

BOOK RECOMMENDATION SYSTEM

NAME : ROHIT RAJ R

ROLL NO : A9929722000989(el)

FORM NO : 5194497

ABSTRACT

Cloud computing is an emerging computational model for delivering services over the internet to the clients at any time and at any location on a pay-per-use basis. Almost every significant business uses recommendation systems to improve the caliber of its offerings. It helps to improve sales and enable higher revenue generation. Recommendation systems help to navigate quickly and receive necessary information.

This project is a quick and intuitive book recommendation system that helps readers to find appropriate book to read next. The system aims at finding the suitable books that suit the interests and preferences of cloud consumers and recommend similar books to the reader based on his interest and item ratings. The model is based on combining content filtering and collaborative filtering method to provide recommendation.

INTRODUCTION OF THE STUDY

Recommendation systems is an software product that use a user's rating of each item and preferences to assist them find books that suit their requirements and interests. Recommendation systems can be found everywhere. Movies on Netflix or scrolled through Instagram, a person will have unknowingly utilized a recommendation system. Nobody has time to shift through thousands of pages in search of a product. Recommendation algorithms result in better service and customer satisfaction, which attracts more traffic to the website. Recommendation systems are crucial in assisting consumers in finding products.

The modern world's expanding technology has given rise to massive book websites. As books are crucial in many people's lives, this allows them to choose the best novels. Various types of books appear on a daily basis. A recommendation system has been built to handle this critical issue, in which recommendations on various publications can be generated based on a research of the users' interests. The book recommendation system is widely used, and it makes use of data sets in search engines.

Recommendation systems are the subordinate class of information filtering system which helps to recommend a particular item to the user. Recommendation is used in a wide range of contexts, including movies, music, news, books, and products. In general, recommendation systems provide a list of suggestions using various methods such as content-based filtering and collaborative filtering.

Through the analysis of big data, the method of recommendation based on users' borrowing records, browsing time and the number of clicks is adopted. These data are statistical information based on the historical records of user accounts, with inaccuracy and limitation. The same user account can be used by more than one person, so the book

borrowing record has no pertinence to the user account. At the same time, new registered users have no record of borrowing books, and the book recommendation system will cause cold start problems.

SYSTEM CONFIGURATION

- Operating System - Windows | Mac
- Browser - Google Chrome
- Software
 - Visual studio
 - Python 3.8.5
 - Anaconda Prompt

TOOLS AND TECHNOLOGY USED

- Python
- Flask
- HTML
- CSS
- MongoDB
- Heroku

OBJECTIVE OF THE STUDY

OVERVIEW

This topic describes about the existing system, drawbacks of the existing system and the proposed system in detail.

RECOMMENDER ENGINES

Recommender engines try to infer tastes and preferences and identify unknown items that are of interest. They are the most immediately recognizable machine learning technique in use today. There are probably seen services or sites that attempt to recommend books based on past actions of users.

EXISTING SYSTEM

The existing systems have different mechanisms with different features.

- The existing book recommendation engines used by the top rated book purchasing websites.
- The existing engines make use of conventional algorithms for recommendations.
- The results from all the recommender techniques are combined and the set for recommended books is generated.

Drawbacks of existing system

- Cannot handle the new users
- Lack of accurate data
- Changing user preferences

- The flow of existing systems and the obstacles faced by users when searching books.

PROPOSED SYSTEM:

A web application is developed in order to overcome the drawbacks in the existing system and handle the new users. It also helps the user to gain some basic information about the user Interest.

The proposed system consists of two major functionalities.

- Content Based Filtering
- Collaborative Based Filtering

LITERATURE REVIEW

OVERVIEW

This topic describes about the various analytical models and techniques, the methodology and approach in detail, used to develop the system.

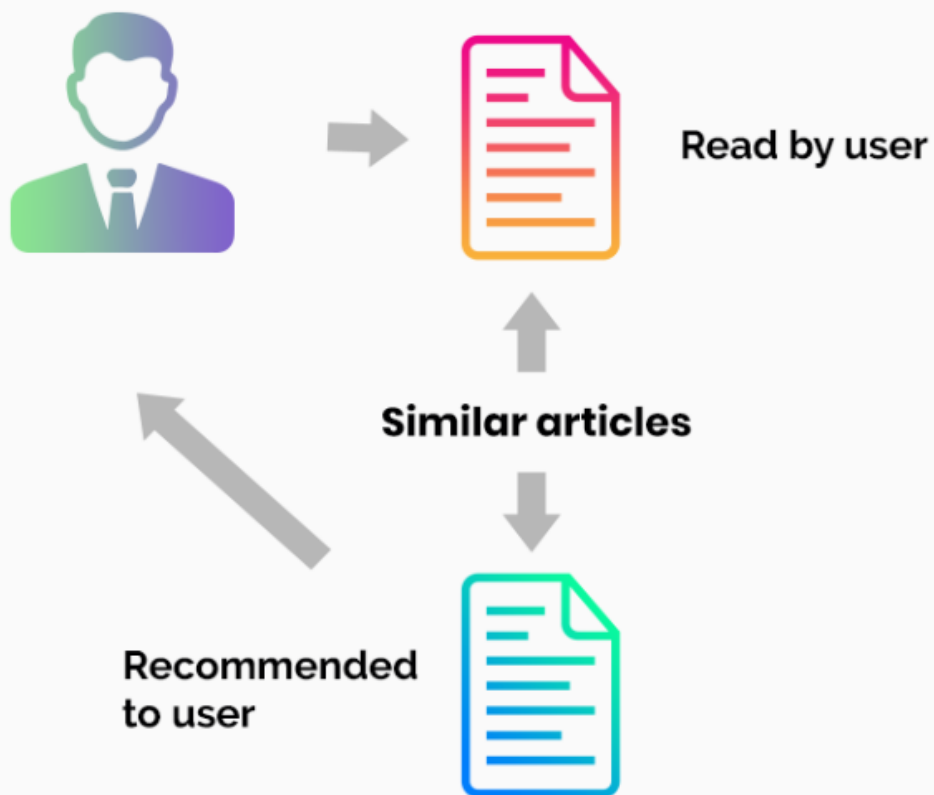
CONTENT BASED FILTERING

Content-based filtering is a type of recommender system that attempts to guess what a user may like based on that user's activity. Content-based filtering makes recommendations by using keywords and attributes assigned to objects in a database and matching them to a user profile.

Content-based algorithm is always the most popular, powerful, and productive method taken at the beginning time of any project. It is the process of selecting items from a huge collection of books that the user is likely to find useful. To discuss current research trends and directions that could lead to the next generation of content-based recommender systems.

The model doesn't need any data about other users, since the recommendations are specific to this user. This makes it easier to scale to a large number of users. The user profile is created based on data derived from a user's actions, such as items searched on a website.

Content-based filtering



Content-based Filtering

METHODOLOGY

Merge the book title, author, category into single string and add to dataframe in new column. Apply count vectorizer to the dataframe. The Count Vectorizer provides a simple way to both tokenize a collection of text documents and build a vocabulary of known words, but also to encode new documents using that vocabulary. The result occurrence of the count vectorizer is converted into a sparse matrix.

	at	each	four	geek	geeks	geeksforgeeks	help	helps	many	one	other	two
document[0]	0	0	0	1	1	0	0	1	0	0	0	1
document[1]	0	0	1	0	2	0	1	0	0	0	0	1
document[2]	1	1	0	1	1	1	0	1	1	0	1	0

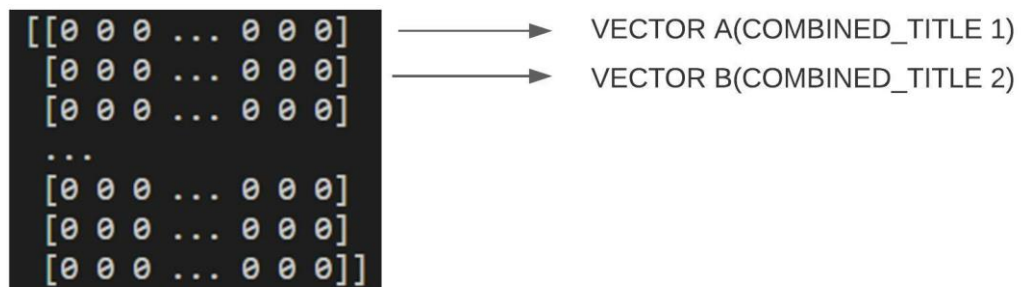
Sparse Matrix

Compute a user similarity matrix, use cosine similarity to find the similarity between two users. Cosine similarity measures the similarity between two vectors of inner product space, It is measured by the cosine of the angle between two vectors.

Formula for content based filtering

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

Content Based Formula



Sparse Matrix For User Ratings

COLLABORATIVE BASED FILTERING

Collaborative Filtering is the process of filtering or evaluating items using the opinions of other people. While the phrase collaborative filtering (CF) has only been around for a few years, it is based on something that humans have done for centuries: sharing thoughts with others.

In recommendation systems, the collaborative filtering method is used to provide recommendations based on ratings submitted by other system users. If a buyer's ratings for one item match, it's likely that their ratings for subsequent goods would match as well.

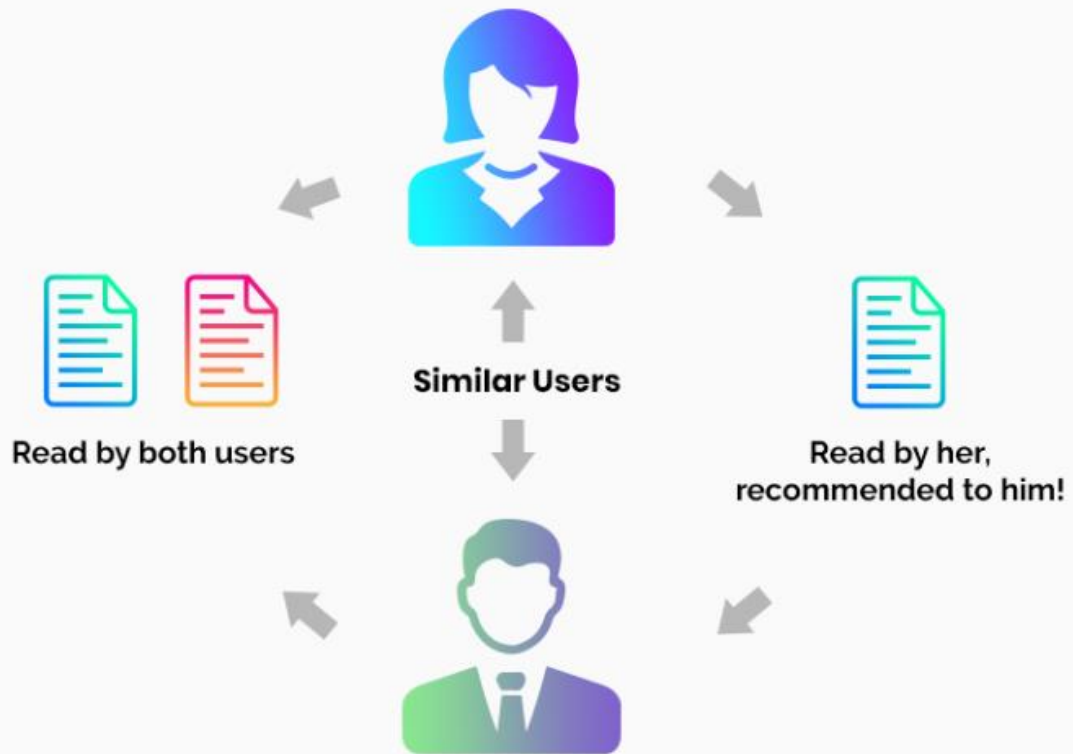
Computers can't judge qualitative criteria like flavor or quality, and recommendations with collaborative filtering are based on human ratings and qualitative parameters. Therefore, Collaborative filtering provides with better results. Proper product recommendation service can contribute to increase purchase by suggesting proper products to the customer.

The fundamental benefit of collaborative filtering is that shoppers are exposed to a wider range of products, which opens up opportunities to persuade shoppers to make repeat purchases.

User-Based Collaborative Filtering is a technique for predicting which items a user would like based on the ratings given to that item by other users with similar tastes to the target user.

- Not required to understand item content:
- No item cold-start problem:
- Captures the change in user interests over time:
- Captures inherent subtle characteristics:

Collaborative filtering



Collaborative-based Filtering

Collaborative filtering systems recommend products to the target customer based on the opinions of other like-minded customers. These systems use statistical techniques to identify a group of customers called neighbor who are similar to the target user. When a neighbor is found, the system employs a number of algorithms to generate recommendations.

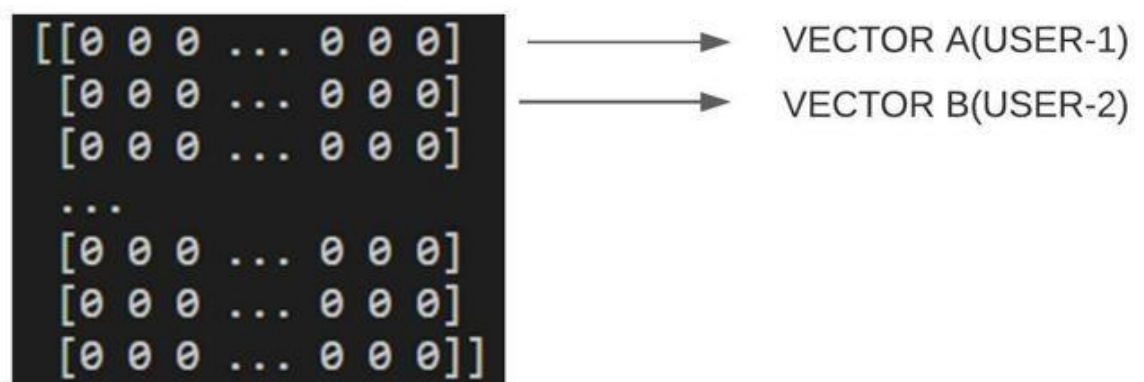
METHODOLOGY

Calculate the user's similarity follow these guidelines, determine how similar two users are, use cosine similarity. The cosine of the angle between two vectors is used to assess cosine similarity between two vectors in inner product space.

Formula for collaborative based filtering

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

Collaborative Filtering Formula



Spare Matrix For Count Vectorizer

RESEARCH METHODOLOGY

OVERVIEW

This topic gives a detailed description about the modules used in the implementation of the system and the visual screens of the website.

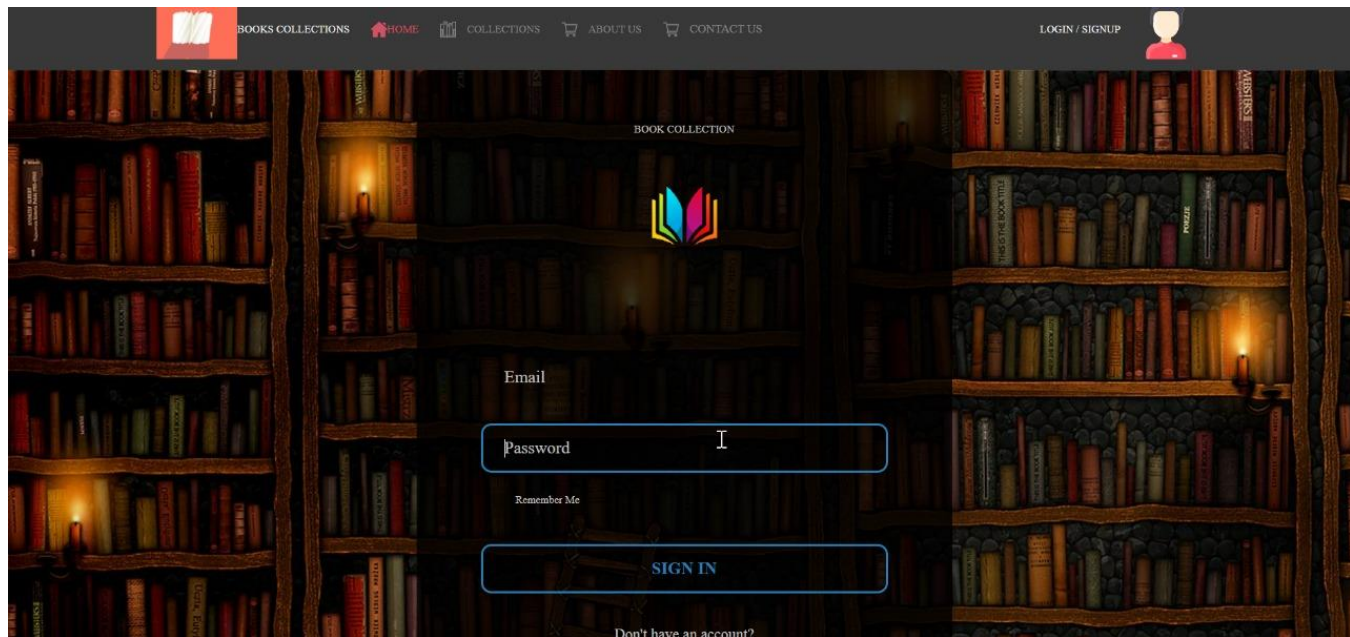
MODULES DESCRIPTION

The first page of the system displayed by the login page module, where the user can login using the credentials. If the user is new, the user is directed to the registration page module, after registration, the user can login using their credentials. After successful sign in, the homepage module is displayed, with top 10 recommendation of books along which includes a search bar, contact us module, about us module. The modules involved in the system are discussed in detail below.

LOGIN:

The login page allows a user to gain access to an application by entering their username and password. The system will display the customer's information once the customer is logged in to the application. The users can access the secured dynamic content once they are successfully logged in to the system. There are two possible results,

- Authentication is successful and the user is directed to the application landing page.
- Authentication fails and the user remains on the login page. If authentication fails, the screen should show an informational or error message about the failure.

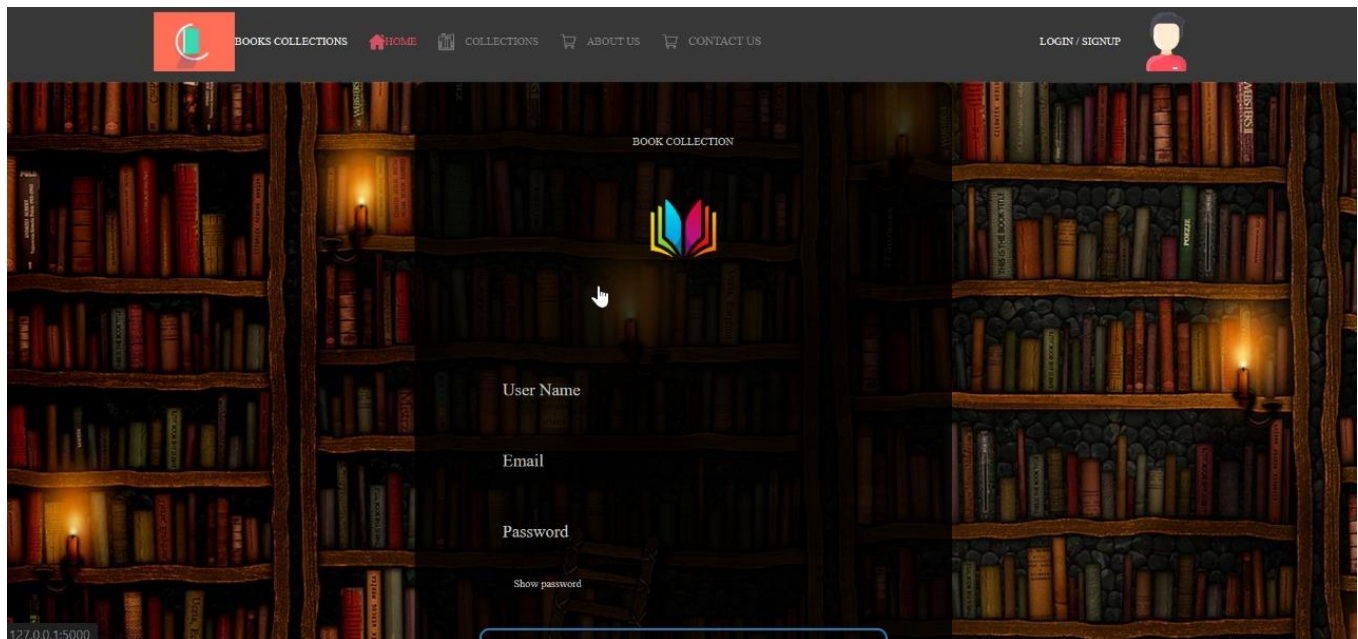


Login Page

REGISTRATION

Registration occurs only the first time a user access the system. It is a way to authorize, check and register the credentials of the user. Every time after initial registration, user can log on to the system using the username and password created at the time of registration.

A registration form is a list of fields that a user will input data into and submit to a company or individual. Companies use registration forms to sign up customers for subscriptions, services, or other programs or plans. A signup page is also known as a registration page. It enables users and organizations to independently register and gain access to your system. It is common to have multiple signup pages depending on the types of people and organizations you want to register.

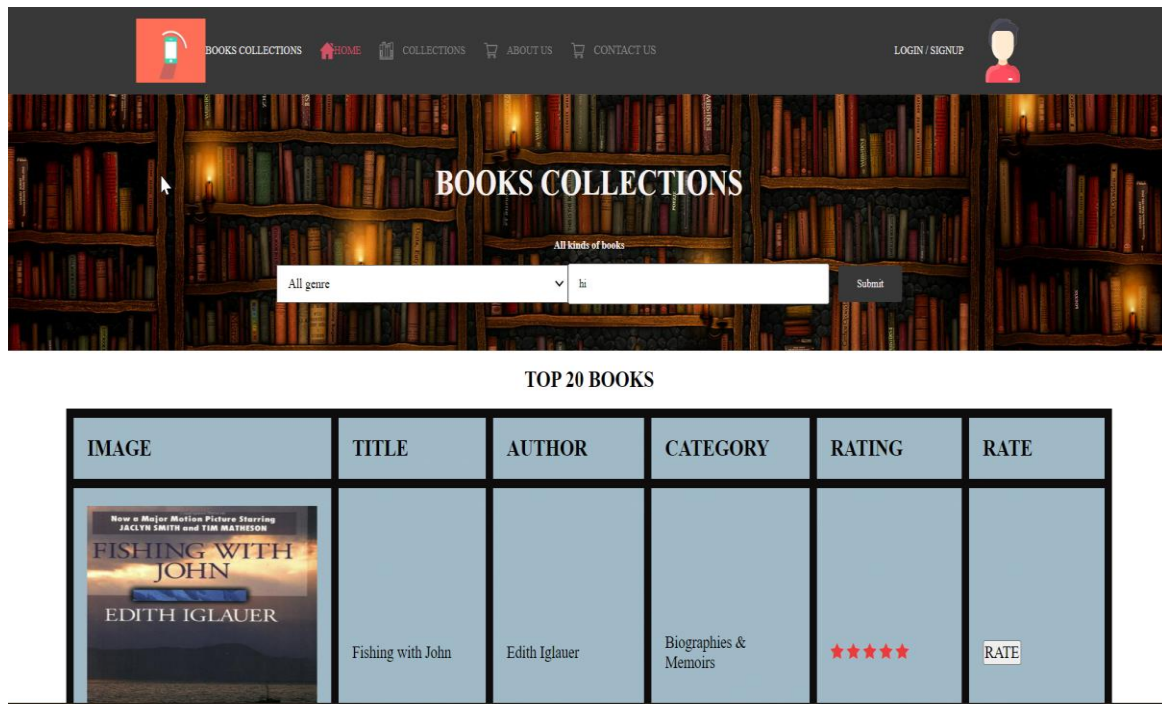


Registration Page

HOME PAGE

A home page is the main web page of a website. The term may also refer to the start page shown in a web browser when the application first opens. The home page will display popular 20 images and the user can search for an option using the search bar to search for any book in general or particularly in a specific category. And other options like Collections, About us, Contact us are shown in the menu.

The homepage is intended to be the first page visitors see when reaching a website's URL. Essentially, it serves as a significant landing page for visitors, capturing their attention and letting them know what the website is all about, whether it's commercial or personal. One of the most common uses of a homepage is to act like a directory, giving visitors valuable information about the website and providing links to specific areas of the particular functions.



Home Page

COLLECTION LIST

This page contains various lists of categories for the books provided, based on specific interest of the user, the books can be selected based on the category. A category listing page (CLP) is a page on a website that presents a list of books based on a category. CLPs are also rich with metadata, and because each product links back to a category page, they can heavily influence SEO rankings and internal link building strategies.

The majority of Ecommerce traffic is sent to product listing pages, either from an email campaign, social media, or advertisement. Therefore, it's important to ensure the CLP experience is carefully designed to

- Promote product discovery
- Encourage user engagement
- Decrease time to purchase

BOOK COLLECTIONS

BOOKS CATEGORY					
Arts & Photography	Biographies & Memoirs	Business & Money	Calendars	Children's Books	Comics & Graphic Novels
Computers & Technology	Cookbooks, Food & Wine	Crafts, Hobbies & Home	Christian Books & Bibles	Engineering & Transportation	Health, Fitness & Dieting
History	Humor & Entertainment	Law	Literature & Fiction	Medical Books	Mystery, Thriller & Suspense
Parenting & Relationships	Politics & Social Sciences	Reference	Religion & Spirituality	Romance	Science & Math

Collection Page

CONTACT US

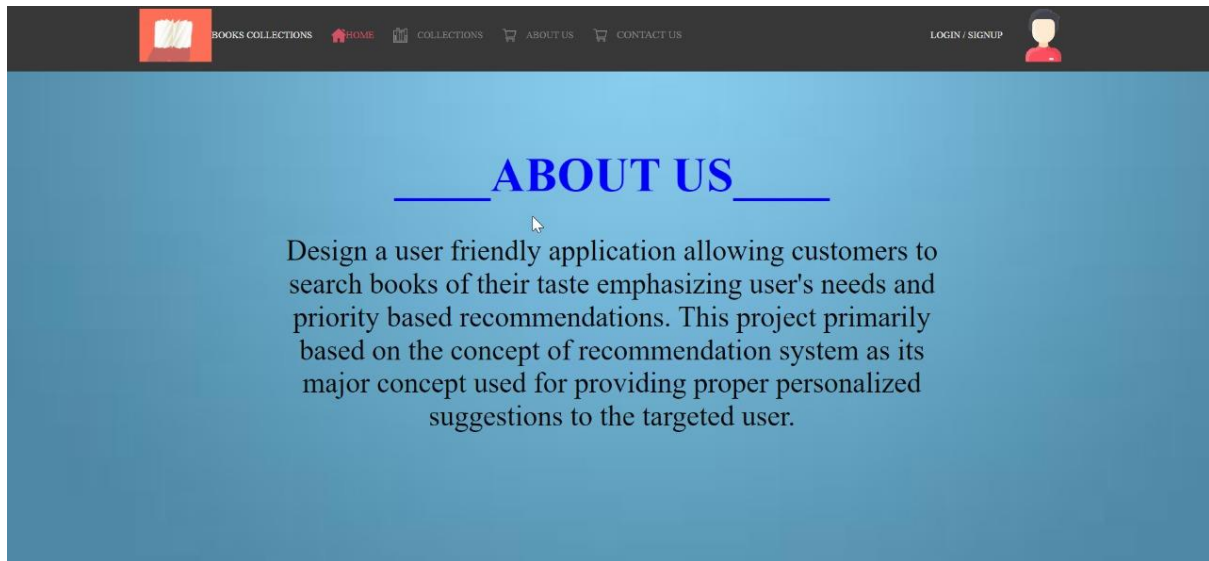
A contact page is a common web page on a website for visitors to contact the organization or individual providing the website. The page contains one or more of the following items:

- Support Team Name
- Phone Number
- E-Mail Address

Contact Us Page

ABOUT US

About us page contain a short detail description about the website and the services provide .



About Us Page

DATA ANALYSIS & INTERPRETATION

OVERVIEW

This topic describes about design of the system and the data flow model in detail.

It includes a visual representation of the data flow in the system

.

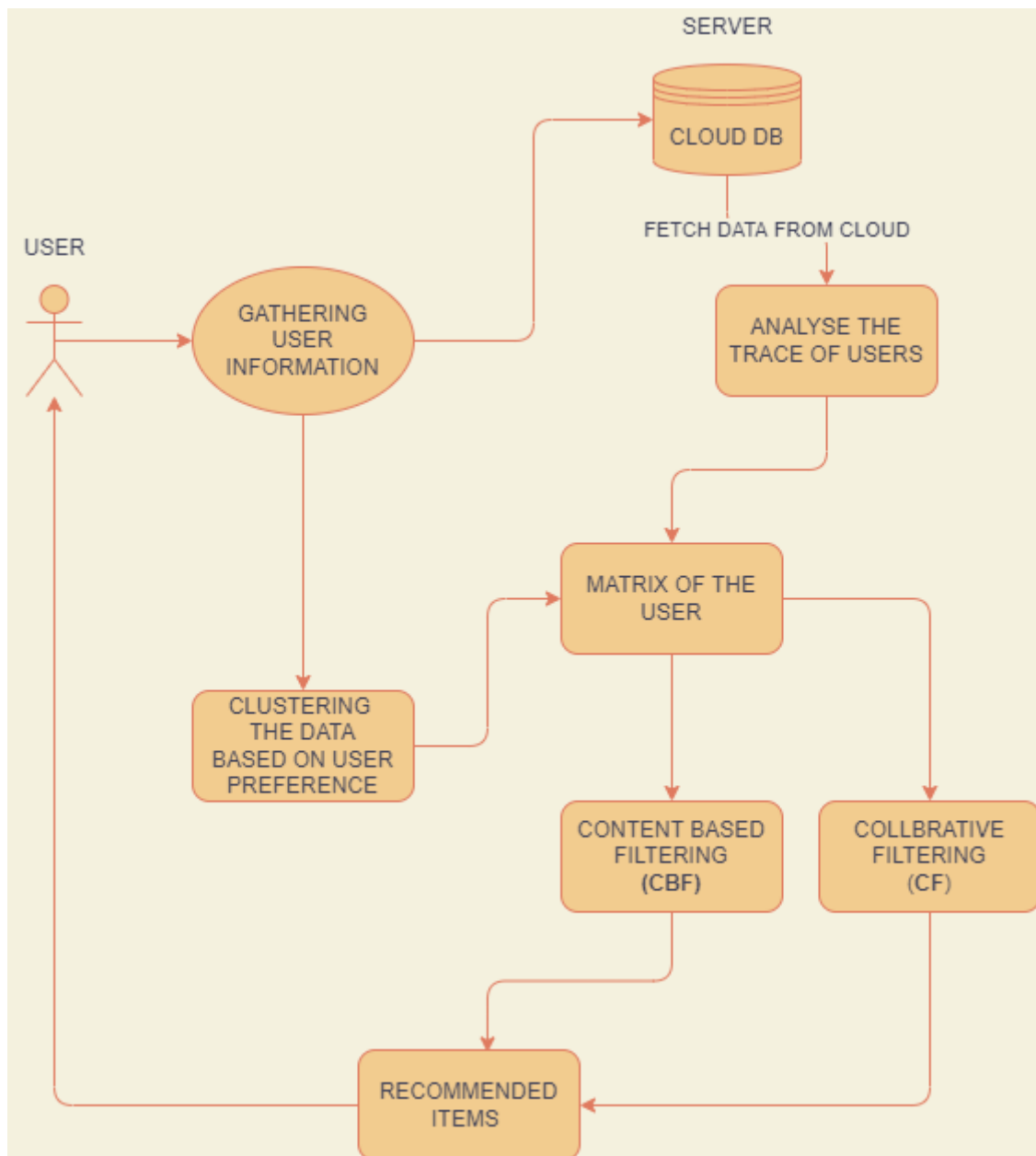
DATA FLOW DIAGRAM

Data flow diagrams are used by information technology professionals and systems analysts to document and show users how data moves between different processes in a system. Analysts usually begin by looking at the big picture before delving into the specifics of each process.

The visual representation makes it a good communication tool between User and system designer. Structure of DFD allows starting from a broad overview and expands it to a hierarchy of detailed diagrams.

The diagram represents gather information from user and store in database. With user information analyze the user data and applying the filtering techniques and recommend item for the user.

- It helps us to understand the functioning and the limits of a system.
- It is a graphical representation which is very easy to understand as it helps visualize contents.
- Data Flow Diagram represent detailed and well explained diagram of system components.
- Data Flow Diagrams can be understood by both technical or nontechnical person because they are very easy to understand.



Data Flow Diagram

RESULT & DISCUSSIONS

RESULTS :

Content-Based Filtering achieves an impressive accuracy rate of 94%, while Collaborative-Based Filtering yields a strong 88% accuracy. The project has been successfully deployed on Heroku, making it accessible worldwide through a global domain.

DISCUSSION :

1. **Improved User Experience:** A well-implemented book recommendation system can significantly enhance the user experience on a book-related website. Users are likely to find books that match their interests and preferences more easily.
2. **Increased User Engagement:** The inclusion of category listings, collections, and contact features can encourage users to engage more with the platform. Users may explore different categories of books, which can lead to more extensive interaction with the website.
3. **Scalability and Performance:** The document mentions the use of content-based and collaborative-based filtering techniques. The discussion can focus on how these methods can scale with a growing user base and provide efficient recommendations.
4. **Challenges and Future Improvements:** The document doesn't delve into challenges or potential areas of improvement. You can discuss challenges like cold-start problems for new users and the need for continuous user data analysis to refine recommendations.
5. **Impact on Sales and Revenue:** Effective book recommendations can positively influence sales and revenue for an e-commerce platform. Discuss how a recommendation system can contribute to increased purchases.

CONCLUSION

Recommender systems are a powerful new technology for extracting additional value for a business from its user databases. These systems help users find items they want to buy from a business. Recommender systems benefit users by enabling them to find items they like. In turn, they benefit the company by increasing sales. Recommender systems are quickly becoming a critical element in online E-commerce. The vast volume of user data in existing corporate databases puts recommender systems under strain, and the increasing volume of user data available on the Web will put them even more under strain. New technologies are needed that can dramatically improve the scalability of recommender systems.

The results have shown that item-based techniques hold the promise of allowing CF-based algorithms to scale to large data sets and at the same time produce high-quality recommendations. A recommendation system connects consumers and products through information services. It will be investigated and developed in the future to provide consumers with a better experience.

BIBLIOGRAPHY

- [1] Huifeng Guo, Ruiming Tang, Yunming Ye, Feng Liu, Yuzhou Zhang, 'A Novel KNN Approach for Session-Based Recommendation', August 2017.

- [2] S.Kumar Jain and C.Rana, "Building a Book Recommender content filtering," WSEAS fTrans. Comput, vol 11, no.2, pp. 27-33, 2012.

- [3] Mathew, P., Kuriakose, B. And Hegde, V. "Book Recommendation System through content based and collaborative filtering method", 2016.

Websites

- [1] <https://research.aimultiple.com/recommendation-system/>

- [2] <https://medium.com>

- [3] <https://heroku.com>