­

|  |  |
| --- | --- |
| Faculty of Computers and Information  Information Systems Department  Graduation Project 2018/2019 | Book Recommendation System  Supervision by:  Dr. Amany Abdo |

Book  
Recommendation  
System

Prepared by:

Abd El-Rahman Anas

Ala’a Mohammed Sayed Rashid

Sayed Ebrahim El-Sayed

Mohammed Tarek Mahmoud

Mostafa Magdy Ahmed

**Abstract**

The online recommendation system has become a trend. Now a days rather than going out and buying items for themselves, reason being, online recommendation provides an easier and quicker way to buy items and transactions are also quick when it is done online. Recommended systems are powerful new technology and it helps users to find items which they want to buy. A recommendation system is broadly used to recommend products to the end users that are most appropriate. Online book selling Web sites now-a-days is competing by considering many attributes. A recommendation system is one of the strongest tools to increase profits and retaining buyer. The existing systems lead to extraction of irrelevant information and lead to lack of user satisfaction. This paper presents Book Recommendation System (BRS) based on combined features of content-based filtering (CBF), collaborative filtering (CF) and association rule mining to produce efficient and effective recommendation. For this we are proposing a hybrid algorithm in which we combine two or more algorithms, so it helps the recommendation system to recommend the book based on the buyer's interest.

**Contents**

1. **Abstract**
2. **Acknowledgment**
3. **Table of contents**
4. **List of Figures**
5. **List of Tables**
6. **Text Chapters  
   Chapter 1: Introduction  
   Chapter 2: Background/Literature Review**  
   **Chapter 3: Methodology  
   Chapter 4: Results and Discussions  
   Chapter 5: Conclusions and Recommendations**

1. Introduction

1.1. Overview ……………………………………………………………….......... 5  
1.2. Objectives ………………………………………………………………......... 5  
1.3. Purpose ………………………………………………………………..............5  
1.4. Scope ……………………………………………………………….................5

2. Project "Planning and analysis"

2.1. Project planning …………………………………………………. 12

2.1.1. General information**:** ………………………….………………….…12  
2.1.2. Technical Information and System Requirements. ……….18

2.4. Analysis of the system ……………………………………… 22

2.4.1. Domain Requirements ………………………………………………. 23  
2.4.2. Functional Requirements ……………………………………………. 24  
2.4.3. Non- Functional Requirements ……………………………………… 25

2.6. Risk and Risk Managements …………………………………...... 27

3. Software Design

3.1. Design of database (ERD or Class) Diagram ……………………. 28

3.1.1. Class Diagram ………………………………………………………. 28  
3.1.2. ERD ………………………………….…………….......................... 29

3.2. Use case ………………………………….………………………. 30

3.2.1. Use case diagram ……………………………………………………. 30  
3.2.2. Use case Scenarios …………………………………………………... 31

3.3. Sequence diagram ………………………………………………. 40  
3.4. Activity diagram ………………………………………………… 42  
3.5. Block diagram …………………………………………………… 44

4. Implementation 45

4.1. Software architecture ……………………………………………. 45  
4.2. User Manual ……………………….……………………………. 48  
4.3. Pseudocode, Flowchart or workflow ……………………………. 61

5. Testing 63

5.1. Unit Testing ………………………………………………………. 63  
5.2. Integrated testing …………………………………………………. 63

6. Results and Discussion 70

6.1. Results ……………………………………………………………. 70

6.1.1. Expected result …………………………………………………........ 70  
6.1.2. Actual results …………………………………………………........... 71

6.2. Discussion ………………………………………………………… 71

7. Conclusion 72  
8. Future work 73

**1. Introduction**

**1.1. Overview**

A recommendation system is a subclass of information filtering system that seeks to predict the "rating" or "preference" a user would give to an item. Our project is Book Recommendation System based on combined features of content based collaborative filtering and association rule mining to produce efficient and effective recommendation.

**1.2. Objectives**

* Facilitating the user's ability to determine what is the next book and what books are important and spread among friends and acquaintances.
* Developing sites that responsible for publishing books, increasing awareness and interaction with different personalities and minds, increasing the number of followers.

**1.3. Purpose**

We can say that the main goal to make this system is saving time and effort for user and help him to choose the best books possible from all books in the site.

Helps the publisher the book to reach to many readers who want this kind of books.

**1.4. Scope**

* To facilitate the publication of books for reading in all fields of science, so that there are some areas need to organize the priority of books, for example that you must read the book (x) before the book (y) to understand the book (y) more and so on.
* The reader needs a guide to easily determine his goal without searching or fatigue

**Acknowledgment**

**DR: Amany Abdu**

**Eng. weal Abbas**

**Host: Pixels**

Comprehensive survey of related work

Amazon

use 4 types of filtering in their recommender system. 1- Recommended for you, it’s based on your history. 2- Your recently viewed items and featured recommendations. 3- Customers who bought this item also bought. 4- frequently bought together.

Netflix

makes sure not to over-personalize. They just recommend shows based on high rate and trends.

Goodreads

learns about your personal tastes from your ratings, then generates recommendations unique to you. You need to rate at least 20 books to get Goodreads Recommendations.

Our System

Recommendation in our system will be based on 3 levels of filtering: 1- Collaborative filtering. 2- Category. 3- highly rated.

**2. Project "Planning and analysis"**

**2.1. Project planning**

**2.1.1. General information**

**Reason for selecting this project**

the online libraries offer you the most popular or top-selling books, is this the best suggestion for the reader, Thanks to the remarkable expansion of data mining Science (the process of discovering patterns in large [data sets](https://en.wikipedia.org/wiki/Data_set) involving methods at the intersection of [machine learning](https://en.wikipedia.org/wiki/Machine_learning), [statistics](https://en.wikipedia.org/wiki/Statistics), and [database systems](https://en.wikipedia.org/wiki/Database_system)), we decided to use it in this field because it is much better and mush helpful for both parties (Reader and online libraries) .

**Reason for selecting this Approach**

This paper presents Book Recommendation System (BRS) based on combined features of content-based filtering (CBF), collaborative filtering (CF) and association rule mining to produce efficient and effective recommendation.

* [**Collaborative**](https://en.wikipedia.org/wiki/Collaborative_filtering) **filtering:**

The system generates recommendations using only information about rating profiles for different users or items. Collaborative systems locate peer users/items with a rating history like the current user or item and generate recommendations using this neighborhood. The user-based and the item based nearest neighbor algorithms can be combined to deal with the cold start problem and improve recommendation results.

* **Content-based filtering:**

The system generates recommendations from two sources: the features associated with products and the ratings that a user has given them. 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.

For this we are proposing a hybrid algorithm in which we combine these two algorithms, so it helps the recommendation system to recommend the book based on the buyer's interest.

**Problem Statement**

* A lot of time is wasted in search of new book to read, many questions to their friends, is that is the best solution?! This should not,
* It is best to use this time to read useful books and learn new things.
* When the online libraries offer you the most popular or top-selling books, is this the best suggestion for the reader? What if it suggested books in the field that he preferred but could not understand simply? Does he know that, there is a book that is easier to understand or that he should read other books before this book that simplify the understanding of this book, etc.

**Project Importance**

* To Reader:
  + The reader needs a guide to easily determine his goal without searching or fatigue
* To site:
  + Optimize targeting for readers
* To the society:
  + Increase the general awareness to improve reading comprehension.

**Project Idea**

Book Recommendation System is a system that seeks to predict the "rating" or "preference" a user would give to an item, based on combined features of content based **collaborative filtering** and **association rule mining.**

**2.1.2. Technical Information.**

**Materials or equipment does the project need**

**• Hardware:**

**Three PCs with medium Capabilities**

**• Software:**

* **Operating system:**

**Cloudera**

* **PL:**

**PHP Programming language(Laravel)**

* **Analytics engine:**

**Apache Spark**

* **Web design:**

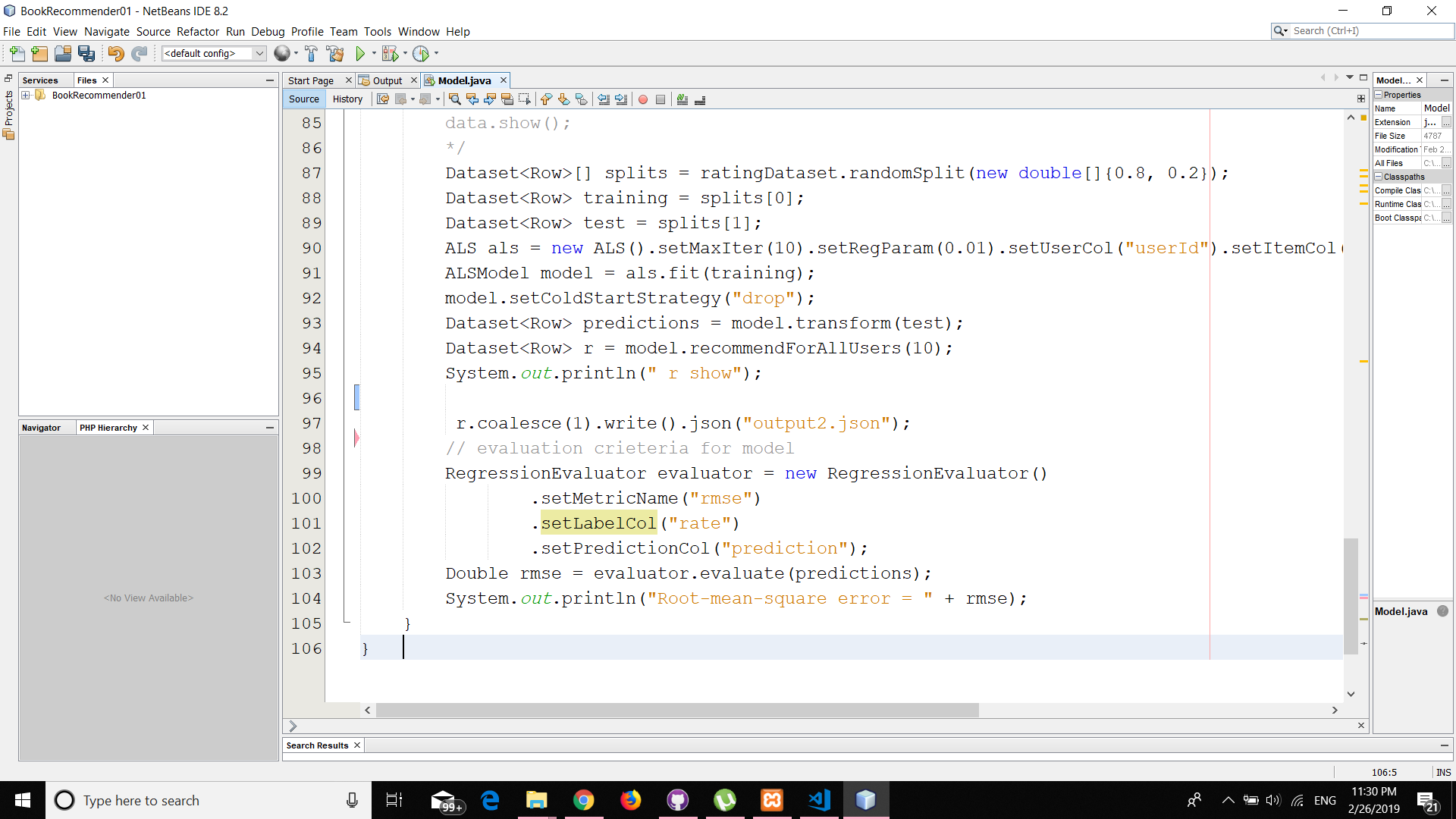
**HTML,**

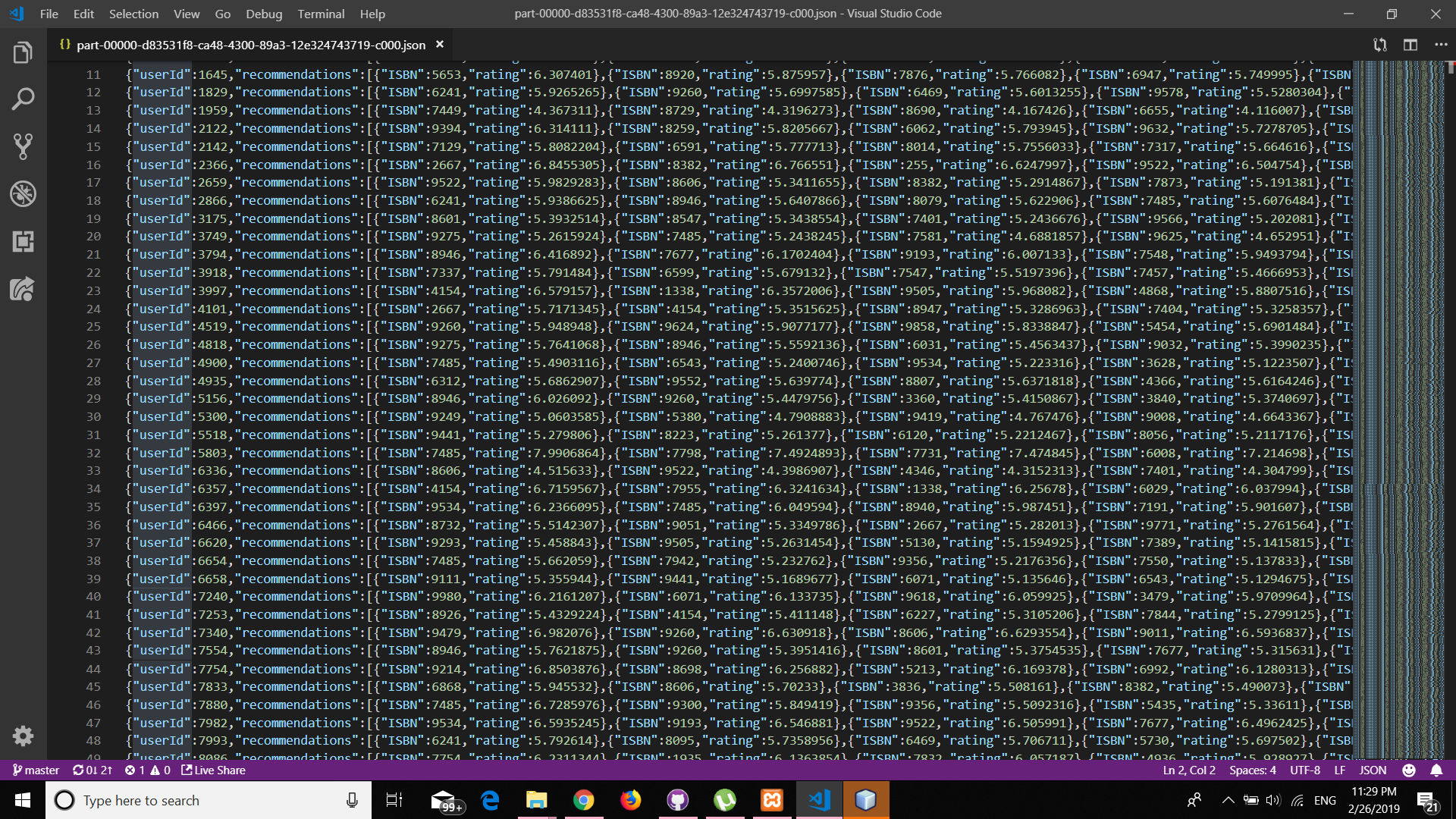
**CSS,**

**JS,**

**Bootstrap.**

Details of partial implementation conforming to the design

****

****

Commands of tools and techniques being used during project implementation

* **ALS algorithm**
* **Hadoop**
* **Spark**
* **PHP(Laravel)**
* **SQL**
* **NetBeans**

**2.4. Analysis of the system**

**2.4.1. Domain Requirements**

* Enough memory should be accounted on the database server to accommodate for any number of users and events they do.
* Reports are required.
* Each user must have a unique user-name/password combination.
* Each user must have a unique home key for their home.
* The database should be backed up every once in a while, in case the original does become corrupt.
* Software must verify all values before making the change in the database.

**2.4.2. Functional Requirements**

**2.4.3. Non- Functional Requirements**

1. **Performance requirements**-In case of opening databases, evaluation there are no delays and the operation is performed in less than 2 seconds for opening computing.  
   −In case of any failure happened in the system, the system shall be able to recover this failure any load or stress happened on the system it shall be able to deal with it.  
   − Measure render response time is generally necessary to include functional test scripts as part of the performance test scenario.  
   −Server response time this refers to the time taken for one system node to respond to the request of another.
2. **Reliability**Our system being reliable if any software or hardware failure is happened the system should be able to know how to detect this failure and know the cause of it. this like if any failure happened in any part of any hardware component the system shall know which part and detect the failure.
3. **Security**−The system shall tell the user to write strong password for his account to protect his account from any hacking.

−In our system we put (home key) for every home to prevent it from  
stranger's access.  
−The first line of defense against hackers is to password protect your  
Wi-Fi router. The factory-given password isn’t good enough.

−Create your own unique pass-phrase, a sequence of words that’s likely to be much stronger than whatever alphanumeric password you may typically use across accounts (an ill-advised practice).  
−Change your router’s network ID name from the default given with  
the device, which will make it harder for attackers to know the kind of  
hardware you use.  
−Turn it off: Disconnect your devices from the internet when they’re  
not in use or if you have no need for that level of networking. Many  
devices only need to connect to other devices on your home network  
and don’t require full access to the World Wide Web.

1. **Usability**In our system we use simple web-based application to help users to use it effectively, and we aspire to make it mobile application also soon.
2. **Availability**The system shall be able to do that:  
   If the internet service gets disrupted while sending information to the server, the information can be send again for verification the system shall be able to do to that.
3. **User Friendly**In the system the user can understand the system seamlessly, as it easy to use.
4. **Support many users**System support most common language, to help many users to use it without need of sense of the system. That made the new system more east than the old system. But the current system doesn’t support this function that make not many users to use it.

**2.6. Risk and Risk Managements**

Risk management is the process of identifying, assessing, reducing and accepting risk. Efforts to avoid, mitigate and transfer risk can produce significant returns. Risk management also leads to a culture of explicitly accepting risk as opposed to hiding in the optimism that challenges and failures aren't possible.

The following are hypothetical expected examples of risk management.

1. **Poor password selection:**

−Every new user should register on our system by user-name, password, etc...  
−Poor password may cause a high risk, so this risk management is:  
−Use a combination of upper-case and lower-case, numbers and special characters  
such as: @ \# \% ( ) , : “ „ etc..  
−Make your password long enough: Between 8 to 20 characters is recommended.  
−Omit duplicates “Use a unique password for each of your accounts. The same password should never be used more than once!”

1. **Unregistered system events**

−In our system events are the data comes from sensors by date and time and saved in database, because of any failure it may be a high risk to un-register that events from our system.

1. **Slow data analysis process**

−Because of the size of data from books, their rating and users, we have huge amount of data and it may cause huge risk, it is possible to slow down the use of algorithm, so algorithm should be used every specific period and be weekly

1. **Hardware Failure**

−We have different types of hardware such as devices, servers, electrical cables, etc....  
−So we can manage it by make alerts if any hardware fail

1. **User could be having not friends and interests**

−It is possible to predict something to anonymous person, so we will suggest to him the most preferred categories and most selling books

There may be other challenges, but these are the most expected challenges