest of the words are given representation using weight and order schemes. Stage two clusters the wording using two methods, Markov and fuzzy C-means. The final stage includes training DBN algorithms in the clusters mentioned above. The research included evaluation of two datasets consisting of 16, 000 linguistic units. The *F-measure* obtained was 91,3% indicating a significantly high text categorization performance (Cai et al., 2018). The methodology is, therefore, can be useful for deep and machine learning based approaches for language recommendation system incorporating sentiment analysis. In addition, Automatic Discretization (AD) is another such deep learning methodology, that uses simple discretization to perform a language analysis (Chen et al., 2017). Various algorithms are used in this technique to discretize automatically. Recurrent Neural Network (RNN) layers are placed on top of each other. The research carried out by Zhang et al, (2018) showed how this algorithm is applied in English and Arabic languages, where L1 which was an output for the first layer was observed to be an input for L2. In this manner, the bidirectional layers get both the input and output sources. At this point, the RNN algorithm can effectively work on either one network or several networks (Britz, 2015). The one-to-one coding is done on a single letter, so that the input and output sequences obtain the same measurements. At the same time, in the one-to-many networks, the input and output base differ in length. The researchers showed that the performance of the former is better than that of the latter, due to its high level of accuracy (Cao et al., 2017). Another study on the impact of the discretization on the levels of noise, network size and importance of post processing step

indicated error rates of 2% and 5.34 % respectively. This methodology is, therefore, seems promising for use in the English and Arabic language recommendation systems.

2.11 Use of Metadata for Digital library recommendations.

Metadata is useful for enhancing the performance of recommender algorithms for any search and retrieval platforms including digital libraries systems. The framework proposed by Ferran-Ferrer is aimed at analysing information of learning from virtual environment to improve the quality of metadata (2007). However, it was found to be difficult to get a complete view of information of students from the metadata obtained, to improve their profile captures. This study did not use any deep learning techniques with metadata for building their model. In another similar work, National Library of Norway is automating Dewey classification, and the Library of Congress used machine learning to predict Library subject headings. There are many approaches have been used to build recommendation systems without metadata capture, and briefly discussed here. Schnick and Knickelbin used lexical measure to measure the reading ability for users (2000). Zhuhadar and Nasraoui combined two recommendation techniques to achieve good performance (2010). They used hybrid recommender system for personalized user experience for reading online learning repository. Hsu (2008) proposed content data and feedbacks to personalized recommender system. Web-based personalized recommender system by using CBF (content based filtering) and Fuzzy algorithms on 10.000 books from different categories have been proposed by Omisore and Samueel (2014), and the experimental evaluation showed that 92.30% efficiency in providing recommendations. Siersdorfer et al, (2010) used user-based CF (collaboration filtering) to determine the user's interest. Moreover, Zhu and Wang (2007), used *apriori* data mining algorithm to optimize book records. Mooney and Roy (2000), used content-based book recommendations to find book description and used machine learning to extract user profiles and find similarity by using Bayesian learning algorithms. Sohail et al, (2013) used sentiment analysis from user reviews to solve the book recommendation problem. Not many works in literature have used metadata for building recommendations. Machin learning and deep learning techniques can also help extract metadata from digital libraries. Cai et al (2003) used SVM to extract metadata from the title of research papers, and showed that SVM leads to good performance in extracting metadata (2003). Recently, Safder et al, (2020), proposed a deep learning based approach for extraction of algorithmic metadata in full text scholarly document by using 93.000 text lines from the document. The result showed that 93.32% accuracy for Pseudo-code extraction. Mai et al (2018) using deep learning model based on MLP, CNN and RNN for subject indexing, and

Safder et al (2018) designed metadata sentence classification using LSTM deep learning models for learning words and sentences by combining CNN (Kim et al , 2014), RNN (Lai et al, 2015), LSTM (Greff, 2016) models. We can also use machine learning models and algorithms by parsing the natural language processing text, and structure the descriptive metadata.

2.12 Research Gap and Innovative Contributions.

As can be seen from the comprehensive review of literature and shown in the Figure 12 and the Figure 13, though there are several methods available for recommendation systems, they have not been specifically targeted for Digital Library systems type applications context and focussed more on English language context for using sentiment analysis in recommendations. Not much research has been reported in multilingual contexts, particularly Arabic contexts.
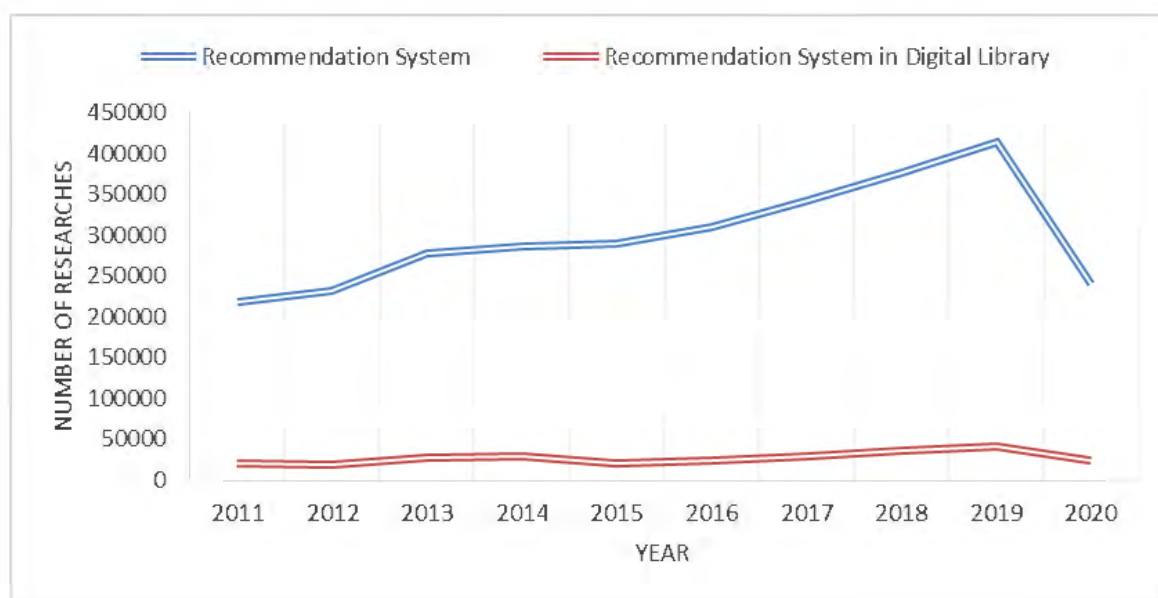
Figure 12: *The difference in research that has been conducted in recommendation system and digital library.*
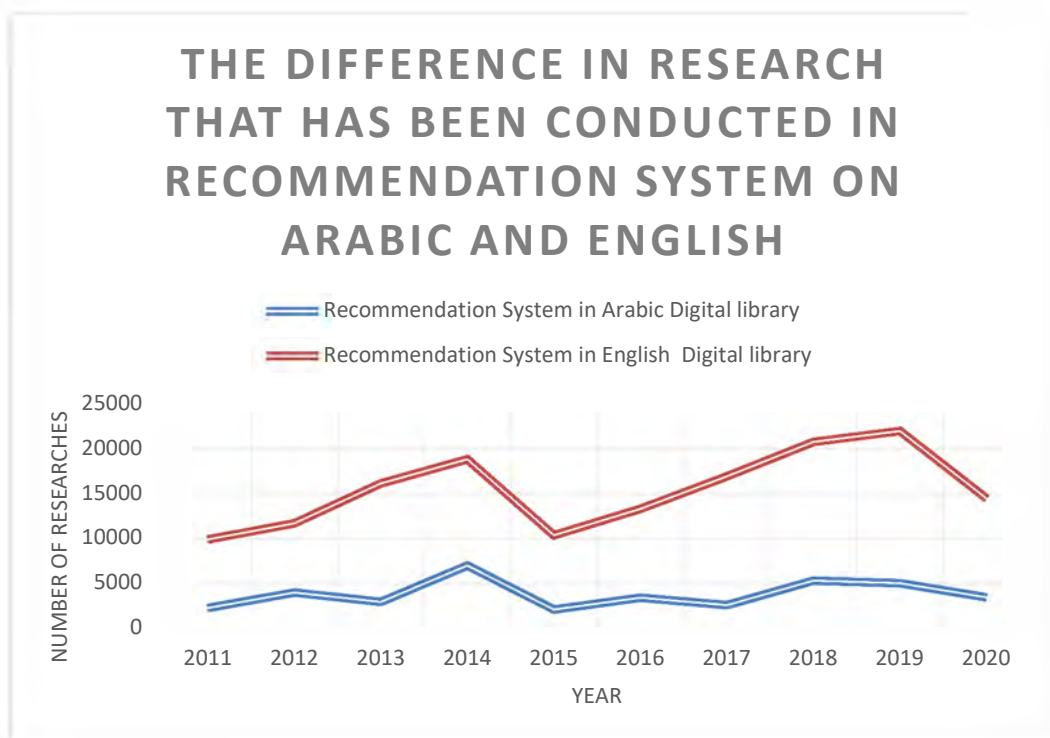
Figure 13:The difference in research that has been conducted in recommendation system on Arabic and English.

Figure 13 shows the difference between the research that has been conducted in the Arabic and English languages. The difference in research that has been conducted in Arabic and English collected by using relevant keywords in recommendation system and digital library field in both languages. The Dimensions website is used to collect the number of research work done. For a specific keyword, the Dimensions is used for a specific period between 2011 to 2020. The results that are retrieved are shown in the top page of the databases result. These results are used in our comparison. It is clear, that there is a big gap between recommendation system as general concept and recommendation system in digital library application context, as well as between the work that has been achieved in Arabic and English. Further, there are not enough studies that use metadata to build digital libraries recommender systems in general, though there is large metadata available. Finally, each of these performance enhancing components of digital libraries system, including recommender components, ratings and sentiment analysis components in multilingual contexts and inclusion of metadata, all have been developed in isolation, for different application scenarios, with different databases, some publicly available open datasets, some private datasets, and some in-house datasets.

This thesis aims to address this gap and proposes a novel integrated computational framework based on several innovative and original contributions. For a baseline comparison, it uses existing approaches and compares with new approaches for assessing the performance improvements achieved.

The main contributions include:

- Development of novel deep learning algorithms for recommendation systems.
- Use of explicit ratings and NLP sentiment analysis for enhancing recommender system performance for multilingual contexts, both English and Arabic language.
- Use of metadata for enhancing the performance of recommendation systems.

The proposed integrated computational framework based on these innovative contributions aim to address the current gap in the research field and can enhance the personalisation capability of information search and retrieval systems in general, and digital library systems in particular, where high quality personalized recommendations can be generated using explicit ratings, sentiment analysis and metadata information in multilingual contexts.

2.13 Chapter Summary.

This chapter has presented a review of the state of the art in data mining, deep learning recommender system techniques, and digital libraries. The review demonstrates that one of the main problems behind building digital library recommendation system based on sentiment analysis or explicit ratings is the isolated development of different approaches, lack of research in the field of multilingual contexts, and lack of availability of datasets in multilingual contexts, particularly the Arabic language. To address this shortcoming a comprehensive integrated computation framework has been proposed and has been presented in rest of the thesis, with each Chapter presenting the methodology, algorithmic implementation, experimental evaluation, and outcomes from each approach used. The next Chapter presents the work towards baseline comparison with a research Testbed/Workbench developed for emulating the digital Library data store.

Chapter 3 *Research Testbed/Workbench for Development of Digital Library Datastore.*

3.1 Introduction

This chapter presents a novel experimental research testbed developed for evaluating the personalization and improved user experience with digital library services, based on open-source technology tools and a data mining approach for personalizing the resources by segmenting the users and their preferences, using *Apriori* algorithm and Frequent pattern mining approach (FP-Growth).

In this work, to investigate the personalisation and user centric modelling capabilities of different approaches, I have set up a research test bed/workbench for evaluating different machining learning algorithms for the proposed scheme. After researching several digital library systems in Australian libraries and Saudi libraries, it was found that most of the systems do not retain the details of the users' borrowing records datasets. This is because of the general law, which states that circulation and registration records identifying the names, addresses, email addresses, and telephone numbers of library users, and the materials borrowed are not public records. Due to this constraint, it is very difficult to develop user centred modelling and personalisation schemes based on data mining and machine learning schemes. Hence, an experimental research test bed was developed, for emulating a digital library system, by developing a web platform based on open-source tools. This web platform emulates real digital library, by using open-source resource *Librarika* content management system, which allows building of cataloguing from online public access catalogue (OPAC) (Librarika, 2017). For preliminary proof-of-concept prototype, a small digital library was built, called the *Knowledge to Action Digital Library* and allowed users to engage in several library scenarios and borrow several resources from this online library platform. A friendly user interface, with multi-level user registration, authentication and login with borrower and administrator privileges was also included. Figure 14 shows the screen shot of the friendly user interface. They can use the library easily by simple interface and by clicking on request button and choose different types of resources, for borrowing, as shown in the Figure 16. This open-source platform was customized for creating reports such as circulation reports. The size of data sets for the study reported in

this chapter involves a sample of 2816 of users' borrowing records. The dataset collection followed the protocols of ethics approval requirements of the University.

Figure 15 shows the Data store design structure, and the Figure 16 shows the user interface screenshots of the web portal and retrieval of data with the proposed data mining approach. The dataset has subjects field corresponding to different disciplines, including Information technology, Science, History, Medicine, Music and Education with each subject given a catalogue number, as well as the multimedia sources.
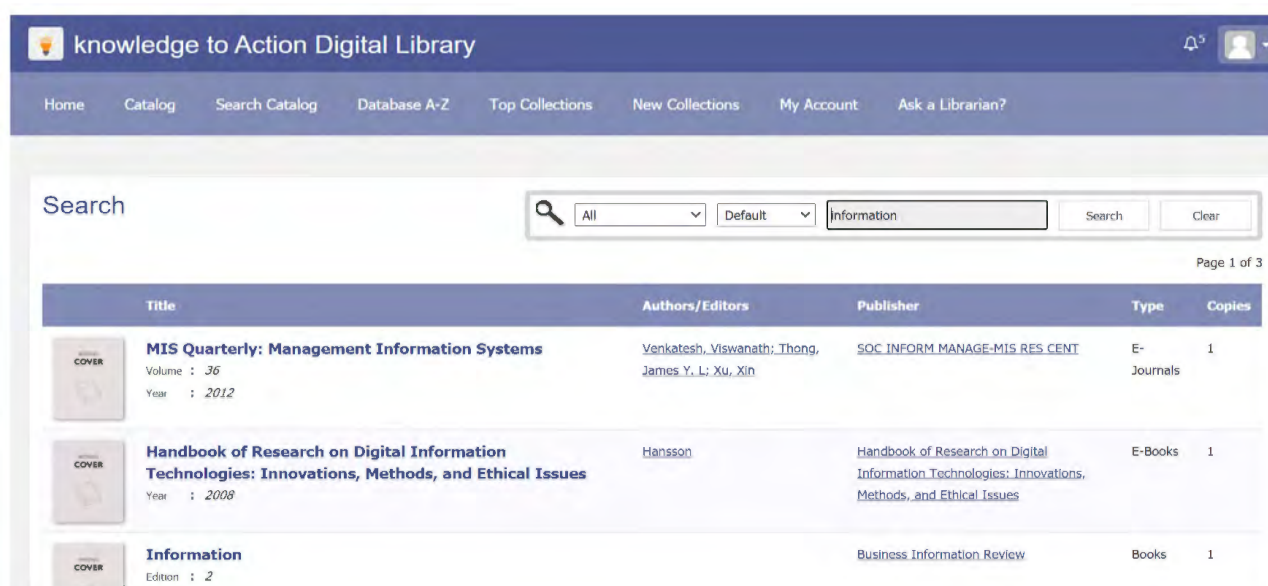
Figure 14:User interface Search page
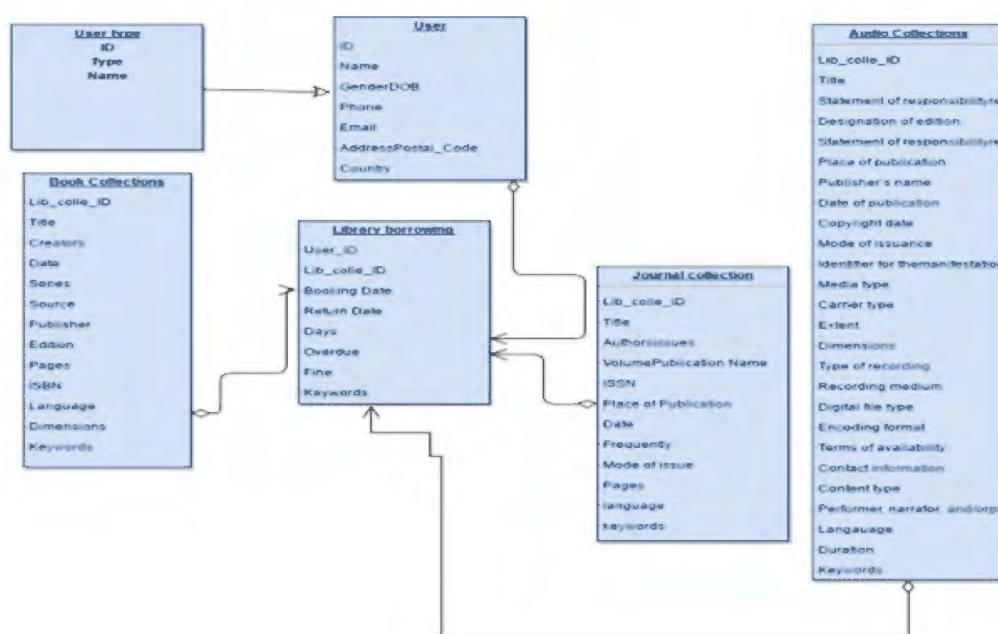
Figure 15:Digital Library Prototype design structure
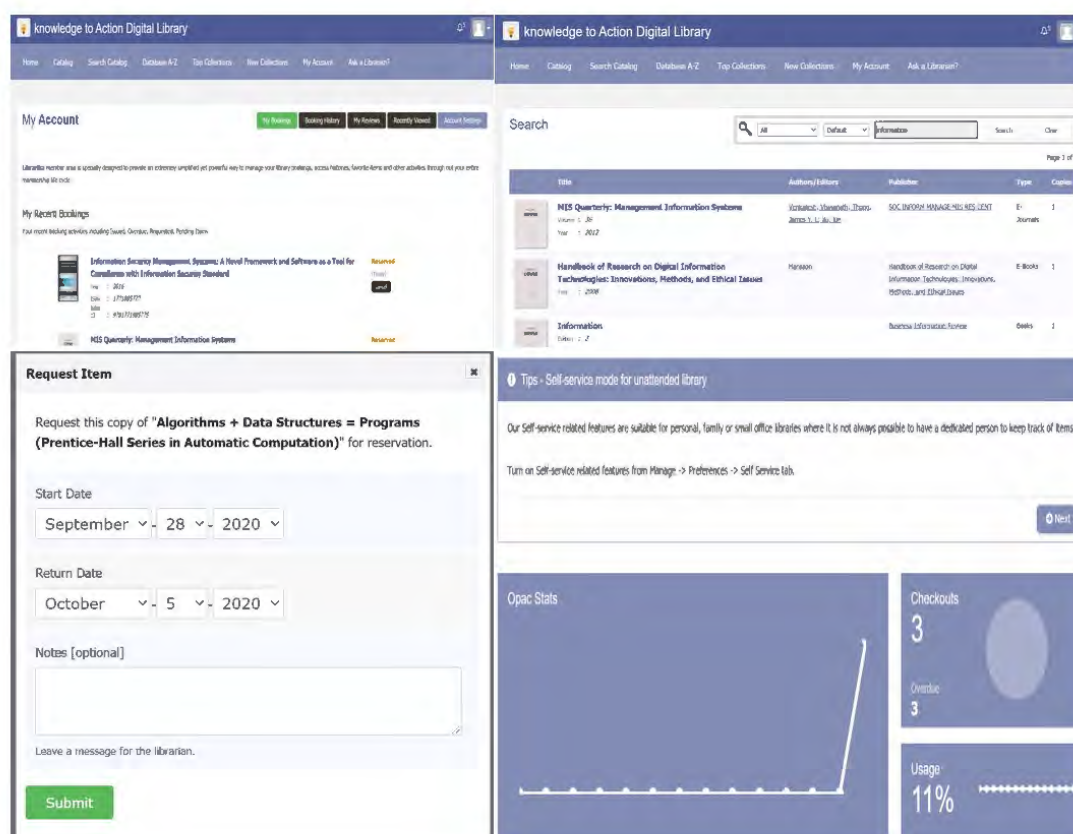
Figure 16: Digital library User interface User interface.

Figure 17:Dataset after pre-processing for mining.

3.2 Association Rules:

Association rules were used for analysing data for frequent patterns and criteria to identify the most important relationships between items in database (Bussaban and Kularbphettong, 2014). The association rules searched for relationships or links that exist between several properties. In addition, association analysis refers to a set of methods used to link purchasing patterns across cross-sections or over time. For example, the market basket analysis uses the information in the goods purchased by consumers, to predict the goods they may purchase in future, if special offers made or if they are identified.

3.2.1 Association Rule algorithms:

Association rules are based on analysing data for frequent patterns and criteria to identify the most important relationships between items in database by analysis of the frequently items appear in the database. The association rules search for relationships or links that exist between several properties. Association analysis refers to a set of methods used to link purchasing patterns across cross-sections or over time. For example, the market basket analysis (a type of correlation) uses the information contained in the goods purchased by consumers to predict the goods they may purchase if special offers are made or if they are identified. There are several association rules algorithms that are developed in data mining field, including *Apriori* algorithm and FP-Growth algorithm, and the differences between these algorithms depend on the support and confidence of algorithms for the association rule formulation. A brief review of the two well-established association rule algorithms the *Apriori* and FP-Growth, by using real life example is presented, including the comparison based on certain performance measures, such as accuracy on circulation records of the digital library.

3.2.2 Basic concept of association rule:

Let *A* be a transaction set, $= I$ *{i1, i2, … i$_m$}* is in *A* composed of collection of all items including *I$_k$ (k=1,2,……m)* called item set. The item set contains the set of *k* items called *k*-items. Also, let *T={t$_1$,t$_2$…..t$_n$}* be a set of all transaction of subset of *I*. Each transaction *t* is assigned a number which is *TID (Transaction ID)*. Different transaction together makeup the transaction set *A* and it contains a transaction association rules that found in database.

Now, let *F* be a set of items. The transaction *t* is said to contain F if and only

$$F \subseteq t.$$

Now, mathematically, an association rule will be an implication of the form

$$F \Rightarrow B$$

Where both F and B are a subset of A and F ∩ B = ∅ ("the intersection of sets A and B is an empty set").

3.2.2.1 *Support degree:*

Support (sometimes called frequency) is simply A probability that a randomly chosen transaction t contains both items sets A and B. Mathematically, support $(A \Rightarrow B)\ t = P\ (A \subset t \land B \subset t)$ = total of transactions of transactions containing both A and B we will use A (Zhou et al. 2014).

3.2.2.2 *Confidence degree:*

Confidence (sometimes called accuracy) is simply a probability that an item set B is purchased in a randomly chosen transaction t given that the item set A is purchased. Mathematically, confidence $(A \Rightarrow B)\ t = P(B \subset t \mid A \subset t)$ total of transactions containing A of transactions containing both A and B we will use a simplified notation that confidence $(A \Rightarrow B) = P\ (B \mid A)$.

3.2.3 *Apriori* algorithms:

Apriori algorithms for association rules, were designed to work on large databases to find the relationship between items. It has been developed in 1993 by Agrawal, Rakesh, Tomasz Imieliński, and Arun Swami (Zhichun and Fengxin, 2008). An advanced variant of these algorithms allows finding association rules on large data and allows implication outcomes that consist of more than one item. The *Apriori* algorithm is based on the fact if a subset B appears T times, any subset F that contains F will appear T times. If F does not pass the minimum support, it is discarded a priori. In this study, I applied *Apriori* algorithm on the dataset.

The results show that the reader who is borrowing the book A also is going to borrow book B according to the rules as shown below:

Table 1: Support and Confident of Rules

Rules	Support	Confident
1278694=>1280470	4.1	13.2
1278677=>1279052	6.3	19.8
1279047=>1411279	3.9	4

	3.5	8.3
1279056=>1279052		
1285064=>1285063	7.1	2.3
1285062=>1285063	1.4	2.5

- 4.1 % of users who borrowed *Algorithms and Data Structures* book, there is 13.2 % possibility of borrowing *Study on the Open Source Digital Library Software* book.

- 6.3% of users who borrowed Data Mining: Concepts and Techniques book, there is 19.8 % possibility of borrowing Data Mining: Practical Machine Learning Tools and Techniques book

- 3.9% of users who borrowed *Computer Viruses and Malware*, there is 4 % possibility of borrowing *User Acceptance of Information Tec*hnology book.

- 3.5 % of users who borrowed Data Mining: Technologies, Techniques, Tools, and Trends book, there is 8.3% possibility borrowing Data Mining: Practical Machine Learning Tools and Techniques book.

- 7.1 % of users who borrowed The Johns Hopkins University Press book there is 2.3 % possibility of borrowing China: Renaissance of the Middle Kingdom (Odyssey Illustrated Guides) book.

- 1.4% of users who borrowed Chinese History Chart (English and Chinese Edition) book, there is 2.5% possibility of borrowing China: Renaissance of the Middle Kingdom (Odyssey Illustrated Guides) book.

3.2.4 FP-Growth Algorithms:

FP-Growth (frequent-pattern growth) algorithm is an improved version of the *Apriori* algorithm. It has been improved by Jiawei Han and others. (Zhichun and Fengxin, 2008). FP-Growth algorithm attempts to find large item sets without candidate generation. This algorithm does not produce the candidate item sets in mining process and it still improves the mining efficiency. It creates FP-tree and contains all the dataset and scanning the datasets twice. (Zeng et al., 2015). There are many concepts related to FP-Growth, the FP-tree, the conditional pattern base and the condition tree. FP-Tree can put data items into a tree after support and inserts data items in each transaction with NULL as its root by descending items of each node. The conditional pattern base contains the set of suffix path in FP-Tree. And condition tree uses principles of the formation of FP-Tree from conditional pattern (Song and Wei 2011).

Application of these algorithms to the specific process of Library borrowing records can be described into the following steps:

- Access the dataset of library borrowing records and set the minimum support which is 2.

- Scan the database DB to get the frequent item sets B and the support for each item.

- Create the root of FP-Tree noted T, and marked as 'Null for all datasets
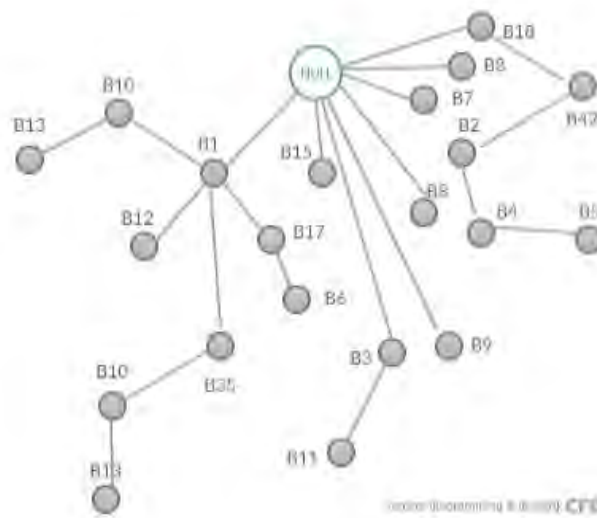
Figure 18:FP-Tree

The condition base pattern starts from the node with minimum support value is [B42] and excludes the node from maximum support value [B11]. The frequent pattern can be identified as:

B5= {B4, B5:2} {B17, B5, 2} {B4, B17, B5:4}

B4 {B5, B4:2}

B3 {B2, B3:2}

3.3 Classification.

Classification is a method of data analysis that extracts important data classes from the model. Such method, called classifiers, predict categorical (discrete, unordered) class labels. For instance, bank loan applications can be built as a classification model as either safe or risky

(Han et al., 2012). The classification model is target function which is useful for descriptive and predictive models.

3.3.1 Classification algorithms:

The classification algorithms investigated for this task, include Naïve Bayes, J48 and *K-NN* and applied on a dataset of 2816 users' borrowing records from Knowledge to action digital library developed. The dataset of users is shown in Figure 19:

ID	Booking D	Return Date	Member No	Media _ID	Media ID	Subject
361108	4/6/2017	4/13/2017	408	B3	1445953	2
361109	4/6/2017	4/13/2017	408	B4	1445827	3
359373	4/5/2017	4/12/2017	406	B6	1432745	5
359381	4/5/2017	4/12/2017	406	B7	1441573	5
354868	########	4/7/2017	405	B10	1317559	6
345715	########	3/30/2017	405	B12	1423064	2
345716	########	3/30/2017	405	B13	1423064	2
285189	########	2/4/2017	241	B22	1285060	2
285190	########	2/4/2017	121	B23	1285062	2
285191	########	2/4/2017	131	B24	1285063	2

Figure 19:Dataset of users borrowing records.

3.3.2 Naïve Bayes:

The Naïve Bayes model is based on the frequency table and Bayes' theorem with independence assumptions between predictors. Let *A* be a data sample whose class label is not known and let *H* be some hypothesis, such that the data sample *A* may belong to a specified class *B*. Bayes theorem is used for calculating the posterior probability *P(B/A)*, from *P(B), P(A),* and *P(A/B)*. Where *P(B/A)* is the posterior probability of target class. *P(B)* is called the prior probability of class. *P(A/B)* is the likelihood, that it is the probability of predictor of given class. *P(A)* is the prior probability of predictor of class.

$$P(A \mid B) = \left| \frac{P(A \mid B) \cdot P(B)}{P(A)} \right.$$

The Naive Bayes classifier works as follows:

1. Let D be the training dataset associated with class labels. Each tuple is represented by n-dimensional element vector, *A=(a1, a2,a3,.....,an).*

2. Consider that there are m classes *B1, B2, B3...., Bm*. Suppose that we want to classify an unknown tuple A, then the classifier will predict that *A* belongs to the class with higher posterior probability, conditioned on *A*. i.e., the Naive Bayesian classifier assigns an unknown tuple *A* to the class B if and only if $P(B_i/X) > P(B_j|A)$ *For 1 ≤ j ≤ m*, and $i \neq j$, above posterior probabilities are computed using Bayes Theorem (Joshua et al.,2016).

The Naïve Bayes algorithm has many advantages, such as good performance and short time needed for evaluating test data. However, there are some disadvantages such as accuracy, which is lower as compared to other classifiers, and needs a large amount of data to get good results (Patil and Sherekar, 2013). After applying the Naïve Bayes algorithms on the dataset of borrowing records, the results show that the accuracy obtained is 95% which can be considered as high accuracy. Other performance measures, including the ROC (Receiver operating characteristic), False positive rate, True positive rate, precision, recall, F-measure, and the confusion matrix are also satisfactory. From figure 20 the results show that the ROC by class is between (0.99 – 1.00). and Recall is between (0.64 – 1.00). As can be seen from the confusion matrix, there are very few missclassifications, given that it is a multiclass unbalanced dataset with few data records.

TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
0.993	0.094	0.930	0.993	0.961	0.909	0.990	0.992	1
0.642	0.004	0.955	0.642	0.768	0.759	0.979	0.901	2
1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	3
1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	4
1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	5
1.000	0.005	0.971	1.000	0.985	0.983	1.000	0.997	6
0.950	0.054	0.951	0.950	0.946	0.916	0.992	0.982	

Figure 20: *Performance of Naïve Bayes Classifier.*

```
=== Confusion Matrix ===

    a    b    c    d    e    f   <-- classified as
 1569   11    0    0    0    0 |   a = 1
  118  233    0    0    0   12 |   b = 2
    0    0   48    0    0    0 |   c = 3
    0    0    0  144    0    0 |   d = 4
    0    0    0    0  296    0 |   e = 5
    0    0    0    0    0  398 |   f = 6
```

Figure 21:Confusion Matrix of Naïve Bayes.

3.4 Performance Analysis.

In this section, the performance analysis of the Naïve Bayes classifier used for digital library dataset developed is discussed. The performance was assessed with different performance metrics, including the ROC (Receiver operating characteristic), False positive rate, True positive rate, precision, recall, F-measure, and the confusion matrix A confusion matrix contains information about actual and predicted classification. (Patil and Sherekar, 2013). For example, when a naïve Bayes classifier, its performance can be assessed by calculating efficiency on correct and incorrect instance classifications and summarised with a confusion matrix. For the confusion matrix as shown in the Figure 21, true positive (TP) for class a (1) which is subject in information technology is 1569 while false positive (FP) is 11. In class b (2) which is subject in science is 233 TP and 120 FP. In class c (3) which is subject in history (TP) is 48 and there is no (FP). In class d (4) which is subject in medicine (TP) is 144 and no (FP). In class e (5) in subject Music (TP) is 296 and no (FP). Last class is f (6) in education (TP) 398 and no (FP).

So, 1569+233+48+144+296+398= 2688 represents the correct instances classified and the incorrect is 11+118+12=141. Now we calculate the rate of TP and FP as below:

$$\text{True positive rate (TPR)} = \frac{diagonal\ element}{Sum\ of\ relevant\ row}$$

$$\text{False positive rate (FPR)} = \frac{Non\ diagonal\ element}{Sum\ of\ relevant\ row}$$

TPR for class a (1) = 1569 (1569+11) = 0.99

FPR for class a (1) =118(118+233+12) = 0.325

TPR for class b (2) =233(118+233+12) = 0.641

FPR for class b (2) =11(1569+11) = 0.006

TPR for class c (3) =48/ (48) = 1

FPR for class c (3) =0/ (1569+11) = 0

TPR for class d (4) =144/ (144) =1

FPR for class d (4) = 0/ (1569+11) = 0

TPR for class e (5) =296/ (296) =1

FPR for class e (5) = 0/ (1569+11) = 0

TPR for class f (6) = 398/ (398) = 1

FPR for class f (6) =0/ (1569+11) = 0

3.5 Decision Tree.

A decision tree is a classifier model that includes a root node, branches, and leaf nodes. The internal node represents a test on an attribute, the branch represents the outcome of a test, and the leaf node holds a class label. The topmost node in the tree represents the root node. The decision tree classification technique has two phases, tree building and tree pruning. The tree building is top-down approach, and trees pruning is button up approach (Joshua et al.,2016). Decision tree has some advantages, including being simple, fast, has high accuracy and takes less storage on memory. The limitation of this algorithm is considered complex for some problems to construct the decision tree. There are many decision tree-based algorithms like ID3, C4.5, C5.0, J48. We applied J48 decision tree algorithm to the Library borrowing records, resulting in an accuracy of 96%, higher than the Naïve Bayes classifier, as shown in the Figure 22. The ROC by class is between (0.99 – 1.00). and Recall is between (0.744 – 1.00). Figure 23 shows the confusion matrix for this classifier.

TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
0.987	0.065	0.951	0.987	0.969	0.928	0.995	0.996	1
0.744	0.008	0.931	0.744	0.827	0.811	0.984	0.926	2
1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	3
1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	4
1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	5
1.000	0.005	0.971	1.000	0.985	0.983	0.999	0.991	6
0.960	0.038	0.959	0.960	0.958	0.933	0.995	0.987	

Figure 22:Accuracy by class for J48 decision tree classifier

```
=== Confusion Matrix ===

    a     b     c     d     e     f    <-- classified as
 1560    20     0     0     0     0 |   a = 1
   81   270     0     0     0    12 |   b = 2
    0     0    48     0     0     0 |   c = 3
    0     0     0   144     0     0 |   d = 4
    0     0     0     0   296     0 |   e = 5
    0     0     0     0     0   398 |   f = 6
```

Figure 23:Confusion Matrix of J48 decision tree classifier

From the confusion matrix shown in figure 23, the correct instances classified are 2617 and the incorrect is 113.

Average TP rate = 0.96

Average FP rate = 0.038

3.6 K- Nearest Neighbour Classification (K-NN).

The K-Nearest Neighbour Classifier algorithm is considered the simplest machine learning algorithms. It is based on a similar principle, relying on close vicinity. It is also called lazy learning algorithm, because it stores all the training samples and does not build a classifier until a new, unlabelled sample needs to be classified.

1-nearest neighbor 2-nearest neighbor 3-nearest neighbor
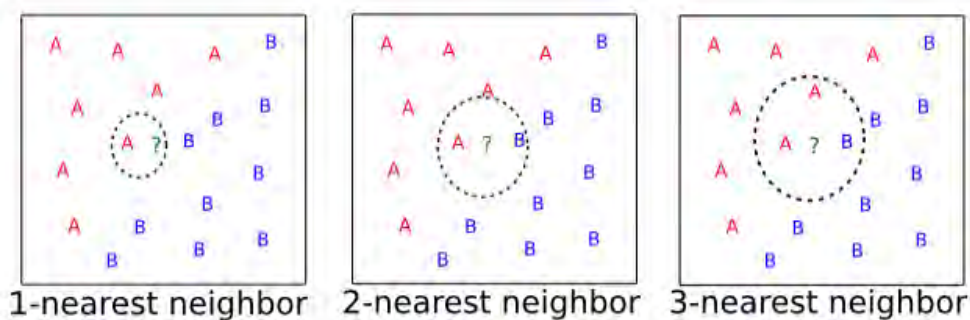
Figure 24:Example of K-NN Classifier (Joshua et al.,2016).

The K-NN classifier works with set value of K (K representing the number of neighbourhood samples), and then calculates the distance between input sample and training samples, sorts the distances, and then takes top K- Nearest Neighbours. (Joshua et al., 2016). The K-NN classifier was applied to the Library borrowing records performance achieved is 96.6% accuracy, with other performance metrics achieved is as shown in the Figure 25 and the confusion matrix in the Figure 26.

TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
0.982	0.041	0.968	0.982	0.975	0.943	0.999	0.999	1
0.843	0.011	0.916	0.843	0.878	0.862	0.996	0.973	2
1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	3
1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	4
1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	5
1.000	0.002	0.985	1.000	0.993	0.991	1.000	0.999	6
0.970	0.025	0.969	0.970	0.969	0.949	0.999	0.996	

Figure 25: Accuracy by class for K-NN classifier.

The average ROC is 0.999. and average Recall is 0.97. From the confusion matrix, the correct instances classified are 2783, and the incorrect is 46 instances.

Average TP rate = 0.97

Average FP rate = 0.025

```
=== Confusion Matrix ===

    a     b     c     d     e     f    <-- classified as
 1552    28     0     0     0     0 |    a = 1
   12   345     0     0     0     6 |    b = 2
    0     0    48     0     0     0 |    c = 3
    0     0     0   144     0     0 |    d = 4
    0     0     0     0   296     0 |    e = 5
    0     0     0     0     0   398 |    f = 6
```

Figure 26:Confusion Matrix of K-NN.

3.7 Performance Comparison:

A comparison of performance achieved by three different machine learning classifiers, the Naïve Bayes, J48 decision tree and K-NN classifier on 2888 library borrowing records dataset, shows that the correct instances classifier by the naïve Bayes are 2618, by J48 are 2617, and by K-NN are 2783 instances, with average F-measure, Precision and Recall across three classifiers shown in the Figure 27.

Classifier	F measure	Precision	Recall
Naïve Bayes	0.946	0.951	0.993
J48 Decision Tree	0.958	0.959	0.960
K-NN classifier	0.969	0.969	0.979

Figure 27:Performance Comparison of three classifiers for digital library dataset.

As can be seen in the Figure 27, the K-NN classifier performs best with average F-measure, Precision and Recall, higher than the Naïve Bayes and J48 Decision Tree classifiers.

Based on the results, it can be concluded that the research into applying association rules algorithms and traditional established classification algorithms on digital library borrowing records have resulted good performance. Also, from the results, we can have insight into the most borrowed book, and the never borrowed book in the records, that can help library management to improve their services and give them constructive advice.

Moreover, the number of copies available for the most borrowed book can be increased to support students' access to learning resources. Further, it was found that the *Apriori* algorithm is slower than the FP-Growth and the counting method iterates through all the transactions each time. On the other hand, the FP-Growth algorithm reads the file twice, as opposed to *Apriori* which reads it once for every iteration. As a result, FP-growth algorithm can implement in a short time and get high accuracy when analysing the association rules of library borrowing records. In addition, the results of K-NN, Naïve Bayes and J48 show that the K-NN has 96.6% accuracy, highest among all three, and J48% has lesser error rate as compared to K-NN and Naïve Bayes. J48 is also easy for humans to understand because its' structure is represented in the form of [IF-THEN] rules. Next Section presents a brief user experience survey of digital library web portal developed.

3.8 Evaluation of overall digital library portal by usability survey:

This section discussed the user experience survey developed for assessing the usability of the data mining based digital library web portal. There were 5 questions in the survey, and 85 people participated in the survey.

Question 1: The question one asked whether the users found the registration page easily. The results show that 40% of respondents disagree however, 60% agreed that it is easy (figure 28).
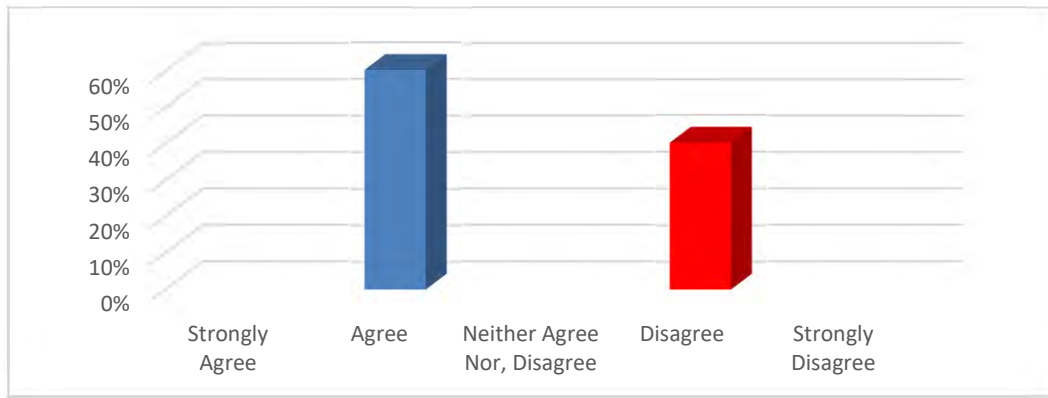
Figure 28:Respondent's Views of how easily they can use reregistration page.

Question 2: The question two asked if they can managing their page navigations easily, and the results show that 70% of users can manage their page easily while 20% disagree and 10% they did not know, as shown on the Figure 29.
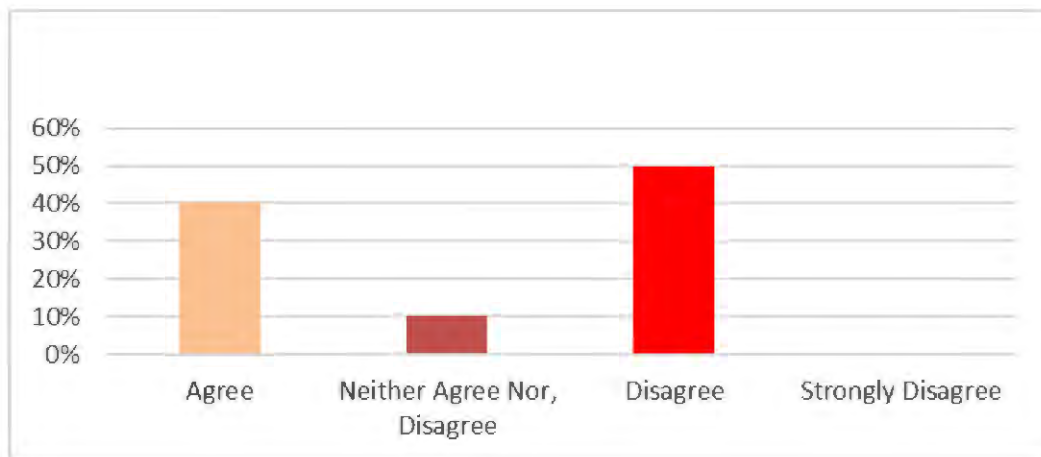
Figure 29:Respondent's Views of how easily they can be managing their page.

Question 3: This question was to find out if the users of digital library can easily borrow items from the digital library. The result shows that 60% found it easily while 30% found it difficult and 10% they did not know. The Figure 30 shows this distribution.
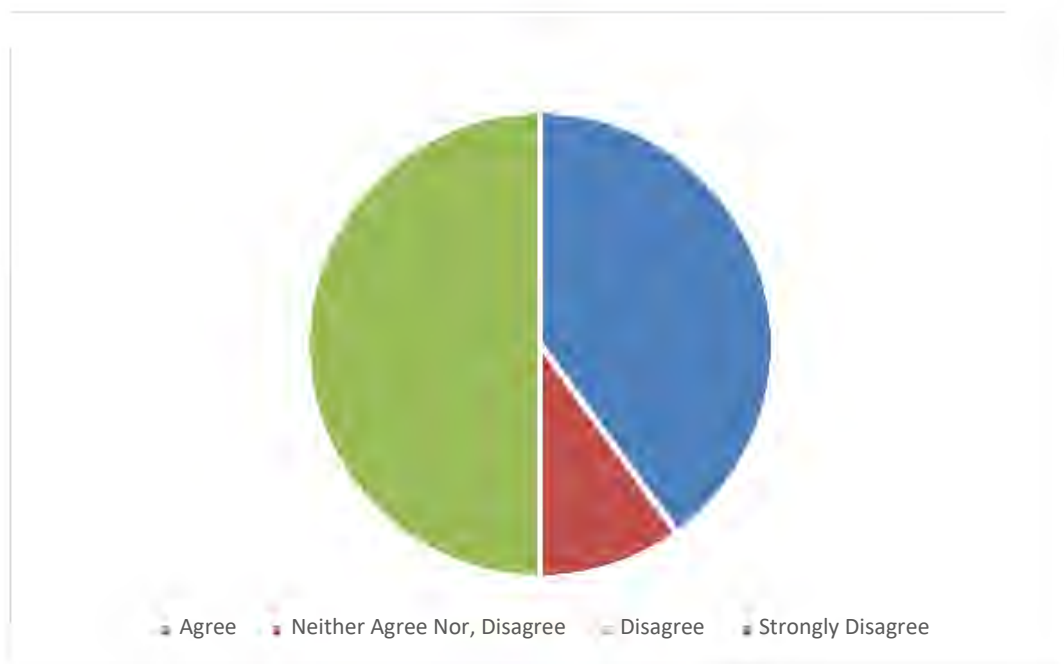
Figure 30:Respondent's Views of how easily they can be borrowing items from library.

Question 4: The purpose of this question was to find out, if the interface graphics ,and colour scheme of the digital library web pages was pleasant, and response showed 40% of the participants agreed that the interface graphics and colour scheme were good, while 50% disagree and 10%, they did know, as can be seen in the Figure 31 below:
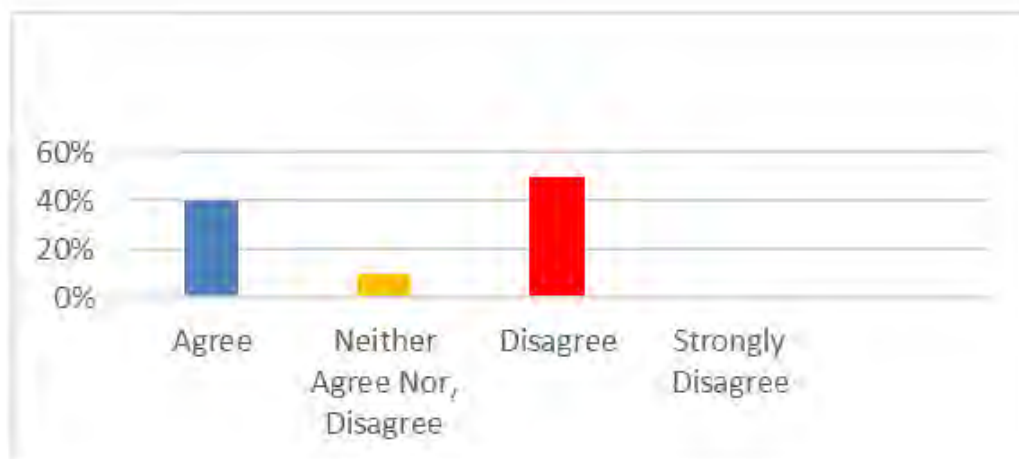
Figure 31:Respondent's Views of the interface.

Question 5: The final question in the survey was to check if the users could seek help easily from the web portal in using it. The response was that 60% of users agreed they could ask librarians if they faced any problems, while 20% disagree and 30% they did not, as depicted in the Figure 32 below:
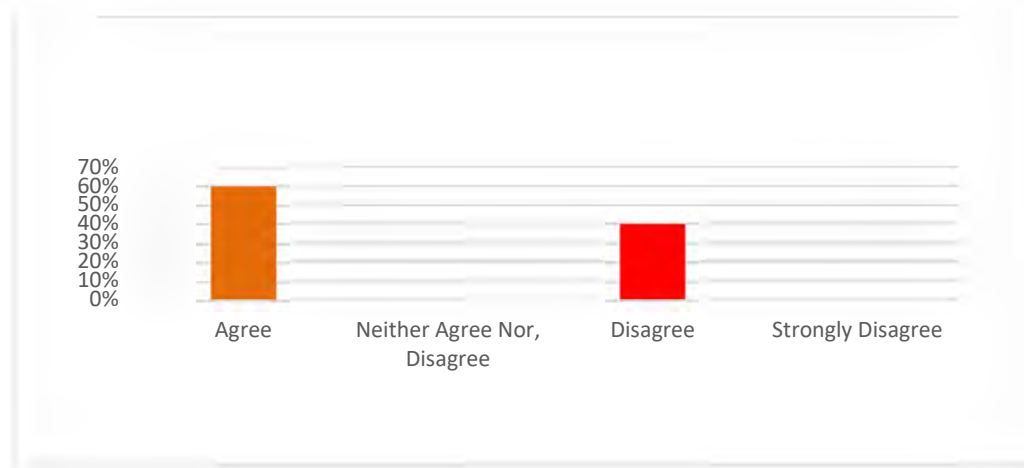
Figure 32:Respondent's Views of they can ask the librarian if they faced any problems.

Overall, the in-house proof of digital library *knowledge to action* web portal has good satisfaction from users, but it needed few improvements on registration process, and other advanced features including personalized recommendations. However, it helped in getting an insight on the challenges involved in design and development of different aspects of complex search and retrieval systems, such as digital libraries, and how different technologies needs to be integrated to provide enhance services in terms of usability, visual interaction, and personalized recommendations.

As the aim of the thesis to enhance the personalisation aspects in terms of providing better recommendations for complex application scenarios, the next step in the investigations was to examine the current state of the art in recommendations, particularly in the real-world digital library systems context for different multilingual settings, including Australia and Saudi Arabia. This has helped in designing the computational framework for recommender systems for this thesis. The next Section presents results from these findings.

3.9 Study on the current status of recommender systems for digital libraries in Australia and Saudi Arabia.

The findings from the study aimed at use of recommender systems in real digital libraries in Saudi Arabia and Australia, and to determine if they deployed the recommendation system and allowed user preferences to be entered, in the deployed digital library systems. The study design involved collection of data from publicly accessible website of Australian university libraries (https://www.australianuniversities.com.au/libraries/), and by accessing their catalogues.

Table 2 shows a comparison of facilities available for different Australian libraries. Further, for Saudi Arabia, the access to digital libraries works in a slightly different manner, as Saudi Arabia has one digital library called Saudi Digital library SDL, which connects all university libraries in one place, as presented in Table 3.

Table 2: Australian Libraries.

Library name	Recommender System	User's preferences
1. Australian Catholic University	NO	NO
2. Australia National University	NO	NO
3. Bond University	NO	Yes, it has personalized profile
4. Central Queensland University	NO	Yes, it has personalized profile
5. Charles Darwin University	NO	Yes, it has personalized profile
6. Charles Sturt University	NO	Yes, it has personalized profile
7. Curtin University	Yes	NO
8. Deakin University	NO	NO
9. Edith Cowan University	NO	NO
10. Federation University	NO	NO
11. Flinders University	NO	NO
12. Griffith University	NO	NO
13. James Cook University	NO	NO
14. La Trobe University	NO	Yes, it has personalized profile

15. Macquarie University	No	Yes, it has personalized profile
16. Monash University	Yes, it has suggestion	NO
17. Murdoch University	NO	NO
18. Queensland University	NO	NO
19. RMIT University	NO	NO
20. Southern University	Yes, it has suggestion	NO
21. Swinburne University	NO	NO
22. University of Adelaide	Yes, it has suggestion	Yes
23. University of Canberra	NO	NO
24. University of Melbourne	Cannot Access	Cannot Access
25. University of New England	NO	NO
26. University of New South Wales	NO	NO
27. University of Newcastle	NO	NO
28. University of Notre Dame	NO	NO
29. University of Queensland	NO	Yes, it has personalized profile
30. University of South Australia	NO	NO
31. University of Southern Queensland	NO	NO
32. University of Sydney	Yes, it has related searches by site	NO
33. University of Tasmania	NO	NO
34. University of Technology Sydney	NO	NO

Library Name	Recommender System	User's Preferences
35. University of Sunshine Coast	NO	Yes, it has personalized profile
36. University of Western Australia	NO	NO
37. University of Wollongong	NO	NO
38. Victoria University	Cannot Access	Cannot Access
39. Western Sydney University	Yes, Suggestion	NO
40. Australian National Library	NO	NO

Table 3: Saudi Arabia Libraries

Library Name	Recommender System	User's Preferences
Saudi Digital Library	NO	NO
King Fahad National Library	NO	NO
King Abdul Aziz Public library	NO	NO
King Fahad Public Library	NO	NO

3.9.1 Results:

After analysing university libraries and public libraries website in Saudi Arabia and Australia, to assess the current status of digital libraries, in terms of availability of providing personalized recommendations and options to include user preferences, it was found that 85% of Australia libraries do not have recommender systems and 15% have suggestions for sources (Figure 33), that relate to previous search results, including Flinders University, Monash University and University of Sydney as shown in the Figure 34 below:

Recommender System in Australian libraries

100

0

YES   NO
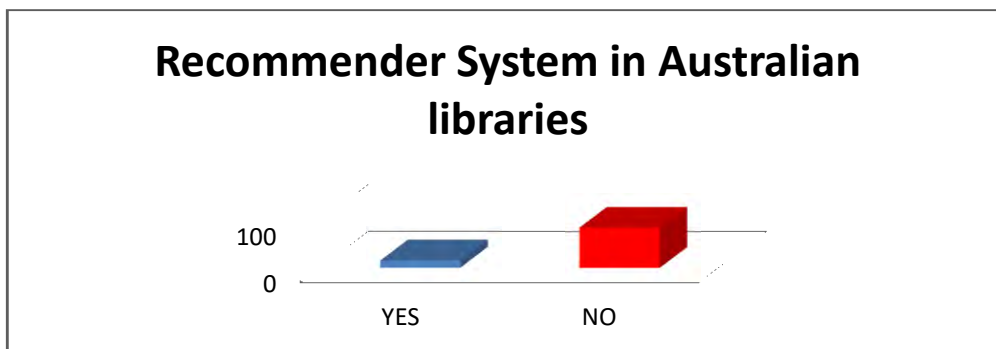
Figure 33:Recommender System in Australian libraries
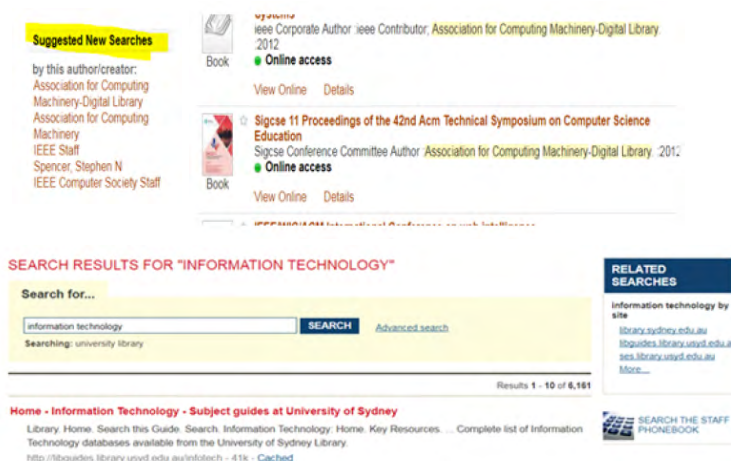
Figure 34:Suggestions provided for related search by users for the Australian library.

For libraries in Saudi Arabia, none of the libraries are equipped with recommender systems and have options for including user preferences for personalized search and retrieval, because the overarching SDL library (under which separate libraries are implemented), does not have recommender system and options for including user's preferences.

However, this type of suggestions cannot qualify as providing recommendations because, it does not include the user profile, and just uses the items and related keyword in the items, and mostly general related topics. Also, 77% of Australian libraries do not have options of including user's preference and 22.5% have personal customized search options, when the user searches for an item on the library website, including Bond University, Central Queensland University, University of Adelaide, and University of the Sunshine Coast, depicted in the Figure 35 and the Figure 36 below:
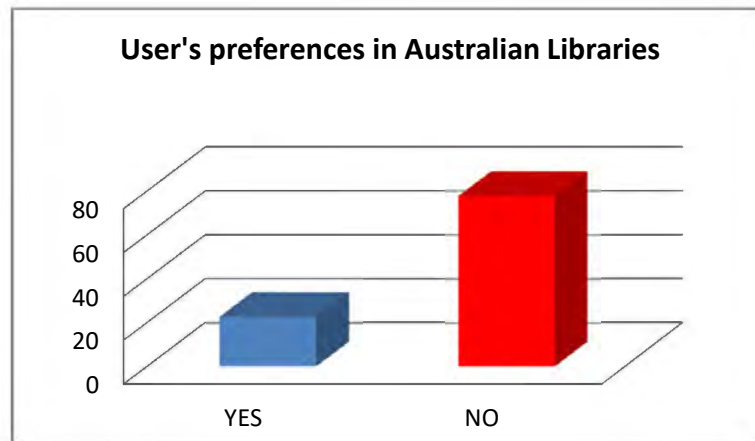
Figure 35:User's preferences in Australian Libraries

Figure 36:Personalization results for a resource search in Australian digital library

The personalized users result is not the same inclusion of user's preferences because user's preferences setup is done early in user profile, and it stays static, and does not take into account, dynamic changes in user search behaviour and updated preferences that need to be included for

every new search activity. Next Section discussion the status of recommendations and inclusion of user preferences for University of Canberra library system.

3.10   University of Canberra Library.

The University of Canberra library has a small digital library that helps students and staff to explore the resources online. The purpose of the digital library of the University of Canberra is to provide a range of high-quality information resources and services to the University community. In technical side, the UC library uses library management system called *Millennium Acquisition System*, which got replaced by a new library management system called *Alma* since December 2017. This system supports library collaboration and helps libraries to optimize user experience and collection with rich analytics. Mr West was confirmed during an interview (C West 2017, personal communication, 6 June), the system does not have recommender system and options for including user's preferences for personalisation, but it better than the previous system.

3.11 Chapter Summary:

This chapter has presented preliminary proof of concept experimental work in understanding the challenges associated with design and development of digital library systems with the recommendation and personalisation features, and status of existing digital library systems deployed for different multilingual International settings. A novel experimental research testbed was developed based on open-source technology tool called *Librarika* and simple and adaptive data mining approaches for retrieving the resources, without any recommendation or personalisation options. The testbed, the prototype web portal implementation, and its evaluation helped provide an insight into complex search and retrieval requirements, and lack of advanced features in existing digital library systems deployments. The results demonstrate that 85% of Australian libraries do not have recommender system and 15% have provide suggestions for similar sources that relate to the search results. In Saudi Arabia, there is no recommender system and options for including user's preference. Very few Australian libraries have very preliminary personalisation features, and not dynamically updated. Hence, it could be confirmed, that currently, no recommendation systems are being used in digital libraries in Saudi Arabia and Australia.

However, this is in contrast with existing large-scale resources search and retrieval systems, including Amazon, Google, Netflix, and other similar platforms, where the status of

recommendations and personalisation embedded in search and retrieval is better than the digital library settings.

Therefore, to be able to develop robust and efficient computational frameworks and algorithmic approaches for better recommendations and personalisation options, it would be worthwhile to investigate these alternate settings, where large databases are publicly available. The improved approaches can then be translated and extended to digital library type settings, as they are similar resources search and retrieval systems.

The rest of the Chapters in this thesis are based on this rationale, with each Chapter focussing on a unique aspect of enhancing the recommendation systems performance and robustness.

Chapter 4 *Recommendation System Framework based on Machine and Deep Learning.*

4.1 Introduction.

This Chapter presents the novel computation framework developed in this thesis for recommendation systems based on Deep Machine Learning. For a baseline comparison, traditional methods for recommender systems based on collaborative filtering and content-based recommendations were also examined. Collaborative Filtering (CF) and Content –Based Recommendation (CBR) are among the most popular recommender system techniques. The use of Collaborative Filtering as a new approach to searching digital libraries introduces many problems in existing search systems. While the CBR search algorithm is based on matching keywords in a user's query to the keywords appearing in the full texts, CF algorithm which is one of the traditional algorithms for implementing recommendations, and analyses user's interests and the area of interesting information from all the information that user can access, share, and evaluate (Webster et al., 2004). Many studies agree that the CF has many advantages for building recommender systems, though they do not necessarily succeed in outmatching user's preference because one point of view will always dominate another in a community. For digital library systems application context, there are many challenges in using CF algorithm. The recommender system needs large dataset and user-item collaboration matrix for collaborative filtering, which could be extremely large and result in deterioration of performance. Moreover, CF method is based on users' past performances. So, a new user cannot take the advantages of recommender system, until the person uses the system, and the system records the individual's interactions with the items, to use them for the next time. Also, CF is based on the discrete set of description quality rather than its rating. Further, many research studies have reported that CF has a problem with a large data sparse rating. However, one of the studies based on clustering approach have reported

promising outcomes for a collaborative recommendation problem (Wang, 2012). Collaborative filtering involves filtering a set of similar information without considering user's information content. According to the empirical analysis of predictive algorithms for collaborative filtering, the CF techniques can be memory-based or model-based techniques (Breese et al., 1998). Memory-based collaborative filtering as described by Breese, Heckerman, & Kadie involves finding users among active users, who apply their preferences to give the rating for popular information (1998). Memory-based collaborative filtering seems to have many advantages, including suitability for large datasets, easy updating of the database, and easy handling of new data. However, there are several limitations of the memory-based collaborative filtering technique, mainly being very slow, because it uses entire database every time and makes data unreliable, leading to inaccurate prediction, if the active user has no items with other users' recommendations (Zhang et al., 2017). Another type of approach used for recommender systems is a hybrid recommender approach, which combines the content based and collaborative filtering approach to obtain better recommender system performance (Agarwal and Chen, 2011). The content-based recommender system mechanism analyses description of items, user's profile, and preferences to find the coordination in content. It uses the heuristic method of classification algorithms to make the recommendations. The content-based method on the other hand requires enough information to build the reliable a classifier, and content-based techniques suffer from the start-up problem. While collaborative filtering can make the recommendations without content information, content-based techniques have problem of overspecialization, which means they can only recommend items with high rating history. Many studies reported that the cold start is the common problem in recommender systems which can be resolved by using hybrid methods (Betru, et al., 2017). Hence traditional methods used for recommender system address problems for a variety of different situations, and can be classified as shown in the Figure 37 below:
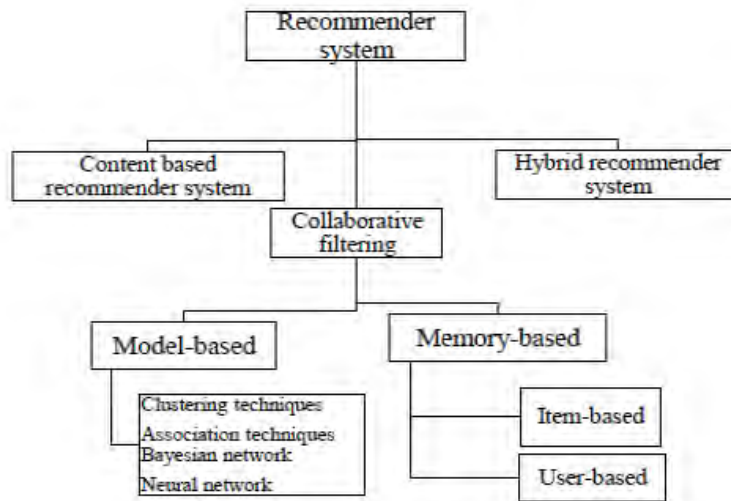
Figure 37:Traditional approaches for Recommendation systems

While the traditional methods do not rely on learning representations and interactions between users and items, using machine learning approaches, particularly recent deep learning can address the problems, corresponding to cold starts and insufficient interaction data. Deep learning approaches can learn representations of data based on supervised or unsupervised approaches, iteratively in several layers of proceedings (Zhang et al., 2017). One study has listed six categories of how machine and deep learning approaches have been used for different recommendation system tasks, including raw retrieval tasks, manual selection, statistical summarization, attribute-based, item-to-item correction, and user-to-user correction tasks (Schafer et al., 2001). The item-to-item correction task uses the content-based approach, and user-to-user correction method uses the collaborative filtering approach. One of the most interesting study is about graph-based recommender approach for digital library that combines content-based and collaborative approaches together. This study found that the system gained improvement with respect to both, precision and recall by using a hybrid approach, or by combining two approaches (Huang et al., 2002). According to Nart and Tasso most of recommender systems use collaborative filtering techniques, which are based on filtering resources based on the users' ratings or feedback (2014). However, there are two issues associated with collaborative filtering techniques, which include

the correlation between several users who rate the items and the number of items. One of the issues is based on the problem with the number of users rating the items, being much larger, than the number of users who maintain sustained similar interest or liking for items over the time. The second problem is cold–start, associated with new items or old items that received very few ratings, and will be unlikely to be recommended in a collaborative system, because they have no ratings (Nart and Tasso 2014). These issues can be addressed by using approaches based on learning representations, including deep learning-based approaches, which can be promising.

Deep learning is the process of learning nonlinear features and functions from complex data, and a recent form of machine learning technique used for learning representations from data. It has resulted in good performance in other similar complex problem solving, including speech recognition, image recognition and face detection problems (Taigman, et al., 2014). Deep learning model development is based on learning of non-linear features and key functionalities in complex data. The deep learning concept surpasses most of the traditional considerations in the management of data (Zhang et al., 2017). The rectilinear activation functions used in the deep learning algorithms are quite powerful in yielding the specific results relative to the expectations (Zheng et al., 2015). Using deep learning-based approaches for solving complex search and resource retrieval problems in Digital libraries application contexts, can allow better recommendation and improved personalisation (Li et al., 2017), (Dwork et al., 2015). Two different approaches for recommendation systems based on Deep Machine Learning have been proposed in this thesis and is described in detail in the rest of the Chapter.

4.2 Deep Learning Based Approaches for Recommendation Systems.

In the recent years, the rate at which data and information is being captured is increasing so fast leading to the problem of information overload (Gantz and Reinsel, 2012).

However, harnessing this large data using efficient big data analysis techniques using novel approaches in machine and deep learning can lead to enhancement in performance of search and retrieval systems (Adomavicius and Tuzhilin , 2005). Wang et al reported a recommendation system as essential component in search and retrieval applications with an aim to analyse use's profiles (2015). However, the aim for recommendation systems in digital library contexts, needs to be more ambitious, and have ability to recommend resources to users, not just based on users' interest, but also recommend unknown resources of interest to users. In addition, digital library recommendation system needs to provide such recommendations by selecting resources form massive datastores of information. Machine learning approaches, the recent deep learning approaches can learn representations from data at multiple levels from these massive data stores and can capture non-linear and non-trivial user-item relationships (Covington et al., 2016), (Mukherjee et al., 2016). However, deep learning method in digital library recommender system has not been well explored yet and most of the real-world systems using collaborative filtering or content-based filtering approaches.

In this Chapter, two different Collaborative Filtering models based on deep learning have been developed, the Multistage Deep Neural Network (CFMDNN) algorithm and Collaborative Filtering based Deep Learning (CFDR) algorithm for addressing the problems with recommender systems and improving their performance is presented. CFDR and CFMDNN learn to model complex structures of user and item interactions using matrix factorization for extracting latent feature representations either separately (CFDR) or jointly (CFMDNN), with joint learning architecture (CFMDNN) with a capability to learn the joint space at a deeper level, through several stages of learning and discovery, without a need for feature engineering or manual handcrafting needed in extracting and combining the latent feature sets. However, they differ in joint modelling capability of interaction between users and ratings, and in turn in their ability for deep discovery in joint latent feature space. Both the models were built used K-fold cross-

validation for training and evaluation and used explicit rating prediction accuracy and the Mean Absolute Error (MAE) as the performance metrics for the evaluation. Also, by using dropout as a regularization technique to model complexity was reduced and overfitting was avoided. By applying densely, fully linked feed-forward neural network, and Adam optimization algorithm (an extension to stochastic gradient descent that has recently seen broader adoption for several deep learning applications), it was possible to achieve improve the performance of the system. The details of the experimental work and the two models are described in the next Section.

4.2.1 Experimental Setup for CFDR and CFMDNN Models.

In this Chapter section, the details of the experimental setup to evaluate the effectiveness of our proposed CFDR and CFMDNN algorithms is described and compared with the traditional Collaborative Filtering (CF) approach (which doesn't use deep learning) as a baseline reference algorithm.

Four different datasets were used, with each dataset with different rating system as outlined below:

- Movielenes 20M dataset with 1- 5 rating scale,
- Amazon digital Music dataset with 1-5 rating-scale,
- Book-Crossing dataset with 0-10 Rating-scale and
- Amazon books dataset with 1-5 rating scales.

When the system has learned the user's previous ratings of items it will give recommendation higher for those items in the future for the user. Besides, the system needs to know the relationship between users and item to make predictions more accurate. It can be determined that by applying the Utility Matrix, giving for each user-item pair, a value that represents the degree of preference of that user for that item. The latent features extracted from the collaborative filtering stage with matrix factorization is processed through a different deep learning algorithmic pipeline for CFMDNN and CFDR architectures shown in the Figures 38 and 39 below:
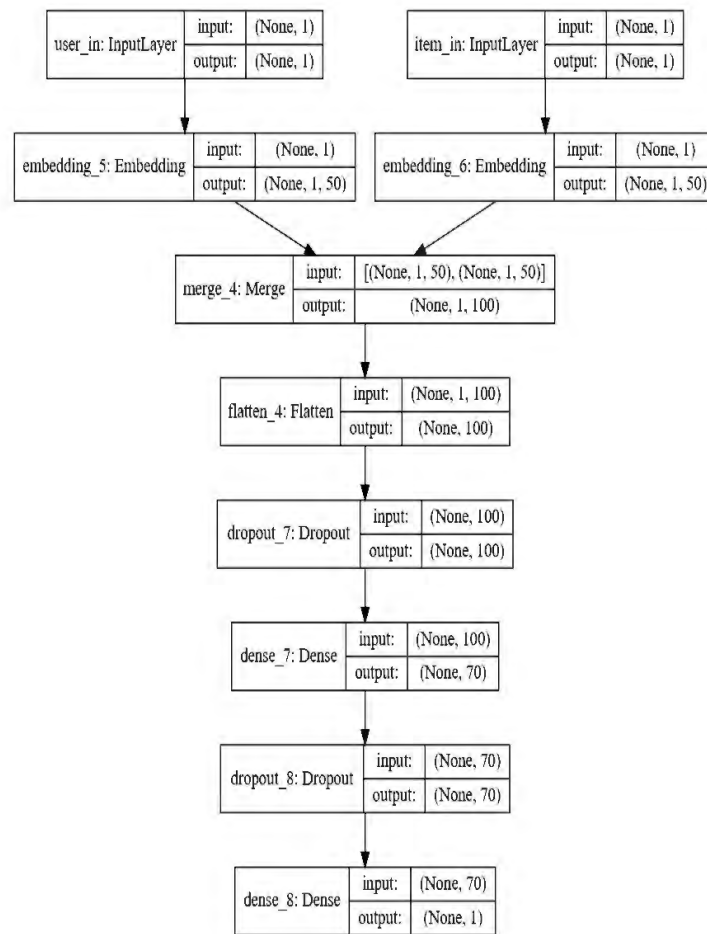
Figure 38:Collaborative Filtering Deep Neural Network Model
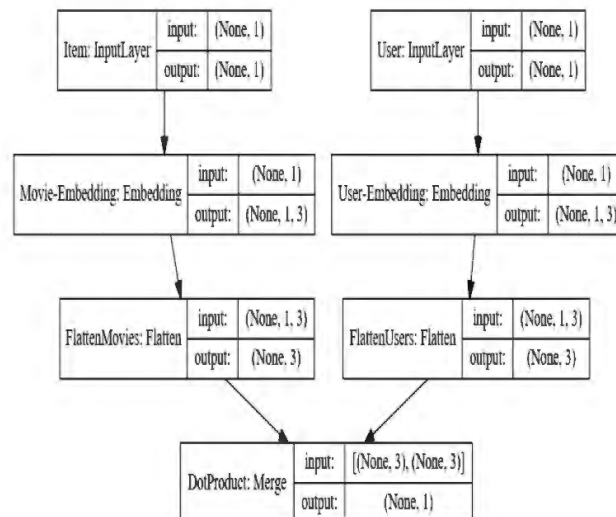
Figure 39:Collaborative Filtering Deep Recommender

A brief description of the two deep learning recommender models is described next. CFMDNN model uses User (u) and Item (i) to create embeddings (low-dimensional) for user and item. Generalized Matrix Factorisation (GMF) combines the two embeddings using the dot product. This is our regular matrix factorisation. We combine/concatenate them to create a joint feature vector which can be passed on to the several deep neural layers. For CFDR architecture, the joint latent feature vector goes through several layers of fully connected feedforward deep neural network, and for CFMDNN architecture, it undergoes additional processing with a second channel of processing with multilayer perceptron and then fused with several deep network layers, with an aim that this extended processing may discover and learn deeper interaction information between users, items and ratings, and can improve recommender system performance. The two deep neural models (CFDR and CFMDNN) were built for each dataset using Keras Deep learning tools, and used an empirical selection of hyperparameters, with leave one out validation with validation split set to 0.1 (90% training data for building the model, and 10% data for testing/validation. For performance comparison, different performance metrics were used including accuracy in ratings class prediction, as well as precision, recall, True Positive and False Positive rates, and the RMSE.

The Root Means Square Error (RMSE) is computed as follows:

$$RMSE = \frac{\sum(u,i) \in Ttest(Tu,i - Tu,i)^2}{|\text{Ttest}|}$$

where Tu,i is the real rating value of user u on item , $T\hat{}u,i$ is the corresponding predicted rating value by the model, which is calculated as.$|Ttest|$ is the number of user item in the test dataset.

4.2.2 Experimental Results:

In this section, the evaluation of the two algorithms described above for different publicly available datasets in terms of different performance metrics including of accuracy, MAE and RMSE are presented.

The details of the datasets are described first, before discussing the evaluation results.

4.2.2.1 *Datasets:*

A. MovieLens20M Dataset:

MovieLens 20M Dataset is a movie recommender system project dataset and details of this dataset is provided in MovieLens web site. Also, it contains 20 million user ratings and 465,000 movie tag which applied to 27,000 movies by 138,000 users and includes 1,100 tags . Table 4 shows the data structure for the rating table, and rating distribution statistics.

Table 4 MovieLens 20M Dataset table structure and ratings distribution

	userid	itemid	ratings	timestamp
635836	4251	1	4.0	973568502
982700	6627	2	3.0	842879122
480565	3259	4	5.0	1230931053
97855	684	5	2.0	997585406
917364	6123	3	4.5	1076944877

B. Amazon Digital Music Dataset:

Amazon Digital Music dataset is from the Amazon product dataset which contains product reviews (reviewer, reviewer name, ratings, text, helpfulness votes, review time) specifically for Digital Music purchase(Wong and Factura, nod) The ratings distribution and data structure for this dataset is as shown in Table 5.

Table 5: Amazon Digital Music Dataset table structure and ratings distribution

	userid	itemid	ratings	Time
0	A2EFCYXHNK06IS	5555991584	5.0	978480000
1	A1WR23ER5HMAA9	5555991584	5.0	953424000
2	A2IR4Q0GPAFJKW	5555991584	4.0	1393545600
3	A2V0KUVAB9HSYO	5555991584	4.0	966124800
4	A1J0GL9HCA7ELW	5555991584	5.0	1007683200

C. Book crossing Dataset:

The book-crossing dataset is collected by Cai-Nicolas Ziegler (Nart & Tasso, 2014).

Table 6: Book- Crossing Dataset structure

	userID	ISBN	bookRating
0	276726	0155061224	5
2	276729	052165615X	3
3	276729	0521795028	6
5	276736	3257224281	8
6	276737	0600570967	6

D. Amazon Books dataset:

Amazon books dataset is from the Amazon product dataset which contains only user's ratings (user, item, rating, timestamp) it includes more than 150,000 user's ratings (Li, 2017).

Table 7: Amazon Books Dataset structure

	userId	itemId	rating	Time
0	AH2L9G3DQHHAJ	116	4	1019865600
1	A2IIIDRK3PRRZY	116	1	1395619200
2	A1TADCM7YWPQ8M	868	4	1031702400
3	AWGH7V0BDOJKB	13714	4	1383177600
4	A3UTQPQPM4TQO0	13714	5	1374883200

4.2.2.2 Experimental Results.

In this section, the results of applying two deep learning algorithms for all datasets are compared with traditional collaborative filtering (without deep learning) approach CF in terms of accuracy, MAE and RMSE.

4.2.2.2.1 Ratings Prediction Accuracy.

Figure 40 shows the effectiveness of models using accuracy measure. CFDR and CFMDNN both have excellent accuracy and it showed how the models can improve the accuracy of recommender system from 59.30% to 100% in Amazon Books dataset and 73.80% to 100% in Amazon Digital Music. However, Book-Crossing dataset has improved with 98% in CFDR and got high accuracy in CFMDNN which is better than CFDR. Also, MovieLens 20M dataset got good accuracy in CFMDNN model rather than CFDR model. For baseline comparison, we also show simple Collaborative filtering approach performance as well, which uses a simple matrix factorization using direct dot product of users and items, without deep learning.

	Amazon Books	Amazon Digital Music	MovieLen20M	Book-Crossing
CF	59.30%	73.80%	48.40%	63.90%
CFDR	100%	100%	74.70%	98%
CFMDNN	100%	100%	99%	100%
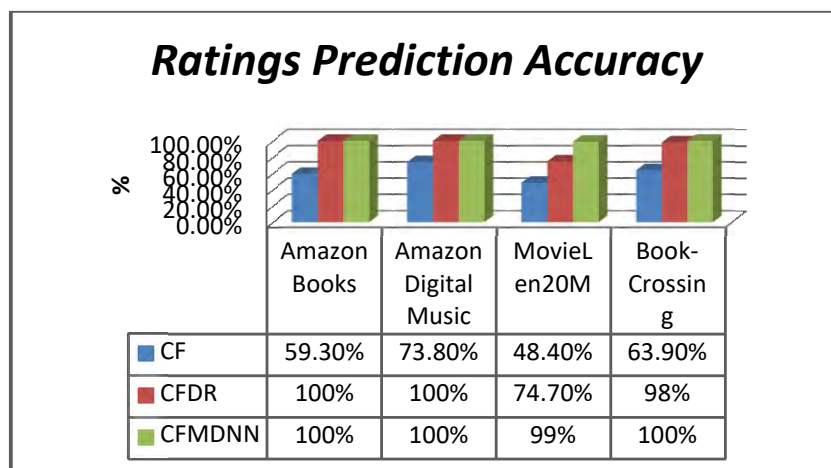
Figure 40:Accuracy of CF, CFDR and CFMDNN.

4.2.2.2.2 Mean Absolute Error (MAE).

Instead of classification accuracy or classification error, the most widely used metric in CF research literature is Mean Absolute Error (MAE), which computes the average of the absolute difference between the predictions and true ratings (Webster et al., 2004). From the Figure 41 it can be seen that the MAE get better in all datasets in CFMDNN models, validating our hypothesis that joint learning space with the CFMDNN model allows better discovery of joint latent features and can lead to a better recommender system rather than non-joint feature learning space as in CFDR model.

Mean Absolute Error(MAE)

	Amazon Books	Amazon Digital Music	MovieLen20M	Book-Crossing
CFDR	0.147	19	0.098	8.4
CFMDNN	0.003	0.1	0.005	0.6
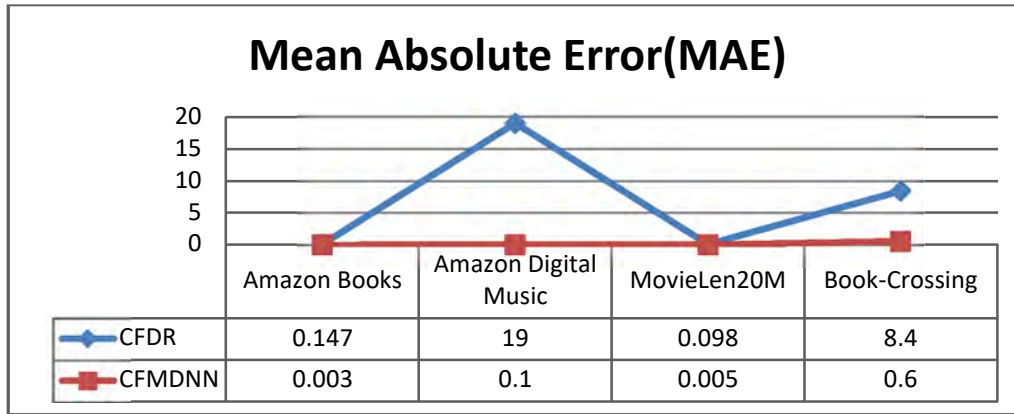
Figure 41:MAE of CF, CFDR and CFMDNN.

4.2.2.2.3 Root Means Square Error (RMSE).

Root Means Square Error (RMSE) is considered one of the most important measurements to test the models. From the Figure 42 we can see that improvement RMSE improvement highly in CFMDNN model. For example, MovieLens 20M dataset improve from 35.5 to 1 with CFMDNN model, whereas with CFDR model, the RMSE achieved it 9.71.

Root Means Square Error (RMSE)

	Amazon Books	Amazon Digital Music	MovieLen20M	Book-Crossing
CF	34.6	29.4	35.5	21.6
CFDR	10	9.71	9.71	76
CFMDNN	1.8	2	1	4.3
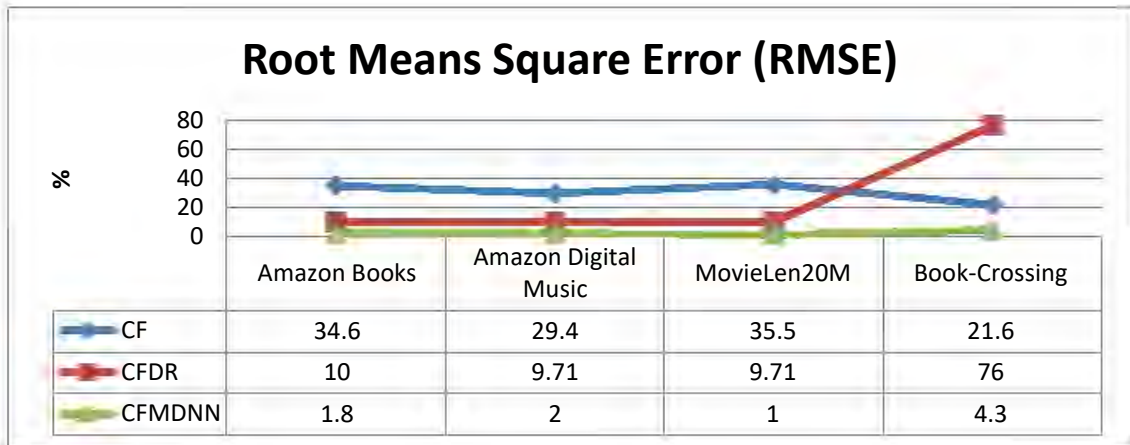
Figure 42:Root Mean Square Error (RMSE) for CF, CFDR and CFMDNN.

4.2.2.2.4 CFMDNN Prediction Model.

Since CFMDNN model turned out be best performing, we examined the ratings prediction accuracy for this model for all datasets. Tables 8 to 11 show the ratings prediction performance for each dataset based on best performing CFMDNN model.

Table 8: *Ratings Prediction with CFMDNN prediction for MovieLens 20M dataset*

User_ID	Movie_ID	Rating (Actual Data)	CFMDNN Prediction Model
1012	187	4	4
0365	236	4	4
0237	0277	4	4
0237	0323	4	4

Table 9: Ratings prediction on Amazon digital Music

User_ID	Music_ID	Rating Actual Data	CFMDNN Prediction Model
2	1	4	4
5	0	5	5
6	2	3	3
7	0	5	5
8	0	5	5

Table 10: Ratings prediction on Book-Crossing Dataset

User_ID	Book_Id	Rating Actual Data	CFMDNN Prediction Model
15	6	9	9
22	3	8	8
23	5	10	10
29	5	10	10

Table 11: Ratings prediction on Amazon Books Dataset

User_ID	Book_Id	Rating Actual Data	CFMDNN Prediction Model
32	324	4	4
222	55	5	5
232	7	2	2
29	2	4	4

4.3 Impact Rating Scales on Recommender System and Used Deep Learning and Neural Network Models to Improve Rating Prediction.

A lot of websites enables users to have the ability to rank products and share their rankings with other users for personalization or social resolutions. In specific recommender systems, personalization recommendations are produced by forecasting ratings for products which users

are insensitive of, grounded on the scores users offered for various commodities. Explicit consumer ratings are composed by way of graphical widgets known as a rating scale. Every website or system usually applies a certain rating scale in several instances opposing from scales applied by other networks in their numbering, granularity, existing of neutral position or visual metaphor. Several tasks in the area of survey design testified on the impact of ranking scales on consumer ratings, these, though, are usually considered impartial tools when in recommender systems. This section offers novel empirical data on the impact of rating scales on consumer ratings. We describe how the rating scale can affect the recommendation system user performance. In addition, we propose a novel approach based on deep learning-based augmentation of the collaborative filtering approach with deep neural networks for discovering the complex and deep interactions in the shared space between users, and ratings/reviews, and provide significant improvement for predicting user ratings. Recommender systems are termed as the data filtering network which addresses the issue of data surplus (Maxwell and Joseph, 2015). They do so through filtering crucial material piece out of the vast quantity of dynamically produced info in line with user's interests, likes or noted behavior about the item. Recommender system can forecast if a certain user may desire a product or not centered on his or her profile. They are crucial to both the users and service provider. They minimize operation expenses of selecting and finding commodities in virtual stores. Also, commendation networks have demonstrated to enhance quality and policy making procedure. In an e-commerce situation, recommender networks improve incomes in an e-commerce environment since they are efficient in selling products. Recommender system in logical archives users by enabling them to go past collection searches. In many recommender systems, the rating scales play an important key to develop the user's needs. The rating scale is performance used by users to rate items such as movies, books, music, and product.

This system used in many fields and websites in different scales such as Amazon, YouTube, and Facebook. The scale could be numbered and Like or dislike. In addition, the effect of rating scales is described in many articles and researches works such as economics, survey design and psychology (Cena et al., 2017). In many recommender systems, the user ratings are considered the valued information for users and system. By using this technique the users increase their preferences for the item by rating them and based on these rating (Adomavicius and Tuzhilin, 2005). In addition, for improving user rating it should improve the rating scale. According to Cena et al, the rating scale affect the user's performance in the recommendation system and that shows in three experimental works (2017). Recommender systems are helping users to find information that they need very quickly and easily through information filtering. These systems are very important because they help to reduce the time for the user is searching. The e-commerce websites and social media platforms enable customers to have the chance to rank content for their benefits or others. For instance, You-Tube enables viewers to rate videotapes as well as share their ranking with other viewers. There is a rating scale that viewers offer their rating which is known as graphical widgets which are associated with certain characteristics. The customer rating is particularly beneficial in recommender networks for both service providers and users.

4.4 **The Impact of Granularity on Rating.**

Through comparison of Movielands, IMDb, and MovieLens with Filmeter and FilmCrave rating they show how granularity affects consumer rating. Criticker, for instance, has the finest granularity hence enables individuals to stricter and precise. MovieLen has 10- point granularity which is implicit via half-star ranking and enables an individual to select a neutral point while IMDb has an explicit 10- point (Cena et al., 2017). Half-star rating is not usually used also Filmeter has the smaller average related with the extensive scoring range and lack neutral icons like the stars hence inhibiting the users rating. It should be noted that all the correlated scales express similar or closely associated granularity and share some objective characteristics. Scales

with a bigger granularity appears to inspire reasoning hence the expression of considerate rating. On the other hand, lower granularity scales result in extreme ranking in line with the state of the task (Preston and Colman, 2000). For instance, in the museum setting, ratings given to the similar product with various rating scales are different and granularity is beneficial to describe the collected information. For scales with the same granularity the rating was closer and where granularity was lower consumer's rating was higher. Another example is Amazon.com which an online endorsement engine which applies scalable product-to-product collaborative filtering methods to suggest virtual items for various customers. Amazon.com utilizes a clear information gathering approach to get information from users. (Koren et al., 2009). Usually, the interface comprises of sections such as rate these products, customer's surfing history, and develop one's recommendations as well as one's profile. User's interests are predicted based on the product he or she has ranked. The customer surfing pattern is compared to the system and the system decides which product or products of preference to suggest to the customer (Lee et al., 2008). Also, Amazon.com has promoted a feature of an individual that has purchased these products.

4.5 Using different rating scales in recommender systems.

In the rating system there are two ways which are explicit and implicit ratings. In this section we focus on explicit rating which are the techniques that allow a user to clearly express her/his interest in an item. Usually, users use the score to these items through a process, such as the 5-star rating system or like/dislike rating system, to indicate their interest in an object (Jawaheer et al., 2010). Moreover, recommender systems collect users' preferences using some of the rating systems such as like/dislike rating system such as social media application. On the other hand, online stores such as Amazon, AliExpress use the star ratings system. (Cena et al., 2017).

Figure 43:Most common explicit rating system.

From the Figure 43 shows the most common explicit rating systems used by users. Many systems use the star ratings system that allows users to indicate which products are of their interest. Moreover, Google+1 is a new feature that Google added to its search engine which help users to evaluate websites that like them. So, they recommend website to their contacts.

4.5.1 Experimental Results:

4.5.1.1 *Datasets.*

We used two different book datasets in data sizes and the number of classes (different ratings scales used). The Book crossing dataset has 10 ratings (10 classes) and the Amazon Book dataset contains 5 ratings (5 classes).

A. Amazon Book Dataset:

Amazon books dataset is from the Amazon product dataset which contains only user's ratings (user, item, rating, timestamp) it includes more than 150,000 user's ratings.

Table 12: Amazon Books Dataset structure

	userId	itemId	rating	Time
0	AH2L9G3DQHHAJ	116	4	1019865600
1	A2IIIDRK3PRRZY	116	1	1395619200
2	A1TADCM7YWPQ8M	868	4	1031702400
3	AWGH7V0BDOJKB	13714	4	1383177600
4	A3UTQPQPM4TQO0	13714	5	1374883200

B. Book crossing Dataset:

The book-crossing dataset is collected by Cai-Nicolas Ziegler. The dataset contains 278,858 users, about 271,379 books and 1,149,780 ratings. The Book-Crossing dataset includes 3 tables: BX-Users (user ID, location, age), BX-Books (ISBN, book title, author, publisher, year of publication) and BX-Book-Ratings (explicit ratings from 1 to 10, implicit ratings expressed by 0) (Nart, D. and Tass, 2014).

Table 13: Book- Crossing Dataset structure

	userID	ISBN	bookRating
0	276726	0155061224	5
2	276729	052165615X	3
3	276729	0521795028	6
5	276736	3257224281	8
6	276737	0600570967	6

4.5.2 Results:

In this section, we describe the details of the experimental setup to evaluate the effectiveness of proposed deep learning models, CFMDNN and CFDR models, and compared with traditional Collaborative Filtering (CF) approach without deep learning. It was shown, how effective is CFMDNN algorithm, due to joint learning and deep discovery of latent interaction features

between users and items with a multi-stage architecture, using different performance evaluation metrics including accuracy, MAE and RMSE. The ratings are often specified on a scale that indicates the specific level of like or dislike of the item at hand (Huang et al., 2015). Two datasets were used, and each dataset has a different rating system. Book-Crossing dataset used 0-10 Rating-scale and Amazon books dataset used 1-5 rating scales.

A. Mean Absolute Error (MAE).

Here we applied CF and CFMDNN models on two datasets to see the improvement of MAE. From figure 44 it can be seen that the MAE got better result on both datasets in CFMDNN model.

Mean Absolute Error(MAE)

	Amazon Books	Book-Crossing
CF	1.65	5.654
CFMDNN	0.003	0.6
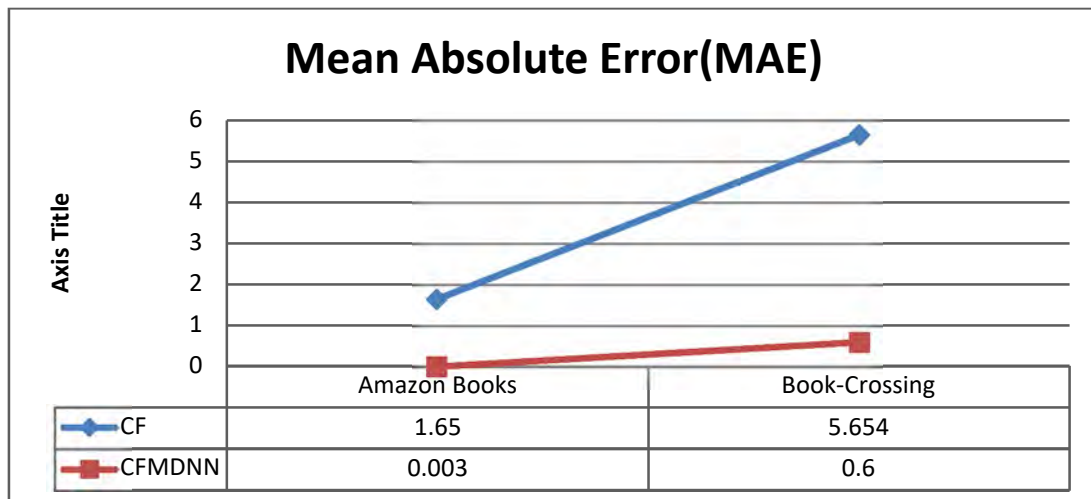
Figure 44:Mean Absolute Error (MAE) between Amazon books and Book-Crossing datasets.

B. Root Means Square Error (RMSE).

Root Means Square Error (RMSE) is considered one of the most important measurements to test the models. From figure 45 we can see that improvement RMSE improvement highly in CFMDNN model. Amazon Books dataset improve from 34.6 to 1.8 with CFMDNN model and Book-Crossing dataset improve from 21.6 to 4.3.

Root Means Square Error (RMSE)

	Amazon Books	Book-Crossing
CF	34.6	21.6
CFMDNN	1.8	4.3
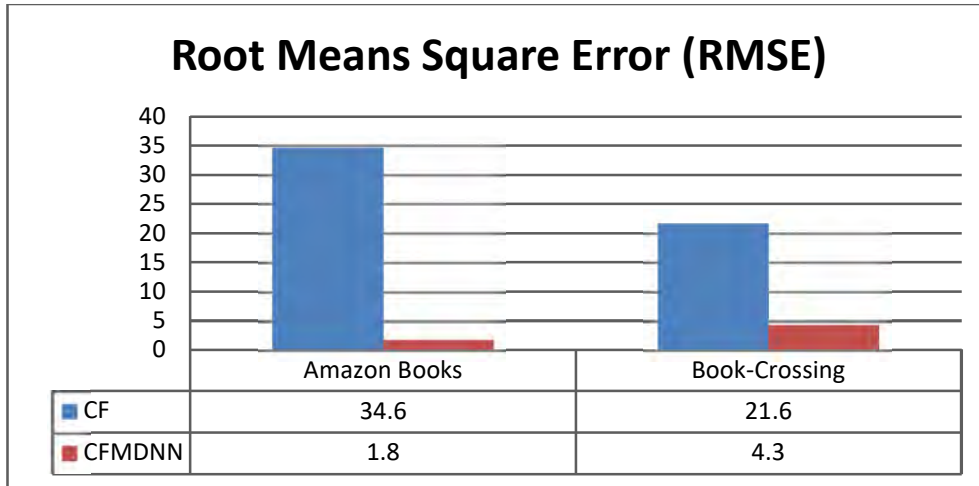
Figure 45:Root Mean Square between Amazon books and Book-Crossing datasets.

4.5.2.1 Ratings Prediction Accuracy.

Figure 46 shows the effectiveness of models using accuracy measure. CFMDNN has excellent accuracy and it showed how the models can improve the accuracy of recommender system from 59.30% to 100% in Amazon Books dataset and 63.90% to 100% in Book-Crossing dataset. For baseline comparison, we also show simple Collaborative filtering approach performance as well, which uses a simple matrix factorization using direct dot product of users and items, without deep learning.
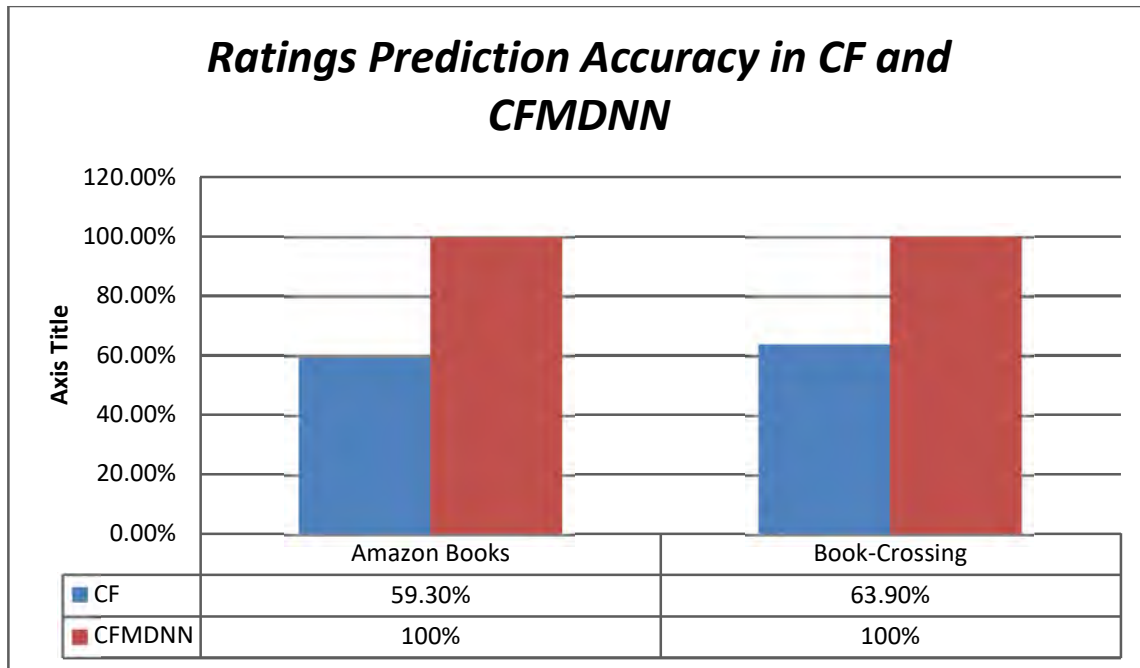
Figure 46:Ratings Prediction Accuracy in CF and CFMDNN.

4.6 Chapter Summary:

This chapter has presented a new deep learning-based framework for augmenting the collaborative-filtering based recommendation systems for capturing the deep and latent hidden interactions between users and items and improves the recommender system performance in terms of rating prediction of unseen items and users. The evaluation of the proposed framework, with two different neural recommender models, on four different publicly available datasets, has shown an improved performance of the system, as compared to the traditional collaborative filtering-based approach based on matrix factorization techniques. Particularly, the multistage deep collaborative neural network model gives good performance and robustness due to its ability for joint modelling and discovery to deep interactions. Next chapter discusses another novel approach for improving the performance of recommendation systems using explicit feedback ratings, in terms of text-based comments, and use of natural language processing techniques for sentiment analysi

Chapter 5 Enhanced Recommendation System Framework based on NLP techniques and Sentiment Analysis of Explicit Ratings.

5.1 Introduction.

In this chapter, an enhanced computation framework for recommendation systems based on analysis of sentiments expressed in explicit feedback from text-based comments in a multilingual context is described. For analysing the text-based comments from explicit feedback, and extracting the sentiments, natural language processing (NLP) techniques were used. The two models based on NLP techniques include a bag of words (BOW) model and Word2Vec model.

For words to be processed effectively in computers, they need some form of numeric representation to ease processing by models such as the bag-of-words and the Word2vec models. The bag-of-words model is an approach for extracting and representing specific features from text for use in modelling. It is commonly applied in machine learning due to its simplicity and flexibility. The Word2vec models, on the other hand, encompass a group of related models that create word embeddings. These models are, however, shallow neural networks that are conditioned to rebuild linguistic contexts of words. These models have been featured prominently in several previous works and past studies. The Bag-of-Words and Word2vec Models have been explored widely when analysing customer reviews by sentimental analysis. (Chakraborty et al., 2018). Dai and other have adopted a standard bag of-features model to forecast the opinion class of film reviews based on consumer's inputs (2015). To defeat challenges with bag-of-words methods during customer reviews through sentiment analysis such as negation, Gabryel proposed the use of hand-written rules to reverse the semantic orientation of a consumer review preceded by a negative review (2018). While such rules were significant in dealing with negative issues (Betru, 2017). Gehrmann and others found it time-consuming to make an algorithm that could

handle dominant human language (2018). Yuan and others showed that the model was inefficient in terms of manual evaluation of lexicon and lend itself to low accuracy (2016). On the other hand, a semi-supervised framework relying on both models was more effective in semantic analysis of consumer reviews (Vinayakumar et al., 2018). Few other approaches were based on paragraph vectors to determine the syntactic and semantic links of a review text using both models (El-Din,2016). The concatenation of review embedding, and product embedding gathered from paragraph vectors outperformed   all previous approaches used in sentimental analysis of customer reviews. Campr and Ježek, compared Bag-of-Words and Word2vec Models on sentiment classification (2016). In terms of accuracy, Word2Vec proved to be most precise and functional word embedding approach. It was highly effective in converting customer reviews into substantive vectors (Zhang et al., 2018). The only problem with the model was that it ignored sentiment data of texts and required huge chunks of reviews for training and yielding the exact vectors. Next Section describes the dataset used for building the recommender system models based on NLP techniques.

5.2 Experimental Work

5.2.1 Datasets

A. Booking Hotel dataset

The dataset is from Booking.com. All data is publicly available. It contains 515,000 customer reviews and scoring of 1493 hotels.

B. Amazon fine food review

This dataset consists of reviews of fine foods from amazon. It contains 500,000 reviews. Reviews include product and user information, ratings, and a plain text review.

C. Arabic Movie Review

This dataset from elcinema.com contains 1524 reviews with positive and negative ratings and feedbacks. This dataset includes negative feedback inside positive sentence with a different accent but in the Arabic language.

Negative_Review	Review_Total_Negative_Word_Counts	Total_Number_of_Reviews	Positive_Review
I am so angry that I made this post available...	397	1403	Only the park outside of the hotel was beauti...
No Negative	0	1403	No real complaints the hotel was great great ...
Rooms are nice but for elderly a bit difficul...	42	1403	Location was good and staff were ok It is cut...

Figure 47:Booking Hotel Dataset.

Summary	Text
Do Not Buy	Do not purchase or support this company in any way until they clean up their act. And for whatever reason Amazon doesn't allow returns of this item,
my cat goes crazy for these!	Best cat treat ever. There isn't anything comparable to the love my cat has for these treats, he snubs away any other kind now.
Very good coffee	I really liked this coffee, it was just as good as everyone claimed it was. Strong, bold and flavorful! I would recommend!
Bigelow Earl Grey Green Tea	Tastes like Earl Grey, but it's green tea so it's healthier.

Figure 48:Food fine amazon Dataset

0	تلك الايام " مغامره سينمائية متميزه إنتظرت "	...	1
1	مسخره خلاقة!..بس الفيلم يسناهل يتشاف! ينجح احم	...	1
2	عناصر نجاح الفيلم سيناريو رائع للسيناريست ياسي	...	1
3	مش "وهمى"!..بس "شغال"! يعود العم "ميل جيبسون	...	1
4	فيلم كوميدى ساخر بنكهة شعبية حريفة جدا! أثار إ	...	1

Figure 49:Arabic Movie review dataset

5.2.2 Data Preparation.

Text data preparation is different for each problem. In each dataset, we clean it up as dataset contained only English reviews and for Arabic reviews contains only Arabic language. All text was converted to lowercase and deleting the white space like commas and brackets. Finally, the text has been split into one sentence per line. However, Arabic language has a different mechanism. We removed punctuations and white space.

5.2.3 Clean Text.

The clean text data is the most important part because different approaches in text cleaning can lead to varied results during model training. In this section we used Tokenization technique which helps to divide the textual information into individual words. In addition, in this experimental work we focused on text mining which is the process of extracting information from unstructured text. It is a multidisciplinary field that draws on data mining, machine learning and information retrieval. All these processes are requiring Pre-processing steps to reduces the size of the input text documents and this called Tokenization. In this study we used the split() function to split the loaded document into tokens separated by white space and we used NLTK to remove English stop words. All datasets include rating scales from 1 to 5 and contain text feedback. Here we are loading the raw data each textual review is grouped into a positive part and a negative part 0 for positive and 1 for negative. We want to use feedback and

score together and test the model performance. We use 'clean_text' function by removing lower

text, remove the punctuation, remove words that contain numbers and remove stop words.

	review	is_bad_review
0	I am so angry that i made this post available...	1
1	No Negative No real complaints the hotel was g...	0
2	Rooms are nice but for elderly a bit difficul...	0
3	My room was dirty and I was afraid to walk ba...	1
4	You When I booked with your company on line y...	0

Figure 50:Booking Hotel Dataset

	Summary	Text	is_bad_review
0	Good Quality Dog Food	I have bought several of the Vitality canned d...	0
1	Not as Advertised	Product arrived labeled as Jumbo Salted Peanut...	1
2	"Delight" says it all	This is a confection that has been around a fe...	1
3	Cough Medicine	If you are looking for the secret ingredient i...	1
4	Great taffy	Great taffy at a great price. There was a wid...	0

Figure 51:Food Fine Amazon Dataset

	text	is_bad_review
0	تلك الايام " مغامرة سينمائية متميزة إنتظرت "...	1
1	مـــحرة خلافة!..بس الفيلم يـدٰاهن يتشاف!ينجح احم...	1
2	عناصر نجاح الفيلم سيناريو رائع للسيناريست باسي...	1
3	مش "اوهسي"!..بس "تشغل"! يعود المح "ميل جيبــون...	1
4	فيلم كوميدى ساخر بنكهه شعبية حريفة جدا! أثار إ...	1

Figure 52:Arabic Movie Review Dataset

5.2.4 Word2Vec Model.

We can predict customers reviews about their feeling for food they bought or movie they watched

or hotel they stayed by adding sentiment analysis features. We used NLTK module called *Vader*

which was designed for sentiment analysis. We use Vader, which is a part of the NLTK module

designed for sentiment analysis. Vader uses a dictionary of words to find which ones are positives or negatives. It also considers the context of the sentences to determine the sentiment scores. We are extracting vector representations for each review by using Gensim module because it creates a numerical vector representation of every word called (Word2Vec). This is performed using neural networks which is similar words that will have similar representation vectors. Then we add the TF-IDF (Term Frequency Inverse Document Frequency). TF count number of times the word appears in the text IDF count the relative importance of this word which depends on how many texts the word can be found.

Table 14: The accuracy of each dataset for testing model.

Datasets	Positive Reviews (0)	Negative Review (1)
Booking Hotel dataset	96%	4%
Amazon fine food review	64%	36%
Arabic Movie Review	0%	100%

A. Wordclouds.

We used *wordclouds* to greater prominence to words that appear more frequently in reviews

*Figure 54:*Wordclouds for Booking Hotel Dataset.

Figure 55:Wordclouds for Arabic Movie Dataset

Most of the words in Amazon fine food dataset are indeed related to the customer experience such as product, coffee, tasting, etc. In addition, in Booking Hotel review dataset is room, location and breakfast. In Arabic movie review dataset is film and prepositions. Arabic reviews have different way to use words and it need different techniques to clean all these problems because that the module does not work with Arabic language. The highest positive sentiment reviews and highest negative sentiment reviews for each dataset show how the test model work.

	Text	pos
472333	GREAT GREAT GREAT GREAT GREAT GREAT GREAT GREA...	0.884
177895	I love the cookbook. Fabulous pictures and de...	0.684
198983	i would buy the mag plus calcium. tastes ok. ...	0.684
356405	i would buy the mag plus calcium. tastes ok. ...	0.684
30645	I have tried many gluten free brownies these ...	0.673
43710	If you are an almond fan like me and like swee...	0.669
347888	If you are an almond fan like me and like swee...	0.669
279287	Fresh, rich and delicious Belgian chocolates. ...	0.668

Figure 56:Highest positive sentiment reviews for Food fine Amazon Dataset.

	review	pos
43101	A perfect location comfortable great value	0.931
211742	Clean comfortable lovely staff	0.907
175551	Friendly welcome Comfortable room	0.905
365085	Good location great value	0.904
109564	Clean friendly and comfortable	0.902
145743	Good value amazing location	0.901
407590	breakfast excellent Clean comfort	0.899
407546	Great place I enjoyed	0.881
218571	Beautiful Quirky Comfortable	0.878
436901	Lovely comfortable rooms	0.877

Figure 57: Highest Positive Sentiment Reviews for Booking Hotel Dataset.

	text	pos
669	...قسة الإبهار رواية قد نبدو للبعض بسيطة وعادية م	0.032
309	...ما بعد الفاع لحد إمتي هنفضل محصورين في هذه ال	0.024
918	...تشويه الاصل كما هي العادة في الدراما المصريه ي	0.013
1098	...اعتنا مشاهدة افلام وانت Tangled ديزني تنصح بـ	0.009
259	...الفتكين...اللي فاكرين كل فيلم ناجح يبقى لازم م	0.008
494	Cold Mou... — لعبة القدر والحرب فيلم (كوك ماونتن	0.003
730	...عندما كانت الديناصورات تحكم دور العرض! كم هو ع	0.000
1168	... محدش فاهم حاجه فيلم "النبابه نطلب البراءة" من	0.000
1517	...السادة المرنتلون وسياسة الانفتاح نميز المخرج عن	0.000
1452	...العبقرى حلمى ـ قد يحرق رايى الفيلم ارجو الحيطة	0.000

Figure 58:Highest Positive Sentiment Reviews for Arabic Movie Review Dataset

	review	neg
323586	I received the wrong item, wrong size. Not hap...	0.720
31777	This stuff is the most vile, disgusting stuff ...	0.533
130741	Bad tasting coffee, no aroma/baked, very dry, ...	0.464
324333	I cant it was so bad, we tossed it away.This w...	0.457
192524	I love bold coffee, but this is seriously over...	0.450
525983	WARNING, WARNING, WARNING....DO NOT BUY !!!!!!...	0.441
431533	Due to very poor packing, 7 of the 14 pouches ...	0.433

Figure 59:Highest negative sentiment reviews for Food fine Amazon Dataset

	review	neg
193086	No dislikes LOCATION	0.831
356368	Nothing Great helpful wonderful staff	0.812
318516	A disaster Nothing	0.804
458794	Nothing Excellent friendly helpful staff	0.799
29666	A bit noisy No	0.796
426057	Dirty hotel Smells bad	0.762
263187	Very bad service No	0.758
443796	Nothing perfect	0.750
181508	Window blind was broken	0.744
175316	Nothing Super friendly staff	0.743

Figure 60:Highest negative sentiment reviews for Booking Hotel Review Dataset.

	text	neg
799	... حر عة أكثر زيدة على نفس الوتيره التى بدأت عام	0.029
1102	...دقيقه فقط بمجموعه من انشهد المنلاحظة والى 40	0.010
78	... كزري من اشهر افلام الرعب الإمريكيه هذا الفيلم	0.009
669	...قمة الإبهار رواية قد تبدو للبعض بسيطه وعادية م	0.006
660	...تور هانج احر سبكورسيزي يبدو أن سبكورسيزي اشدا	0.006
736	Gravita... في الفضاء لا احد يسمع صراخك ..Gravity	0.004
494	...Cold Mou — نعبة القدر والحرب فيلم (كوند ماونتن	0.000
1517	...السادة المرتشون وسياسة الانفتاح نمير انفجرت عن	0.000
1452	...العبقري حلمي — قد يحرف رايى القيلم ارجو ،لحبطة	0.000
239	... عمل رائع لا بعد العمل الفني نجح الا ادا كان	0.000

Figure 61:Highest Negative Sentiment Reviews for Arabic Movie Review Dataset.

Chapter 5   93

We can see from the tables above the module is working perfectly for datasets with English language, but it does not work for Arabic language.
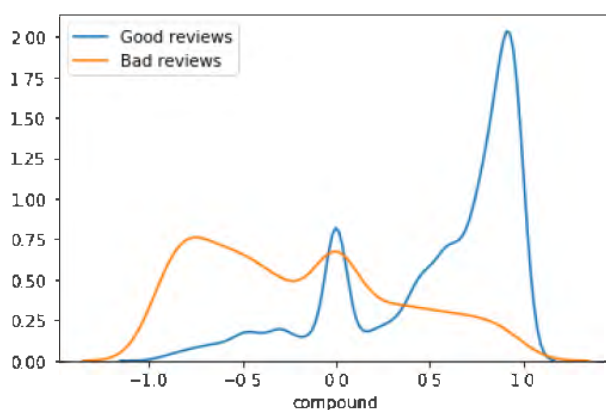
Figure 62:Sentiment distribution for positive and negative reviews for Booking hotel dataset
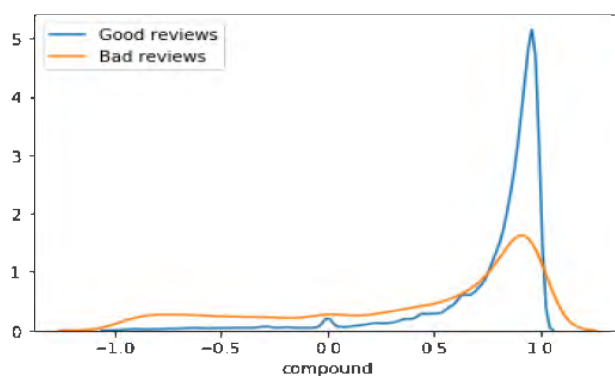
Figure 63:Sentiment distribution for positive and negative reviews for Food fine Amazon dataset
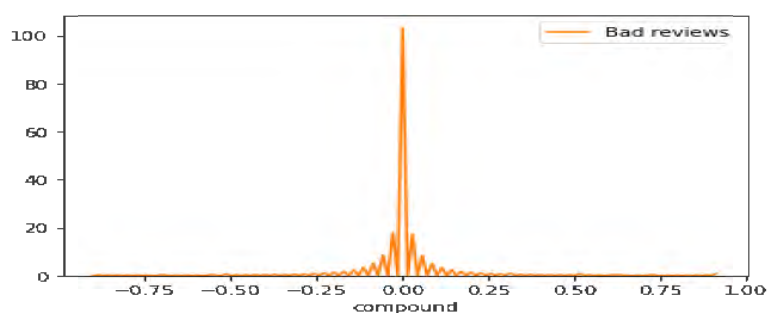
Figure 64:Sentiment distribution for positive and negative reviews for Arabic movie review dataset

From the Figure 64, we can result that, the model Word2Vec does not have good performance for any languages such as Arabic language. To address this problem, Random Forest classifier was used for Arabic dataset to improve the performance.

5.2.5 Random forest (RF).

Random Forest (RF) is considered one of the most flexible and easy machine learning method because it can be used for both classification and regression tasks. The RF classifier averages different decision trees from random samples of the database. A decision tree partitions the dataset into smaller subsets and simultaneously builds the tree with decision nodes and leaf nodes (Chakraborty et al., 2018). In Arabic movie dataset is shows the most important features are indeed the ones that come from the previous sentiment analysis. The feature of the words is from verbs and it does not make any meaning if it is positive or negative and because that the importance of the word is zero.

Figure 65:Most important features for Arabic Movie Dataset.

5.2.6 Discussion.

From the Figure 64, it can be seen, that the Sentiment distribution for Arabic language gives bad feedback but in the real dataset it contains positive feedback, but the system cannot predict as shown in the Figure 65 by using (RF) method because all words that appear are verbs. In addition, comparing the results of Word2Vec model in three datasets with different languages Word2Vec model does not work for Arabic language. On the other hand, the results obtained by Word2Vec are highly effective in sentiment analysis for English language datasets. In addition, Arabic language might be having a different and complex structure that affects the results. Firstly, the morphology and structure are very complex. Secondly, Arabic language has challenges because of its history, region, grammar, and culture. In addition, Arabic language has a different way of writing. It written from right to left. It contains 28 letters which 25 letters are consonants, and 3 letters are vowels. Arabic language used diacritical mark which is considered a challenging part for computer systems to recognize and for sentiment analysis by machine. Moreover, Arabic language has complex Morphology where one-word leads to many important meanings. For example, the word (لَبِس) means (Wearing) and word (لَبْس) means (Confusing). So, the diacritical marks can affect Arabic sentiment analysis as is evident from the experimental results.

5.3 Chapter Summary.

In this Chapter, an enhanced recommendation system framework based on using text-based feedback comments, and NLP processing techniques in a multilingual context was presented. This chapter established that the Word2vec Model has attracted extensive research. However, Word2Vec was found to be working well for English language text processing better because of availability of better language dictionaries for computer-based processing, whereas it did not lead to any improvements for Arabic language contexts. This could be due to complex structure and morphology of Arabic language and lack of sufficient tools in terms of language dictionaries for

computer-based processing. Also, Arabic sentiment analysis research is very limited while sentiment analysis has many kinds of research in English. Next Chapter continues this line of research and proposes an improved model for Arabic language. The novel approach used for improving the performance of recommendation system is based on a combination of Bag of Words and Word2Vec methods. By using hybrid deep learning methods, based on natural neural networks for Arabic sentiment analysis with SVM and MLP, CNN and MCNN architectures, an attempt was made to improve prediction of user's reviews on both language Arabic and English recommendation system contexts.

Chapter 6 Improvement Recommendation System Framework Based on Deep Learning and Sentiment Analysis.

6.1 Introduction:

This Chapter continues with a line of investigation from the previous Chapter and attempts to enhance the Recommender systems for multilingual contexts using deep learning and sentiment analysis approaches. With increased adoption of Artificial Intelligence AI, machine learning and data science technologies recently, it has become possible to develop real world computational platforms for e-Commerce environments, that are more user and customer centric. Recommendation systems are one such customer-centric computational platforms, that allow commercial eCommerce platforms to provide better services to customer service and user experience. They are evolving continuously, and progressively getting more efficient with improved personalization capabilities in offering   recommendations, by incorporating a multitude of tools to capture and collect different types of human input, in terms of user ratings, opinions, votes, sentiments, qualitative and quantitative feedback and customer reviews. As discussed in previous chapters, a recommender system, or a recommendation system (often used synonymously) is a sub class of information processing and filtering system that attempts to predict the "rating" or "preference" a user would give to an item. Sometimes instead of 'system', they are also referred to as recommender platform or engine) and are extensively used in commercial applications and e-Commerce environments (Betru et al.2017; Boudad et al.2018). Moreover, recommender system platforms that are reflective, user inclusive and customer-centric provide better and personalized recommendations, and can empower users in finding what they need, and at the same time can help businesses to provide better customer service and user experience with their platform. Some of the tasks involved in providing recommendations in these platforms include, analysis of behaviour and engagement of users, with the products, items or services being made available, and provide personalized recommendation, that addresses their needs, and can lead to better customer satisfaction. This can lead to improvement in prediction of ratings, and dependent on many factors, including their purchase history, online browsing behaviour with items, likes and preferences provided, their demographic profiles and life-style related aspects, and that of other similar users with similar profiles or life-styles. By including the user feedback, their votes, sentiments and

feelings about a product or a service, recommendation systems can help businesses and organizations to provide better service to customers, enhance revenue streams, reduce operational expenses, and identify new customer segments, and build targeted marketing campaigns, and customer retention strategies, with timely recognition of customer loyalty with incentives and rewards.

Development of reflective, personalized and customer-centric e-Commerce computational platforms equipped with better recommendation capabilities, involve extensive use of individual customer profiles and similar cohorts, with either similar demographic profiles or purchase behaviour. These are the primary mechanisms for capturing the preferences, and this knowledge is used to target the products and services based on these preferences. Many large businesses and service providers such as Netflix, Amazon, YouTube and Facebook, have been using such computational platforms, equipped with recommendation tools at different scales for capturing user engagement and preferences with their products and service offering, based on extensive use of their individual profiles and behaviour, as well as that of other similar cohorts (Heikal et al., 2018). Some of the tools used for providing personalized recommendations include, individual profile capture, tracking of the customer shopping sessions of the individual users, and combine with user feedback from a similar cohort, with similar history with products, items or services, and match their demographics and behaviours (Lulu and Elnagar, 2018). The quality and personalization capability of the recommendation platforms depends on the way they capture and collects the user ratings, votes, sentiments, or feedback. This could be either explicit or implicit data collection. The collection of explicit ratings and feedback from a user, involves a concrete rating scale (such as Likert scale for rating a movie from one to five stars). On the other hand, the implicit ratings capture involves collection of information implicitly and automatically, as the user engages with the system, and could involve certain activities, such as logging the actions of a user, or tracking the browsing behaviour of the user on the system. Explicit ratings collection is easier as the ratings that is captured from a user directly, can be directly interpreted as the user's preferences, and by using similar explicit ratings from other similar users, it is easier to extrapolate on the ratings predictions, and provide personalized product or service recommendation (Mandal and Maiti, 2018). The drawback with explicit collection, however, is that the responsibility of gathering ratings is on the user or the customer, who may not be interested, or has the time to provide a rating or feedback. that it puts the responsibility of data collection on the user, who may not want to take time to enter ratings. In contrast, the implicit collection of feedback is implicit and does not involve the user or expects any extra effort on the part of the user, and it is possible to

collect large quantity of information implicitly. Implicit data is also easy to collect in large quantities without any extra effort on the part of the user. However, there are several complexities involved with both the explicit and implicit feedback collection, and some of the reasons are as outlined below:

1.   Ratings and feedback collection can just record the actions of a user and may not know the person engaging with the system, particularly, if more than one person uses the same system with shared logins.

2.   Explicit ratings with the collection of votes, and capture of sentiments, may not accurately represent true user preferences, as user is not sure and not deeply engaging with the system. Some studies in this area show that some ratings by users are impulsive and differ based on their mood or state of mind.

3.   Implicit collection of information, without the knowledge of the user raises several privacy and ethical issues, and it is overstepping on the part of corporate businesses and computational platforms to engage in implicit ratings collection without consent from the users or customers.

Ideally, a combination of two, both explicit and implicit ratings would lead to better outcomes, with a recommendation system using explicit ratings strategy to provide more personalized recommendations and use implicit ratings for providing more pragmatic and practical recommendations. In this paper, we used explicit ratings based on publicly available datasets.

The traditional methods for capturing and embedding these user ratings into the recommendation provided by the system are called collaborative filtering or content filtering-based approaches or a hybrid combination of the two. is based on a concept called collaborative filtering technique. The collaborative filtering technique is more popular out of the two common approaches and relies on a technique called matrix factorization technique. Though the collaborative filtering and matrix factorization are more common and popular, there are several problems associated with collaborative filtering techniques, due to the complex correlation between users who rate the items and the number of items, the cold-start problems involving new items, and insufficient ratings with some old items, particularly with lesser user engagement, and causing problems in modelling the user interaction with items. Many previous studies of similar complex nature, including speech recognition, computer vision and natural language processing, have proposed data driven machine learning and deep learning-based methods for capturing the complex and deep hidden interactions between users and items, and

offer much promise in improving the performance of online recommendation systems (Albawi et al., 2017). Deep Learning is a subfield of Machine learning and Artificial Intelligence and allows end-to-end fully automated computer-based approaches for problem solving and automation, often requiring minimal or no human intervention. There are several deep learning architectures proposed in literature, depending on the processing task and application context involved. Some of the popular architectures used for other similar areas including computer vision, speech processing and natural language processing include Feedforward Networks (FFNs), Convolutional neural networks (CNN), recurrent neural networks (RNN), recursive neural networks (RecNN), long short-term memory (LSTM) and memory network (MemNN). In this chapter, we propose a novel approach based on using deep machine learning to address problems. The work extends some of the previously reported work and shows the impact of using two different types of explicit feedback, both the ratings scale, and free text comments for collecting user experience. While the ratings score provides a quantitative feedback, the use of open text comments allows qualitative feedback and informs the sentiments associated with user-item interaction. Further, the proposed deep learning-based architecture includes sentiment analysis with multilingual support (English and Arabic language) support for capturing the open text qualitative comments on the items or service. The multilingual aspect of proposed recommendation system architecture allows different types of qualitative feedback to be collected and can allow better personalization to different demographic segments. In addition, Arabic language is considered the top six of the world's major languages. It includes 28 letters: 25 consonants and only 3 vowels. In addition, diacritical marks are used in the Arabic script as short vowels placed either above or below the letters for correct pronunciation and clarification of the meaning (Lulu and Elnagar, 2018). The absence of short vowels in the majority of MSA texts cause a lexical ambiguity problem that challenges computational systems. For example the undiacritized word شعر may mean (شِعْرٌ poetry), (شَعْرٌ hair) or (شَعَرَ to feel).In the Arabic language, several morphological aspects exist for a word: derivation, inflection, and agglutination (Cena et al. 2017; Boudad et al. 2018).

6.2 Multi-layer perceptron MLP for Arabic Language Sentiment Analysis.

Multi-layer perceptron (MLP) is a neural network using simple models for solving difficult task such as prediction models. MLP allows robust algorithms to be applied to solve difficult problems (Brownlee, 2017). We used MLP, as it has the ability to learn the representation from the training data with the layered neural network model. For analysis text-based sentiment text for Arabic language, we built MLP models trained on a pre-trained word vector representation.

Two different datasets were used for building the Arabic Movie dataset and Booking Hotel dataset. The textual review in each of the two datasets used has positive and negative feedback comments, denoting the sentiment inferred from each review. For the English dataset, positive review has a 0 label, and 1 for negative review, For Arabic language dataset, 1 was used for positive review and 0 for negative reviews. This was done by original dataset designers, which we used as is. We want to use these feedback text comments to predict positive or negative sentiments using MLP model and evaluate/ test the model performance. To evaluate the performance of the model we used the different performance metrics (i.e. Precision, Recall and F1-Score) for both the datasets. The Precision metric measures the number of positive class predictions that belong to the positive class. The Recall metric, on the other hand measures the number of positive class predictions made of all positive examples in the dataset. Finally, F-Measure provides a single score that balances both the concerns of precision and recall in one number. The results showed that MLP has good performance on Arabic dataset, as shown on table 15 , and the Arabic dataset performs with Word2Vec and MLP representation for Arabic language sentiment analysis, and can predict the user's feedback from text as shown in the figure 66.   User engagement is a high priority for any recommendation system. The question is how to determine which users like which of the items and if the system presents an appropriate result based on their likings. The primary function of MLP is to classify and cluster information with multiple factors taken into consideration, and these features are precisely what you need for user profiling.

Table 15: MLP Performance measures

Datasets	Accuracy	Precision	Recall	F1-Score
Booking Hotel Dataset	96%	0.96	0.98	0.97
Arabic Movie Dataset	87%	0.56	0.50	0.53
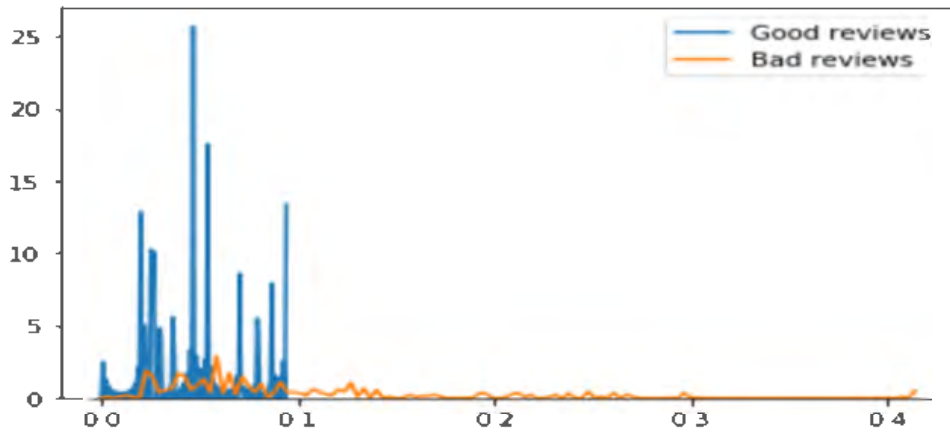
Figure 66:Sentiment distribution for positive and negative reviews for Arabic movie review dataset by using MLP model.

6.3 Collaborative Filtering based Multistage Deep Neural Network architecture (CFMDNN).

This deep learning model, the Multistage Deep Neural Network architecture (CFMDNN), uses multiple layers of Feed-Forward Neural networks (FFNs) jointly with collaborative filtering for modelling complex structures of user and item interactions and captures can extract latent features. The capability to learn the joint space at a deeper level, through several stages of learning and discovery with multiple FFN stages, does not require feature engineering or manual handcrafting for extracting and combining the latent feature sets.

Due to the joint modelling capability of the proposed deep learning model based on CFDMDNN architecture, it is possible to capture the interactions between users and feedback, at a deeper level in the joint latent feature space, with multiple layers of fully connected feed forward neural network layers (FFN). The models were built using Keras deep learning tools, and the validation and testing were done using K-fold cross-validation, with prediction accuracy used as the performance metric for the performance evaluation. Also, to prevent overfitting, the model complexity was reduced by using dropout as a regularization technique, along with Adam optimizer for iterative updating of network layer weights. Chapter 4 proposed CFMDNN architecture.

6.4 Multistage Convolutional Neural Networks Model.

The second model for the proposed deep learning framework is the MCNN which uses multiple Convolution neural networks (CNN), instead of Feed-Forward Neural networks (FFNs) for implementing the CFMDNN architecture. The Convolutional Neural Network (CNN) is

considered as one of the most popular deep neural networks. The name 'convolution' in CNN comes from mathematical linear operation between matrixes called convolution (Al-sallab et al., 2015). Many studies proposed the use of CNNs for processing images and extracting features for complex Computer Vision and Image Classification tasks. Recently, some of the studies have also applied CNNs to problems corresponding to Natural Language Processing tasks and achieved some interesting results. CNN contains multiple layers, and includes a convolutional layer, non-linearity layer, pooling layer and fully connected layer.

The MCNN model allows qualitative feedback in terms of short comments in two different languages, English and Arabic to be included, and allows a multilingual information to be captured for analysing user sentiments about the products or service, and embedding it embedded in recommendation system predictions. Figure 70 show the MCNN model for proposed deep learning architecture.

The first channel in MCNN model includes NLP processing, and involves following steps for processing the text comments captured from users:

- Sentences or documents are represented as a matrix.
- Each row of the matrix corresponds to one token which is a word, and a word2vec approach is used for word-embedding.

For example, if we have 5 words and used 100 Embedding, we then have a matrix with 5×100. The CNN stage used has multiple layers that do different tasks.

Input Layer is the first one in CNN stage, and it reads the text or image, and is followed by a convolution stage for extracting features from the input, and the extracted CNN features are fed to a linear classifier. In general, use of several CNN stages and the Pooling stages for building the CNN based models allows better learning and extraction of latent features and hidden interactions between inputs.

For example, if processing NLP texts, the combination CNN and Pooling layers have the advantage of reducing the size of the feature (Sentence Matrix) as shown in the Figures 68 and 69.

After combining several CNN and Pooling layers, the final layer for our NLP processing stage includes fully connected layer followed by the SoftMax and the classification layer. The Figure 67 shows this NLP channel processing for the proposed MCNN model and for the proposed deep learning-based recommender system framework. The Figure 68 and 69 show the word2vec NLP processing for English and Arabic language comments fed as input to this stage.
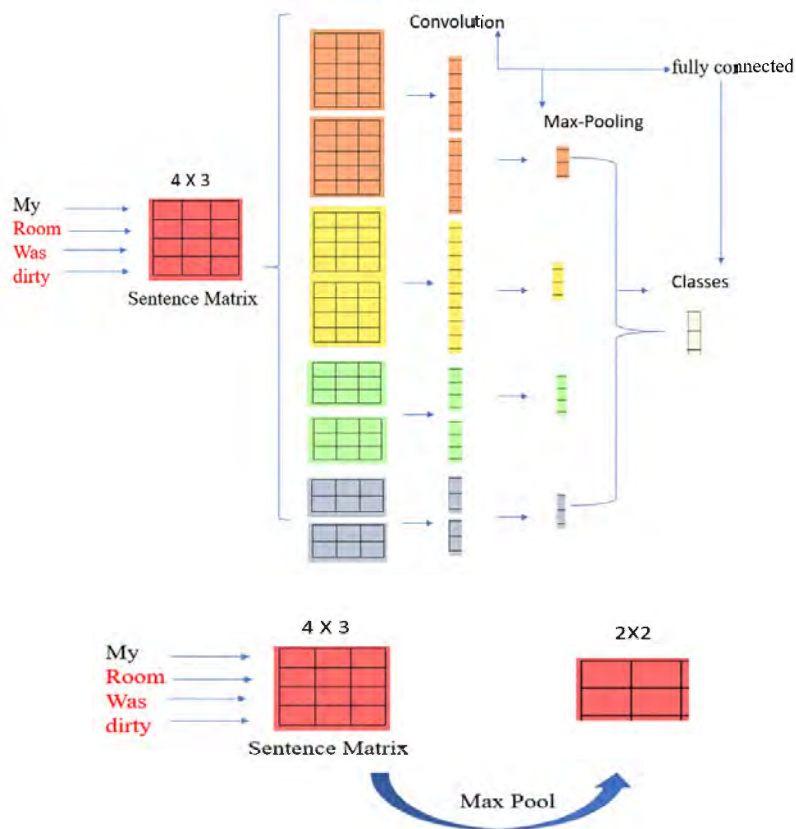
Figure 67:Illustration of the multiple-channel convolutional neural network for text

Figure 68:Sentence Matrix for English Language

Figure 69:Sentence Matrix for Arabic language

The CNN based NLP stage for modelling the sentiments captures as text feedback comments in either English or Arabic language is combined with the user and items based processing stages with CNN, and this multistage CNN model allows collaborative filtering based on both qualitative feedback in terms of text comments and quantitative feedback in terms user ratings to be modelled. The use of CNNs allows the latent interactions between the users and the items to be captured at a deeper level, leading to improved recommender system performance. The

Figure 70 shows the architecture for this multistage CNN (MCNN) model for the proposed deep learning framework.
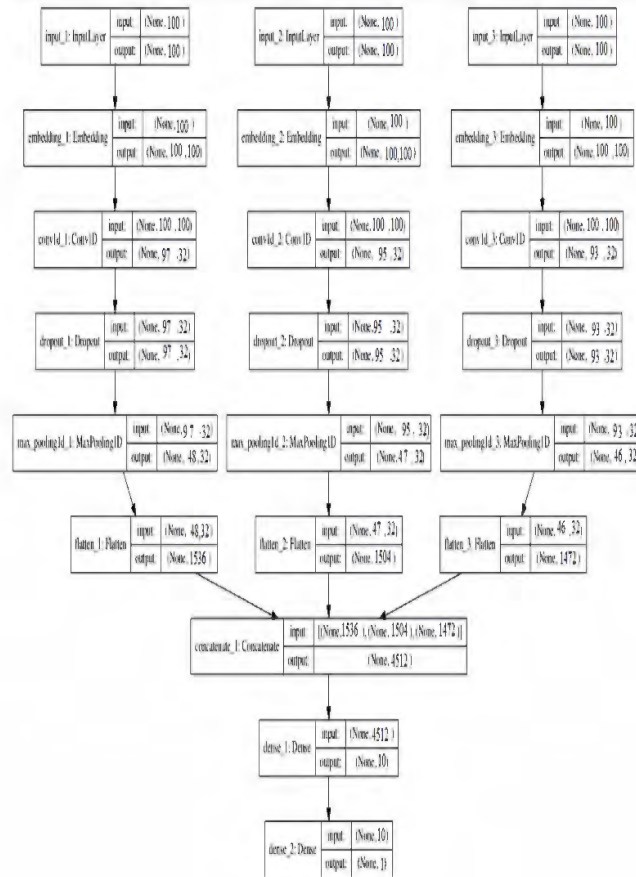
Figure 70:Multichannel Convolutional Neural Network (MCNN) Model.

6.5 Experimental Evaluation.

Different datasets were used for evaluation of two deep learning models (CFMDNN and MCNN) for the proposed deep learning framework. This section presents details of the experimental evaluation done, starting with description of publicly available datasets used.

    A. Datasets:

         • **Booking Hotel dataset**

The dataset is from Booking.com. All data is publicly available. It contains 515,000 customer reviews and scoring of 1493 hotels.

Negative_Review	Review_Total_Negative_Word_Counts	Total_Number_of_Reviews	Positive_Review	R
I am so angry that i made this post available...	397	1403	Only the park outside of the hotel was beauti...	
No Negative	0	1403	No real complaints the hotel was great great ...	
Rooms are nice but for elderly a bit difficul...	42	1403	Location was good and staff were ok It is cut...	

Hotel review Word2Vec.ipynb

Figure 71: Booking Hotel dataset Design Structure.

	text	is_bad_review
0	I am so angry that i made this post available...	1
1	No Negative No real complaints the hotel was g...	0
2	Rooms are nice but for elderly a bit difficul...	0
3	My room was dirty and I was afraid to walk ba...	1
4	You When I booked with your company on line y...	0

Figure 72:Booking Hotel dataset Design Structure after processing

- **Arabic Hotel Review dataset**

This dataset from Booking.com contains 93700 hotel reviews with positive and negative ratings and feedbacks. This dataset includes negative feedback inside positive sentence with different accent but in Arabic language.

	text	polarity
0	... المكان الذي يمكنك فيه مراجعة الذات والتفكر هو	1
1	...n\r موقع رائع وحديقة رائعة ويستحق نجمة إضافية	0
3	... بدون إدارة n\r بدون روح كأنه فندق ثلاثة نجوم	0
5	...فندق جميل، منظر رائع من بركة السباحة (على السط	1
6	...عثرت على n\r مكان لطيف وهادئ بعيدا عن الزحام	1

Figure 73: *Arabic Hotel review dataset Design Structure after processing*

```
0    0.956798
1    0.043202
Name: is_bad_review, dtype: float64
```
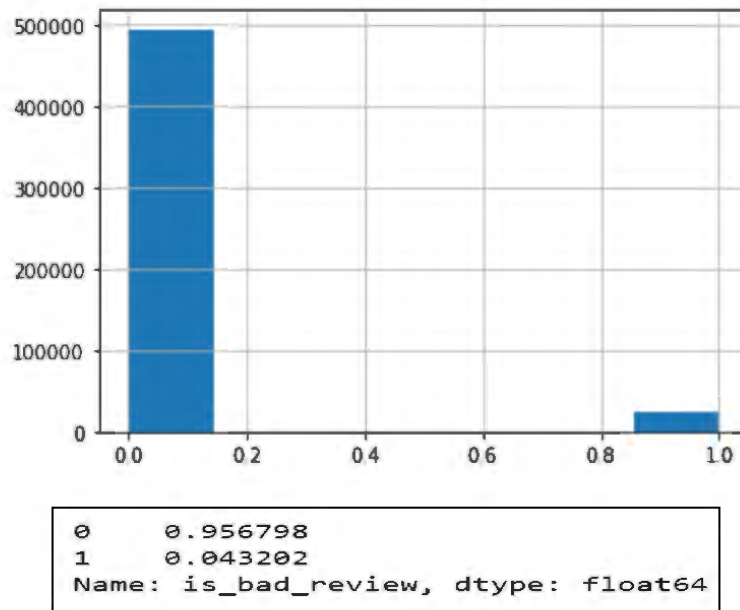
Figure 74:Number of Negative and Positive reviews after processing for English language

```
1    0.833656
0    0.166344
Name: polarity, dtype: float64
```
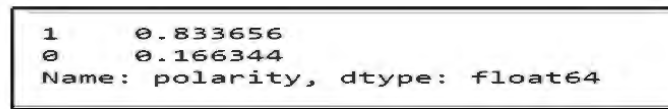
Figure 75:Number of Negative and Positive reviews after processing for Arabic language

6.5.1 Data Processing for NLP stage.

Text data preparation for NLP processing stage is different for each language. We clean datasets up by keeping original languages. All text was converted to lowercase, white space was deleted, including punctuation marks and brackets. Also, stop words from files for each dataset were deleted. The next step involves tokenization, which helps in dividing the textual information into individual words. We used NLTK kit for pre-processing the text, and using different NLTK functions, removed the stop words and tokenized each word in each dataset. Finally, the text was split into one sentence per line. Availability of clean text data is the most important part in extracting the sentiments from the text data captured in comments either in the portal, because different approaches in text cleaning can lead to very different results and influence model building. Some of the NLTK functions used for text processing involved *split()* and *clean_text()* functions (Yuan et al., 2016). Apart from this pre-processing step, no further NLP feature

engineering and other language modelling related steps were included in the NLP processing. It was left to CNN to extract the latent text features from minimally processed words from the text comments. Each textual review in the dataset used has positive and negative annotation, denoting the sentiment inferred from each review. For English dataset positive review has a 0 label, and 1 for negative review, as shown in the Figure74. For Arabic language dataset used, 1 was used for positive review and 0 for negative reviews, as shown in Figure 75. We want to use feedback and rating score together to build the collaborative filtering based deep neural model and test the model performance.

6.5.2 Results.

In this section, we describe the performance of CFDMNN and MCNN models for different datasets.

The MCNN model performs better than the CFMDNN model for English language-hotel bookings dataset, with 96% ratings prediction accuracy for MCNN model, and 95% accuracy for CFMDNN model. The improvement in accuracy for English language dataset with multichannel CNN model (MCNN)could be due to inclusion of third channel for embedding the textual reviews, or due to the use of CNN architecture. For Arabic Hotel dataset, the CFMDNN model performed better than the MCNN model, with 86% ratings prediction accuracy for CFMDNN and 83% for MCNN model. The Arabic Hotel dataset is smaller than English language dataset, with 515,000 reviews for English language dataset, and 93700 reviews for Arabic dataset. Further, the language structure for Arabic language and vocabulary is completely different as compared to the structure of English language vocabulary, and without any NLP feature engineering, the CNN in the NLP stage might not have captured the structural features specific to Arabic language. Figure 76 shows the ratings prediction accuracy for the CFDMNN and MCNN models for different datasets. As can be seen in Table 16 and the Figure 76, the rating prediction accuracy is improved for MCNN model as compared to CFDMM model for English language dataset (hotel-bookings dataset, as the CNN layer used in MCNN is able to extract the NLP features better as compared to FFN used in CFDMM. On the other hand, as can be seen in Table 17 and the Figure 76 for Arabic dataset, the FFN used for CFDMNN network can extract the latent features corresponding to interaction between the users and items better than the CNN used in MCNN, in spite of the limited data availability for building this model. However, the addition of NLP channel for embedding sentiments

expressed via text comments, even with sophisticate CNN processing layer does not seem to improve the overall model performance.

Table 16:Accuracy of Testing Prediction model for English dataset

ID	Text	Is_Bad_Review	Prediction Model
0	My room was dirty, and I was afraid to walk	1	Negative
1	Rooms are nice but for elderly a bit difficult	0	Positive
2	I am so angry	1	Negative

Table 17:Accuracy of Testing Prediction model for Arabic dataset

Text	Is_Bad_review	Prediction model
فندق جميل منظر رائع من بركة السياحة	1	Positive
المكان الذي يمكنك فيه مراجعة الذات والفكر	1	Positive
أسوأ فندق اقمت فيه على الإطلاق	0	Negative
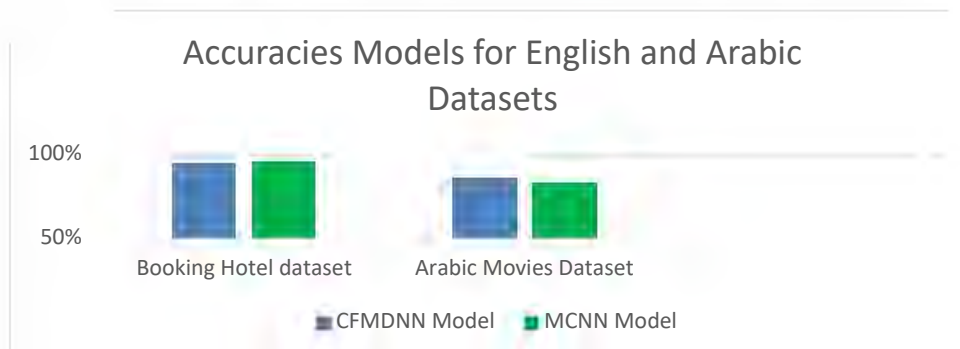
Figure 76 : *Models Accuracies for English and Arabic Language datasets*

6.6 Chapter Summary.

In this chapter we propose a deep learning based collaborative filtering framework for building recommender systems, that combines the user ratings and sentiments captured through text

comments for English and Arabic language contexts. Two different deep neural models which are the CFDMNN model, based on feedforward neural network layer, and MCNN based on multichannel CNN were developed for enhancing the performance for recommendation systems. The deep neural models with minimal feature engineering done have resulted in better performance in terms of ratings prediction accuracy, as compared to traditional approach based on collaborative filtering. The evaluation of the proposed deep learning based collaborative filtering framework, and two models (MCNN and CFDMNN) for English language and Arabic language dataset show better recommendation system performance, when traditional collaborative filtering layer is augmented with deep learning layers, and additional channels of input in terms of embedding sentiments expressed in the textual reviews is included. Next chapter discusses another novel approach to enhance the performance and robustness of recommendation systems, based on applying metadata, as digital library systems are equipped with large meta data information, and shows the benefits of using metadata.

Chapter 7 Deep Machine Learning for Digital library recommendation system based on Metadata.

7.1 Introduction.

Nowadays, massive information about anything available in the public domain, makes the finding of specific information about an item, very difficult for users. Digital libraries are such massive search and resource retrieval systems, it is important to provide tailored information retrieval for users of such systems and enhance their user experience. It helps users to find what they need directly, on can make the digital libraries, a valuable information and knowledge warehouse. FRBR (Functional Requirements for Bibliographic Records) is a model that describes the grouping of different entities in the bibliographic environment and their arrangement in a hierarchical format, illustrating the bibliographic relationship between the different entities within the registers. It is a conceptual model of the bibliographic environment which includes the data that the library wants to make available to the users (Strader, 2017). It should be noted, that the FRBR has a digital objective with RDA (Resource Description and Analysis) as its complementary component, and they work together to provide a strong resource descriptor to the items. With the rapid advancements in information and communication technologies, the range and quality of services offered by digital libraries have improved significantly, with better user experience, improved tailored and personalized recommendations, and sensitivity to user preferences.

In this Chapter, we discuss a new approach proposed in this research that uses word embedding techniques such as Word2Vec and FastText for both Arabic and English datasets, for pre-processing and helping to convert language text vectors to real numbers. By combining the strengths of the deep learning models with that of word embedding techniques for enhancing the metadata classification, is the key contribution presented in this topic for enhancing the performance of recommendation system. By considering metadata to build better recommendations, it is possible to help many users to discover resources and tailored and personalized information needed. Chapters 3,4,5 and 6 showed how machine learning and deep learning algorithms can improve recommendation system, and lead to high retrieval accuracy, by taking into consideration explicit ratings from users, sentiments expressed in feedback

comments in terms of free text. In this chapter we investigate algorithmic approaches to incorporate metadata information to enhance the performance of recommendation system, and provide better references. There are a lot of advantages of using metadata, for improving the system performance, as it available easily for any digital resource search and retrieval system, and more so in digital library application scenario, as libraries use detailed cataloguing standards, that inherently contain rich metadata, and this can solve the cold start problem for recommendation systems with much ease. With rich metadata, we do not need users to use the system, for item recommendations to be created, or to recommend to them similar items they are searching for, because metadata stored for the item being searched, will help in comparing the bibliographic results, and then it can recommend resources based on the analysis of metadata. By combining additional support from metadata along with explicit ratings and sentiment analysis of feedback text, it could be possible to enhance the recommender system performance for complex application contexts and address the short-comings such as cold-starts. The rest of the chapter discussed this approach in detail.

7.2 Digital Library Metadata.

Metadata is providing information about information. It means information which is comprehensible to the computer to define, locate, or describe web resources. For a long time, metadata has been involved by library by a card catalogue. They called it as cataloguing rules, controlled vocabulary Indexing format etc. For machines they have developed. Metadata is standard bibliographic information summaries, indexing terms and abstracts for discovery, evaluation, fitness for use, access, transfer, and citation. It also can be Information about authenticity availability and accessibility, digital signature, copyright, reproduction. E-resources, textual information graphics and any electronic works are the application of metadata. DLs contain a wealth of information and resources that users can use to help them find what they need. Users usually try to ascertain other user's opinions that are found online. As a result of that, much opinion and sentiment about specific topics could be collected and analysed from these websites. Therefore, the need to automate the process of text sentiment analysis has now arisen. It will be helpful for users to be able to access opinions and sentiments about a specific resource in a reasonable manner.

7.2.1 Types of metadata.

There are many kinds of metadata, mainly divided into three main categories. The first one is administrative metadata which is contain rights metadata that contains information related to rights and use information. It used to provide resource management information, e.g. when and how the resource was developed. Also, technical metadata which contains technical details such as file size. Preservation metadata which contains information about object. Second type is descriptive metadata which is describes a resource. Third type is structural metadata which describes how the information fits together and it information required to document an item 's internal layout so it can be rendered to the user in a sensible form (for example, a book needs to be presented in the order of its page). This type of metadata is required because it may often consist of several (sometimes thousands) files. (University Library, 2020).

7.2.2 Metadata role.

Metadata plays a significant role in the digital information system with important purposes like data description, data browsing, data transfer (Solodovnik, 2011).

- Metadata increases accessibility.
- Metadata's main function is to search for resource discovery and its location.
- Metadata for Interoperability
- Compatibility of metadata with the information structures for information retrieval and exchange.
- Metadata for Multi-Versioning
- The creation of multi-versions for the same object has many purposes: preservation research, dissemination / product development purpose.
- Metadata for right management
- Metadata lets depositors monitor the various layers rights and reproductions of information that exist for information objects and their different versions.
- Metadata for system Improvement

Metadata is also beneficial for systems evaluation and refinement to make them more effective from a technological and economical point of view. The data can also be used to design new system.

7.3 Semantic web.

It is one of the modern technologies in the field of web technology and developed in the field of educational technology which is based on artificial intelligence in the processes of classification, research and management of websites. By indexing what is placed in it and reconciling it with its synonyms, then the ability to distribute that information for use in more than one context. This technology is the third generation for the Web (or Web 3.0). In addition, one of the most useful techniques of semantic web is semantic search. It is one of the recent trends that rely on semantic web applications to retrieve information from the system, depending on the significance of the terms that the beneficiary wants to get results about, instead of the system retrieves results based on the common arrangement of sites, the focus in the semantic search is to provide results consistent with the meanings of words. Thus, when the user directs a query to the retrieval system that includes a word or phrase, the semantic search mechanism aims to provide the results most relevant to his inquiry, according to the meaning of the words that the user is searching for. Furthermore, semantic similarity used mathematical tools to evaluate the strength of the semantic relationship between units by a numerical description based on assessment of information that related to meaning (Harispe et al.,2015). Hence, we propose to use this application with metadata to build recommendation system, and the experimental results show that the semantic similarity gives good results.

7.4 Support Vector Machine (SVM).

SVM is a supervised machine learning model that uses classification algorithms as follow:

$$h(x_i) = \begin{cases} +1 & if \; w \cdot x + b \geq 0 \\ -1 & if \; w \cdot x + b < 0 \end{cases}$$

Several researchers have used classification for classifying system search results into different topic such as sport, games, cooking, travel, etc, and help users to identify results quickly in the area of interest. In addition, we used SVM to classify subjects of digital library metadata by present prototype system that has been developed to perform categorization of digital library recommendation system results. Also, we applied this automatic classifier based on a supervised machine learning algorithm, support vector machine (SVM) for classifying recommendation system results based on library metadata, user's review and ratings. This help to improve prediction of item for users based on their needs and interests.

7.5 Similarity Measures.

7.5.1 TF-IDF Term Frequency Inverse Document Frequency similarity.

TF-IDF is a feature vector that helps to compute semantic similarities between text. TF-IDF is an abbreviation for Term Frequency-Inverse Document Frequency, is algorithm that used to transform a text into numbers.

$$tfidfweight = \sum_{i \in d} tf_{i,d} * log(\frac{N}{df_i})$$

where, tfi,d is the number of occurrences of i th term in document d, dfi is the number of documents containing i th term, N is the total number of documents (Sitikhu et al, 2019). We used fidfVectorizer which is transforming text to feature vectors that can be used as input to the estimator. We applied it on Abstract and title as shown below:

Figure 77: Fi-DF Vectorizer for Arabic language.

Figure 78:TF-IDF Matrix for Arabic Language.

7.5.2 Cosine similarity.

The Cosine Similarity is finding the similarity between items. It is a symmetrical algorithm, which means that the result from computing the similarity of source A to source B is the same as computing the similarity of source B to source A

$$similarity(A,B) = \frac{A \cdot B}{\|A\| \times \|B\|} = \frac{\sum_{i=1}^{n} A_i \times B_i}{\sqrt{\sum_{i=1}^{n} A_i^2} \times \sqrt{\sum_{i=1}^{n} B_i^2}}$$

7.5.3 Jaccard similarity (JS).

Is finding textually similar documents in a large corpus or data such as news. Jaccard similarity is looking to find text at character-level similarity not meaning. (Niwattanakul et al.,2013). Jaccard similarity measurement by coefficient between two datasets is the result of division between the number of features divided by the number of properties as shown below:

$$JS = \frac{|A \cap B|}{|A \cup B|}$$

For example, Consider two sets A = {8, 9, 2, 4, 6} and B = {0, 8, 6, 5, 7, 2}

$$JS = \frac{|\{8,6\}|}{|\{0,8,6,5,7,2,4,9\}|} = \frac{2}{8} = 0.25$$

7.5.4 Semantic Similarity.

Semantic similarity measure is one of the most important tools in the recent years. It has a great interest Natural Language Processing (NLP). It is an important role in the field of linguistics, especially those related to the similarity of words meaning based on paradigmatic relations. Semantic similarity measures compute the similarity between ideas included in knowledge sources by measure distance between items based on the likeness of the meaning. There are many tools for semantic similarity such as WordNet and Word2Vec. In addition, Word2Vec is an open sourced word embedding training toolkit. Word2vec was created and published in 2013 by Tomas Mikolov at Google (Mikolov et al., 2013).

Relationship	Example 1	Example 2
France - Paris	Italy : Rome	Apple : Iphone
Big - Bigger	Small : Larger	Kona : Hawaii
Miami - Florida	Baltimore : Maryland	USA : Pizza
Einstein - Scientist	Messi : Midfielder	Obama : Barack
Sarkozy - France	Google : Android	Quick : Quicker

Figure 79:Examples of the word pair relationships by Mikolov (2013).
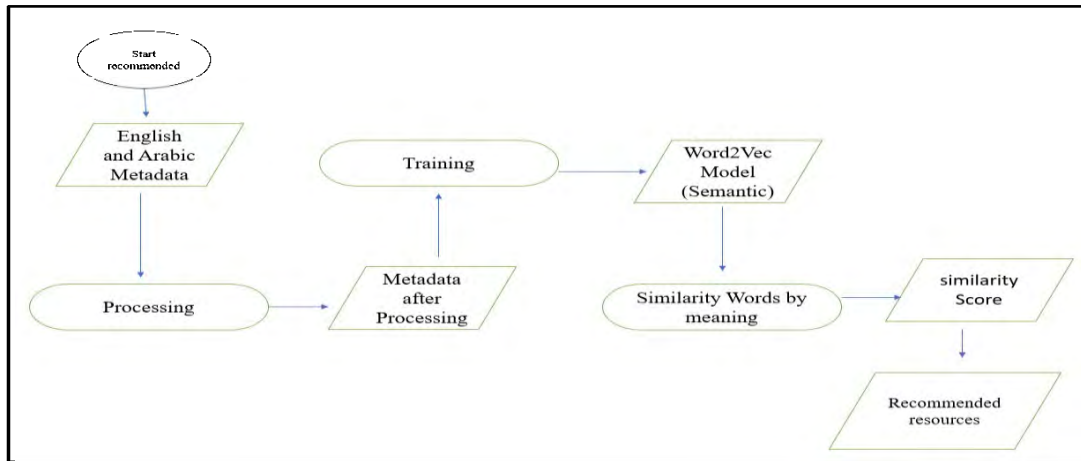
Figure 80: Flowchart of Semantic Recommender System Based on Metadata.

7.6 Top N Accuracy metrics.

The evaluation of the performance of these methods was done by measuring the top-N accuracy. Top N accuracy is the conventional accuracy in which the answer from the model must be the expected answer. (Sitikhu et al.,2019). So, the subjects of each source the most similar sources were checked and compared. If sources are in the same subjects then it is noted as correctly classified, otherwise not. We used title, abstract for creating model and comparing results.

7.7 Experimental works.

In our experimental work, we are going to extract metadata and datasets from Saudi digital library website. We used Support Vector Machine classification-based method for metadata extraction (SVM) algorithm. In addition, Dublin Core it used as a metadata standard and defines 15 elements for resource description: Title, Creator, Subject, Description, Contributor, Publisher, Date, Type, Format, Identifier, Source, Relation, References, Is Referenced By, Language, Rights and Coverage (Han et al. ,2003). We classify the Subject, Title and Abstract

of the metadata into 28 classes for Arabic dataset and 14 for English dataset. Then we applied similarity techniques on metadata to find the relationship between users and resources for both languages Arabic and English. Moreover, in our experimental work, we used cosine similarity, Jaccard similarity, TF-IDF similarity and semantic similarity to find the relationship between source title and Abstract. We used these four methods to find firstly the best method to representing words. After that we used Convolutional Neural Network (CNN) to improve the text classification and gives good result for both datasets. Additionally, we applied RNN and Long Short-Term Memory (LSTM) because RNN shows poor result when it used alone LSTM gives good result as shown in result section. Besides, in this experimental study we produced a high-performance recommender system for digital library by using hybrid deep leaning techniques based on metadata for both languages Arabic and English.

7.7.1 Datasets.

Dataset is from Saudi Digital Library (SDL) which was established to provide an advanced information services, including making available digital sources of information, and making it accessible to all researchers and students. In addition. SDL contains the largest gathering of digital information sources in the Arab world, as it currently contains more than (446) digital books with full texts and (169) global and Arabic data base that includes the full texts of millions of academic articles and more than (5,200,000) million university theses and (461) A thousand of multimedia includes images and scientific films in various scientific disciplines, which fall within the scope of interest of educational institutions and were obtained through more than 300 international publishers. We used Saudi Digital library resources dataset by using data mining techniques and extracting data from library website. The data is extracting for both Arabic and English languages. We have 10.000 resources for both languages. The datasets have different categories which are (Computer science, Education, Electronic computer, Human computer interaction, Technology, Children and Islamic History. Furthermore, dataset contains Metadata of each resource such as title, Author, abstract, categories and description. As shown below:

	Title	Author	Journal Title	ISSN	ISBN	Publication Date	Volume	Issue	First Page	Page Count	Accession Number
0	العوامل الخمسة الكبرى للشخصية وبيئة التعلم ...كما	عماد خالد العمري: نصر يوسف مقابلة	Journal of Al-Quds Open University for Educati...	23074647	NaN	2019	10	28	15.0	24.0	139384612
1	أساليب التعلم لدى طالبات الكليات الصحية بجامعة...	عبير محمد العراج	Journal of Education - Sohag University	16872649	NaN	Oct-19	66	NaN	225.0	38.0	140364284
2	الآلات والتعلم العميق...من الذكاء الى الذكاء ا...	NaN	NaN	edszan	NaN	2020	NaN	NaN	NaN	NaN	edszan.9f4e386a-b30f-4d1c-a94e-904242522afd

DOI	Publisher	Doctype	Subjects	Keywords	Abstract	
10.5281/zenodo.3474102	Al-Quds Open University	Article	NaN	Deep Learning Style; Learning Environment; Sur...	NaN	http://sdl.edu.sa/middlew：
10.12816/EDUSOHAG.2019.47183	Journal of Education - Sohag University	Article	NaN	Achievement; learning methods; students of hea...	The aim of the current research is to identify...	http://sdl.edu.sa/middlew：
NaN	لغة العصر	Image	تجارة/ثقافة/ شؤون عسكرية/ شركات/علوم ...وتكنولوجيا	NaN	NaN	http://sdl.edu.sa/middlew：

Figure 81: Dataset metadata for Arabic language.

	Title	Authors	Source	Publisher_Information	Publication_Year	Collection	Subject_Terms	Description	Document_Type
0	Gender trends in computer science authorship	Wang, Lucy LuStanovsky, GabrielWeihs, LucaEtzi...	NaN	NaN	2019.0	Computer Science	Computer Science - Digital LibrariesComputer S...	A comprehensive and up-to-date analysis of Com...	Working Paper
1	'Computer Science for all': Concepts to engage...	Spieler, BernadetteGrandl, MariaEbner, MartinS...	NaN	NaN	2019.0	Computer Science	Computer Science - Computers and Society	Knowledge in Computer Science (CS) is essentia...	Working Paper
2	Gender Balance in Computer Science and Enginee...	Marzolla, MorenoMirandola, Raffaela	Proceedings of the 13th European Conference on...	NaN	2019.0	Computer Science	Computer Science - Computers and Society97P70K...	Multiple studies have shown that gender balanc...	Working Paper
3	FAIR and Open Computer Science Research Software	Hasselbring, WilhelmCarr, LeslieHettrick, Simo...	Proceedings of the 13th European Conference on...	NaN	2019.0	Computer Science	Computer Science - Software Engineering	In computational science and in computer scien...	Working Paper

Figure 82: Dataset metadata for English language.

7.7.2 Data processing.

In each dataset, we clean it up as dataset contained only English and Arabic metadata. All text converted to lowercase, delete the white space like commas and brackets. Moreover, the text

has been split into one sentence per line. However, Arabic language has a different mechanism. We removed punctuation, comma and white space. All these processing was in, Title, Abstract, Subjects and Journal title. Furthermore, we clean text which is the most important part because different approaches in text cleaning can lead to very various results during model training. In addition, we used Subjects features which contain keywords we build this function to transform the Subjects into a list of keywords. Moreover, we computed similarity between two documents as flowing steps: first we gave the title similarity a weight of 3.5 then abstract similarity a weight of 4.0. After that subject's similarity a weight of 1.5 and the journal similarity a weight of 1. The similarity between two documents is then:

$$Similirty\ matrix$$
$$= title\ similarity + abstract\ similarity + subjects\ similarity$$
$$+ journal\ similarity$$

If a feature does not exist, we do not consider it in the final score, for instance a document does not have a title, then the similarity score will be computed following:

$$Similirty\ matrix = abstract\ similarity + subjects\ similarity + journal\ similarity$$

All these processes are required Pre-processing steps to reduces the size of the input text documents and this called Tokenization which helps to divide the textual information into individual words. We used the split() function to split the loaded document into tokens separated by white space and we used NLTK to remove English and Arabic stop words. Besides, we used word Embedding methods such as Wors2Vec and Fasttext. Also, we used python libraries for many functions.

7.7.3 Results:

7.7.3.1 *SVM Classification*

In Arabic dataset we used 28 classes and in English dataset we used 14 classes that are going to extract the text from Abstract and title to classify a document into a Subject automatically as shown below:
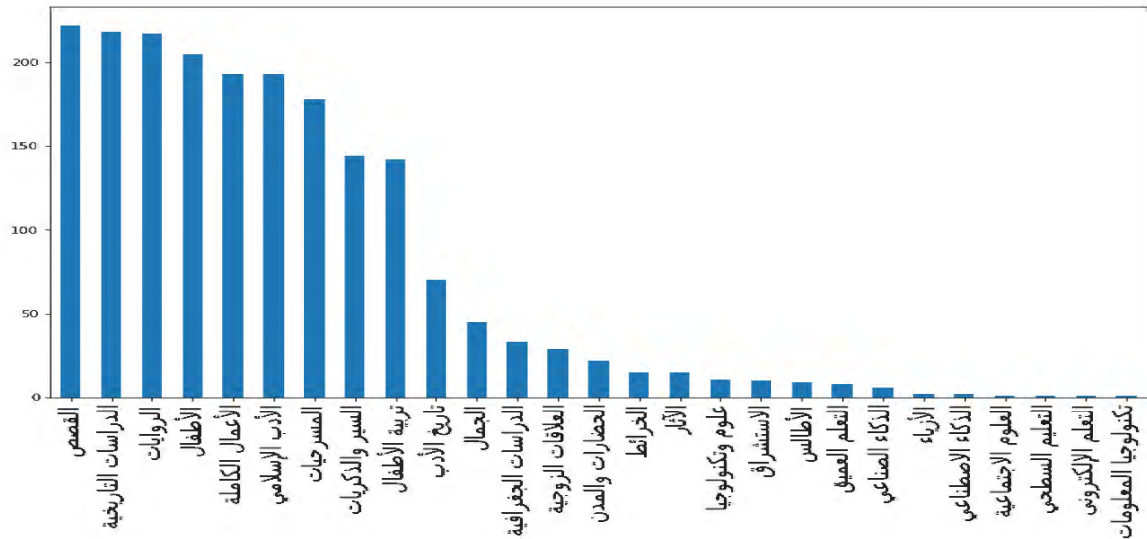
Figure 83:Classification classes for Arabic Dataset.
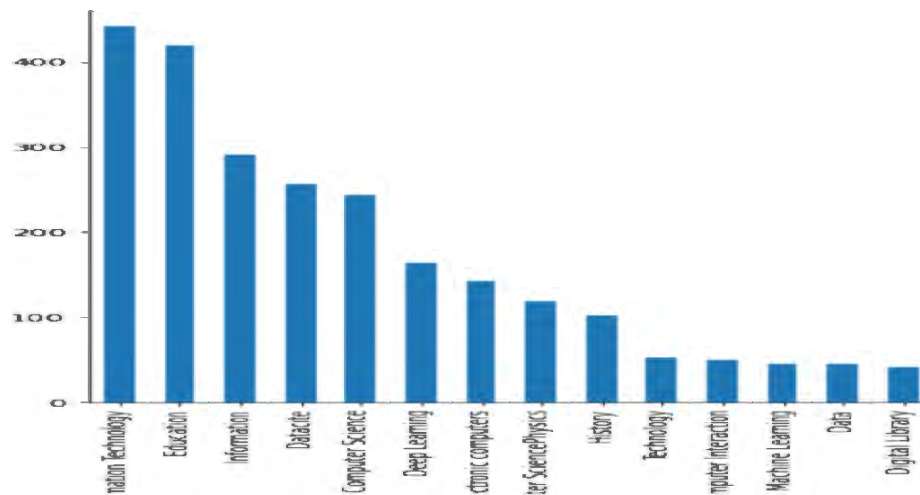
Figure 84:Classification classes for English Dataset.

Moreover, some classes are underrepresented, they have only 1 member and in order to split the dataset we need to add at least one example to these classes.

	Title	Journal Title	Subjects	Abstract
1992	...الدور التربوي النفسي لبعض وسائل الإعلام وتأثير	نور للنشر	تربية الأطفال	...استجابة للاهتمام العالمي بالطفولة باعتبارها ا
1993	...ئر اختلاف عنصري قوائم الشارات في تعلم قائمة عل	مجلة البحث العلمي في التربية	التعلم الإلكتروني	...استهدفت الدراسة الحالية بحث أثر اختلاف عنصري ت
1994	...العوامل الخمسة الكبرى للشخصية وبيئة التعلم كمن	مجلة جامعة القدس المفتوحة للأبحاث والدراسات ال	التعليم السحابي	
1995	...العوامل الخمسة الكبرى للشخصية وبيئة التعلم كمت	nan	العلوم الاجتماعية	...هدفت الدراسة إلى الكشف عن القدرة التنبؤية للعو
1996	...توظيف تضيفات تكنولوجيا المعلومات والاتصال في	مجلة كلية التربية	تكنولوجيا المعلومات	

	Title	Subjects	Abstracts
2454	Creation of a consulting tool and implementati...	LCC:Computer engineering. Computer hardwareLCC...	In this paper, a manual ontology for a Compute...
2455	Computer Science and Metaphysics: A Cross-Fert...	Computer ScienceMathematics	Computational philosophy is the use of mechani...
2456	Educated to Learn : How to enhance the educati...	UPSALLA1	The very nature of computer science with its c...
2457	Training future teachers in natural sciences a...	Computer ScienceMathematicsPhysics (Other)	The monograph defines the conditions of traini...
2458	Survey on Perception of People Regarding Utili...	LCC:Technology (General)LCC:Science	this research explores the manipulation of bio...

Figure 85:Saudi Digital library metadata for Arabic and English datasets.

Performance is evaluated using accuracy for both datasets based on title and abstract classifiers. Arabic dataset for title classifier gives 91% accuracy more than English dataset which gives 70%. Besides, Arabic dataset for abstract classifier gives 93% accuracy more than English dataset which gives 73% as shown on table 18. In our view, these results are English dataset classes of subjects are overlapping which make accuracy low and it cannot handle by SVM.

Table 18:SVM classification model accuracy

	title_classifier	abstract_classifier
Arabic Dataset	91%	93%
English Dataset	70%	73%

In addition, we Compute title vector by averaging the words within title and compute abstract vector by averaging the words within abstract. From accuracy results we can see that both datasets got good accuracy and the test classifier gives right prediction as shown below:

```
Classification of document with title: الدعوة الإسلامية فى القرن الحالى
--------------------
The prediction label according to title is: الأعمال الكاملة
The prediction label according to abstract is: الأدب الإسلامي
```

Figure 86: SVM prediction model for Arabic Dataset based on title and abstract.

```
Classification of document with title: DEEVA: A Deep Learning and IoT
Based Computer Vision System to Address Safety and Security of
Production Sites in Energy Industry
--------------------
The prediction label according to title is: Deep Learning
The prediction label according to abstract is: Machine Learning
```

Figure 87:*SVM Prediction Model for English Dataset based on Title and Abstract.*

Moreover, SVM classifier has shown significant dimensionality reduction and accuracy improvement in text classification.

7.7.3.2 *Similarity Methods.*

In similarity methods. We used the same title for all methods to compare the similarity scores and accuracy based on Top N accuracy. We recommended three sources similar to source title in Arabic dataset.

(الآلات والتعلم العميق).

The cosine similarity gave score as show in the table below:

Table 19:*Cosine similarity results for Arabic dataset.*

Title	Subjects	Similarity Score
الذكاء لاصطناعي والتعلم الالي وا لتعلم العميق .. كيف نميز بينها ؟	علوم وتكنولوجيا / Science and Technology	0.34
احدث اتجاهات الذكاء الاصطناعى و التعلم العميق	تجارة/تنمية/ثقافة/جرائم وحوادث/شركات / Companies/Crimes and Accidents/Culture/Development/Trade	0.11
مركز تدريب للذكاء الاصطناعى وال تعلم العميق فى مصر	تنمية/عمل وعمال / Development/Work and workers	0.09

From Table 19, we can see that the highest score is 0.34. In Jaccard similarity the highest score is 0.34 the same as cosine similarity, however, the other titles gave score high than cosine similarity as show below:

Title	Subjects	Similarity Score
الذكاء لاصطناعي والتعلم الالي وا لتعلم العميق .. كيف نميز بينها ؟	علوم وتكنولوجيا / Science and Technology	0.34
احدث اتجاهات الذكاء الاصطناعى و التعلم العميق	/ تجارة/تنمية/ثقافة/جرائم وحوادث/شركات Companies/Crimes and Accidents/Culture/Development/Trade	0.31
مركز تدريب للذكاء الاصطناعى وال تعلم العميق فى مصر	تنمية/عمل وعمال / Development/Work and workers	0.28

TD-FID similarity gives better score which is 0.52 as shown in the table below:

Table 21:TD-FID similarity results for Arabic Dataset

Title	Subjects	Similarity Score
الذكاء لاصطناعي والتعلم الالي وا لتعلم العميق .. كيف نمير بينها ؟	علوم وتكنولوجيا / Science and Technology	0.52
احدث اتجاهات الذكاء الاصطناعى و التعلم العميق	/ تجارة/تنمية/ثقافة/جرائم وحوادث/شركات Companies/Crimes and Accidents/Culture/Development/Trade	0.34
مركز تدريب للذكاء الاصطناعى وال تعلم العميق فى مصر	تنمية/عمل وعمال / Development/Work and workers	0.34

The last method is Word2Vec semantic similarity which gives a good score compared to other which is 0.63 which is considered good score and distance as shown in the table below:

Table 22:Word2Vec semantic similarity results for Arabic Dataset

.

Title	Subjects	Similarity Score
الذكاء لاصطناعي والتعلم الالي وا لتعلم العميق .. كيف نميز بينها ؟	علوم وتكنولوجيا / Science and Technology	0.63
احدث اتجاهات الذكاء الاصطناعى و التعلم العميق	/ تجارة/تنمية/ثقافة/جرائم وحوادث/شركات Companies/Crimes and Accidents/Culture/Development/Trade	0.58
مركز تدريب للذكاء الاصطناعى وال تعلم العميق فى مصر	تنمية/عمل وعمال / Development/Work and workers	0.56

In English dataset we used recommend item based on this title (*Gender trends in computer science authorship)* the cosine similarity gave 0.06 to 0.09 for recommender sources as shown below:

Table 23:Cosine similarity results for English Dataset

Title	Subjects	Similarity Score
Advance gender prediction tool of first names and its use in analysing gender disparity in Computer Science in the UK, Malaysia and China	Science and Technology	0.06
Investigating the Intersection of Science Fiction, Human-Computer Interaction and Computer Science Research	Computer Science - Human-Computer Interaction computer science	0.06
Effective Assessment of Computer Science Capstone Projects and Student Outcomes	Computer Science	0.09

Jaccard Similarity gave better result than cosine similarity which is 0.4 as shown below:

Table 24:Jaccard Similarity results for English Dataset

Title	Subjects	Similarity Score
Advance gender prediction tool of first names and its use in analysing gender disparity in Computer Science in the UK, Malaysia and China	Science and Technology	0.4
Investigating the Intersection of Science Fiction, Human-Computer Interaction and Computer Science Research	Computer Science - Human-Computer Interaction computer science	0.2
Effective Assessment of Computer Science Capstone Projects and Student Outcomes	Computer Science	0.3

In addition, TD-FID similarity gave 0.5 which is better than JS as shown below:

Table 25:TD-FID similarity results for English Dataset

Title	Subjects	Similarity Score
Advance gender prediction tool of first names and its use in analysing gender disparity in Computer Science in the UK, Malaysia and China	Science and Technology	0.48
Investigating the Intersection of Science Fiction, Human-Computer Interaction and Computer Science Research	Computer Science - Human-Computer Interaction computer science	0.2
Effective Assessment of Computer Science Capstone Projects and Student Outcomes	Computer Science	0.5

Besides, semantic similarity based on Word2Vec gave best score which is 0.7 similarity to source title that given as shows below:

Title	Subjects	Similarity Score
Advance gender prediction tool of first names and its use in analysing gender disparity in Computer Science in the UK, Malaysia and China	Science and Technology	0.66
Investigating the Intersection of Science Fiction, Human-Computer Interaction and Computer Science Research	Computer Science - Human-Computer Interaction computer science	0.5
Effective Assessment of Computer Science Capstone Projects and Student Outcomes	Computer Science	0.7

7.7.3.3 *Top N Accuracy.*

We calculated accuracies based on total correct recommended sources for both datasets as shown below:

$$Accuracy = \frac{\text{total no of correctly recommended sources}}{\text{Total no of sources}} * 100$$

We used SDL datasets for Arabic and English languages and recommended items based on Title, Subject and Abstract as metadata of digital library. We 2447 items under computer sciences categories for English dataset and 2098 items for Arabic dataset. In addition, each similarity method has different accuracy and semantic method gave a great result than other because it depends on meaning rather than antonyms and synonyms as shown in the table below:

Table 27:Accuracy results for English and Arabic Datasets for similarity methods

Dataset	Cosine Similarity	Jaccard Similarity	TD-FID	Semantic similarity
SDL Arabic	30%	20%	20%	96%
SDL English	4%	55%	14%	84%

From Table 27, semantic similarity for Arabic dataset has good accuracy which is 96% and 84% for English dataset. Arabic dataset used a specific Arabic language library for Python which helps to remove unnecessary characteristics and remove diacritics. The recommender system based on semantic similarity has been recommend items more than other methods because other similarity methods depend on the weight of words not on the meaning and there are many items on dataset have different words which related to the same meaning which other

methods cannot detect. For example, we recommended sources based on (*Gender Balance in Computer Science and Engineering in Italian Universities*) Title and the highest recommended item is (*When Human-Computer Interaction Meets Community Citizen Science*) if we are looking on the abstracts and categories of both titles and used the Cloud word for each abstract of title and recommended item we can see that, the (*Science word and computer are the most used*). However, the meaning of both articles is about applying computer science on community and how human computer interaction impact the community and society. Hence, the semantic method is depending on the meaning rather than the words distance. As shown below:
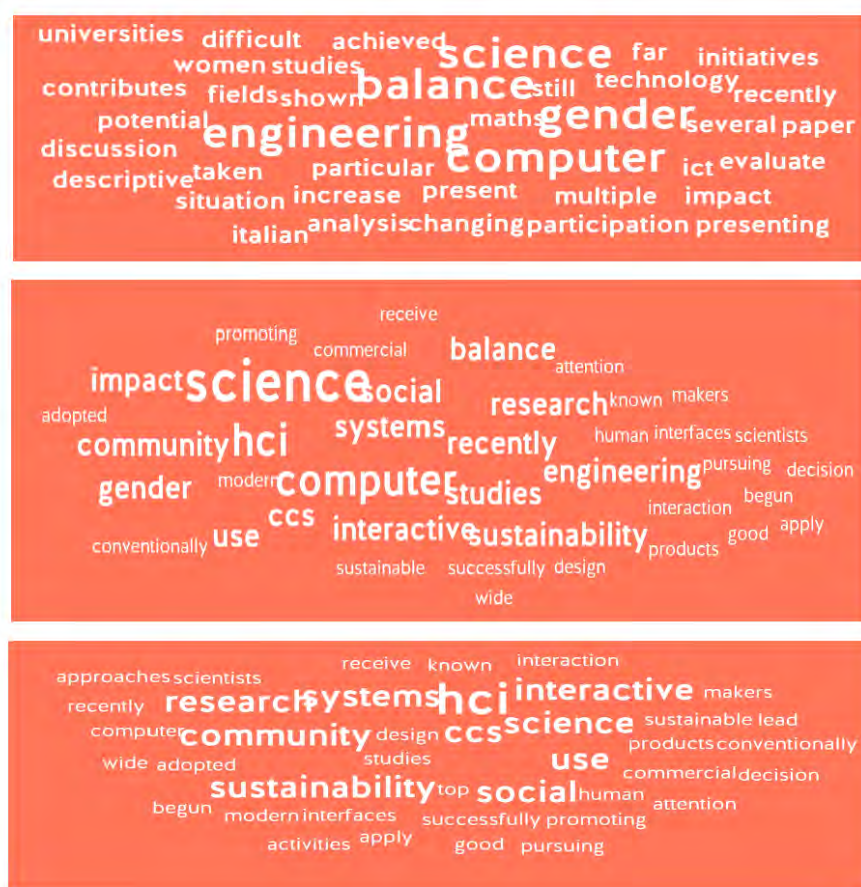
Figure 88: Cloud Words for English Titles.

Moreover, for Arabic language we use title [الآلات والتعلم العميق] the highest recommended is [الذكاء الاصطناعي والتعلم الالي والتعلم العميق .. كيف نميز] if we look on the abstracts of both items we found that both articles describe how can the intelligent artificial effect deep learning and how we can make differences between them and the most words used are [Deep learning(والتعلم العميق) and intelligence(الذكاء)] as show in the cloud words below:

Figure 89:Cloud Words for Arabic Titles.

From the results, applying machine learning and deep algorithms with metadata can improve digital library recommendation system and provide a new way to use digital library recommendation system.

7.7.3.4 *Convolution neural networks (CNN) Text classification.*

We used the same datasets and classes of SVM classifier. CNN is a class of deep, feed-forward artificial neural networks can improve the metadata classification for production to build digital library recommendation system. We are using accuracy to evaluate the model. From results CNN text classification for English dataset based on title classifier gives 93.70% accuracy than Arabic dataset which gives 88% which is better than SVM classification model. Also, abstract

classifier for English dataset improved the accuracy and gives 90% rather than SVM classification which gave 70% as shown in Table 18. From my point of view, CNN has approved good results in many fields on computer sciences and shows good performances on many experimental works. SVM works relatively well when there is clear margin of separation between classes because that is giving good result for Arabic dataset.

Table 28:Result of CNN text classification for English and Arabic datasets

	title_classifier	abstract_classifier
Arabic Dataset	88%	93%
English Dataset	93.70%	90%

We test the classifier for both dataset and gives good prediction as shown below:

```
Classification of document with title: تنمية مهارات القراءة التحليلية والتعلم العميق لدى تلاميذ الصف الخامس الإبتدائى
--------------------
The prediction label according to title is: التعلم الإلكترونى
The prediction label according to abstract is: التعلم الإلكترونى
```

Figure 90:CNN prediction model for Arabic Dataset.

```
Classification of document with title: Pedagogical problems encountered by teachers of English to Computer Science students in the Indonesian context
-------------------
The prediction label according to title is: Human-Computer Interaction
The prediction label according to abstract is: Human-Computer Interaction
```

Figure 91: CNN Prediction Model foe English Dataset.

7.7.3.5 *RNN and LSTM.*

RNN model shown poor result for both datasets because RNN cannot capture information from input text and the ratio of classes is highly different as well as it the Abstracts are long. Hence, we used LSTM to improve the results. LSTM gives good performance for Arabic dataset more than English dataset as shown in Table 29. For Arabic dataset title classifier give accuracy 86% and 44% for English dataset. Also, abstract classifier gives 81% accuracy for Arabic dataset

and 45% for English dataset. We test the model to see if it is working for predication. The Figures 92 and 93 show that the model is working well.

Table 29:Result of LSTM model for English and Arabic datasets.

	title_classifier	abstract_classifier
Arabic Dataset	86%	81%
English Dataset	44 %	45%

We test the classifier for both dataset and gives good prediction for title but not very accurate for Abstract because the performance of RNN is poor on full text, but it is strong on titles as shown below:

```
Classification of document with title:  تصمم بنظم الذكاء الصناعي
والتعلم العميق وتوظف في استخدامات أمنية وطبية وتجارية تقنيات متطورة
للتعرف على الوجوه
--------------------
The prediction label according to title is: الذكاء الصناعي
The prediction label according to abstract is: الذكاء الصناعي
```

Figure 92: RNN and LSTM prediction model for Arabic Dataset.

```
Classification of document with title: Gender trends in computer science authorship
--------------------
[15]
The prediction label according to title is: Information Technology
The prediction label according to abstract is: Information Technology
```

Figure 93: RNN and LSTM prediction model for English Dataset.

7.8 Chapter Summary.

This chapter presents the methodology we followed to construct our metadata digital library recommender system. First, the metadata for SVM, CNN, RNN and LSTM classifications were parsed for titles, subjects and Abstract and stored in a text file to be distributed among nodes.

Then the classifiers process begins by preparing the system for title and abstract. In addition, text pre-processing is applied to the title, subjects and Abstract, by applying cleaning, segmenting and stop-word removal techniques for both Arabic and English datasets. However, text pre-processing for Arabic is different from English each dataset has its process. The third step involves applying similarity method, called the semantic similarity, because it shows good performance accuracy for both datasets as compared to the Cosine, TD-IDF and Jaccard similarity. Next Chapter summarises the results and findings from this thesis, discusses the challenges and proposes directions for future research, and the thesis concludes with references and bibliography related to this work.

Chapter 8 *Conclusion and Future Works*

The main objective of this thesis is to enhance performance for digital library recommendation systems. This thesis also provides an alternative way to alleviate the problems with recommendation problems involving cold starts, by using metadata and deep machine learning techniques. The novel computation framework proposed algorithmic approaches to examine the relationship between users and resources based on ratings, sentiment analysis and metadata. The main idea is to understand how the novel recommendation system developed is likely to generate what users want, according to their relationship with system, in terms of previous behavior, and other similar users' behavior. Also, while most of the investigations focused on English language, multilingual contexts were explored as well, particularly the Arabic language context. While there are several works available for English Language, here is a limited amount of work in the case of the Arabic language, and some of the findings in this thesis can contribute significantly to the body of knowledge in this field, as the Arabic language is one of the rich morphology based languages that need more and efficient approaches for analysing the sentiments from Arabic text processing, and ratings classification for building multilingual digital library recommendation system in future.

This thesis has proposed a new computational framework and algorithmic model to use explicit ratings, user's feedback expressed as text-based sentiments and metadata to describe knowledge regarding user information needs. The novel algorithmic models have been proposed to enhance performance of recommender systems by processing multiple channels of feedback, including the use of new users' preferences, when there is limited information about these users by using deep machine learning techniques and used user's' feedbacks from explicit and implicit ratings about sources. The contributions made in this thesis allow extending the data mining techniques for improve digital library borrowing system services and personalize the services, by capturing

user's preferences. Further, by including multilingual contexts, as well as RDA and metadata information, and using open source tools and datasets, the user experience and personalization with large scale search and retrieval systems, such as digital library systems, can be extended to several other languages and dialects in the world. The novel deep learning models proposed in this work improve the quality of item recommendations. The combination of user feedbacks from rating and sentiment analysis with the item metadata representation, based on the deep machine learning lead to improved performance of digital library recommendation systems. The experimental evaluation was done on the different real world publicly available data sets collected from the BookCrossing community, the Amazon website such as (Amazon digital music, Food fine Amazon and Amazon book), MovieLen20M, Arabic Movie, Booking hotel, as well as in-house dataset with Librarika open source platform, and a private dataset obtained with necessary approvals (The Saudi Digital Library -SDL datasets) .

To summarise, this thesis proposes four novel deep learning approaches to improve the performance of the digital library recommendation system and reduce the new user (cold-start) problem, which are the CFMDNN, the MCNN, and use of multilingual sentiment text, and the Metadata in the deep learning Models. The experimental results were compared with the state-of-the-art baseline approaches and other experimental results reported in the literature. It was found that the proposed recommendation models can quality of digital library recommendations, and improve the state of the art, particularly for multilingual contexts.

8.1 Research Challenges.

Recommendation system for digital libraries needs to be more flexible and dynamic because models will need to develop according to changing in terms of time and space. This thesis work focussed on making contributions to this end. However, datasets were difficult to find, because libraries do not keep user records, as most digital library systems are based on American systems and due to privacy concerns. This problem is more acute if sentiment-based

recommendations in multilingual context are needed, such as for the Arabic language contexts, as publicly available datasets were so limited.

8.2 Future works.

The research presented in this thesis has opened new opportunities for future research, some of which are summarised below:

- The main idea of this thesis is to make use of the relationship between users and items based on different deep machine learning techniques, and building recommendations for digital library like application contexts, by including user feedbacks, preferences, explicit ratings, sentiments, and metadata. This relationship between users and items should be explored further for improving the quality of recommendations and user profiling, without compromising the privacy issues. To this end, some of the recent privacy preserving data mining and deep learning approaches could be the future directions coming out of this work. Particularly, in multilingual contexts, novel word-embeddings that are privacy preserving, could be promising to improve the quality and performance of recommender systems. Therefore, this thesis work can be extended further by focussing on privacy preserving deep learning techniques for improving the quality of recommendations for complex application scenarios, such as multi-lingual digital library systems, that may include implicit and explicit ratings, sentiment text feedback as well as the metadata information.

- Digital libraries provide information services for users who have diverse needs. Due to the large amount of data that archived in digital library systems, including text, and multimedia resources, and information about different users, it is quite challenging to maintain the conflicting requirements of privacy and user-experience, as inherently focus towards privacy, can make it less user-friendly. As main aim of recommendation system is to provide personalized experience to users in large scale search and retrieval environments, such as multi-lingual digital library systems, some of the future research efforts need to be towards the appropriate redesign of cataloguing standards, used in the current modern digital library systems, so that quality of the user experience is improved with better and personalized recommendations, without compromising the privacy aspects.

- One of the key elements of any projects that use data mining techniques and deep learning are the amount and quality of the data available. It would be useful if more publicly available datasets are available for complex application scenarios as discussed in this thesis.

- The findings in Chapters 4, 6, 7 have shown that the deep machine learning, data mining and neural networks are ideal computational approaches to predict user's opinions from the text in multilingual contexts, and used for improving the quality of recommendations. While this work has focussed on Arabic and English languages, which have different morphologies and structures, this could be useful to extend the same models to other languages, as well, leading to a global recommendation system application, that can interpret user feedback expressed in multiple different languages. Although the combination of techniques used in this thesis has produced very good results.

- Further research is planned in collaboration with the Saudi Digital library and university libraries in Saudi Arabia and Australia to obtain actual datasets of libraries and users and extend the proposed models and computational framework proposed in this thesis.

References

Aslanian, E., Radmanesh, M., & Jalili, M., 2016, 'Hybrid recommender systems based on content feature relationship'. IEEE Transactions of Services Computing, p. 1. https://doi.org/10.1109/TII.2016.2631138

Abbasi, A,. Chen, H., & Salem, A., 2008, 'Sentiment analysis in multiple languages: Feature selection for opinion classification in Web forums', *ACM Transactions on Information Systems (TOIS)*, vol. 26, no. 3, pp. 1–34.

Alharbi, A.S.M., and de Doncker, E., 2019. Twitter sentiment analysis with a deep neural network: An enhanced approach using user behavioral information. *Cognitive Systems Research*, *54*, pp.50-61

Agarwal, D., and Chen, B.C., 2011. Machine learning for large scale recommender systems. *Tutorial at ICML*.

Aly, M.A. and Atiya, A.F., 2013. LABR: a large scale Arabic book reviews dataset, Presented at the ACL (2), pp. 494–498.

Antonie, M.L., and Zaiane, O.R., 2002, December. Text document categorization by term association. In *2002 IEEE International Conference on Data Mining, 2002. Proceedings.* (pp. 19-26). IEEE.

Al Sallab, A., Hajj, H., Badaro, G., Baly, R., El-Hajj, W. and Shaban, K., 2015, July. Deep learning models for sentiment analysis in Arabic. In *Proceedings of the second workshop on Arabic natural language processing* (pp. 9-17).

Albawi, S., Mohammed, T.A., and Al-Zawi, S., 2017, August. Understanding of a convolutional neural network. In *2017 International Conference on Engineering and Technology (ICET)* (pp. 1-6). IEEE.

Adomavicius, G., and Tuzhilin, A., 2005. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE transactions on knowledge and data engineering*, *17*(6), pp.734-749.

Abid, F., Alam, M., Yasir, M., and Li, C., 2019. Sentiment analysis through recurrent variants latterly on convolutional neural network of Twitter. Future Generation Computer Systems, 95, pp.292-308.

Bussaban, K. and Kularbphettong, K., 2014. Analysis of Users' Behavior on Book Loan Log Based On Association Rule Mining. International Journal of Computer, Control, Quantum and Information Engineering, 8(1), pp.18-20.

Burrows, T., 1999. Electronic texts, digital libraries, and the humanities in Australia. *Library hi tech*.

Brownlee, J., 2017. *Deep Learning for Natural Language Processing: Develop Deep Learning Models for your Natural Language Problems*. Machine Learning Mastery

Bharati, M. and Ramageri, M., 2010. *Data mining techniques and applications*. Indian Journal of Computer Science and Engineering Vol. 1 No. 4 301-305

Bansal, T., Belanger, D., & McCallum, A., 2016, 'Ask the GRU: multi-task learning for deep text recommendations'. In *RecSys'16: Proceedings of the 10th ACM Conference on Recommender Systems*, pp. 107–114. https://doi.org/10.1145/2959100.2959180

Batmaz, Z., Yurekli, A., Bilge, A., & Kaleli, C., 2018, 'A review on deep learning for recommender systems: challenges and remedies'. *Artificial Intelligence Review*, vol. 52, pp. 1-37. https://doi.org/10.1007/s10462-018-9654-y

Boudad, N., Faizi, R., Thami, R.O.H., and Chiheb, R., 2018. Sentiment analysis in Arabic: A review of the literature. Ain Shams Engineering Journal, 9(4), pp.2479-2490.

Bendersky, M,. Wang, X., Metzler., D & Najork, M., 2017, February, 'Learning from user Interactions in Personal Search via Attribute Parameterization', in *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining* (pp. 791-799), ACM.

Betru, B.T., Onana, C.A., and Batchakui, B., 2017. 'Deep learning methods on recommender system: a survey of state-of-the-art.' International Journal of Computer Applications, 162(10), pp.17-22.

Britz, D., 2015. Recurrent neural networks tutorial, part 1–introduction to rnns. *online] www. wildml. com, Available at: http://www. wildml. com/2015/09/recurrent-neural-networkstutorial-part-1-introduction-to-rnns/.[Accessed 02 May. 2019].*

Breese JS., Heckerman, D., & Kadie, C., 1998, 'Empirical analysis of predictive algorithms for collaborative filtering,' in Uncertainty in Artificial Intelligence.

Barry, J., 2017. 'Sentiment analysis of online reviews using bag-of-words and LSTM Approaches.' In AICS (pp. 272-274).

Bisio, F., Meda, C., Gastaldo, P., Zunino, R. and Cambria, E., 2017. Concept-level sentiment analysis with SenticNet. In *A Practical Guide to Sentiment Analysis* (pp. 173-188). Springer, Cham.

Bansal, B., and Srivastava, S., 2018. 'Sentiment classification of online consumer reviews using word vector representations.' Procedia Computer Science, 132, pp.1147-1153

Beel, J., Gipp, B., Langer, S., & Breitinger, C., 2016, 'Research-paper recommender systems: a literature survey'. International Journal on Digital Libraries, vol. 17, pp. 305–338. https://doi.org/10.1007/s00799-015-0156-0.

Bin, C., 2013, August. Study on Data Mining in Digital Libraries. In International Conference on Information Computing and Applications (pp. 282-291). Springer, Berlin, Heidelberg.

Chakraborty, K., Bhattacharyya, S., Bag, R., and Hassanien, A.A., 2018b. 'Sentiment analysis on a set of movie reviews using deep learning techniques.' Social Network Analytics: Computational Research Methods and Techniques, p.127.

Chen, S.S., 2003. Preserving digital records and the life cycle of information. *Adv. Comput.*, *57*, pp.69-107.

Chen, G., Liu, D. and Li, J., 2001, December. Influence and conditional influence-new interestingness measures in association rule mining. In *10th IEEE International Conference on Fuzzy Systems.(Cat. No. 01CH37297)* (Vol. 3, pp. 1440-1443). IEEE.

Cullen, K., 2005. Delving into Data. *Library journal*, *130*(13), pp.30-33.

Cena, F., Gena, C., Grillo, P., Kuflik, T., Vernero, F. and Wecker, A.J., 2017. How scales influence user rating behaviour in recommender systems. Behaviour & Information Technology, 36(10), pp.985-1004.

Costabile, M.F., Esposito, F., Semeraro, G. and Fanizzi, N., 1999. An adaptive visual environment for digital libraries. *International Journal on Digital Libraries*, *2*(2-3), pp.124-143.

Christakopoulou, K, Beutel, A, Li, R, Jain, S, & Chi, EH 2018, Q & R : a two-stage approach toward interactive recommendation. In: *Proceedings of KDD'18, ACM*, viewed at <http://alexbeutel.com/papers/q-and-r-kdd2018.pdf>

Callan, J., Smeaton, A., Beaulieu, M., Borlund, P., Brusilovsky, P., Chalmers, M., Riedl, J., Smyth, B., Straccia, U. and Toms, E., 2003. Personalisation and Recommender Systems in Digital Libraries Joint NSF-EU DELOS Working Group Report. Sophia Antipolis, France: ERCIM. Retrieved January, 15, p.2008.

Chang, C. & Chen, R., 2006, "Using data mining technology to solve classification problems: A case study of campus digital library", Electronic Library, vol. 24, no. 3, pp. 307-321.

Chen, C.C. and Chen, A.P., 2007. Using data mining technology to provide a recommendation service in the digital library. *The Electronic Library*.

Covington, P., Adams, J. and Sargin, E., 2016, September. Deep neural networks for youtube recommendations. In *Proceedings of the 10th ACM conference on recommender systems* (pp. 191-198).

Cambria, E., Das, D., Bandyopadhyay, S. and Feraco, A., 2017. Affective computing and sentiment analysis. In *A practical guide to sentiment analysis* (pp. 1-10). Springer, Cham.

Cao, S, Yang, N, & Liu, Z., 2017, Online news recommender based on stacked. In: *2017 IEEE/ACIS 16th international conference on computer and information science* (ICIS), pp 721–726. https://doi.org/10.1109/ICIS.2017.7960088.

Cai, X, Han, J, & Yang L., 2018, Generative adversarial network based heterogeneous bibliographic network representation for personalized citation recommendation generative adversarial network based heterogeneous bibliographic network representation. *AAAI*.

Cai, C.Z., Han, L.Y., Ji, Z.L., Chen, X. and Chen, Y.Z., 2003. SVM-Prot: web-based support vector machine software for functional classification of a protein from its primary sequence. *Nucleic acids research*, *31*(13), pp.3692-3697.

Chen, J, Zhang, H, He, X, Nie. L, Liu, W, & Chua, T-S., 2017, Attentive collaborative filtering: multimedia recommendation with item- and component-level attention. In: *Proceedings of the 40th international ACM SIGIR conference on research and development in information retrieval—SIGIR'17*, pp. 335–344. https://doi.org/10.1145/3077136.3080797.

Campr, M., and Ježek, K., 2015. 'Comparing semantic models for evaluating automatic document summarization.' In International Conference on Text, Speech, and Dialogue (pp. 252-260). Springer, Cham.

Dwork, C., Feldman, V., Hardt, M., Pitassi, T., Reingold, O., & Roth, AL., 2015, June, 'Preserving Statistical Validity in Adaptive Data Analysis', in *Proceedings of the forty-seventh annual ACM symposium on Theory of computing* (pp. 117-126), ACM.

Dai, H., Wang, Y., Trivedi, R., & Song, L., 2016, Deep coevolutionary network: embedding user and item features for recommendation. In: *Proceedings OfACM conference, Halifax, Canada, August 2017 (KDD'17)*, 10.

Dai, A.M., Olah, C. and Le., Q.V., 2015. 'Document embedding with paragraph vectors.' arXiv preprint arXiv:1507.07998

Dehghani, M., Boghrati, R., Man, K., Hoover, J., Gimbel, S.I., Vaswani, A., Zevin, J.D., Immordino-Yang, M.H., Gordon, A.S., Damasio, A. and Kaplan, J.T., 2017. 'Decoding the neural representation of story meanings across languages.' Human Brain Mapping, 38(12), pp.6096-6106.

Dushay, N., 2002, July. Localizing experience of digital content via structural metadata. In *Proceedings of the 2nd ACM/IEEE-CS joint conference on Digital libraries* (pp. 244-252).

Donkers, T., Loepp, B., & Ziegler, J., 2017, Sequential user-based recurrent neural network recommendations. In: *Proceedings of the eleventh ACM conference on recommender systems—RecSys'17*, pp. 152–160. https://doi.org/10.1145/3109859.3109877.

De Nart, D. and Tasso, C., 2014. A personalized concept-driven recommender system for scientific libraries. *Procedia Computer Science*, *38*, pp.84-91.

Dushay, N. (2002). *Models and tools for generating digital libraries: localizing experience of digital content via structural metadata*. In Proceedings of JCDL 02 Conference, 244-252.

El-Beltagy, S.R., and Ali, A., 2013, March. Open issues in the sentiment analysis of Arabic social media: A case study. In *2013 9th International Conference on Innovations in Information Technology (IIT)* (pp. 215-220). IEEE.

El-Din, D.M., 2016. Enhancement bag-of-words model for solving the challenges of sentiment analysis. *International Journal of Advanced Computer Science and Applications*, *7*(1).

Elkahky, AM., Song, Y., & He, X., 2015, A multi-view deep learning approach for cross domain user modeling in recommendation systems. In: *Proceedings of the 24th international conference on World Wide Web—WWW'15*, pp 4–11. https://doi.org/10.1145/2736277.2741667.

Ferran-Ferrer, N., Casadesús, J., Krakowska, M., and Minguillón, J., 2007. Enriching e-learning metadata through digital library usage analysis. *The electronic library*, *25*(2).

Guenther, K., 2000. Applying data mining principles to library data collection. Computers in libraries, 20(4), pp.60-63.

Ghallab, A., Mohsen, A., and Ali, Y., 2020. Arabic Sentiment Analysis: A Systematic Literature Review. Applied Computational Intelligence and Soft Computing, 2020.

García, A.A.R., and Monroy-Muñoz, A., 2017. RDA description of electronic and digital resources in the digital library. *Qualitative and Quantitative Methods in Libraries*, *4*(3), pp.611-623.

Ghosh, R., Ravi, K., and Ravi, V., 2016, March. A novel deep learning architecture for sentiment classification. In *2016 3rd International Conference on Recent Advances in Information Technology (RAIT)* (pp. 511-516). IEEE.

Gabryel, M., 2018. 'The bag-of-words method with different types of image features and dictionary analysis.' J. UCS, 24(4), pp.357-371.

Greff, K., Srivastava, R.K., Koutník, J., Steunebrink, B.R., and Schmidhuber, J., 2016. LSTM: A search space odyssey. *IEEE transactions on neural networks and learning systems*, *28*(10), pp.2222-2232.

Gehrmann, S., Dernoncourt, F., Li, Y., Carlson, E.T., Wu, J.T., Welt, J., Foote Jr, J., Moseley, E.T., Grant, D.W., Tyler, P.D. and Celi, L.A., 2018. 'Comparing deep learning and concept extraction based methods for patient phenotyping from clinical narratives.' PloS one, 13(2), p.e0192360.

Gantz, J., and Reinsel, D., 2012. The digital universe in 2020: Big data, bigger digital shadows, and biggest growth in the far east. *IDC iView: IDC Analyze the future*, *2007*(2012), pp.1-16.

Huang, J., Osorio, C., and Sy, L.W., 2019a. 'An empirical evaluation of deep learning for ICD-9 code assignment using MIMIC-III clinical notes.' Computer Methods and Programs in Biomedicine, 177, pp.141-153.

Heikal, M., Torki, M., and El-Makky, N., 2018. Sentiment analysis of Arabic Tweets using deep learning. *Procedia Computer Science*, *142*, pp.114-122.

Hunter, P., and Guy, M., 2004. Metadata for Harvesting: The Open Archives Initiative, and how to find things on the Web. *The Electronic Library*.

Huang, J., Feris, R.S., Chen, Q., and Yan, S., 2015. Cross-domain image retrieval with a dual attribute-aware ranking network. In *Proceedings of the IEEE international conference on computer vision* (pp. 1062-1070).

Huang, Z., Chung, W., Ong, TH & Chen, H., 2002, 'A graph-based recommender system for digital library,' in Proceedings of the 2nd ACM/IEEE-CS joint conference on digital libraries. ACM, pp. 65-73

Hongliang, C., and Xiaona, Q., 2015, The video recommendation system based on DBN. In: *Proceedings—15th IEEE international conference on computer and information technology, CIT*. https://doi.org/10.1109/CIT/IUCC/DASC/PICOM.2015.154.

Hunter, P., &Guy, M., 2004, Metadata for harvesting: the Open Archives Initiative, and how to find things on the Web. The Electronic Library, 22,-p169.

Hsu, M.H., 2008. A., personalized English learning recommender system for ESL students. *Expert Systems with Applications*, *34*(1), pp.683-688.

Hassan, A., and Mahmood, A., 2017, April. Deep learning approach for sentiment analysis of short texts. In *2017 3rd international conference on control, automation and robotics (ICCAR)* (pp. 705-710). IEEE.

He, Z., Cai, Z., Sun, Y., Li, Y., & Cheng, X., 2017, 'Customized Privacy Preserving for Inherent Data and Latent Data,' *Personal and Ubiquitous Computing*, vol. 21, no. 1, pp. 43-54.

Harper, F.M., and Konstan, J.A., 2015. The movielens datasets: History and context. *Acm transactions on interactive intelligent systems (tiis)*, *5*(4), pp.1-19.

Harispe, S., Ranwez, S., Janaqi, S., and Montmain, J., 2015. Semantic Similarity from Natural Language and Ontology analysis preprint.

Han, J., Kamber, M., & Pei, J., 2012, *Data mining: concepts and techniques, third edition* 3rd ed., Morgan Kaufmann Publishers, Waltham, Mass.

Hart, A., 2010., The RDA Primer: "A guide for the occasional cataloger". Santa Barbara, Calif Linworth.

Jegan 2017, 'Difference between machine learning, deep learning and data mining,' *TVS Next*, 6 February, viewed 31 August 2017, http://tvsnext.io/blog/machine-learning/.

Jadhav, S.R., and Kumbargoudar, P., 2007. Multimedia Data Mining in Digital Libraries: Standards and Features. Proc. READIT, p.54.

Joshua, J.V., Alao, O.D., Adebayo, A.O., Onanuga, G.A., Ehinlafa, E.O. and Ajayi, O.E., 2016. Data mining: A book recommender system using frequent pattern algorithm. *Journal of Software Engineering and Simulation*, *3*(3), pp.01-13.

Joulin, A., Grave, E., Bojanowski, P., and Mikolov, T., 2016. Bag of tricks for efficient text classification. *arXiv preprint arXiv:1607.01759*.

Jawaheer, G., Szomszor, M. and Kostkova, P., 2010, September. Comparison of implicit and explicit feedback from an online music recommendation service. In *proceedings of the 1st international workshop on information heterogeneity and fusion in recommender systems* (pp. 47-51).

Jadhav, S.D., and Channe, H.P., 2013, Comparative Study of K-NN, Naive Bayes and Decision Tree Classification Techniques

Krishna, V., Mujjiga, S., Chakravarthil, K. and Vijayananda, J., 2019. 'Distributional semantics of clinical words.' In 2019 IEEE 13th International.

Koren, Y., Bell, R. and Volinsky, C., 2009. Matrix factorization techniques for recommender systems. Computer, 42(8).

Krishna, V., Mujjiga, S., Chakravarthil, K. and Vijayananda, J., 2019. 'Distributional semantics of clinical words.' In 2019 IEEE 13th International Conference on Semantic Computing (ICSC) (pp. 106-109). IEEE.

Kim, H., Choo, C.Y. and Chen, S.S., 2003, August. An integrated digital library server with OAI and self-organizing capabilities. In *International Conference on Theory and Practice of Digital Libraries* (pp. 164-175). Springer, Berlin, Heidelberg.

Kim, H., 2005. *Developing semantic digital libraries using data mining techniques*. University of Florida. (pp. 55-64).

Kim, H.K., Kim, H. and Cho, S., 2017. 'Bag-of-concepts: Comprehending document representation through clustering words in distributed representation.' Neurocomputing, 266, pp.336-352.

Kovacevic, A., Devedzic, V., and Pocajt, V., 2010. Using data mining to improve digital library services. The electronic library, 28(6), pp.829-843.

Kaur, R., 2019. 'Distributed knowledge based clinical auto-coding system.' In Proceedings of the 57th Conference of the Association for Computational Linguistics: Student Research Workshop (pp. 1-9).

Librarika 2017, viewed 15 August 2017, https://librarika.com/

Lynch, C., 2012," What's Become of the Digital Library?",Educause Library, https://library.educause.edu/resources/2002/1/whats-become-of-the-digital-library.

Li, Y., Jiang, ZL., Wang, X., Yiu, SM & Liao, Q., 2017, June, 'Outsourced Privacy-Preserving C4. 5 Algorithm over Arbitrarily Partitioned Databases', in *Data Science in Cyberspace (DSC), 2017 IEEE Second International Conference on* (pp. 124-132), IEEE.

Li, W., Han, J., and Pei, J., 2001, November. CMAR: Accurate and efficient classification based on multiple class-association rules. In *Proceedings 2001 IEEE international conference on data mining* (pp. 369-376). IEEE.

Libraries Australia , 2015.
https://www.nla.gov.au/librariesaustralia/sites/nla.gov.au.librariesaustralia/files/search-manual.pdf
Viewed on 10 July 2020.

Lone, T.A., and Khan, R.A., 2014. Data Mining: Competitive Tool to Digital Library. DESIDOC Journal of Library & Information Technology, 34(5).

Lee, T.Q., Park, Y., and Park, Y.T., 2008. A time-based approach to effective recommender systems using implicit feedback. *Expert systems with applications*, *34*(4), pp.3055-3062.

Lulu, L., and Elnagar, A., 2018. Automatic Arabic dialect classification using deep learning models. *Procedia computer science*, *142*, pp.262-269.

Lai, S., Xu, L., Liu, K., and Zhao, J., 2015, February. Recurrent convolutional neural networks for text classification. In *Twenty-ninth AAAI conference on artificial intelligence*.

Mnih, A., and Salakhutdinov, R.R., 2008. Probabilistic matrix factorization. In *Advances in neural information processing systems* (pp. 1257-1264).

Mandal, S., and Maiti, A., 2018, December. Explicit feedbacks meet with implicit feedbacks: a combined approach for recommendation system. In *International Conference on Complex Networks and their Applications* (pp. 169-181). Springer, Cham.

Mooney, R.J., and Roy, L., 2000, June. Content-based book recommending using learning for text categorization. In Proceedings of the fifth ACM conference on Digital libraries (pp. 195-204).

Mourad, A., and Darwish, K., 2013, June. Subjectivity and sentiment analysis of modern standard Arabic and Arabic microblogs. In *Proceedings of the 4th workshop on computational approaches to subjectivity, sentiment and social media analysis* (pp. 55-64).

Mountassir, A., Benbrahim, H. and Berrada, I., 2012, October. An empirical study to address the problem of unbalanced data sets in sentiment classification. In *2012 IEEE international conference on systems, man, and cybernetics (SMC)* (pp. 3298-3303). IEEE.

Mikolov, T., Chen, K., Corrado, G. and Dean, J., 2013. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

Mishra, R.N., and Mishra, A., 2013. Relevance of Data Mining in Digital Library. International Journal of Future Computer and Communication, 2(1), p.10

Mukherjee, D., Banerjee, S., Bhattacharya, S., and Misra, P., Tata Consultancy Services Ltd, 2016. *Method and system for context-aware recommendation*. U.S. Patent 9,510,050.

MOHE, 2010. Establishing the Saudi digital library. Retrieved from http://www.mohe.gov.sa/ar/news/Pages/30_october_2_2010.aspx.

McGinty, L., & Smyth, B., 2006, "Adaptive Selection: An Analysis of Critiquing and Preference-Based Feedback in Conversational Recommender Systems", International Journal of Electronic Commerce, vol. 11, no. 2, pp. 35-57.

Mai, F., Galke, L., & Scherp, A., 2018, 'Using Deep Learning for Title-Based Semantic Subject Indexing to Reach Competitive Performance to Full-Text', in *Proceedings of the 18th ACM/IEEE on joint conference on digital libraries*, ACM, pp. 169–178.Nürnberg, P.J., Furuta, R., Leggett, J.J., Marshall, C.C., and Shipman III, F.M., 1995, June. Digital Libraries: Issues and Architectures. In *DL* (p. 0).

Niwattanakul, S., Singthongchai, J., Naenudorn, E., and Wanapu, S., 2013, March. Using of Jaccard coefficient for keywords similarity. In Proceedings of the international multiconference of engineers and computer scientists (Vol. 1, No. 6, pp. 380-384).

Ng, Y.K., and Jung, U., 2019, November. Personalized Book Recommendation Based on a Deep Learning Model and Metadata. In *International Conference on Web Information Systems Engineering* (pp. 162-178). Springer, Cham.

Nazir, S., Yousaf, M.H., Nebel, J.C., and Velastin, S.A., 2018. 'A bag of expression framework for improved human action recognition.' Pattern Recognition Letters, 103, pp.39-45.

National Library of Australia. Available: http://www.nla.gov.au

Najafabadi, M.M., Villanustre, F., Khoshgoftaar, T.M., Seliya, N., Wald, R., & Muharemagic, E., 2015, "Deep learning applications and challenges in big data analytics", Journal of Big Data, vol. 2, no. 1, pp. 1-21.

Nürnberg, P.J., Furuta, R., Leggett, J.J., Marshall, C.C., and Shipman III, F.M., 1995, June. Digital Libraries: Issues and Architectures. In DL (p. 0).

Özgöbek, Ö., Shabib, N., and Gulla, J.A., 2014. Data Sets and News Recommendation. In *UMAP Workshops*.

Omisore, M.O., and Samuel, O.W., 2014. Personalized Recommender System for Digital Libraries. International Journal of Web-Based Learning and Teaching Technologies (IJWLTT), 9(1), pp.18-32.

Oliver, C., 2010. Introducing RDA. a guide to the basics. Chicago: American Library Association, 13;2; 65. U.S. RDA Test Coordinating Committee. (2011). Report and Recommendations of the U.S. RDA Test Coordinating Committee, 10, 179-182.

Portugal, I., Alencar P., & Cowan D., 2018, 'The use of machine learning algorithms in recommender systems: a systematic review'. Expert Systems with Applications, vol. 97, pp. 205–227. https://doi.org/10.1016/j.eswa.2017.12.020

Perkowitz, M., and Etzioni, O., 1998, July. Adaptive web sites: Automatically synthesizing web pages. In *AAAI/IAAI* (pp. 727-732).

Perkowitz, M., and Etzioni, O., 2000. Towards adaptive web sites: Conceptual framework and case study. *Artificial intelligence*, *118*(1-2), pp.245-275.

Papernot, N., Abadi, M., Erlingsson, Ú., Goodfellow, I., & Talwar, K., 2016, 'Semi-supervised knowledge transfer for deep learning from private training data', *arXiv preprint arXiv:1610.05755*.

Preston, C.C., and Colman, A.M., 2000. Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. *Acta psychologica*, *104*(1), pp.1-15.

Patil, T.R., and Sherekar, S.S., 2013. Performance analysis of Naive Bayes and J48 classification algorithm for data classification. International Journal of Computer Science and Applications, 6(2), pp.256-261.

Poria, S., Cambria, E., & Gelbukh, A., 2016, 'Aspect extraction for opinion mining with a deep convolutional neural network'. *Knowledge-Based Systems*, vol. 108, pp. 42–49.

Ragini, J.R., Anand, P.R., and Bhaskar, V., 2018. 'Big data analytics for disaster response and recovery through sentiment analysis.' International Journal of Information Management, 42, pp.13-24.

Rezaeinia, S.M., Ghodsi, A. and Rahmani, R., 2017. 'Improving the accuracy of pre-trained word embeddings for sentiment analysis.' arXiv preprint arXiv:1711.08609

Refaee, E,. and Rieser, V., 2016, June. iLab-Edinburgh at SemEval-2016 Task 7: A hybrid approach for determining sentiment intensity of Arabic Twitter phrases. In *Proceedings of the 10th international workshop on semantic evaluation (SEMEVAL-2016)* (pp. 474-480).

Rustogi, S., Sharma, M., and Morwal, S., 2017. Improved parallel apriori algorithm for multi-cores. *International Journal of Information Technology and Computer Science (IJITCS)*, *9*(4), p.18.

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

Strader, C.R., 2017, "From User Tasks to User Services: Placing the Functional Requirements for Bibliographic Records Models Into a Larger Framework", Technical Services Quarterly, vol. 34, no. 4, pp. 347.

Soni, S., and Sharaff, A., 2015, March. Sentiment analysis of customer reviews based on hidden markov model. In *Proceedings of the 2015 International Conference on Advanced Research in Computer Science Engineering & Technology (ICARCSET 2015)* (pp. 1-5).

Sitikhu, P., Pahi, K., Thapa, P., and Shakya, S., 2019, November. A Comparison of Semantic Similarity Methods for Maximum Human Interpretability. In 2019 Artificial Intelligence for Transforming Business and Society (AITB) (Vol. 1, pp. 1-4). IEEE.

Siersdorfer, S., Chelaru, S., Nejdl, W., and San Pedro, J., 2010, April. How useful are your comments? Analyzing and predicting YouTube comments and comment ratings. In *Proceedings of the 19th international conference on World wide web* (pp. 891-900).

Schafer, J.B., Konstan, J.A., and Riedl, J., 2001. E-commerce recommendation applications. *Data mining and knowledge discovery*, *5*(1-2), pp.115-153.

Shokri, R., & Shmatikov, V., 2015, October, 'Privacy-Preserving Deep Learning', in *Proceedings of the 22nd ACM SIGSAC conference on computer and communications security* (pp. 1310-1321). ACM.

Stojanovic, J., Gligorijevic, D., Radosavljevic, V., Djuric, N., Grbovic, M. and Obradovic, Z., 2017. 'Modeling healthcare quality via compact representations of electronic health records.' IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB), 14(3), pp.545-554.

Solodovnik, I., 2011. Metadata issues in Digital Libraries: key concepts and perspectives. JLIS. it: Italian Journal of Library, Archives and Information Science. Rivista italiana di biblioteconomia, archivistica e scienza dell'informazione, 2(2), pp.4-27.

Sandoval, SV., 2015, Novelty and diversity evaluation and enhancement in recommender systems. PhD Dissertation, Universidad Autonoma de Madrid, viewed at <https://saulvargas.es/phd-thesis.pdf>

Song, Y. and Wei, R., 2011, July. Research on application of data mining based on FP-growth algorithm for digital library. In Mechanic Automation and Control Engineering (MACE), 2011 Second International Conference on (pp. 1525-1528). IEEE.

Safder, I, Hassan, S-U., Visvizi, A., Noraset, T, Nawaz, R., & Tuarob, S., 2020, 'Deep Learning-based Extraction of Algorithmic Metadata in Full-Text Scholarly Documents', Information processing & management, vol. 57, no. 6, p. 102269–.

Safder, I., Batool, H., & Hassan, S.-U., 2018, Deep feature engineering using full-text publications. In ICADL Poster Proceedings. Hamilton, New Zealand: The University of Waikato

Schnick, T., and Knickelbine, M.J., 2000. *The Lexile framework: An introduction for educators*. MetaMetrics.

Sohail, S.S., Siddiqui, J. and Ali, R., 2013, August. Book recommendation system using opinion mining technique. In *2013 international conference on advances in computing, communications and informatics (ICACCI)* (pp. 1609-1614). IEEE.

Taigman, Y., Yang, M., Ranzato, M.A. and Wolf, L., 2014. Deepface: Closing the gap to human-level performance in face verification. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1701-1708).

Tillett, B., 2005. What is FRBR? A conceptual model for the bibliographic universe. The Australian Library Journal, 54(1), pp.24-30.

Tang, J., & Wang, K., 2018, Personalized top-N sequential recommendation via convolutional sequence embedding. In: *Proceedings of the eleventh ACM international conference on web search and data mining—WSDM'18*, pp 565–573.

Thai, M.T., Wu, W., and Xiong, H., eds., 2016. *Big Data in Complex and Social Networks*. CRC Press.

Tsuji, K., Kuroo, E., Sato, S., Ikeuchi, U., Ikeuchi, A., Yoshikane, F. and Itsumura, H., 2012, September. Use of library loan records for book recommendation. In Advanced Applied Informatics (IIAIAAI), 2012 IIAI International Conference on (pp. 30-35). IEEE

Thabtah, F.A., Cowling, P., and Peng, Y., 2004, November. MMAC: A new multi-class, multi-label associative classification approach. In *Fourth IEEE International Conference on Data Mining (ICDM'04)* (pp. 217-224). IEEE.

Uppal, V., and Chindwani, G., 2013. An empirical study of application of data mining techniques in library system. International Journal of Computer Applications, 74(11).

University library, 2020, viewed on 11 July 11, 2020, https://guides.library.ucsc.edu/c.php?g=618773

Vinayakumar, R., Soman, K., & Poornachandran, P 2018, 'Evaluating deep learning approaches to characterize and classify malicious URL's', *Journal of intelligent & fuzzy systems*, vol. 34, no. 3, pp. 1333–1343.

Véras, D., Prota, T., Bispo, A., Prudêncio, R., & Ferraz, C., 2015, 'A literature review of recommender systems in the television domain'. Expert Systems with Application, vol. 42, no. 22, pp. 9046–9076. https://doi.org/10.1016/j.eswa.2015.06.052Review

Velasquez, D., & Campbell-Meier, J., 2015, 'Digital library development in Australia'. *Australian Library and Information Association.* Retrieved from http://information-online.alia.org.au/content/digital-library-development-australia.

Vateekul, P., and Koomsubha, T., 2016, July. A study of sentiment analysis using deep learning techniques on Thai Twitter data. In *2016 13th International Joint Conference on Computer Science and Software Engineering (JCSSE)* (pp. 1-6). IEEE.

Webster, J., Jung, S., & Herlocker, J., 2004, 'Collaborative filtering: a new approach to searching digital libraries,' *New Review of Information Networking*, vol. 10, no. 2, pp. 177-191.

Walkowiak, T., Datko, S., and Maciejewski, H., 2018. 'Bag-of-words, bag-of-topics and word-to-vec based subject classification of text documents in Polish-a comparative study.' In International Conference on Dependability and Complex Systems (pp. 526-535). Springer, Cham.

Witten, I.H., and Frank, E., 2002. Data mining: practical machine learning tools and techniques with Java implementations. *Acm Sigmod Record*, *31*(1), pp.76-77.

Willemsen, M., 2016, 'Anonymizing Unstructured Data to Prevent Privacy Leaks during Data Mining'. in *Proceedings of 25th Twenty Student Conference on IT*.

Wang, P., 2012, 'A personalized collaborative recommendation approach based on clustering of customers,' *Procedia Physics*,vol. 24, pp. 812-816. doi.org/10.1016/j.phpro.2012.02.121

Wong, A., and Factura, B., Amazon Digital Music: Sentiment Analysis and Text Mining. *Amazon Digital Music: Sentiment Analysis and Text Mining*.

Whitney, C., and Schiff, L.R., 2006. The Melvyl recommender project: Developing library recommendation services. California Digital Library.

Wang, H., Wang, N., and Yeung, D.Y., 2015, August. Collaborative deep learning for recommender systems. In Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining (pp. 1235-1244).

Yin, X., and Han, J., 2003, May. CPAR: Classification based on predictive association rules. In *Proceedings of the 2003 SIAM International Conference on Data Mining* (pp. 331-335). Society for Industrial and Applied Mathematics.

Yassir, A., and Nayak, S., 2012. Issues in data mining and information retrieval. International Journal of Computer Science and Communication Networks, 2(1), pp.93-8.

Yuan, D., Richardson, J., Doherty, R., Evans, C., and Altendorf, E., 2016. 'Semi-supervised word sense disambiguation with neural models.' arXiv preprint arXiv:1603.07012.

Yonetani, R,. Boddeti, VN., Kitani, KM & Sato, Y., 2017, 'Privacy-Preserving Visual Learning Using Doubly Permuted Homomorphic Encryption, *arXiv preprint arXiv:1704.02203*.

Zhang, L., Wang, S., and Liu, B., 2018. Deep learning for sentiment analysis: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *8*(4), p.e1253.

Zhang, Q, Yang, LT, Chen, Z, Li, P & Deen, MJ., 2017, 'Privacy-Preserving Double-Projection Deep Computation Model with Crowdsourcing on Cloud for Big Data Feature Learning', *IEEE Internet of Things Journal*.

Zhang, Q., Yang, LT., and Chen, Z., 2016, 'Privacy Preserving Deep Computation Model on Cloud for Big Data Feature Learning', IEEE Transactions on Computers, vol. 65, no. 5, pp. 1351-1362.

Zhang, X., and LeCun, Y., 2015. Text understanding from scratch. *arXiv preprint arXiv:1502.01710.*

Zhang, M., 2011. Application of Data Mining Technology in Digital Library. *JCP*, *6*(4), pp.761-768.

Zhuhadar, L., and Nasraoui, O., 2010. A hybrid recommender system guided by semantic user profiles for search in the e-learning domain. *Journal of Emerging Technologies in Web Intelligence*, *2*(4).

Zhu, Z., and Wang, J.Y., 2007, August. Book recommendation service by improved association rule mining algorithm. In *2007 international conference on machine learning and cybernetics* (Vol. 7, pp. 3864-3869). IEEE.

Zeng, Y., Yin, S., Liu, J., and Zhang, M., 2015. Research of improved FP-Growth algorithm in association rules mining. *Scientific Programming*, *2015*.

Zhichun, L., and Fengxin, Y., 2008. An improved frequent pattern tree growth algorithm. *Applied Science and Technology*, *35*(6), pp.47-51.

Zhou, Y., Jiang, L.P. and Yang, L.I.U., 2014. Data fusion algorithm based on support degree under interferential environment. Fire Control Comm. Control (03), pp.12-14.