**Mini Project Report on**

**E-Commerce Products Recommendation System**

Mini Project Report on

TITLE

Submitted in partial fulfillment of the requirement for the award

of the degree of

Mini Project Report on

TITLE

Submitted in partial fulfillment of the requirement for the award

of the degree of

**Submitted in partial fulfillment of the require for the award of the degree of**

**BACHELOR OF TECHNOLOGY**

**IN**

**COMPUTER SCIENCE & ENGINEERING**

**Submitted by:**

**Rishabh Nauni 2319402**

**Under the Mentorship of**

**Dr. Amit Gupta**

**Associate Professor**

****

**Department of Computer and Engineering**

**Graphic Era Hill University**

**Dehradun, Uttarakhand**

**January 2025**

****

CANDIDATE’S DECLARATION

I hereby certify that the work which is being presented in the project report entitle “E-Commerce Products Recommendation system” in partial fulfilment of the requirement for the award of the Degree of Bachelor of Technology in Computer Science and Engineering of the Graphic Era Hill University, Dehradun shall be carried out by the under the mentorship of Amit Gupta, Associate Professor, CSE Department, Graphic Era Hill University, Dehradun.

Rishabh Nauni 2319402

**Table of Contents**

Chapter No. Description Page No.

Chapter 1 Introduction 1

Chapter 2 Literature Survey 2-3

Chapter 3 Methodology 4-8

Chapter 4 Result and Discussion 9-10

Chapter 5 Conclusion 11

References 12

**Chapter 1**

Motivation- While thinking about what my mini project should be, I noticed my father shopping online for shoes. As he searched, he kept getting more and more recommendations for similar shoes. This made me realize how common recommendation systems are in apps like YouTube, Amazon, and Spotify. So, I thought, why not build an e-commerce product recommendation system for my project?

Recommendation system is very important or crucial for various aspect like personalize user experience, help customers what they need quickly and efficiently. In this report we are discussing about the E-commerce product recommendation system using different technique like content-based filtering, collaborative filtering and hybrid method

**Introduction**

E-commerce has evolved the retailing world by allowing businesses to reach a global audience and destroy geographical barriers. As Jeff Bezos appropriately said, "E-commerce is the future of retail." There are several reasons why this business model has widely succeeded, with one of them being the personalization of shopping experience. Personalized shopping has emerged as one of the most important features of the digital era. To meet the variable needs and preferences of individual customers, e-commerce websites use recommendation systems to filter through an immense amount of information from across the internet and present relevant product suggestions to users. Advanced algorithms and machine learning techniques can take a recommendation system a long way to improve user experience and generate higher sales. This research explores different recommendation techniques; these include content-based filtering, collaborative filtering, and hybrid approaches. Content-based filtering will go through product attributes and preferences of a user to make recommendations similar to those selected and appreciated by the user before. On the other hand, collaborative filtering depends on user behavior data, that is ratings and purchase history information, to determine similar users and develop personalized recommendations. Hybrid approaches combine strengths between both techniques to offer more accurate and diverse recommendations. A deeper understanding of these practices will help beget meaningful recommendation systems that track the changing needs of the e-commerce consumers.

**Chapter 2**

**Literature Survey**

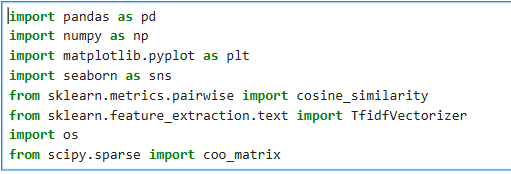
Electronic commerce has transformed the way consumers shop online, and recommendation systems have been incorporated into this process. Users can receive relevant product recommendations from these systems to ensure enjoyable experiences. This field has seen a great deal of research and development over the past few years, leading to the potential for diverse techniques and approaches.' Content-Based Filtering. Filtering based on content Content-based filtering works by suggesting items similar to those that the user has previously acted with. Rather than solely relying on product attributes like genre, keywords, or descriptions, it suggests items that share similar traits. The approach may work well when it comes to personalized recommendations, but for new users with limited interaction history, the selection of items is restricted. However, there are downsides. Collaborative Filtering. Collective filtering represents recommendations based on the preference of other users just like their own. It commonly involves examining user-item interaction data, such as ratings or purchase records to identify users with similar preferences. This approach may be effective in capturing user preferences, but it can cause problems such as low ratings and difficulty with recommendation to new users or items with limited ratings. Hybrid Approaches. Combining the strengths of content-based and collaborative filtering techniques, hybrid recommendation models provide recommendations that are more diverse and precise. This can help overcome limitations in other methods, such as the cold-start problem and limited item pool.[A]. A few frequently employed hybrid techniques include: By combining the predictions of content-based filtering models and collaborative filters, a weighted hybrid approach is established that assigns performance ratings to each model. The use of feature-augmented collaborative filtering adds to the effectiveness of recommendations, particularly for cold-start items, by adding feature attributes to collaborative model filters. By taking into account contextual information, such as the user's location, time, and device, this approach can provide context-based recommendations. The goal is to make these recommendations more relevant. Challenges and Future Directions. The progress made in recommendation systems has not eliminated several significant issues: The scarcity of user-item interaction data results in a significant reduction in collaborative filtering functions' data efficiency. The Cold-Start Problem: It's not easy to recommend items to new users or those with limited ratings. Serenity: It can be a challenge to merge personalized recommendations with unexpected revelations. Concerns About Privacy: User privacy is of utmost importance when collecting and processing personal information. The next steps in research concerning recommendation systems are as follows: Through reinforcement learning, contextual Bandits can dynamically adapt recommendations to user behavior and feedback. Deep learning techniques: Using models of deep learning to extract complex patterns from large datasets. Social Influence is the utilization of social network information to influence recommendations. Creating transparent and trusting AI that can support recommendations. By exploring new approaches that can enhance the effectiveness of recommendation systems and improve online shopping, these issues may be addressed. Additionally, there are some suggestions for improving the experience of shoppers.

**Chapter 3**

**Methodology**

**1. Load Package and Libraries-**

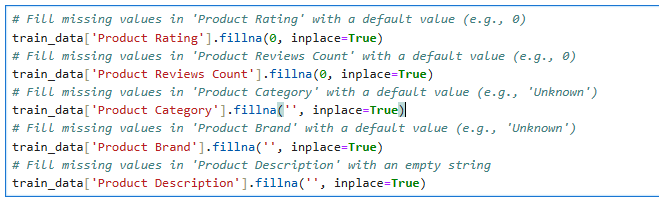
In this mini project, we use many libraries:  
**Pandas:** Applying data manipulation and performing operations on the data.  
**NumPy:** Applying array operations.  
**Matplotlib and Seaborn:** These libraries are used to visualize data present in the TSV file with the help of graphs.



**2. Data Loading and Preprocessing -:**

First, we should load the dataset. In this project, I am using a TSV file named train\_data = pd.read\_csv('marketing\_sample\_for\_walmart\_com-walmart\_com\_product\_review\_\_20200701\_20201231\_\_5k\_data.tsv', sep='\t'). This dataset has 5000 products and 32 features.

  
  
We also have to check for missing and duplicate data, remove duplicates, and fill in the missing values.



We can also shorten the current column names for easier use.

**EDA (Exploratory Data Analysis)-**

 is the process that describes the understanding of the dataset, the pattern, , and relationship of data.

Firstly, we can plot a bar graph of the most popular items.

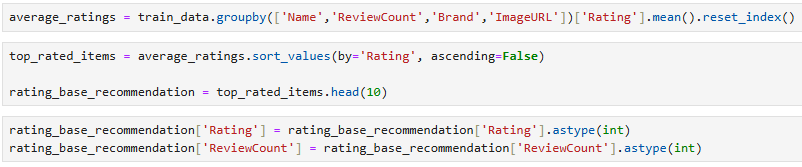


Secondly, we can also plot a graph to show the most frequently used ratings.

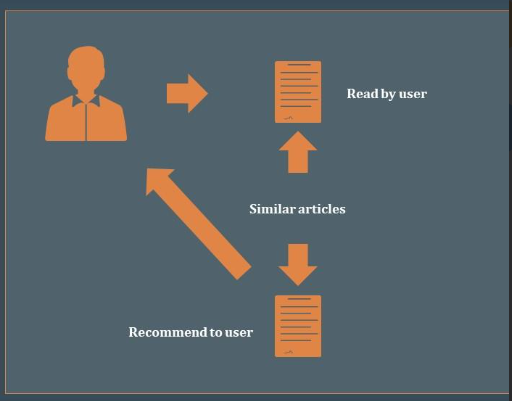


## **Model Development-**

# **Rating Base Recommendations system-** when we are first time opening the e-commerce website we don’t know about the preference of the user, so we just gave the top most rated item to the feed. This can be done by the rating base recommendation system.



# **Content Base Recommendation-**

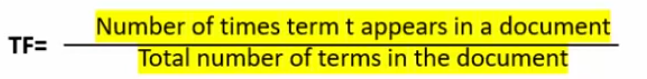


Content -based recommendation systems suggest items to users based on the attributes or features of the items and the user’s previous preferences.

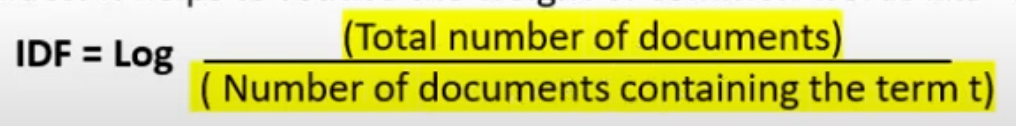
**Similarity Calculation:** Use techniques like cosine similarity or TF-IDF to measure the similarity between items based on their features.

**1. TF-IDF Vectorizer-** It helps in transforming text data into numerical data, which is useful for machine learning algorithms.

**Term Frequency (TF):** Measures how frequently a word appears in a document.



**Inverse Documents Frequency (IDF)-** Measures the importance of a word.



**TF-IDF Score-** A higher TF-IDF score indicate the word is important in that documents but not common across documents.



**2. Cosine Similarity-** is a metric used to measure how similar two vector are, regardless of the magnitude.



I hereby certify that the work which is being presented in the project report entitled “Title of

the project” in partial fulfillment of the requirements for the award of the Degree of

Bachelor of Technology in Computer Science and Engineering of the Graphic Era (Deemed

to be University), Dehradun shall be carried out by the under the mentorship of Mentor

Name, Designation, Department of Computer Science and Engineering, Graphic Era

(Deemed to be University), Dehradun

I hereby certify that the work which is being presented in the project report entitled “Title of

the project” in partial fulfillment of the requirements for the award of the Degree of

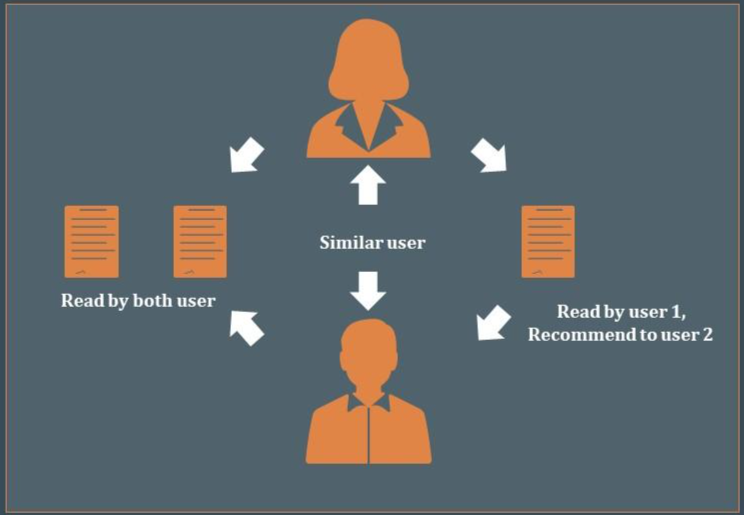
Bachelor of Technology in Computer Science and Engineering of the Graphic Era (Deemed

to be University), Dehradun shall be carried out by the under the mentorship of Mentor

Name, Designation, Department of Computer Science and Engineering, Graphic Era

(Deemed to be University), Dehradun

**Collaborative Base Recommendation-**

****

Collaborative filtering is a technique used in recommendation to make predictions about a user’s preference based on the preferences of many users.

**Steps:**

**1. Compute Similarity Between Users**

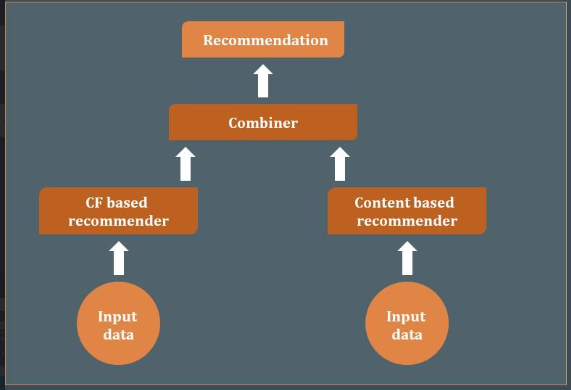
**2. Select Neighbours**

**3.Recommend items**

**Cosine Similarity-** is a metric used to measure how similar two vector are, regardless of the magnitude.



**Hybrid Base Recommendation-**



Hybrid recommendation systems combine two or more recommendation techniques to leverage the strength of each and reduce their weaknesses.

**Chapter 4**

**Result and Discussion**

**Content Base Recommendation-**

**Advantages-**

**Accuracy-** In this filtering accuracy is very high because it chooses the item from past experiences.

**Relevance-** This algorithm is used able to suggest the products which is relevant.

**Disadvantages-**

It cannot be scalable for large numbers of products.Quality may be sensitive in this filtering.

**Collaborative Base Recommendation-**

**Advantages-**

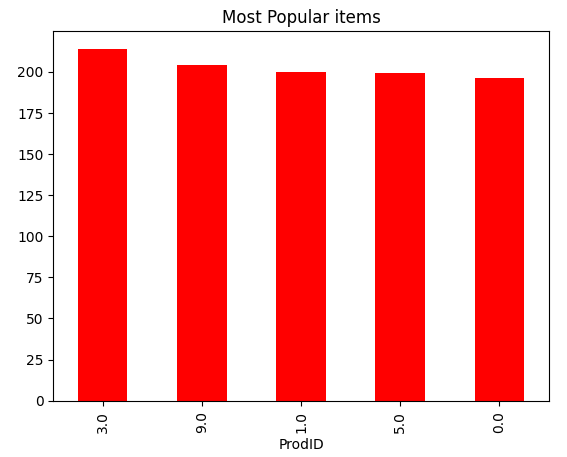
It helps to find the products where attributes are not defined well enough.

It provides personalized recommendation based on collective preference of uses.

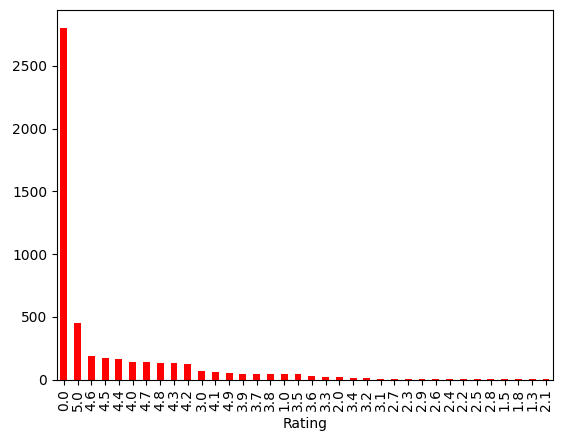
**Disadvantages-**

**Cold start problem**- if the user is new so it highly difficult to find the similarity from the other user.

**Scalability-**As the number of user and items grows complexity is increasing to calculate the similarity.

The data we processed gave us some results. During the EDA, we found the bar graph of the most popular items in the dataset.****

Secondly, we can also plot a graph to show the most frequently used ratings.

****

**Chapter 5**

**Conclusion**

In this report, we detail the design and implementation of an e-commerce recommendation system. Indeed, we believe that such systems can potentially revolutionize the online shopping experience, helping e-commerce vendors realize increased sales and higher levels of customer satisfaction.

Recommendation systems can help customers find product recommendations that match their preferences by personalizing product suggestions, which in turn may increase the engagement and loyalty of customers. Consequently, the significance of a good recommendation system to an ever-expanding e-commerce industry will only grow.

For future research, new strategies could include the implementation of social influence, contextual information, and advanced algorithms in machine learning to further enhance accuracy and relevance in recommendations. Through their continuing emphasis on new technologies, e-commerce companies are able to exploit recommendation systems for successful competitiveness in the rapidly changing digital marketplace.

# **References**

|  |  |
| --- | --- |
| [1] | N. Saeed, “medium,” 2024. [Online]. Available: https://medium.com/@611noorsaeed/building-an-e-commerce-recommendation-system-with-flask-and-machine-learning-506fdfc99ddf. |
| [2] | J. K. J. R. Ben Schafer, “Research Gate,” [Online]. Available: https://www.researchgate.net/publication/2507550\_Recommender\_Systems\_in\_E-Commerce. |
| [3] | “Slide Team,” [Online]. Available: https://www.slideteam.net/recommendations-based-on-machine-learning-powerpoint-presentation-slides.html. |

Hybrid recommendation systems combine two or more recommendation techniques to leverage the strength of each and reduce their weaknessesI hereby certify that the work which is being presented in the project report entitled “Title of

the project” in partial fulfillment of the requirements for the award of the Degree of

Bachelor of Technology in Computer Science and Engineering of the Graphic Era (Deemed

to be University), Dehradun shall be carried out by the under the mentorship of Mentor

Name, Designation, Department of Computer Science and Engineering, Graphic Era

(Deemed to be University), Dehradun

I hereby certify that the work which is being presented in the project report entitled “Title of

the project” in partial fulfillment of the requirements for the award of the Degree of

Bachelor of Technology in Computer Science and Engineering of the Graphic Era (Deemed

to be University), Dehradun shall be carried out by the under the mentorship of Mentor

Name, Designation, Department of Computer Science and Engineering, Graphic Era

(Deemed to be University), Dehradu