

SAMRIDDHI COLLEGE
Tribhuvan University
Institute of Science and Technology



A PROJECT REPORT
ON
“BOOK RECOMMENDATION SYSTEM
USING COLLABORATIVE FILTERING”

In Partial Fulfillment of Requirements for the Bachelor Degree in Computer Science and
Information Technology

Submitted To:

Department of Computer Science and Information Technology,

Samriddhi College

Submitted By:

Rabin Basan (T.U. Roll No. 7609/072)

Sabina Dhital (T.U. Roll No. 7617/072)

September, 2019



Samriddhi College
[T.U. Affiliated]
Lokanthali-16, Bhaktapur

Supervisor's Recommendation

I hereby recommend that this project prepared under my supervision by RABIN BASAN an SABINA DHITAL.

Entitled "A Project Report on BOOK RECOMMENDATION SYSTEM" in partial fulfillment of the requirements for the degree of B.Sc. CSIT Computer Science and Information Technology be processed for the evaluation.

.....

SUPERVISOR
Faculty Member,
B.Sc. CSIT Department, Samriddhi College
Lokanthali, Bhaktapur



Samriddhi College
[T.U. Affiliated]
Lokanthali-16, Bhaktapur

Letter of Approval

This is to certify that this project prepared by RABIN BASAN and SABINA DHITAL entitled “A Project Report on BOOK RECOMMENDATION SYSTEM” in partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Information Technology has been well studied. In our opinion it is satisfactory in the scope and quality as a project for the required degree.

<p>.....</p> <p>Mr. Sandeep Shrestha Principal Samriddhi College</p>	<p>.....</p> <p>Mr. Ramesh Shing Saud Supervisor</p>
<p>.....</p> <p>Signature of the External Examiner</p>	<p>.....</p> <p>Signature of the Internal Examiner</p>

ABSTRACT

With the increase in the use of e-commerce sites, it has become very easy for the users to find the items of their interest without wasting a lot of time. People are busier now a days. They don't have time to search the book they want to read. There are large number of books and all the books are not reachable to every person. At past the books were collected and well arranged in library but at current time books are collected and arranged in the computer and it has become vast. People may not have sufficient time to search the book they want to read. So, to solve this problem we design the book recommendation system that helps the user to get the book they want to read and also to recommend the other books they may like. Book recommendation system is an online application that reduce user's time to search books by recommending book automatically.

The developed system is a "Book Recommendation System Using Collaborative Filtering" which deals with developing a website for a reader as well as a bookstore. The users are recommended based on previous users rating using matrix factorization technique. BRS the online application that provide the list of books to the user according to the rating provided by user. The system as a real-life application has to quickly and successfully process user's data for systematic recommendation facilities. This system can play the significant role in many applications relevant to system such as library management and e-library and other areas.

ACKNOWLEDGMENTS

The completion of project would not have been possible without the support and guidance of many individuals. We would like to express our gratitude to our project supervisor Mr. Ramesh Shing Saud sir who took interest in our project and guided us through the project by providing necessary ideas, information and knowledge for developing guiding us in different phases of book recommendation. This system is based on matrix factorization. It automatically provides the books list to the user according to the user's rating.

We are thankful and fortunate enough to get constant support for our colleagues and teaching staff of B.Sc.CSIT department, which helped us, complete our project. We would also like to extend our regards to all the non-teaching staff of B.Sc.CSIT department for their timely support.

Rabin Basan (T.U. Roll No. 7609/072)

Sabina Dhital (T.U. Roll No. 7617/072)

Table of Content

SUPERVISOR’S RECOMMENDATION	i
LETTER OF APPROVAL	v
ABSTRACT.....	iv
ACKNOWLEDGEMENT.....	v
LIST OF FIGURE.....	vii
LIST OF TABLES.....	ix
LIST OF ABBREVIATION.....	x
CHAPTER 1- INTRODUCTION.....	1
1.1 Over View	1
1.2 Statement of Problem.....	2
1.3 Objective.....	2
1.4 Scope and Limitation.....	3
1.4.1 Scope.....	3
1.4.2 Limitations.....	3
1.5 Report Organization.....	5
CHAPTER 2- LITERATURE REVIEW.....	6
CHAPTER 3 - SYSTEM REQUIREMENT ANALYSIS.....	8
3.1 Requirement Analysis.....	8
3.1.1 Preliminary Analysis.....	8
3.1.2 Functional Requirement.....	8
3.1.3 Non-Functional Requirement.....	10
3.1.4 Software Requirement.....	11
3.2 Feasibility Analysis.....	11
3.2.1 Technical Feasibility.....	11
3.2.2 Operational Feasibility.....	11
3.2.3 Economic Feasibility.....	11
3.3 Data Modeling.....	12

3.4 Process Modeling.....	13
3.4.1 Context Diagram.....	13
3.4.2 Level-1 DFD.....	14
CHAPTER 4- SYSTEM DESIGN.....	16
4.1 System Flowchart.....	16
4.2 System Architecture.....	17
4.3 Sequence Diagram.....	18
4.4 Class Diagram.....	19
4.5 Activity Diagram.....	20
CHAPTER 5 - IMPLEMENTATION AND TESTING.....	22
5.1 Module Description	22
5.2 Algorithm Implementation	23
5.3 Tool Used	25
5.4 System Testing	27
5.4.1 Unit testing	27
5.4.2 Integrating Testing	32
5.4.3 System Testing	32
CHAPTER 6- CONCLUSION AND RECOMMENDATION.....	33
6.1 Conclusion	33
6.2 Recommendation	33
CHAPYER 7 - REFRENCES	34

LIST OF FIGURES

Figure 3.1 Use Case Diagram of BRS.....	9
Figure 3.2 ER-Diagram of BRS.....	12
Figure 3.3 Context Diagram BRS.....	13
Figure 3.4 DFD level-1 of of BRS.....	14
Figure 4.1 Flow Chart of BRS.....	16
Figure 4.2 Architecture Diagram og BRS.....	17
Figure 4.3 Sequence Diagram of BRS.....	18
Figure 4.4 Class Diagram of BRS.....	19
Figure 4.5 Activity Diagram of BRS.....	20
Figure 5.1 Unit Test case for Successful login.....	28
Figure 5.2 Unit Test case for Unsuccessful login.....	29
Figure 5.3 Unit Test case for Valid Ratting.....	29
Figure 5.4 Unit Test case for Invalid Ratting	30
Figure 5.5 Unit Test case for Searching.....	31

LIST OF TABLES

Table 3.1 Use Case Diagram Table for User.....	10
Table 3.2 Use Case Diagram Table for Admin.....	10
Table 5.1 Unit Test case for Successful login.....	28
Table 5.2 Unit Test case for Unsuccessful Login.....	28
Table 5.3 Unit Test case Valid Rating.....	29
Table 5.4 Unit Test case for Invalid Rating.....	30
Table 5.5 Unit Test case for Searching.....	31
Table 5.6 Integrating Testing.....	32

LIST OF ABBRIVATION

1. BRS = Book Recommendation System
2. CF = Collaborative Filtering
3. UML = Unified Modeling Language
4. CD = Context Diagram
5. MF = Matrix Factorization
6. DFD= Data Flow Diagram
7. RDBMS= Relational Database Management System
8. SQL = Structured Query Language
9. HTML = Hypertext Markup Language
10. UI = User Interface
11. PHP = Hypertext Preprocessor
12. CSS = Cascading Style Sheet

CHAPTER 1- INTRODUCTION

1.1 Overview

With the rapid development of internet technologies, the number of online websites for books has increased which enhanced the competition among them. This online website for book store helps to Recommendation system, which is one of the stronger tools to increase number of users using website. The book recommendation system must recommend books that are of user's interest. Recommendation systems are widely used to recommend products to the end users that are most appropriate. This system uses features of collaborative filtering to produce efficient and effective recommendations. Collaborative recommendation is probably the most familiar, most widely implemented and most mature of the technologies. Collaborative recommender systems aggregate ratings of objects, recognize commonalities between users based on their ratings, and generate new recommendations.

The system makes use of numerical ratings of books provided by the active user to assess the similarity between users' profiles to predict recommendations of unseen books to active user. The system makes use of Matrix Factorization technique, which allow us to discover the latent features underlying the interactions between user and books. In the system, there is a group of users and a set of books, given that each users have rated some books. The algorithm will predict how the users would rate the books that they have not yet rated, such that proper recommendations are made for the users. In this case, all the information we have about the existing ratings can be represented in a matrix and the task of predicting the missing ratings can be considered as filling in the blanks such that the values would be consistent with the existing ratings in the matrix. The intuition behind using matrix factorization to solve this problem where there should be some latent features that determine how a user rates a book. For example, two users would give high ratings to a certain book if they both like the plot and characters in the book, or if the book is a drama, which is a genre preferred by both users.

1.2 Statement of Problem

People spend a lot of time searching books on the internet. It is difficult to find out the book people want to read on the internet. The internet is puzzle house that has a lot of similar and dissimilar contents next to each other at random. There may arise different problem while searching the book user wants. Here, we propose a BRS that can recommend books according to the user's rating or the popularity of the book. The system will make easier to search and find the book according to their interest and also saves their time in searching.

1.3 Objective

The general objective of the application is to minimize the time of the user to search book. Specific objectives of our project are:

- To provide the easy user interface to user.
- To recommend book.
- To provide platform to view the books online.
- Built user's profile for a recommender system by monitoring dynamic user behaviors. The user profile must adapt to changing user's interest.

1.4 Scope and Limitation

1.4.1 Scope

Book recommendation can further be used and implemented in various ways:

- This system can be used on e- libraries, e- commerce sites.
- This system can be implemented in android and IOS based system.
- This system can provide large number of choices of book and also recommend for books.
- This system will help users to find book easily and saves their precious time.

1.4.2 Limitation

There are certain limitations of the system. Understanding these limitations is important in order to build a successful recommendation system. The limitations of proposed system are:

- As the number of users grow, the algorithms suffer scalability issues, you would have to create a sparse matrix with one trillion elements.
- The lack of right data Input data may not always be accurate because humans are not perfect at providing ratings.
- User's taste must be represented as a learnable function these content features.
- May be unable to exploit quality judgements of other users.
- Problem may arise when there is lack of information of users and items.

1.5 Report Organization

The purpose of a report is to inform the reader about the system and the different content included in the report. There are many chapters included in the system report. The first chapter is “Introduction” that includes the basic information about the system. This chapter includes the brief introduction about the BRS. The problem statement, objective of our system is also included in this chapter. Problem statement include the information about why we needed the system and the objective include the main facts for developing the system. Similarly, this chapter also include the scope and the limitation of BRS. There may be many limitations of the system that the system may not be able to fulfil which is included in this system. Next, chapter is “Literature Review”. A literature review is a search and evaluation of the available literature in your given subject or chosen topic area. A literature review includes about the different areas from where the system information is gathered.

“System Analysis” is included in chapter three. It included many information about the system. It includes the requirement analysis, feasibility study. Requirements analysis focuses on the tasks that determine the needs or conditions to meet the new or altered product or project. Requirements are of two types Functional and non-functional requirement. Functional requirements define the specific functions that the system performs, along with the data operated on by the functions. Whereas non-functional requirement deals with the quality and performance of the system. It focuses on security, advancements and performance mechanism of the system. Similarly, feasibility study is an analysis that takes all of a project's relevant factors into account—including economic, technical, legal, and scheduling.

Nest chapter is “system design” that in the detail explanation about the system design and its architecture. System design include high-level design and the high level includes System flow Chart, Architecture design, Sequence Diagram, Class Diagram and Activity diagram. System flow chart is the symbolic representation of system that describe the work flow of the system between different user and system. Architecture design is a concept that focuses on components or elements of a structure. Similarly, Sequence Diagrams are interaction diagrams that detail how operations are carried out. They capture the interaction between objects in the context of a collaboration. Activity diagram is another important behavioral diagram in UML diagram to describe dynamic aspects of the system. Activity diagram is essentially an advanced version of flow chart that modeling the flow from one activity to

another activity. And the last one is class diagram, a class diagram in the Unified Modeling Language (UML) is a type of static structure diagram that describes the structure of a system by showing the system's classes, their attributes, operations (or methods), and the relationships among objects.

“Implementation and Testing” describe about different modules used in the system. It also describes about the implementation of the algorithm in the system. Which algorithm is used and what modules are used in the algorithm are used in system is describe in this phase. It describes about collaborative filtering. Similarly, it also includes different tools used in the system development. And also, system testing is done in this phase. The last chapter included in the report is “Conclusion and Recommendation”. Conclusion is the last part of the report, its result. And the recommendation enhanced the future improvement that can be done with the system and to overcome the limitation of the system.

To sum up the report contain different parts which has its own importance and every part describe about different phase in system development.

CHAPTER 2- LITERATURE REVIEW

Sushama R, Darshana B and Pooja Malhotra [1], proposed a system “Book Recommendation System” which presents a new approach of recommending books to the buyers. The system had a combined feature of content filtering, collaborative filtering and association rule mining to produce efficient and effective recommendations. The system has considered many parameters like content of the book and quality of the book by doing collaborative filtering of ratings by the other buyers. This recommender system also used associative model to give stronger recommendations.

Prem M and Raymond J. Mooney and Ramadass N [2], make a research to improve recommendation system. They studied about content-based recommendation and collaborative filtering. They found that Content-based methods can uniquely characterize each user, but CF still had some key advantages over them.

Basu C, Hirsh H, Cohen W [3] discussed about content-based recommendation systems, i.e., systems that recommend an item to a user based upon a description of the item and a profile of the user’s interests. They found that the Content-based recommendation systems may be used in a variety of domains ranging from recommending web pages, news articles, restaurants, television programs, and items for sale. Although the details of various systems differ, content-based recommendation systems share in common a means for describing the items that may be recommended, a means for creating a profile of the user that describes the types of items the user likes, and a means of comparing items to the user profile to determine what to recommend. The profile is often created and updated automatically in response to feedback on the desirability of items that have been presented to the user.

C. Pan, W. Li [4] designed Research Paper Recommendation System with topic analysis in Computer Design and Applications. A research paper recommender system greatly helped researchers to find the most desirable papers in their fields of endeavor. Due to the textual nature of papers, content information was integrated into existed recommendation methods. the alphabet by comparing between two images. The processes are starting from the acquisition process, filtering process, threshold the image, clustering the image of alphabet and lastly recognize the alphabet. This system prototype could give several advantages to the users, but this system prototype is still

facing a number of limitations. So that, further research could be done by other person to improve the system prototype into a better system.

Zhao, Z. and Shang, M. [5] proposed a user-based collaborative filtering algorithm on Hadoop which involved the implementation of user-based collaborative filtering algorithms on a cloud-based platform (case study of Hadoop). The work aimed at solving the scalability problems of collaborative filtering and was able to establish the fact that a simple method that partitions users into groups according to two basic principles namely the tidy arrangement of mapper number to overcome the initiation of mapper and the partitioning of tasks such that all processors finish task at the same time can actually result in a linear speedup of performance.

Cheng, Z., and Hurley, N. [6] proposed Region KNN: a scalable hybrid collaborative filtering algorithm for personalized web service recommendations. The above study considered the difference between web service recommendations and product recommendations used in e-commerce websites. Region KNN was designed and presented for large scale web services recommendation. Region KNN used an approach that employs the characteristics of QoS by building efficient region model that quickly generates web service recommendations using modified memory-based collaborative filtering algorithm. From the result of the study, experiments show the high scalability of the Region KNN.

Linden, G., Smith, B., and York, J. [7] proposed Amazon.com recommendations: an item-to-item collaborative filtering. This approach solved recommendation problems with the use of an algorithm different from the traditional collaborative filtering algorithm, cluster models algorithm and search-based algorithm methods. The item-to-item collaborative filtering algorithm applied to Amazon's online shop computations was able to scale independently of the number of customer's and items in the product catalog.

Hence, from the above reviews we collect some ideas to design our system and it is going to work. As these reviews have their own explanation on algorithm for designing a system and its component like image pre-processing, detecting number plate area from image captured, character segmentation and character recognition, we will be using template matching algorithm for character recognition.

CHAPTER 3- SYSTEM REQUIREMEN ANALYSIS

Requirements analysis in systems development, encompasses those tasks that go into determining the needs or conditions to meet for a new or altered product, taking account of the possibly conflicting requirements of the various entity i.e. system, users. Requirements Analysis is the process of defining the expectations of the users for an application that is to be built or modified. Requirements must be actionable, measurable, testable, related to identified business needs or opportunities, and defined to a level of detail sufficient for system design. Requirements can be functional

3.1. Requirement Analysis

Requirement analysis holds the process of reviewing and determining the system needs, functional requirements and non-functional requirements that a system must meet. For requirement analysis, following approaches are followed:

3.1.1 Preliminary Analysis

- Problem analysis was done and titled as problem statement.
- Performed literature review to know about algorithm used in existing system.
- Analysis of system planning and design.

3.1.2. Functional Requirement

Functional requirement identifies the provision of the system and the system's reaction to certain input and how the system should behave in day to day basis. Functional requirements define the specific functions that the system performs, along with the data operated on by the functions. The functional requirements are presented in scenarios that depict an operational system from the perspective of its end users.

The functional requirement for the 'Book Recommendation System' includes following tasks:

- The system should incorporate mechanism to authenticate its users.
- The system should verify and validate all user input and should notify in case of error detection and should help the user in error correction.
- The system should recommend the book to users based on user's preferences.
- System should not take long time to recommend.

Use Case Diagram

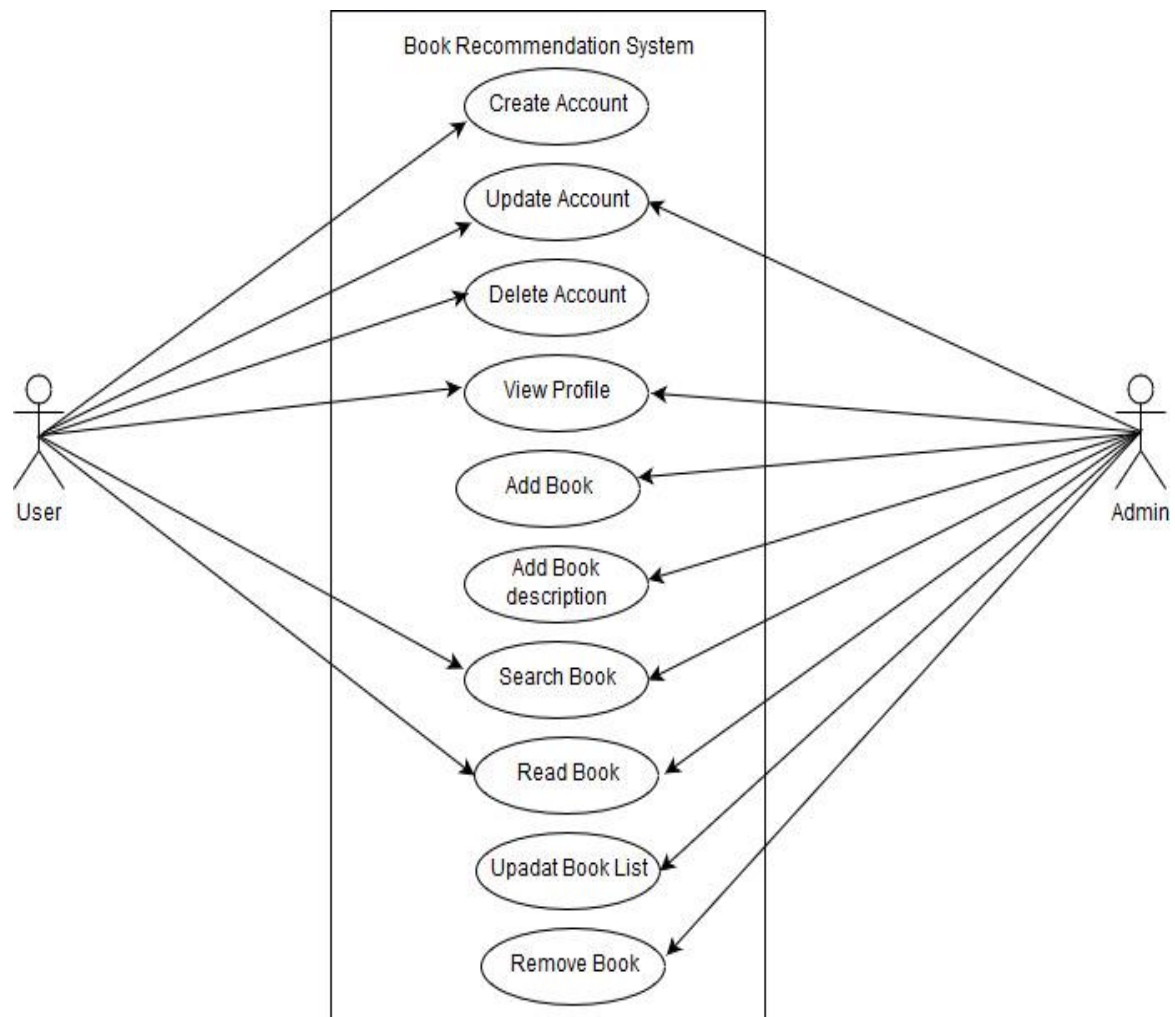


Figure 3.1: Use Case diagram of Book Recommendation System

Use-case Description

A use case diagram is a graphic depiction of the interactions among the elements of a system. A use case is a methodology used in system analysis to identify, clarify, and organize system requirements. The Use Case diagram of BRS is given above. The Use Case diagram of the system demonstrates the different ways that the user and admin might interact with the system. The main components of the Use case diagram are actors, system and goal. Here are two actors 'user' and 'admin'. The system specifies a specific sequence of the action and interaction between the actor and the system. Actors perform different task which is represented in ovals shape and known as use cases. A line is drawn between the actors and use cases which represents the association between them.

For example, in the above Use Case diagram of the system user perform different actions like creating account, updating account, searching book and many more. Given table describe the activities of user and admin of above use-case diagram.

Table 3.1: Use Case Diagram Table for User

Action	Job description
Create Account	User register their account
Update Account	User change their information like name, password
Delete Account	User delete their account permanently
View Profile	User can view their profile, account detail
Search Book	User search the book they want to read
Read book	User can read the book according to their interest

Table 3.2: Use Case Diagram Table for Admin

Action	Job Description
Update Account	Admin can update their account
View Profile	Can view users and own profile
Add Book	Admin can add the book in system database
Add Book Description	Admin can add book description to the books
Read Book	Read the book available in the system
Update Book list	Admin update the book list
Search Book	Admin can search book

3.1.3. Non-functional Requirement

Non-functional requirement deals with the quality and performance of the system. It focuses on security, advancements and performance mechanism of the system. The non-functional requirements of the 'Book recommendation System' includes following points:

- There must be high accuracy while recommending.
- The system should be usable, reliable and effective to the user.
- The system should only perform in terms of valid input provided by the user.
- The system should provide security in data stored in database.
- The system should provide graphical user interface to user.

3.1.4. Software Requirement

- MySQL database
- OpenCV Library
- Microsoft Visual Studio 2019 professional
- Jupyter Notebook
- JetBrains PyCharm 2018 community edition

3.2. Feasibility Analysis

3.2.1. Technical Feasibility

This is concerned with the specifying equipment and software that successfully satisfy the requirements. The system is technically feasible as it can be developed easily with the help of available technology. The system requires HTML (Jinja template), CSS, Bootstrap, JavaScript which is used as front-end, Django framework as back-end and python as both

3.2.2. Operational Feasibility

The system is highly user friendly and it is much easier to interact with the user. User does not need special training to operate the system. Therefore, the system will provide maximum easiness. The system will have simple user interface, so that any non-technical user can operate with the system.

3.2.3. Economic Feasibility

Economic analysis is the most frequently used technique for evaluating the effectiveness of the system. The tangible benefits proposed that the manual work and burden is reduced maximum as possible, resulting the reduction in manpower requirement and cost incurred on manpower as well. The system provides many benefits that can't be measured in terms of money for e.g. user friendliness, more efficient user response, maintenance of database, etc.

3.3. Data Modeling

Data modeling is the process of creating a data model for the data to be stored in a Database. Data model emphasizes on what data is needed and how it should be organized instead of what operations need to be performed on the data. Data Model is like architect's building plan which helps to build a conceptual model and set the relationship between data items. This data model is a conceptual representation of

- Data objects
- The associations between different data objects
- The rules

ER (Entity Relation) diagram is used as data modeling technique. The ER- diagram of the BRS is given below:

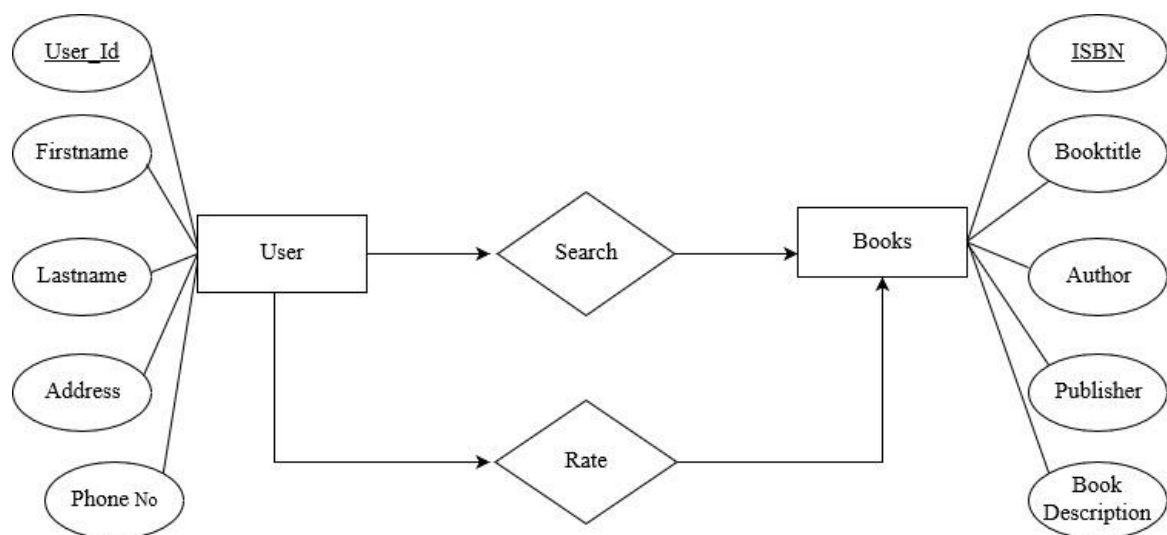


Fig 3.2: ER-diagram of Book Recommendation System

ER- diagram is a high-level conceptual data model diagram. Entity-Relation model is based on the notion of real-world entities and the relationship between them. The ER- diagram of BRS is given which show the relationship among the entities. Entities with their attributes in the BRS are

- Users (User_id, Firstname, Lastname, Address, Phone no)
- Book (ISBN, Booktitle, Author, Publisher, Book Description)

ER- diagram shows the relation between these entities using the relation. Entities are denoted by the rectangle and in this system, entities are “users” and “book”. Every entity

has its own attributes. Attributes are the properties of the entity. Users entity have attributes like User_id, firstname, Lastname, Address and Phone no. The primary key attribute is underlined. Similarly, Book also have attributes like ISBN, Booktitle, Author, Publisher, Book description. The relation is denoted by diamond symbol and it is used to show the relationship between entities. In BRS users can search and rate the book. Here, search and rate are relation.

So, above ER diagram of the Book Recommendation System shows the relation among the entities with different attributes.

3.4 Process Modeling

Process models are processes of the same nature that are classified together into a model. Thus, a process model is a description of a process at the type level.

3.4.1 Context Diagram

A context diagram (CD) is a diagram that defines the boundary between the system, or part of a system, and its environment, showing the entities that interact with it. The Context Diagram shows the system under consideration as a single high-level process and then shows the relationship that the system has with other external entities Context diagram is also known as Level-0 Data Flow Diagram. The Context diagram of BRS is as follow:

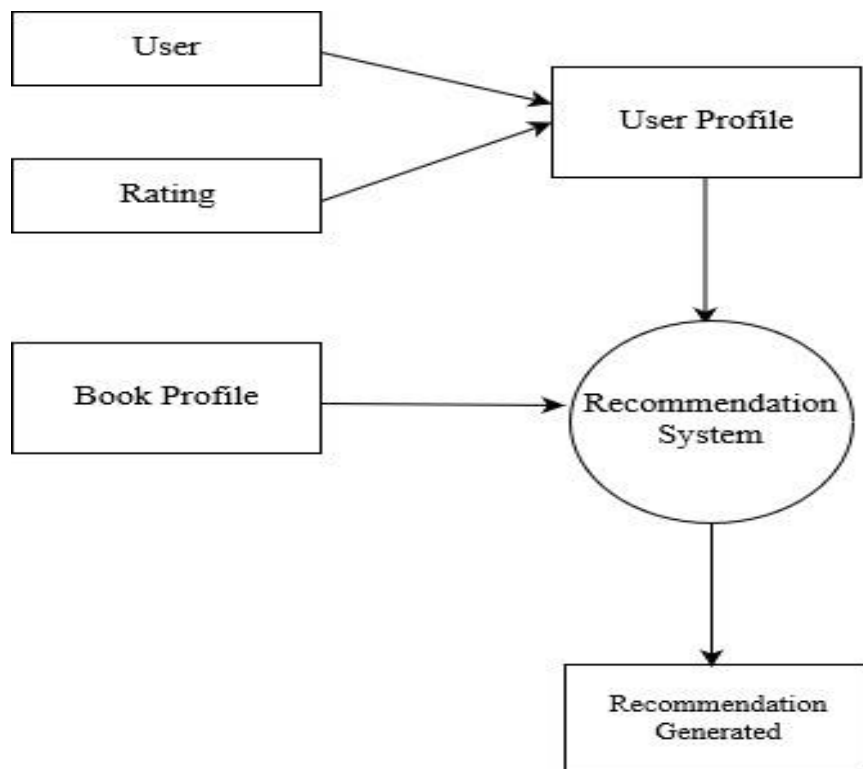


Fig 3.3: Context Diagram of Book Recommendation System

A Context Diagram (CD) is a graphical visualization of the movement of data through an information system. The context diagram of the system is given above. It shows the data flow among different entities. In above diagram there are entities like user profile, book profile. User profile contain the information about user and the rating provided by user to the books. The rating is provided to the system which will generate the recommendation and provide to user. Book profile include the information about books.

3.4.2 Level -1 DFD

A data flow diagram (DFD) maps out the flow of information for any process or system. It uses defined symbols like rectangles, circles and arrows, plus short text labels, to show data inputs, outputs, storage points and the routes between each destination. Level 1 diagram shows all the processes that comprise a single process on the level 0 diagram. It also shows how information moves from and to each of these processes and more detail the content of higher-level process. DFD level-1 of BRS is as follow

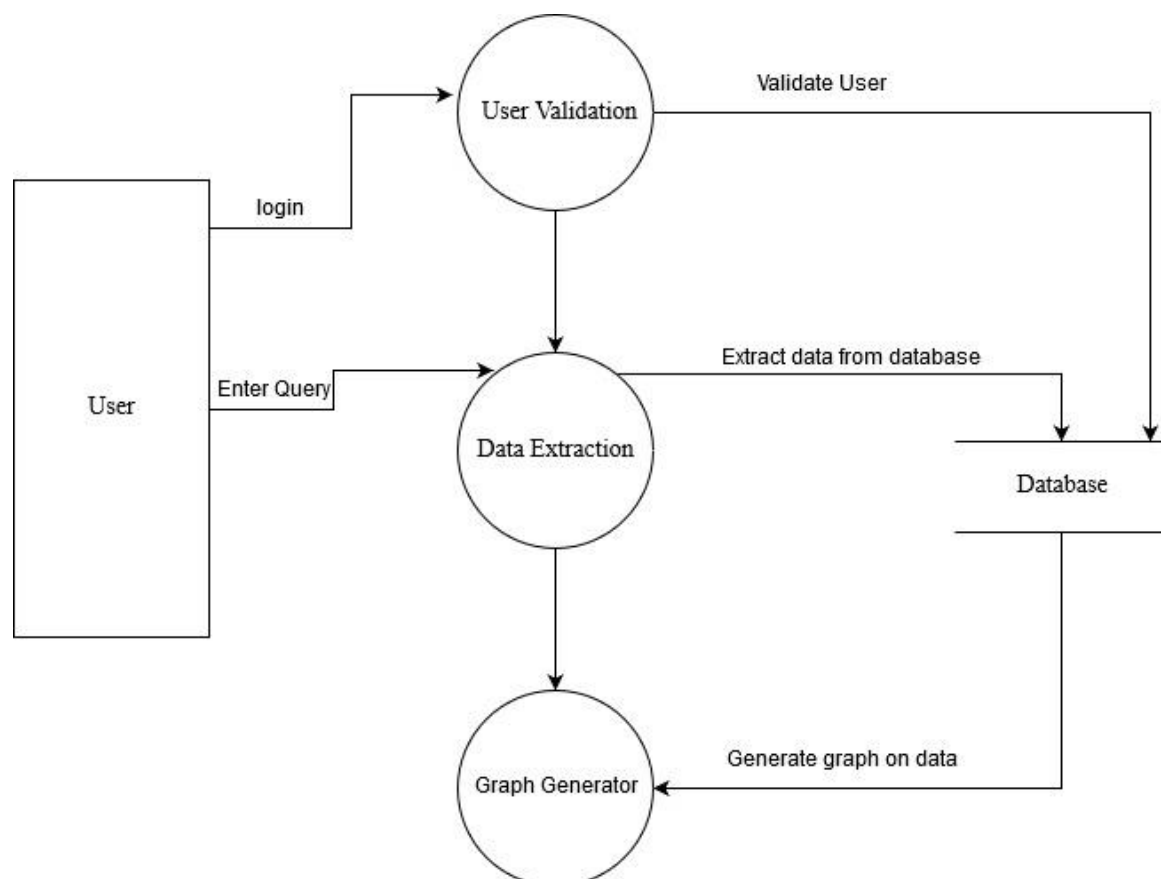


Fig 3.4: DFD level -1 of Book Recommendation System

A Data Flow Diagram (DFD) is a traditional way to visualize the information flows within a system. DFD level -1 of the Book Recommendation System is given above which uses different symbols to describe the flow of the information among the system or process. At first the User will login with his/her account, then the validation process is done to validate the user and that is represented by the circle symbol which means process. After the user is validated then the user is allowed to enter the query. After the query, another process is carried out and that is Data Extraction. The system extracts the data according to the query given by the user. The data is extracted from the database and provide it to the user. Similarly, the Graph Generator process represented by circle symbol generate the graph on extracted data. And after this the process ends.

In this way, DFD level-1 diagram shows all the processes that comprise a single process on the level 0 diagram and also show how the information moves from and to each of the processes.

CHAPTER 4- SYSTEM DESIGN

4.1. System Flowchart

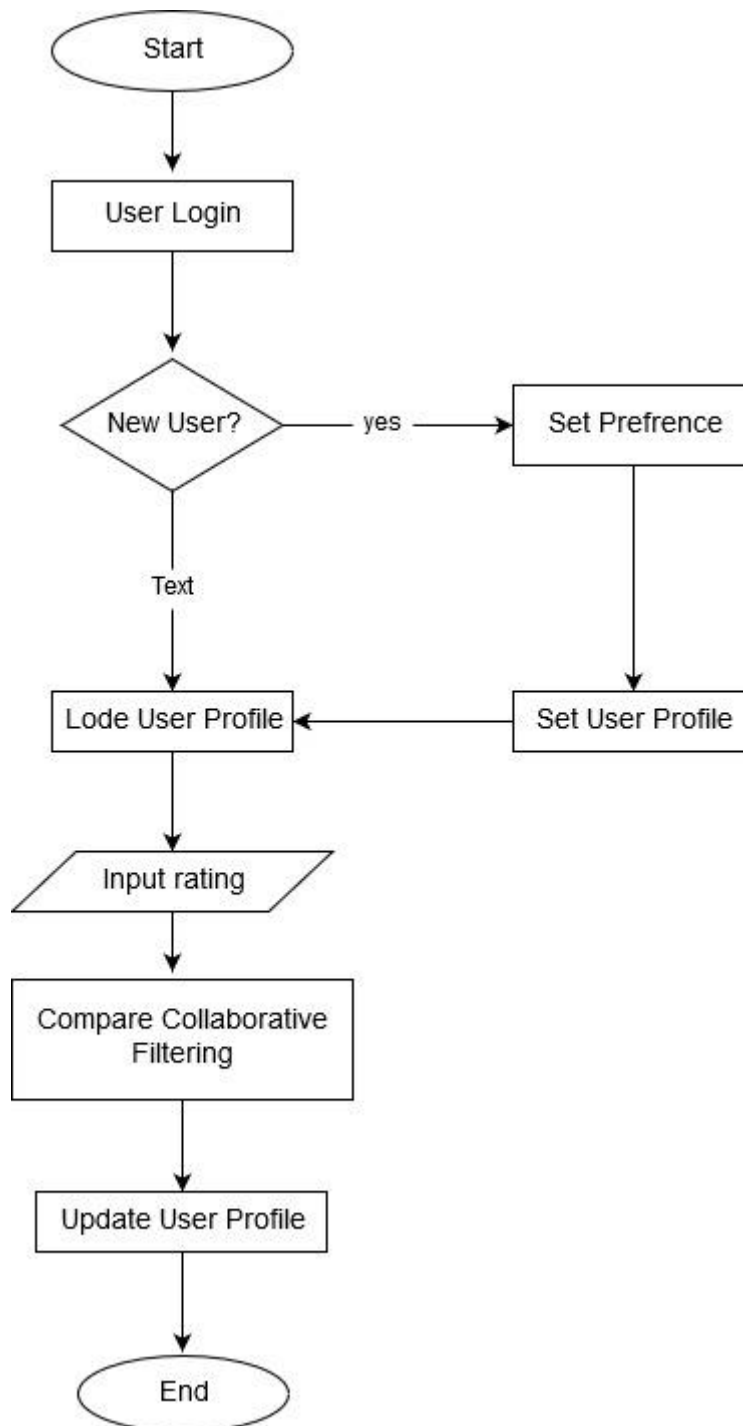


Fig 4.1 : Flowchart of Book Recommendation System

Flow chart of the Book Recommendation System is given above. Flow chart is the diagrammatic representation of the algorithm or the workflow of the system. The flowchart of the Book Recommendation system is given above. The flowchart represents the user's activity on the system. First of all, the user will be logged in to the system. The system will check whether the user is new or not. If the user is new, he/she will be asked to create new account else the user's profile will be loaded in the screen. User will rate the books and then application will use collaborative filtering algorithm based on the rating given by user and then recommend other similar books to user and update user's profile. In this way the flow chart defines the working steps in recommendation process.

4.2. System Architecture

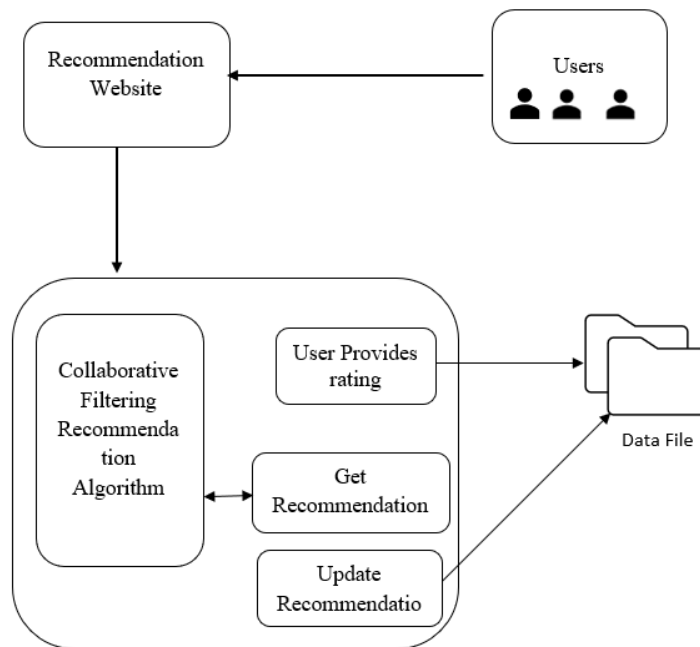


Fig 4.2 : Architecture design of Book Recommendation System

Architectural design is a concept that focuses on components or elements of a structure. The figure above illustrates the architecture of Book Recommendation System. The process includes in recommending a book to the user is defined by the picture. At first the user searches the book or enter his/her choices. The input data given by the user are known as keyword and that is received by the main system. Those keywords are matched with the keyword mapping database which stores all the data. After mapping the keywords, they are sent to book recommendation system which analysis the keyword as well as the past history of the user and then recommend the book according to the algorithm. The information is then sent to the main system which will provide it to the user.

4.3 Sequence Diagram

Sequence Diagrams are interaction diagrams that detail how operations are carried out. They capture the interaction between objects in the context of a collaboration. Sequence Diagrams are time focus and they show the order of the interaction visually by using the vertical axis of the diagram to represent time what messages are sent and when.

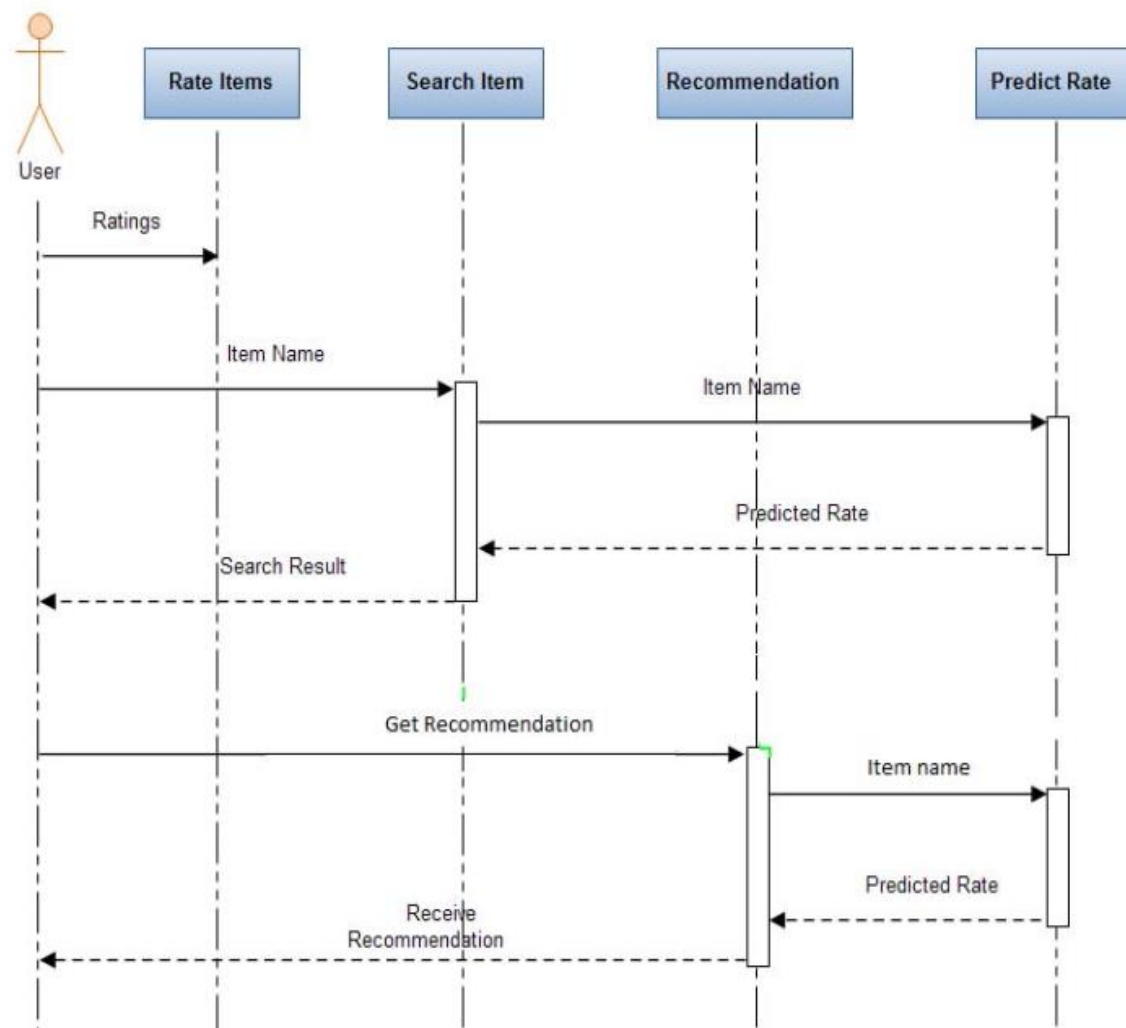


Fig 4.3 : Sequence Diagram of Book Recommendation System

The sequence diagram of the BRS is given above. It shows the detail operation carried out in the system. It sequentially shows the steps in recommending the book to the user. At first user provide the username and password and the system validate the user according to the data stored in the database. Then the user enters the query and the system request the data from the database. System search the query in the database and then it analysis the ratting given by user. Then it uses collaboritive filtering algorithm to recommend book to user according to query.

4.4 Class Diagram

A class diagram in the Unified Modeling Language (UML) is a type of static structure diagram that describes the structure of a system by showing the system's classes, their attributes, operations (or methods), and the relationships among objects. The class diagram of BRS is as follow which include five classes.

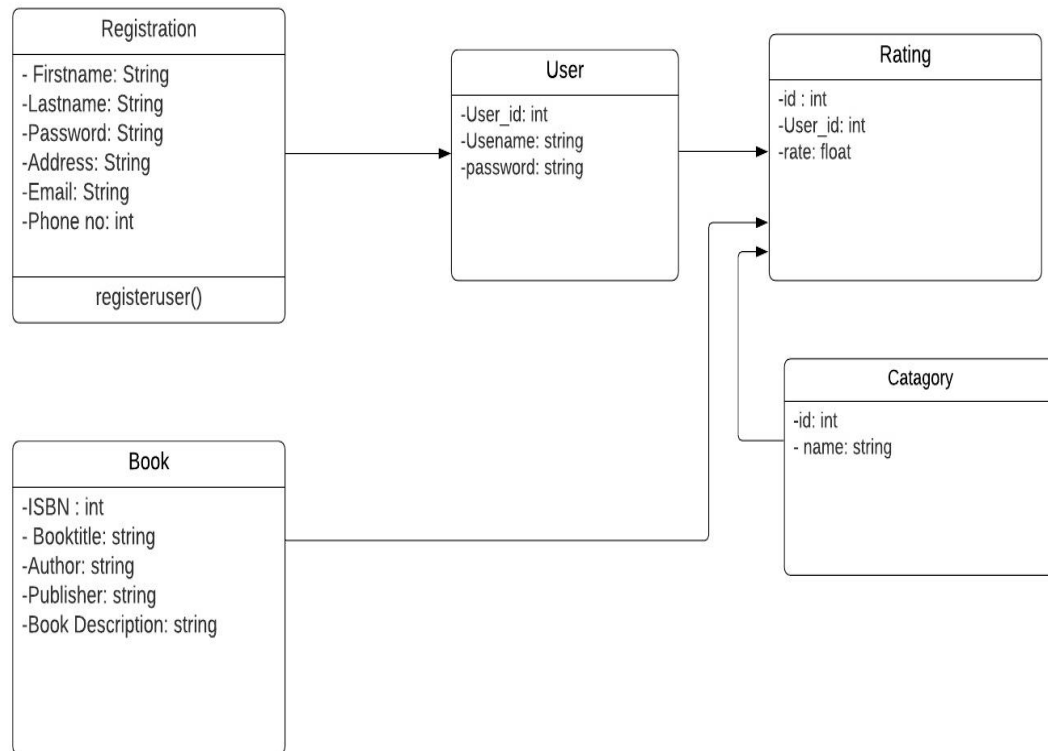


Fig 4.4 : Class Diagram of Book Recommendation System

There are five class in the class diagram of BRS. They are, Book, registration, user, rating and category.

Registration class include all the list that the user field to register their account. It includes firstname, lastname, password, address, email and phone no. which user need to enter while createing their account.

User class include user_id, username and password. When a user tries to log into the application, it is checked if the username and password match the value in the database and if the value matches then the user is allowed to log into the system.

Book class consists the list of ISBN, booktitle, Author, Publisher, Book description. Rating class include id, user_id and rate. And the category class include id and name.

4.5 Activity Diagram

Activity Diagrams describe how activities are coordinated to provide a service which can be at different levels of abstraction. Typically, an event needs to be achieved by some operations, particularly where the operation is intended to achieve a number of different things that require coordination, or how the events in a single use case relate to one another, in particular, use cases where activities may overlap and require coordination. The activity diagram of the BRS

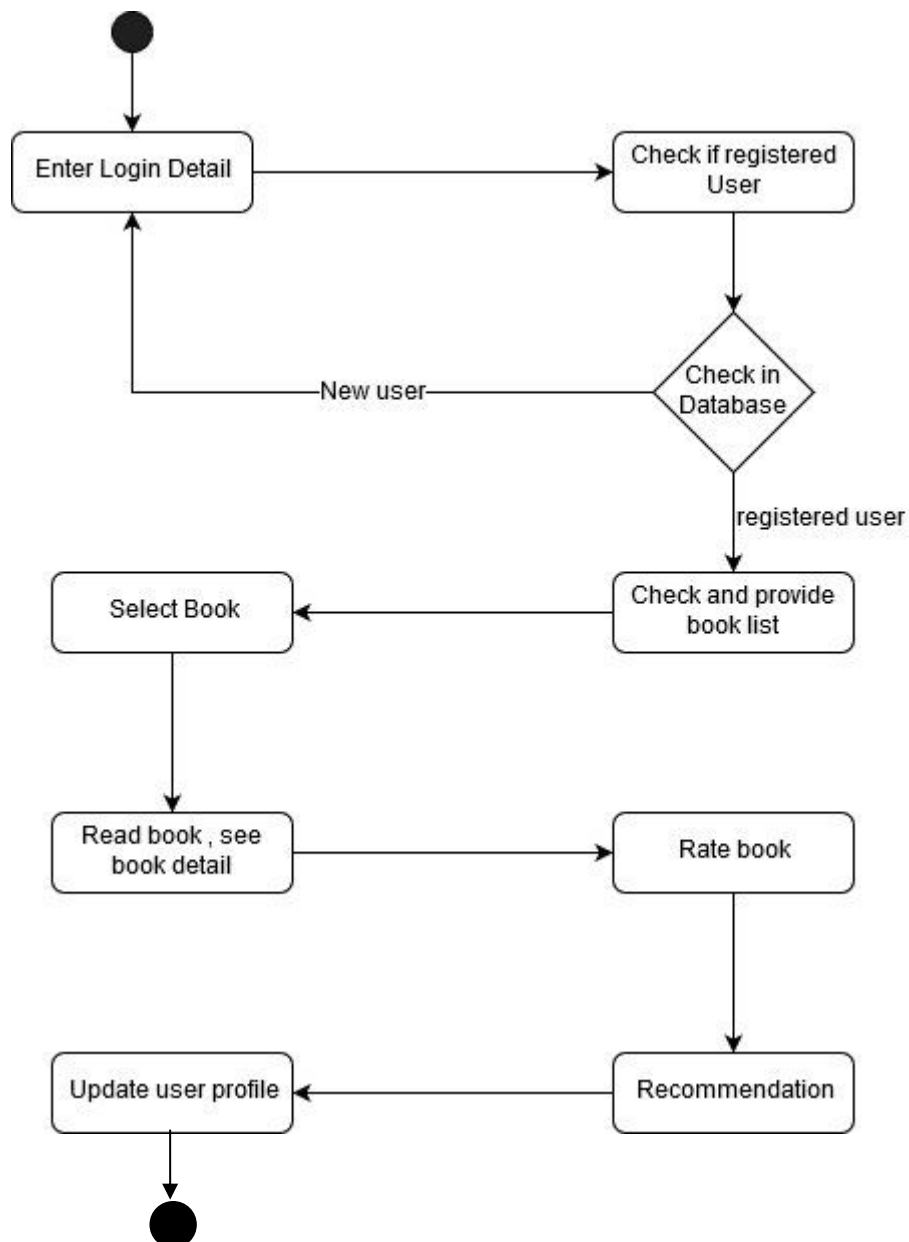


Fig 4.5 : Activity diagram of BRS

The Figure 9 explains activity of the application. At first, when the user opens the application for the first time, they need to register to the application. Then they are required to log into the application with their valid user credentials. The user can now rate different books. The ratings of the users are saved into the database. The user can also choose to get recommendations. When the user searches for the recommendation, the similarity is calculated and the results are displayed on the screen.

CHAPTER 5 – IMPLEMENTATION AND TESTING

5.1 Module Description

Module is a logical separation of a functionality within a project. Different modules are used in the system. BRS includes following modules:

1. User management
2. Rating
3. Search Module
4. Book etc.

Different modules are used in the system like user management, ratings, search module etc. These modules have their own functionality.

User management Module is used to manage the user in the system. User may be new or old and they have their respective accounts. The user management module is used to manage the user their account and to keep the record of the user and which can be further used for recommendation process. User management helps the system to recommend the book to the user according to their ratings or records. Similarly, the rating module is also one of the modules used in the system. Rating module helps in rating process. Which user have rate the book and which type of book is being rated and many more information are manage by the rating module. It takes the rating from the user and these rating are used for recommending the similar book to the user that they may like.

Similarly, search module is another module used in the BRS. The search information is obtained by the search module. It keeps the record of the book that the user search for and also keeps the record of the user who search for the book. And this process is also used for making recommendation of the book to the user. The book module keeps the record of all type of book in the system. The absence and presence of the book, types of book and many information are recorded by the book module.

So different modules are used by the BRS which help in making recommendation of the book to the user.

5.2 Algorithm Implementation

The algorithm used in this system is Matrix Factorization Algorithm. It is a class of collaborative filtering algorithms used in recommender systems. Matrix factorization algorithms work by decomposing the user-item interaction matrix into the product of two lower dimensionality rectangular matrices. In the system it is used for recommending the books to the customers. It recommends the product that are highest rated by the customers. When the customers want to get the recommended books, this algorithm recommends the books that are highest rated among all the available books in the system. This helps to save time of customers as only the best books are recommended. The idea behind matrix factorization is to represent users and items in a lower dimensional latent space.

Technique:

Singular Value Decomposition (SVD++) It is model-based approach that factorize the user-item rating matrix as the product of two lower dimensional matrices, the first one has a row for each user, while the second has a column for each item. The row or column associated to a specific user or item is referred to as latent factors. It compresses user-item matrix into a low-dimensional representation in terms of latent factors. One advantage of using this approach is that instead of having a high dimensional matrix containing abundant number of missing values we will be dealing with a much smaller matrix in lower-dimensional space.

Algorithm:

Inputs: Total users and Total ratings are taken in data frame along with user and product id.

Output: Sorted rating indexes and list of recommended products.

Process:

Myrecommend()

Step 1: Normalizing Y and R where Y=total rating and R=Total users stored in database.

Step 2: An array X and Theta is created with specified shape and filling with random values.

Step 3: X and Theta are converted into 1D array.

Step 4: Cost function and Gradient function is calculated

Step 4: Result array is obtained by optimizing coficostFunc, Myflat, cofiGrad and arguments. (Y,R,MyNu,MyNf,MyLambda) with specified iteration.

Step 5: The obtained result is reshaped into 1D array resX, resTheta

Step 6: Prediction_matrix is returned with dot product of resX and resTheta. Ymean and prediction_matrix is returned from Myrecommend().

Step 7: Calculate my_prediction array by using flatten() function on Ymean

Step 8: Sorting my_prediction array value using np.argsort to form pred_indx_sorted

Step 9: pred_indx_sorted is the sorted 1D array of ratings indexes which is then evaluated by slicing each element with slicing function from bottom to top.

Step 10 : Condition like “case” and “when” is used to check if the correct product is returned Step 11: Product_list is displayed if the product is present in Product_list and satisfied by the condition of Step 10.

Pros:

- **No domain knowledge necessary :-** We don't need domain knowledge because the embeddings are automatically learned.
- **Serendipity :-** The model can help users discover new interests. In isolation, the ML system may not know the user is interested in a given item, but the model might still recommend it because similar users are interested in that item.
- **Great starting point :-** To some extent, the system needs only the feedback matrix to train a matrix factorization model. In particular, the system doesn't need contextual features

Cons:

- **Cannot handle fresh items :-** The prediction of the model for a given (user, item) pair is the dot product of the corresponding embeddings. So, if an item is not seen during training, the system can't create an embedding for it and can't query the model with this item. This issue is often called the **cold-start problem**.
- **Hard to include side features for query/item :- Side features** are any features beyond the query or item ID. For movie recommendations, the side features might include country or age. Including available side features improves the quality of the model.
- **New Item Problem :-** New item problem arises mainly due to the fact that collaborative systems relies on similar user rating on one item for prediction. If one item is not rated by enough users, collaborative system results can be very much biased.

5.3 Tools used

Different tools are used for developing the system. Following are the tools used in developing BRS.

1. Front End Tools

The front end of a BRS is the part that users interact with. Everything that the user sees when user is navigating around the Internet, from fonts and colors to dropdown menus and sliders, images text etc. Our system used many front-end tools for making the system more interactive to user. BRS use following tools as front-end tools

- **Django**

Django was used to design the UI for our system. Since our system is designed using python. So, it was decided to use a python framework itself to design the UI too. Django helps users when they need to move fast, deploy fast, and also make changes as you move ahead. Django is an open source framework and easy to use too.

- **HTML (hypertext markup language)**

Hypertext Markup Language, a standardized system for tagging text files to achieve font, color, graphic, and hyperlink effects on World Wide Web pages. It includes tags. An HTML tag is commonly defined as a set of characters constituting a formatted command for a Web page. At the core of HTML, tags provide the directions or recipes for the visual content that one sees on the Web.

- **CSS (Cascading Style Sheets)**

CSS is a presentation language created to style the appearance of content. CSS provides style to the document like text, table, navigation bar header etc.

- **JavaScript**

JavaScript (JS) is a scripting language, primarily used on the Web. It is used to enhance HTML pages and is commonly found embedded in HTML code. JavaScript is an interpreted language. JavaScript renders web pages in an interactive and dynamic fashion.

- **Bootstrap**

Bootstrap is a free and open-source CSS framework directed at responsive, mobile-first front-end web development. Bootstrap is the most popular HTML, CSS, and JS framework for developing responsive, mobile first projects on the web.

2. Backend Tool

The back end of a BRS consists of a server, an application, and a database. A back-end developer builds and maintains the technology that powers those components which, together, enable the user-facing side of the website to even exist in the first place. Our system uses different back- end tools.

- **Python**

Python language was used to design Neural Network since it is efficient in field of Machine learning and Artificial Intelligence. Python is a great object-oriented, interpreted, and interactive programming language. Python combines remarkable power with very clear syntax. It has modules, classes, exceptions, very high-level dynamic data types, and dynamic typing. There are interfaces to many system calls and libraries, as well as to various windowing systems. Python is also usable as an extension language for applications written in other languages that need easy-to-use scripting or automation interfaces.

- **Php**

PHP (recursive acronym for *PHP: Hypertext Preprocessor*) is a widely-used open source general-purpose scripting language that is especially suited for web development and can be embedded into HTML. PHP code may be embedded into HTML code, or it can be used in combination with various web template systems, web content management system and web frameworks.

3. Database

A database is a collection of information that is organized so that it can be easily accessed, managed and updated. Computer databases typically contain aggregations of data records or files. Data within the most common types of databases in operation today is typically modeled in rows and columns in a series of tables to make processing and data querying efficient. The data can then be easily accessed, managed, modified, updated, controlled, and organized. Our system use following database:

- **MySQL**

MySQL is an Oracle-backed open source relational database management system (RDBMS) based on Structured Query Language (SQL). MySQL is an important component of an open source enterprise stack. MySQL server is available as a separate

program for use in a client-server networked environment and as a library that can be embedded (or linked) into separate applications.

4. Documentation Tool

Documentation has always been a key to success. Understanding the importance of documentation will take project a step forward than other. There are different formats and editors that can be used to write a well-structured documentation. Different documentation tools are used in project development.

- **Microsoft Word**

Microsoft Word is a word processor developed by Microsoft. Microsoft word is used to prepare project report in our project. MS -word provide many features in making report which makes more easier in documentation.

- **Microsoft PowerPoint**

Microsoft PowerPoint is the presentation software developed by Microsoft. PowerPoint is used to make presentation of the project which include many slides. Every slide contains the information about the project.

5.4 System Testing

System Testing is a level of application testing where a complete and integrated application is tested. The purpose of this test is to evaluate the system's compliance with the specified requirements. Different steps are conducted to test the application.

5.4.1 Unit Testing

Unit testing is the level of software testing of the project's system in which the smallest testable parts of a system called unit is individually tested. For unit testing each and every functional module is tested and the result is analyzed to ensure that the information flows in and out of the program. Each unit of the software is tested to verify that the detailed design for the unit has been correctly implemented.

Table 5.1: Unit Test case for Successful login

Test Case	Test case for successful login
Objective	To check the user is valid
Description	Check the user name and password entered by the user.
Expected Result	To be directed into the application
Actual Result	User valid and enter into the system

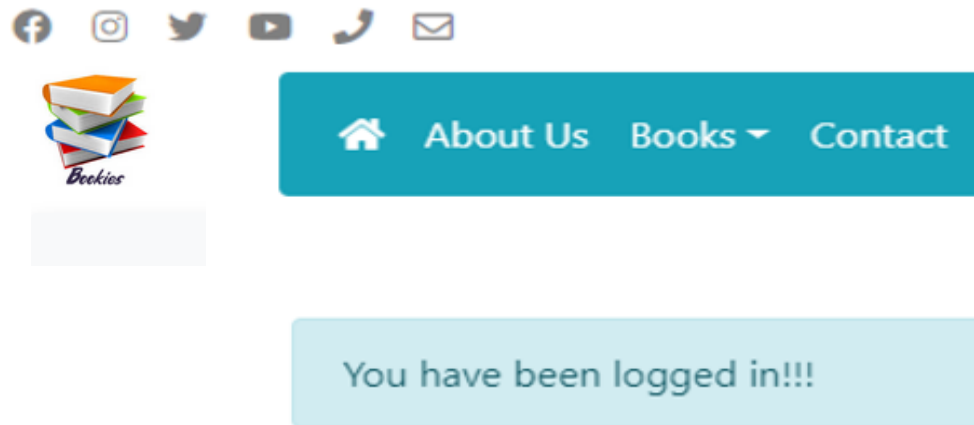


Fig 5.1: Unit Test case for Successful Login

Table 5.2: Unit Test case for unsuccessful login

Test Case	Test case for invalid login
Objective	To check the user is valid or not
Description	Check the user name and password entered by the user.
Expected Result	To stop user to directed into the application
Actual Result	User invalid and re-enter username and password

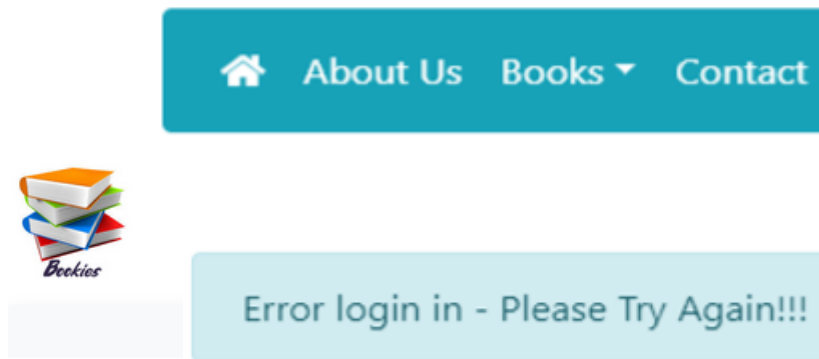


Fig 5.2: Unit Test Case for Unsuccessful Login

Table 5.3: Unit Test case for valid rating

Test Case	Test case for valid rating
Objective	To check the rating provided by user
Description	User rate the books according to their choice
Expected Result	To use the rating given by user for recommending process
Actual Result	To take rating as the input through the user for recommending process



Your Rating is submitted



Fig 5.3: Unit Test Case for valid rating

Table 5.4: Unit Test Case for invalid rating

Test Case	Test case for invalid rating
Objective	To check rating provided by user
Description	User provide rating
Expected Result	To provide the error message on screen
Actual Result	To take rating checked it and send invalid message to user

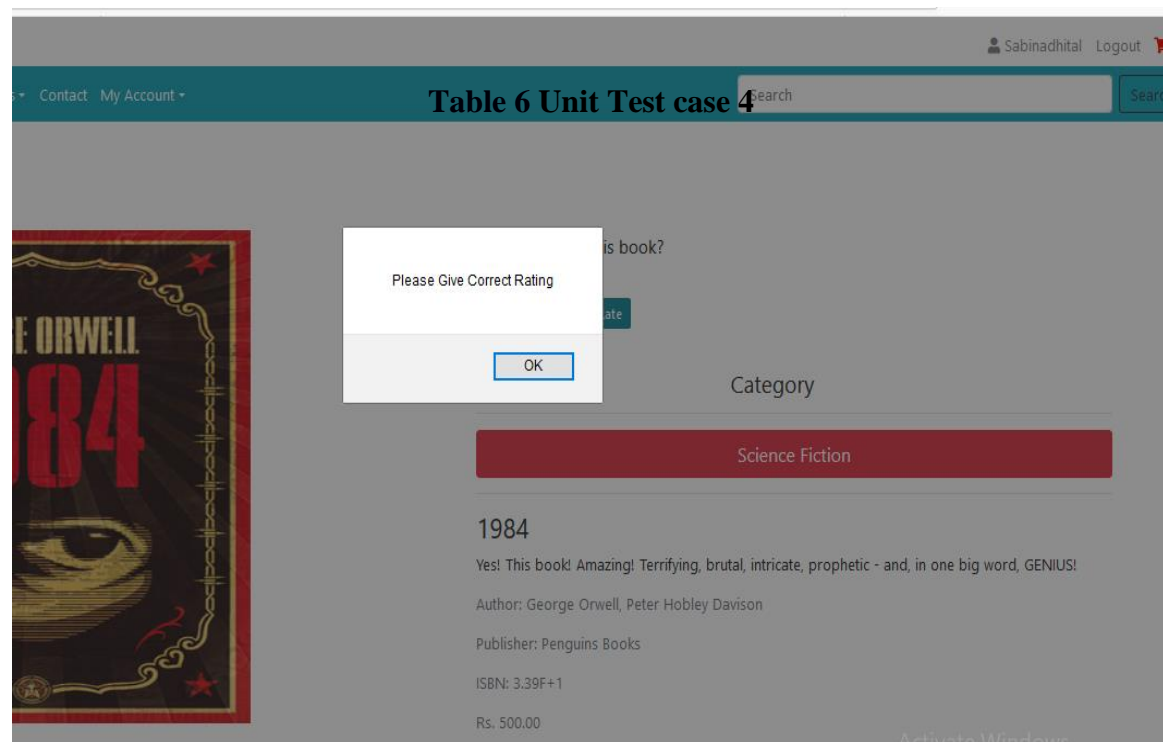


Fig 5.4: Unit Test Case for invalid rating

Table 5.5: Unit test case for Searching

Test Case	Test case for searching book
Objective	To search book according to user choice
Description	The system takes search information and provide search result
Expected Result	To provide search result to user
Actual Result	To take searching information as the input through the user and provide result to that search

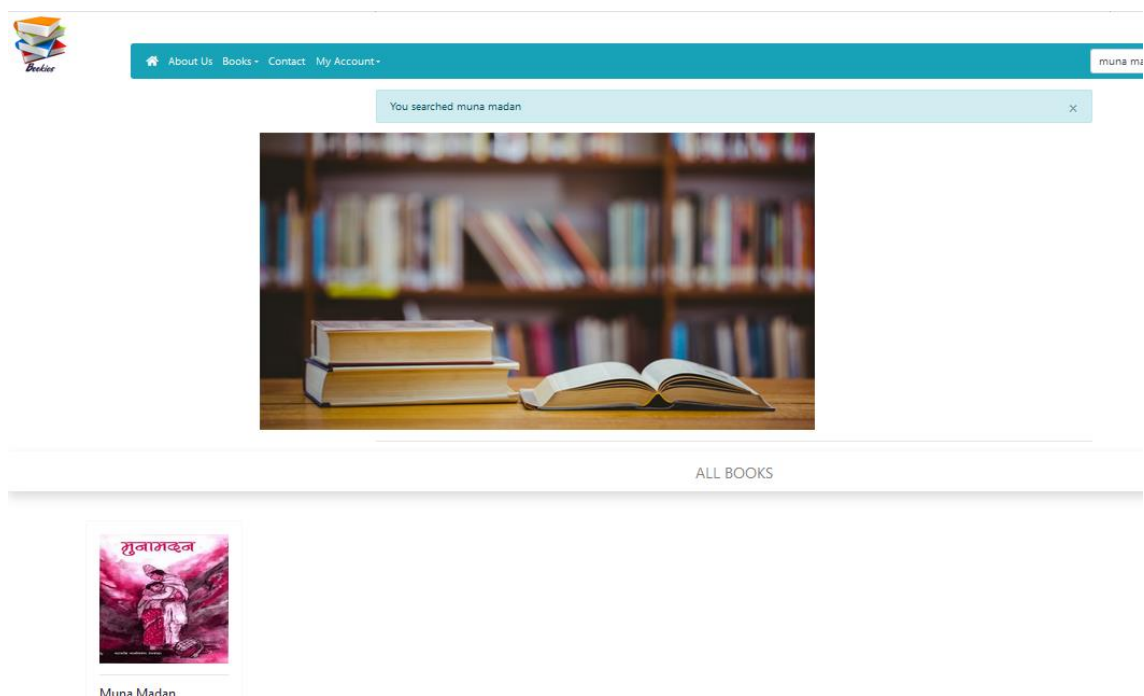


Fig 5.5: Unit Test Case for searching

5.4.2 Integration testing

Table 5.6: Integration testing

Test case	Objective	Description	Expected Result	Actual Result	Remarks
1	To recommend books to user according to the user's rating	To check user, obtain rating from user and used them for recommending.	Book recommend to user according to their interest	Whenever user login the book is recommend to user according to rating	This will help in searching book the user like

5.4.3 System Testing

Test case 1

Objective: To execute the overall system properly.

Descriptive: The all the modules should work properly in order to run the system where the user and the rating provided by the user is checked properly and recommend the book according to that rating.

Expected Result: To recommend the book according to user interest.

Actual Result: The system shows books on the basis of rating provided by user.

CHAPTER 6 - CONCLUSION AND RECOMMENDATION

6.1 Conclusion

The project has been developed successfully and the performance of the system has been found good. The system provides platform to view the books online and can purchase them. The books are recommended to the users using matrix factorization technique. The system is efficient in maintaining customer's details, reduces the work load of the shopkeeper to know the quantity of books available and also keep the records of how many books are purchased and sold.

6.2 Recommendation

BRS is able to fulfill all its objective. It can easily recommend books to user; it is able to provide user interface to make communication between system and user. Although the objectives specified for the development system have been achieved, there are some areas that can improve the system's functionality and usability. The following improvements are the most important:

- 1. Improve user's profile:** The user's profile has been developed using the categories and words of the books evaluated by the user. Further work can be done to make a significant improvement in profile accuracy by using demographic data profiles such as age and sex. In addition, the user's profile can be extended to take into consideration the search term provided by the user. Also, it would be beneficial to introduce more classification of books into the user profile, for example by author and publisher, instead of only book's categories.
- 2. Scalability issues in collaborative filtering:** The collaborative filtering algorithm should be adapted to address the scalability issue as the number of users increases and their profile size becomes large.
- 3. Extending learning module:** The learning module of the system can be extended by allowing the users to provide textual comments on the books, which provides additional information for other users to read. The system can filter and then record the textual comments provided by the user.
- 4. Evaluate content-based filtering approach:** Future work can be done to evaluate the accuracy of the content-based filtering recommendations. Then, an investigation should be conducted to identify which existing evaluation matrix could evaluate the accuracy between different recommender system techniques.

CHAPTER 7 - REFERENCES

- [1] Sushama R, Darshana B and Pooja M (2015), study and Design Book Recommendation System. Available at: <http://www.ijirst.org/articles/IJIRSTV11I1135.pdf>
- [2] Prem M and Raymond J, Mooney and Ramadass N (2002). Content-Based Collaborative Filtering for Improved Recommendation. Available at <https://www.aaai.org/Papers/AAAI/2002/AAAI02-029.pdf>.
- [3] Basu, C., Hirsh, H., Cohen W.: Recommendation as Classification: Using Social and Content-Based Information in Recommendation. In: Proceedings of the 15th National Conference on Artificial Intelligence, Madison, WI (1998) 714-720
- [4] C. Pan, W. Li Research Paper Recommendation with topic analysis in Computer Design and Applications IEEE, 4 (2010)
- [5] Zhao, Z. and Shang, M. (2010). User-based Collaborative Filtering Recommendation Algorithms on Hadoop. Third International Conference on Knowledge Discovery and Data Mining. IEEE journal. 478-481. doi: 10.1109/WKDD.2010.54.
- [6] Cheng, Z., and Hurley, N. (2010). Effective diverse and obfuscated attacks on model-based recommender systems. RecSys '09: Proceedings of the Third ACM Conference on Recommender Systems, ACM, New York, NY, USA, 141–148
- [7] Linden, G., Smith, B., and York, J. (2003). Amazon.com recommendations: Item-to-item collaborative filtering. IEEE Internet Computing, 7(1), 76–80.
- [8] Schafer JB, Konstan J, Riedl J. Recommender system in e-commerce. In: Proceedings of the 1st ACM conference on electronic commerce; 1999. p. 158–66.
- [9] Mooney, R. J., and Roy L. 2000. Content-based Book Recommendation using learning for text categorization. In Proceedings of the Fifth ACM Conference on Digital Libraries, 195–204
- [10] Bogers, T. and Bosch, A. (2007). Comparing and Evaluating Information Retrieval Algorithms for News Recommendation. Tilburg University. Minneapolis, Minnesota, USA.

- [11] Goldberg, D., Nichols, D., Oki, B and Terry, D. (1992).Using collaborative filtering to weave an information tapestry. *Communications of the ACM*, 35(12): pages 61-70.
- [12] Richard Johnsonbaugh and Marcus Schaefer.*Algorithms*, volume 2. PearsonEducation, 2004.
- [13] Internet movie database.<http://www.imdb.com>. Accessed: 2013-05-12.
- [14] G. Salton and M.J. McGill.*Introduction to modern information retrieval*. McGraw-Hill computer science series. McGraw-Hill, 1983.
- [15] Amit Singhal, Chris Buckley, and Mandar Mitra. Pivoted document length normalization. In *SIGIR*, pages 21–29, 1996.
- [16] C. D. Manning, P. Raghavan, and H. Sch  tze.*Introduction to Information Retrieval*. Cambridge University Press, 2008.
- [17] Ayb  ke   zt  rk.*Textual Summarization of Scientific Publications and Usage Patterns*. PhD thesis, Ume   University, 2012.
- [18] Anil K. Jain, M. Narasimha Murty, and Patrick J. Flynn. Data clustering: A review. *ACM Comput. Surv.*, 31(3):264–323, 1999.
- [19] William P. Jones and George W. Furnas. Pictures of relevance: a geometric analysis of similarity measures. *J. Am. Soc. Inf. Sci.*, 38(6):420–442, November 1987.
- [20] A. Huang. Similarity measures for text document clustering. pages 49–56, 2008.