# **APPENDIX 1**

**BOOK RECOMMANDATION SYSTEM**

**END TERM REPORT**

***By***

**Name of students roll no**

Yogesh Kumar 08

Aman Kumar Anand 10

Pankaj Dhunwa 20

Manoj Vyas 22

SECTION : k18kY



**Department of Intelligent Systems**

**School of Computer Science Engineering**

**Lovely Professional University, Jalandhar**

**April-2020**

**APPENDIX 2**

**Student Declaration**

This is to declare that this report has been written by me/us. No part of the report is copied from other sources. All information included from other sources have been duly acknowledged. I/We aver that if any part of the report is found to be copied, I/we are shall take full responsibility for it.

Yogesh Kumar

Roll no: 08

Aman Kumar Anand

Roll no: 10

Pankaj Dhunwa

Roll no: 20

Manoj Vyas

Roll no :22

Lovely Professional University

Date:

**APPENDIX 3**

(A typical specimen of table of contents)

**TABLE OF CONTENTS**

**TITLE PAGE NO.**

**1. Background and objectives of project assigned ..................... 5**

**1.1 objective…………........................................................... 5**

**1.1.1 Recommendation system.......................................... 6**

**1.1.2 content-based filtering………….............................. 6**

**1.1.3 collaborative filtering............................................... 7**

**2 Description of Project ... ............................................................ 7**

**2.1 code part of proje............................................................ 7**

**2.2 description of work division…....................................... 14**

**2.3 Technology and framework to be used………………. 15**

**2.4 SWOT Analysis achieved in project ………………… 15**

**APPENDIX 4**

BONAFIDE CERTIFICATE

Certified that this project report “…BOOK RECOMMANDATION SYSTEM…….” is the

Bonafede work of “YOGESH KUMAR, AMAN KUMAR ANAND, PANKAJ DHUNWA,

MANOJ VYAS .…………” who carried out the project work under my supervision.

<<Signature of the Supervisor>>

<<Name of supervisor>>

<<Academic Designation>>

<<ID of Supervisor>>

<<Department of Supervisor>>

1. **Background and Objective of the Project assigned :**

The booming technology of the modern world has given rise to the enormous book websites. This maker the buyers to choose the best books to read as books play a vital role in many people’s life. The various kinds of books come into existence on day to day basis. So, in order to eliminate this critical situation, the recommendation system has been introduced in which the suggestion on the various books can be provided based on the analysis of the buyer’s interest. The Book Recommendation System is an intelligent algorithm which reduces the overhead of the people. This provides benefit to both the seller and the consumer creating the win-win situation. The E-commerce site to network security, all demands the need for the recommended system to increase their revenue rate. The content filtering, association rule mining and collaborative filtering are the various decision-making techniques employed in the recommendation system as it helps buyers by the strong recommendations as there are various books, buyers sometimes cannot find the item they search for. The Book Recommendation System is widely implemented using search engines comprising of data sets.

* 1. **Objective:**

The Book Recommendation System aims to provide the best suggestion to the

user by analysing the buyer’s interest. The quality and the content are taken into

consideration by employing content filtering, association rule mining and

collaborative filtering.

**RECOMMANDATION SYSTEM:**

Before going to the book recommendation system first lets understand about the recommendation system and how it works ?

There is an extensive class of Web applications that involve predicting user responses to options. Such a facility is called a recommendation system. However, to bring the problem into focus, two good examples of recommendation systems are:

**1.** Offering news articles to on-line newspaper readers, based on a prediction of reader interests.

**2.** Offering customers of an on-line retailer suggestions about what they might like to buy, based on their past history of purchases and/or product searches. Recommendation systems use a number of different technologies. We can classify these systems into two broad groups.

* + 1. **CONTENT-BASED FILTERING :**

Another common approach when designing recommender systems is content-based filtering. Content-based filtering methods are based on a description of the item and a profile of the user's preferences. These methods are best suited to situations where there is known data on an item (name, location, description, etc.), but not on the user. Content-based recommenders treat recommendation as a user-specific classification problem and learn a classifier for the user's likes and dislikes based on product features.

In this system, keywords are used to describe the items and a user profile is built to indicate the type of item this user likes. In other words, these algorithms try to recommend items that are similar to those that a user liked in the past or is examining in the present. It does not rely on a user sign-in mechanism to generate this often-temporary profile. In particular, various candidate items are compared with items previously rated by the user and the best-matching items are recommended. This approach has its roots in information retrieval and information filtering research.

To create a user profile, the system mostly focuses on two types of information:

1. A model of the user's preference.
2. A history of the user's interaction with the recommender system

### **COLLABORATIVE FILTERING :**

One approach to the design of recommender systems that has wide use is collaborative filtering. Collaborative filtering is based on the assumption that people who agreed in the past will agree in the future, and that they will like similar kinds of items as they liked in the past. The system generates recommendations using only information about rating profiles for different users or items. By locating peer users/items with a rating history similar to the current user or item, they generate recommendations using this neighbourhood. Collaborative filtering methods are classified as memory-based and model-based. A well-known example of memory-based approaches is the user-based algorithm

1. **Description :**

Building recommender systems today requires specialized expertise in analytics, machine learning and software engineering, and learning new skills and tools is difficult and time-consuming. In this post, we will start from scratch, covering some basic fundamental techniques and implementations in Python. In the future posts, we will cover more sophisticated methods such as content-based filtering and collaborative based filtering.

**Data:**

[Book-Crossings](http://www2.informatik.uni-freiburg.de/~cziegler/BX/) is a book ratings dataset compiled by Cai-Nicolas Ziegler. It contains 1.1 million ratings of 270,000 books by 90,000 users. The ratings are on a scale from 1 to 10.

Code:-

import pandas as pd  
import numpy as np  
import matplotlib.pyplot as pltbooks = pd.read\_csv('BX-Books.csv', sep=';', error\_bad\_lines=False, encoding="latin-1")  
books.columns = ['ISBN', 'bookTitle', 'bookAuthor', 'yearOfPublication', 'publisher', 'imageUrlS', 'imageUrlM', 'imageUrlL']  
users = pd.read\_csv('BX-Users.csv', sep=';', error\_bad\_lines=False, encoding="latin-1")  
users.columns = ['userID', 'Location', 'Age']  
ratings = pd.read\_csv('BX-Book-Ratings.csv', sep=';', error\_bad\_lines=False, encoding="latin-1")  
ratings.columns = ['userID', 'ISBN', 'bookRating']

## **Ratings data:**

The ratings data set provides a list of ratings that users have given to books. It includes 1,149,780 records and 3 fields: userID, ISBN, and bookRating.

Code:-

A screenshot of a social media post

Description automatically generated

## **Ratings distribution:**

The ratings are very unevenly distributed, and the vast majority of ratings are 0.

Code:-

A screenshot of a social media post

Description automatically generated

## **Books data:**

The books data set provides book details. It includes 271,360 records and 8 fields: ISBN, book title, book author, publisher and so on.

Code:-

A screenshot of a computer

Description automatically generated

## **Users data:**

This dataset provides the user demographic information. It includes 278,858 records and 3 fields: user id, location and age.

Code:-

A screenshot of a computer

Description automatically generated

## **Age distribution:**

The most active users are among those in their 20–30s.

Code:-

A screenshot of a social media post

Description automatically generated

## **Recommendations based on rating counts:**

A screenshot of a social media post

Description automatically generated

The book with ISBN “0971880107” received the most rating counts. Let’s find out what book it is, and what books are in the top 5.

Code:-

A screenshot of a computer

Description automatically generated

The book that received the most ratings in this data set is Rich Shapero's Wild Animus. Something in common among these five most rated books - they are fictions or novels. The recommender suggests that novels and fictions are popular and likely receive more ratings. And if someone likes "Wild Animus", probably we should recommend him(her) "The Lovely Bones: A Novel".

## **Recommendations based on correlations:**

Find out the average rating and the number of ratings each book received.

A screenshot of a computer

Description automatically generated

#### **Observation:**

In this dataset, the book that received the most ratings is not highly rated at all. So if we were set to use recommendations based on rating counts, we would definitely make mistakes here.

#### To ensure statistical significance, users with less than 200 ratings, and books with less than 100 ratings are excluded.

A screenshot of a social media post

Description automatically generated

### **Rating matrix**

Convert the table to a 2D matrix. The matrix will be sparse because not every user rate every book.

A screenshot of a computer

Description automatically generated

Let's find out which books are correlated with the 2nd most rated book "The Lovely Bones: A Novel". To blatantly quote from the Wikipedia: It is the story of a teenage girl who, after being raped and murdered, watches from her personal Heaven as her family and friends struggle to move on with their lives while she comes to terms with her own death.

A screenshot of a social media post

Description automatically generated

We obtained the books' ISBNs, but we need to find out the names of the books to see whether they make sense.

A screenshot of a social media post

Description automatically generated

Let's select three books to examine from the above highly correlated list "The Nanny Diaries: A Novel", "The Pilot's Wife: A Novel" and "Where the heart is".

"The Nanny Diaries" satirizes upper class Manhattan society as seen through the eyes of their children's caregivers.

Written by the same author of "The Lovely Bones", "The Pilot's Wife" is the third novel in Shreve's informal trilogy to be set in a large beach house on the New Hampshire coast that used to be a convent.

"Where the Heart Is" dramatizes in detail the tribulations of lower-income and foster children in the United States.

These three books sound right to me to be highly correlated with "The Lovely Bones". Seems our correlation recommender system is working.

**Description of Work Division in terms of Roles among Students :**

* We all have contributed equally in this project.
* Firstly, we need to known the basic knowledge about recommendation system ,and need to known how actually it works.
* Everybody in our group gather the information about recommendation system and now we have to implement in our project.
* Everybody have to gather the large amount of book data as csv file all have done quite well.
* Then we choose right csv file according to us to use in our project.

**Yogesh Kumar :**

Yogesh have to find the distribution of rating in out book data csv file , which include at least large numbers of books data .

And many of have null ratings so he made a possible graph so, we can easily calculate the maximum no of ratings give in books.

**Aman Kumar Anand :**

Aman helps us to find the users location and their age distribution , in which the maximum no of user are between in which age group, on the basis of Yogesh rating calculation.

And the book which are found of and made a graph of such no of users as based on age distribution, so it really helps in the recommendation system engine.

**Pankaj Dhunwa:**

Pankaj helps us in the recommendation based on rating count. On the basis of Yogesh and Aman rating results Pankaj find the top 5 best rating.

And after that find the topmost 5 rated book from our large amount of data.

**Manoj Vyas :**

Manoj helps us to find another similar book from which Pankaj have found the top-rated books.

On the basis of their ratings ,age distribution ,type Manoj trying to compare from another good rated book. And we found good results.

**Technology and framework used:**

* We have used our Jupiter notebook software.
* Python idle verson3.0.
* Trying to get knowledge from the google, Wikipedia, GitHub.
* We get our book data csv file from some blogger.

**SWOT Analysis achieved in project :**

While making this project we make many analyses done. We think and discuss a lot how this project could be better, but we are only able to make a small book recommendation system. while making we faces many challenges and trying to resolve our quarries.

Book recommendation system is the best way to find our choice book on the basis of which u have given the rating ,through the writers u like to read most.

U just need the microseconds to find the best book u like the most using the book recommendation system. Which give many opportunities to students or any person which are trying to get knowledge.

It may be some threat to the minimum level if a person want knowledge for bad purposes, we have many types of book written in our world or in our book database system .if we found the large amount of data we can easily access through the outside of the project so , if we can make good secure book recommendation system engine definitely we get the better result.