Project Synopsis
on

Recommendation System using ML

Submitted as a part of the course curriculum for

Bachelor of Technology
in
Computer Science

Submitted by
UTKARSH MISHRA (2000290120180)
SURYA PRATAP SINGH (2000290120169)
ADRIKA TRIPATHI (2000290120015)

Under the Supervision of
Ms. Akanksha
(Assistant Professor)

KIET Group of Institutions, Ghaziabad
Department of Computer Science
Dr. A.P.J. Abdul Kalam Technical University

2022-2023

DECLARATION

We hereby declare that this submission is our work and that, to the best of our knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the award of any other degree or diploma of the university or other institute of higher learning, except where due acknowledgment has been made in the text.

Signature of Students
Name: Utkarsh Mishra
Roll No.: 2000290120180
Date: 14/11/22

Name: Surya Pratap Singh
Roll No.: 2000290120169
Date: 14/11/22

Name: Adrika Tripathi
Roll No.: 2000290120015
Date: 14/11/22

ACKNOWLEDGEMENT

It gives us a great sense of pleasure to present the synopsis of the B.Tech Mini Project undertaken during B.Tech. Third Year. We owe a special debt of gratitude to Ms Akanksha (Assistant Professor), Department of Computer Science, KIET Group of Institutions, Delhi- NCR, Ghaziabad, for her constant support and guidance throughout the course of our work. Her sincerity, thoroughness and perseverance have been a constant source of inspiration for us. It is only his/her cognizant efforts that our endeavours have seen the light of the day.

We also take the opportunity to acknowledge the contribution of Dr. Ajay Kumar Shrivastava, Head of the Department of Computer Science, KIET Group of Institutions, Delhi- NCR, Ghaziabad, for his full support and assistance during the development of the project. We also do not like to miss the opportunity to acknowledge the contribution of all the faculty members of the department for their kind assistance and cooperation during the development of our project.

Last but not the least, we acknowledge our friends for their contribution to the completion of the project.

Signature
Name: Utkarsh Mishra
Roll No.: 2000290120180
Date: 14/11/22

Signature
Name: Surya Pratap Singh
Roll No.: 2000290120169
Date: 14/11/22

Signature
Name: Adrika Tripathi
Roll No.: 2000290120015
Date: 14/11/22

ABSTRACT

Along with the continuous scale expansion of the e-commerce market, it provides a wide variety of goods as well as choices for the users. The consequent " information overload " is causing information producers to worry about how to make their products stand out, while also worrying information consumers about how to get the information they want. This paper puts forward a product recommendation system based on electronic commerce. It collects e-commerce transaction data, analyses and mines the rich data information, and recommends goods based on user preference and interest. It reduces the time to select interesting goods and improves the efficiency of purchasing for the customers, and improves the success rate of sales to create a greater economic benefit for the enterprises.

TABLE OF CONTENTS

	Page No.
TITLE PAGE	1
DECLARATION	2
CERTIFICATE ….....	3
ACKNOWLEDGEMENT.....	4
ABSTRACT.....	5
CHAPTER 1 INTRODUCTION	7-8
1.1. Introduction ……....	
1.2 Problem Statement.…….....	
1.2. Objective……	
CHAPTER 2 LITERATURE REVIEW…….....	9-14
CHAPTER 3 PROPOSED METHODOLOGY ……....	15-19
3.1 Flowchart	
3.2 Algorithm Proposed	
CHAPTER 4 TECHNOLOGY USED …….…….…….	20
CHAPTER 5 CONCLUSION ….....	21
REFERENCES….....	22

Chapter 1: INTRODUCTION

With the development of information network technology, electronic commerce is a form of business to exchange goods. The main forms of e-commerce are online shopping, online communication and transaction, online electronic payment, and
the accompanying business finance and comprehensive service have gradually become a new mode of operation. The continuous growth of the e-commerce market brings rich and diversified commodities to consumes and increases the
selectivity of consumers in the electronic field.

The recommendation system analyses the retrieval, purchase, and other information left by consumers on e-commerce websites, so as to understand the current needs of consumers. When the user logs in again, the product recommendation will be carried out on the homepage of the website, and the recommendation content will be more diversified, to effectively save consumers' time

Problem Statement

In the e-commerce recommendation system, there is a problem with accuracy as the recommended products are not always relevant. And also, the problem of cold start (for the new user having less data) and the gray sheep problem in which the user choices are not matching with any major group items or users.

Objective

The main objective of our project is to increase the accuracy of the e-commerce recommendation system and will work on the precision of the set of data for finding better results.

To create a machine learning model to recommend relevant books to users based on popularity and user interests.

Aim to predict users' interests and recommend product items that are quite likely interesting for them.

Chapter 2 LITERATURE REVIEW

Movie Recommendation System

In [8], Sharma, N.& Dutta, M proposed that There are 4 recommender system technologies considered: Collaborative Approach (behaviors of the user) Content-based Approach (Information about the movie watched by the user) Demographic Approach (Information about the user i.e. age, gender, location Hybrid Approach (More than one type of recommendation system approach is used) Cold start Problem (new Product launches in the market)Data sparsity (when a user does not rate the product) Scalability (when the system is unable to compute high levels of data) Privacy (your data is not private)

There are many ideas floating around regarding movie recommendation systems like working on emotion-based recommendations, which can be developed to improve the accuracy of movie recommendation systems. Also, with the advancements in neural networks and deep learning, there is hope to include time constraints in the recommendation system.

Personalized Book Recommendation to Young Readers

In [7] , Urdaneta-Ponte, M.C., etal. proposed that, The objectives of this study are twofold: 1) to develop two book recommender prototypes for young readers 2) to evaluate the effectiveness of the recommenders through a user experiment System logs of the 16 participants were extracted for calculating MRR. As a baseline, we consider the recommendation satisfactory when a book selected by a user is ranked among the first half in the list. That is, the baseline has an MRR score of ⅕ = 0.2. The average MRR score across all users for Recom. 1 is 0.283 and that for Recom. 2 is 0.406.

In this study, two book recommender prototypes were developed, and a preliminary user evaluation was conducted to compare them. Results demonstrate that the mean reciprocal rank (MRR) scores of both recommenders were higher than the baseline, and users deemed both recommenders useful, user-friendly, and appealing to
young readers.

Personality – Aware Product Recommendation system

In [9], Chen, Y etal. Proposed that Product recommendation systems are divided into two main classes. Based on the right product to the right user. Collaborating filtering (new product to the user based on his/her previous search history) Content filtering (Similarity with previous buying product)

It incorporates the user's personality traits to predict his topic of interest. It matches the user's personality fact with the associated item.

Recommendations in the Marketplace

In [6] Mehrotra, R., Carterette, B. states that, Multi-sided marketplaces are steadily emerging as viable business models in many applications (e.g., Amazon, Airbnb, YouTube), wherein the platforms have customers not only on the demand side (e.g., users) but also on the supply side (e.g., retailers). Such marketplaces involve interaction between multiple stakeholders among which there are different individuals with assorted needs.

The aim of this research is to understand recommendations in a marketplace so that they can adapt their systems accordingly. In addition to the algorithmic overview, we intend to provide details on various toolkits and existing datasets that could be leveraged to perform experiments.

Improving the prediction accuracy using extreme gradient boosting and encoding approaches.

In [1] , Shahbazi, Z., etal. States that, The Lack of accuracy in the recommendation of a product. Recommendation systems have immensely been utilized by several domains including e-commerce to recommend products. (Collaborative Filtering, Data Mining, time-series analysis, statistical analysis, product analysis, user interest analysis, SVR, and Linear Regressor)

The Collaborative Filtering approach assumes that two like-minded people are most likely to exhibit a similar likeness pattern. However, if classification accuracy is good then prediction accuracy declines and vice versa.
XG booster-based collaborative filtering product recommendation collectively used the user profile and click information to improve the accuracy of the recommendation system.

Aware Product Recommendation system based on user Interest mining and Meta path Discovery

In [2], Dhelim, S.,Ning, H., Aung, N., Huang, R., & Ma,J. States that The previous legacy recommendation system relies only on the user's previous behavior. So, problems are recommendation redundancy and predictability concerning new users and items.
Product recommendation systems are divided into two main classes based on the right product for the right user.
Collaborative Filtering
Content filtering or content-based filtering
Product recommendation system based on interest mining and Meta path discovery and the system predicts the user's needs and the associated items. Incorporates the user's personality facts with the associated items.
Automatic personality recognition system which can detect the user personality traits.

Identifying Grey Sheep Users By The Distribution of User Similarities In Collaborative Filtering

In [3], Zheng, Y., Aganani, M., & Singh,M., proposed that, The problem of the “Grey Sheep" user, is still under investigation. “Gray Sheep" users are a group of users who have special tastes and they may neither agree nor disagree with the majority of the users. The identification of them becomes a challenge in collaborative filtering, since they may introduce difficulties to produce accurate collaborative recommendations.

Recommender system is one of the information systems which assist user’s decision making by recommending a list of appropriate items to the end users tailored to their preferences. It has been successfully applied to a number of applications, such as e-commerce (e.g., Amazon, eBay), online streaming (e.g., Netflix, Pandora), social networks (e.g., Facebook, Twitter), tourism and restaurant (e.g., Yelp), etc.

Approach to identify Grey Sheep users based on the distribution of user similarities or correlations. The proposed approach in this paper is much easier than the previous methods in terms of the complexity. In our future work, we will compare the proposed approach with existing ones to explore whether our approach is significantly better. Furthermore, the same approach can also be used to identify Grey Sheep items in addition to the Grey Sheep users.

Hybrid Popularity Model for solving Cold-Start Problem in Recommendation System

In [4] , Ifada, N., Ummamah, & Kautsar, M., states that Recommender systems are a sort of information filtering technology that aims to offer information items that are likely to be of interest to the user. The cold start problem occurs when the system is unable to form any relation between users and items for which it has insufficient data.

Cold-start and scalability are common issues in recommendation systems. Cold-start happens when the system has no record or rating history of the target user . Therefore, predicting user interest is puzzling since no personalized data can be used as a reference. Meanwhile, the scalability of a recommendation system is challenged when the complexity gradually increases due to the size of data.

The results indicate that the employment of the user popularity model is always more beneficial, for solving the cold start problem in recommendation system than the item popularity model.
To advance the findings in this study, we plan to test our proposed H- Pop model on a non-cold-start recommendation system as well as on other datasets. Additionally, it is also worthwhile to study the impact of scalability problem solving by implementing other clustering techniques.

Improved Tourism Recommendation System

In [5], Pu, Z., Du, H., Yu,S., & Feng, D., states that Users are unable to find the right place and they are missing some historical visits and there are several problem related with the availability of things in that particular area.
In recent years, with the rapid development of the Internet, Internet of Things, Cloud Computing and

Artificial Intelligence, the value of information recommendation through the Internet has been increasing, which has brought a lot of convenience to people's work and life.
In terms of tourism, users can search for travel information through the Internet, judge the purchase of tourism products and services based on their own hobbies, choose their preferred tourist cities or even travel routes, and enjoy the convenience brought by information technology.
This model made two innovations on the original model. Firstly, we updated the similarity calculation formula in the traditional sense. The weights were introduced, and the formula was successfully developed as a weighted similarity calculation formula. Secondly, for an individual user, this article improved the calculation formula for product scores, thus improving the recommendation results.

Item Recommendation Using Hybrid Method

In [6], Cheng, T. states that Recommender System provides various choices of the user preferences for suggesting the product/service to purchase. Collaborative filtering is one of the techniques in the Recommender system used to find reviews and ratings of the users for similar products or users.
Recommender System's work is based on user's information from various sources and provides recommendations of good products. The ultimate goal is to bring out suggestions to users to make better decisions from many alternatives available over the web.
Using of ratings gives clarity to the user for the products. The system focuses on recommendation of items /products to the users. Proposed systems for item recommendation using Hybrid method have competence to attract new customers and retain existing ones.

Chapter 3: METHODOLOGY

Objective: - The main objective of our project is to increase the accuracy of the e-commerce recommendation system and also will work on the precision of the set of data for finding better results.

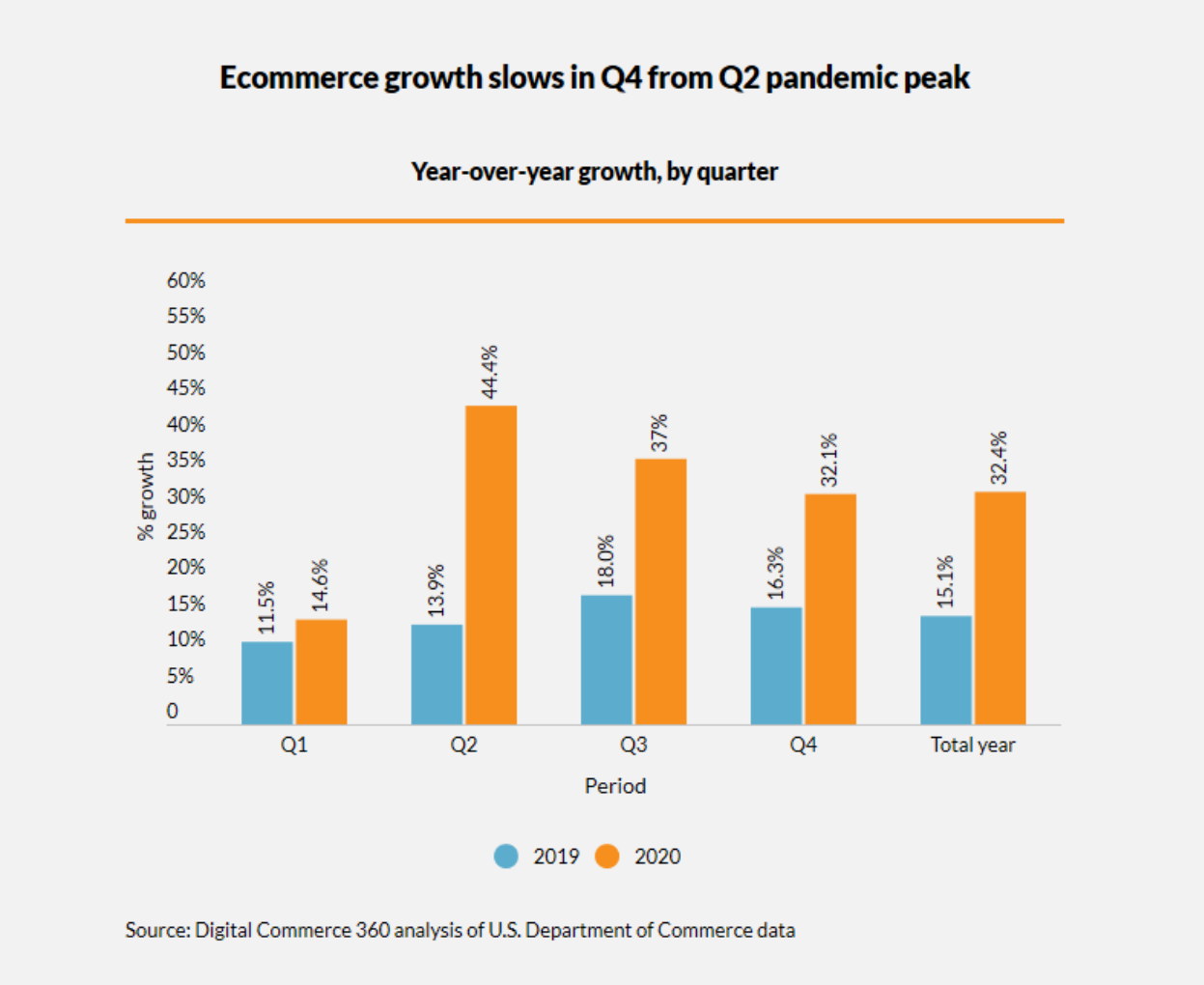

fig3.1 Bar graph of *US e-commerce sales in the period 2019-2020*

.In the above bar graph we can clearly see that the sales rate increases from 15.1% to 32.4% over the year 2019 to 2020 using the recommendation system.

Algorithm used:

Collaborative Filtering

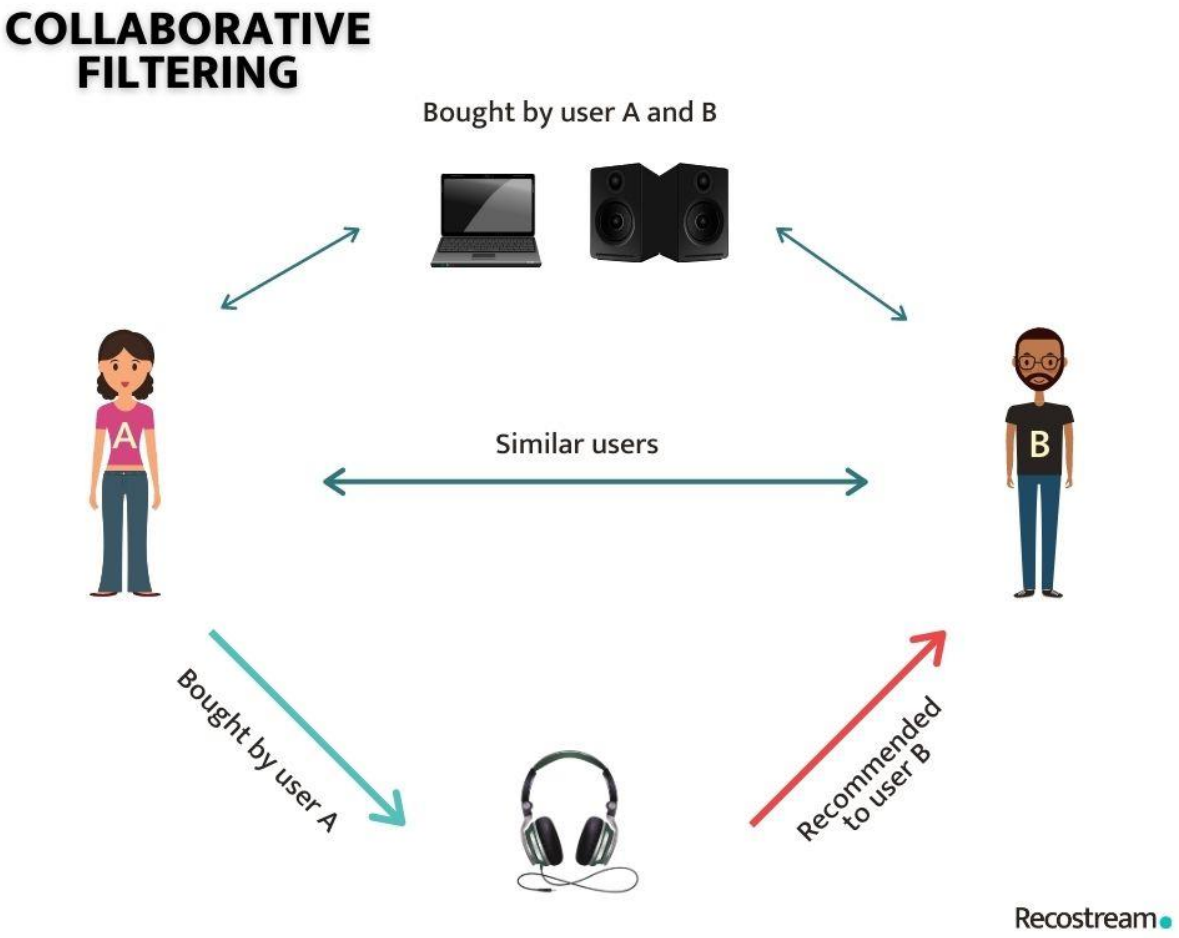

Fig 3.2 collaborative filtering

Content Based filtering:

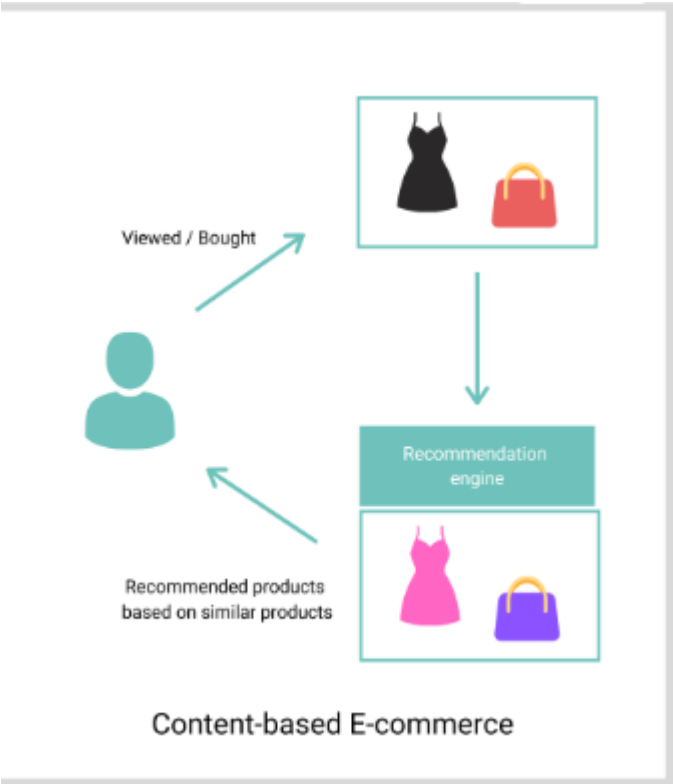

Fig 3.3 content-based E-commerce

- Demographic filtering:

  For tackling the problem like cold start (new user with less data) we will use the demographic filtering algorithm

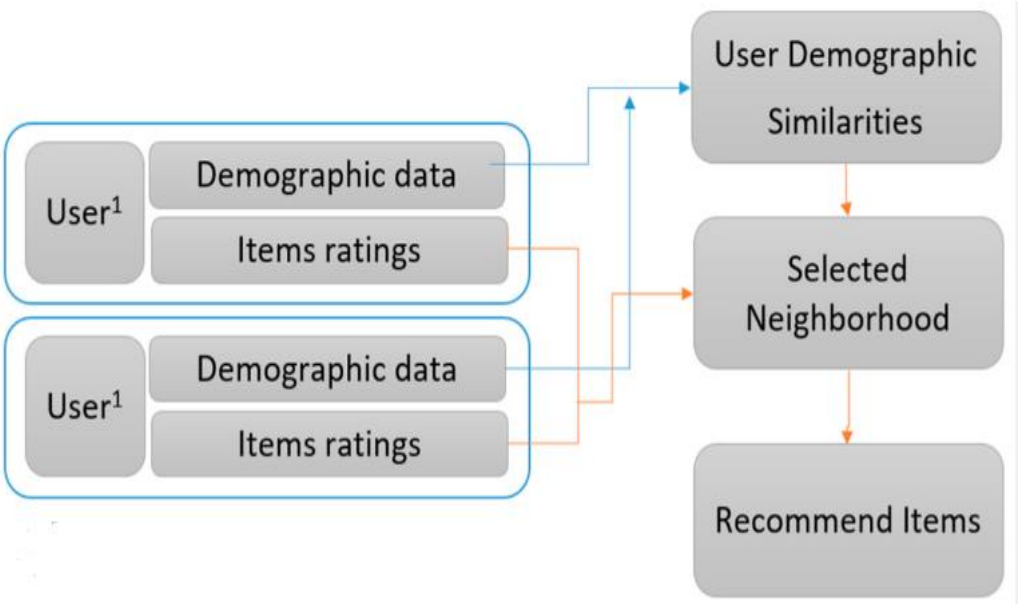

Fig 3.4 Demographic data is based on Location, Age and Gender.

KNN (K- nearest neighbor):

For handling gray sheep user problem (user who neither agree or disagree with majority group)

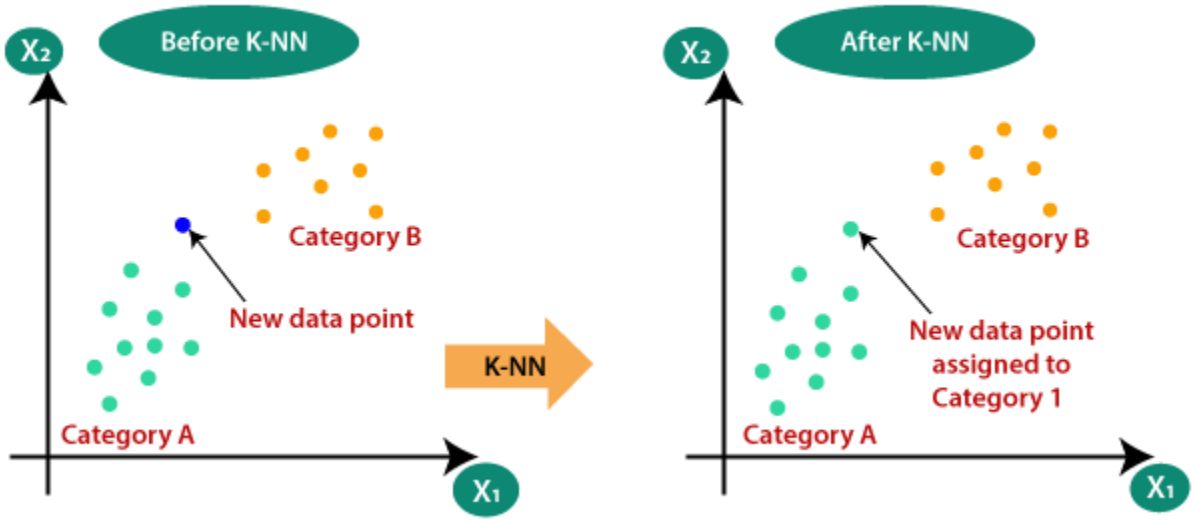

Fig 3.5 KNN (K- nearest neighbor)

Flow chart of working model :-

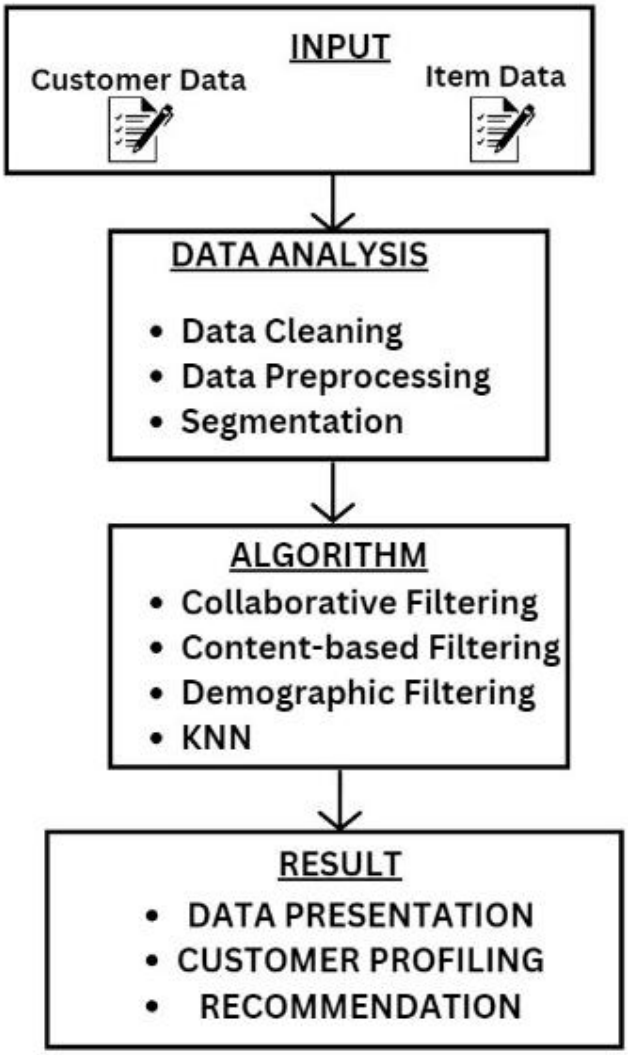

Fig 3.6 Flow chart

Chapter 4: TECHNOLOGY USED

Collaborative Filtering

We use Collaborative Filtering to understand what users may like based on the articles and information two users like in common. This provides a basis for stronger connections with every piece of content you share with the user and increases the chances of engagement. In the e-commerce world, collaborative filtering can be used to make the following recommendations.
You might also like people who bought this also bought would you like to add ____ to your cart? People who viewed this product also viewed Often bought together And a whole lot more.

Semantic Analysis

Semantic analysis is based on text consumption and sharing patterns, our AI core helps you identify what topics a user may be more interested in reading depending on their interactions, likes, shares and comments on a piece of information.

Popularity Analysis

AI and ML engine keeps track of the most popular content in your field and helps incorporate this into recommendations based on your audience's interests.

Communication Platform

An in-built communication platform helps you compose emails, newsletters, SMS, onsite, in-app and push messages to your users using an easy-to-use drag-and-drop template. Save the time and effort you spend designing templates and spend more time understanding your users.

Chapter 5: CONCLUSION

By collecting e-commerce user information and data, analyzing and mining e-commerce user data characteristics and shopping records, and according to the analysis results, using appropriate recommendation algorithm to recommend commodities for users, a commodity recommendation system based on e-commerce is realized.

The methods of data collection and preprocessing in e-commerce are studied; clustering and collaborative filtering algorithm are combined in the system to reduce the real-time response time and enhance extensibility, network topology, and system software architecture is designed to implement.

REFERENCES

1. Shahbazi, Z., Hazra, D., Park, S., & Byun, Y. C. (2020). *Toward Improving the Prediction Accuracy of Product Recommendation System Using Extreme Gradient Boosting and Encoding Approaches. Symmetry, 12(9), 1566.*

2. Dhelim, S., Ning, H., Aung, N., Huang, R., & Ma, J. (2021). Personality-Aware Product Recommendation System Based on User Interests Mining and Metapath Discovery. IEEE Transactions on Computational Social Systems, 8(1), 86–98.

3. Zheng, Y., Agnani, M., & Singh, M. (2017). *Identifying Grey Sheep Users By The Distribution of User Similarities In Collaborative Filtering. Proceedings of the 6th Annual Conference on Research in Information Technology - RIIT '17.*

4. Ifada, N., Ummamah, & Kautsar, M. (2020). *Hybrid popularity model for solving cold-start problem in recommendation system. Proceedings of the 5th International Conference on Sustainable Information Engineering and Technology.*

5. Pu, Z., Du, H., Yu, S., & Feng, D. (2020). *Improved Tourism Recommendation System. Proceedings of the 2020 12th International Conference on Machine Learning and Computing.*

6. Mehrotra, R., & Carterette, B. (2019). *Recommendations in a marketplace. Proceedings of the 13th ACM Conference on Recommender Systems - RecSys '19.*

7. Urdaneta-Ponte, M.C.; Mendez-Zorrilla, A.; Oleagordia-Ruiz, I. Recommendation Systems for Education: Systematic Review. Electronics **2021**, 10, 1611.

8. Sharma, N., & Dutta, M. (2020). *Movie Recommendation Systems. Proceedings of the 8th International Conference on Computer and Communication Management*

9. Chen, Y., Chen, B., He, X., Gao, C., Li, Y., Lou, J.-G., & Wang, Y. (2019). *λOpt. Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining.*

10. Chung M B. Active Recommendation Method Based Mobile Network Using Book Management Application [J]. International Journal of Machine Learning and Computing, 2019, 9(2)