Shri Vile Parle Kelavani Mandal's
Dwarkadas J. Sanghvi College of Engineering
(Autonomous College Affiliated to the University of Mumbai)
NAAC Accredited with "A" Grade (CGPA : 3.18)

PROJECT SYNOPSIS

on

E-COMMERCE RECOMMENDATION SYSTEM

SUBMITTED BY

	Name	SAP ID
1.	Aarya Shah	60002190003
2.	Hetankshi Vora	60002190043
3.	Jill Karia	60002190052
4.	Khushi Gandhi	60002190054

Under the guidance of

Dr. Sanjay B. Deshmukh

Assistant Professor

DEPARTMENT OF

ELECTRONICS AND TELECOMMUNICATION ENGINEERING

Academic Year : 2022 – 23

Shri Vile Parle Kelavani Mandal's

Dwarkadas J. Sanghvi College of Engineering

Plot No. U - 15, JVPD Scheme, Bhaktivedanta Swami Rd,

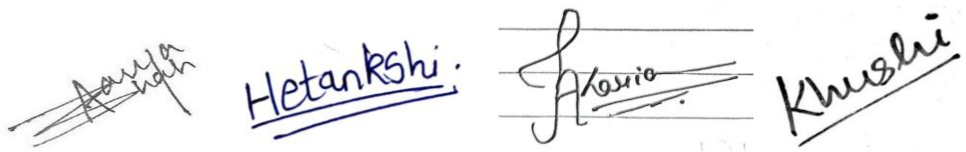
Vile Parle, Mumbai, Maharashtra 400056

TABLE OF CONTENTS

DECLARATION

We hereby declare that this submission is our own 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



Name

Aarya Shah

Hetankshi Vora

Jill Karia

Khushi Gandhi

SAP ID

60002190003

60002190043

60002190052

60002190054

Date

Department of Electronics and Telecommunication Engineering

This is to certify that Project Report Stage – II

"E-Commerce Recommendation System"

Submitted by

1. Aarya Shah (60002190003)
2. Hetankshi Vora (60002190043)
3. Jill Karia (60002190052)
4. Khushi Gandhi (60002190054)

Students of **Electronics and Telecommunication Engineering** have successfully completed their **Project Stage - II** required for the fulfilment of **SEM VIII** as per the norms prescribed by the **University of Mumbai** during the first half of the year 2023. The project synopsis has been assessed and found to be satisfactory.

(Internal Guide) **(External Guide)**

(Head of Department) **(Principal)**

(Internal Examiner) **(External Examiner)**

ACKNOWLEDGEMENTS

We would like to express sincere gratitude to our project guide Dr. Sanjay Deshmukh who has given valuable contributions for our project and technical report. This project would not have been achievable without his excellent assistance and sufficient mentoring, which helped us at every stage of the process. We would like to thank our team members for their assistance and insight during the research.

ABSTRACT

Electronic commerce, sometimes known as e-commerce, is a technique of purchasing and selling goods and services online via electronic platforms such as the Internet. Businesses can boost customer engagement, conversion rates, and satisfaction by offering highly personalized product recommendations and insights into customer behavior. In general, an E-commerce Recommendation System is a highly technologically sophisticated tool essential to the success of online shopping. The system often uses consumer behavior data such as browsing history, search history, purchase history, demographic data, location, and time of day to identify user preferences. The proposed approach examines product attributes like keywords, descriptions, or tags using a content-based filtering method to suggest products similar to the ones in which the consumer has previously shown interest. They provide personalized recommendations to improve user satisfaction. Overall, an E-commerce Recommendation System is a software application that makes product recommendations for users with the help of algorithms and data mining approaches. In this project, we tested algorithms such as VGG16, DenseNet121, and ResNet50 which are deep learning convolutional neural network (CNN) architectures that train on large image datasets.

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviation	Full Form
BPR	Bayesian Personalized Ranking
CBF	Content-based Filtering
CF	Collaborative Filtering
CNN	Convolutional Neural Network
DBF	Demographic-based Filtering
DenseNet	Dense Convolutional Network
HDFS	Hadoop Distributed File System
IDF	Inverse Document Frequency
IDF-W2V	Inverse Document Frequency - Word2Vec
IRS	Intelligent Recommendation System
KBF	Knowledge-based Filtering
ResNet	Residual Network
RS	Recommendation System/Recommender System
RSME	Root Mean Square Error
SVM	Support Vector Machine
TF-IDF	Term Frequency - Inverse Document Frequency
VGG16	Visual Geometry Group-16

CHAPTER 1

INTRODUCTION

Recommendation Systems (RS) is a relatively new area of study in the field of information systems. The objective of E-Commerce is to present related products based on previous search history. Collaborative filtering, content-based recommendations, and hybrid recommendations are the three major categories for recommendation systems. The system may make recommendations for products that are comparable to and complementary to those in which the user has already expressed an interest by using content-based filtering approaches, which involve examining product properties like keywords, tags, or product descriptions. As an alternative, the system may examine the behaviour of other users with comparable likes and preferences in order to produce recommendations with the help of collaborative filtering strategies. In order to make recommendations that are more precise and pertinent, the system may additionally integrate collaborative filtering with content-based filtering. Creating a user profile that incorporates explicit and implicit feedback, such as ratings, reviews, and purchase history, is a typical step in the hybrid approach. From this profile, suggestions are then created based on the user's preferences and behavior.

In this thesis we offer a content-based recommendation system for fashion items that, when fed with a user's request for a specific kind of apparel or accessories, recommends similar fashion products. Before the system gives a recommendation, the model performs image processing on the accepted image to check if a fashion item features on it or not. The main objective of this work is to develop a fashion ecommerce recommendation system that offers solutions to inquiries regarding apparel shopping. Similar products results can be retrieved from various E-Commerce websites using the related search tags. Three algorithms, VGG16, DenseNet121, and ResNet50, were tested. Deep learning convolutional neural network (CNN) architectures VGG16, DenseNet121, and ResNet50 all train on large image datasets. A CNN architecture VGG16, consists of 16 convolutional layers with three fully connected layers. Its architecture is simple and has the ability to extract meaningful features from images. DenseNet121, a variant of the DenseNet architecture, is a CNN architecture based on densely connecting layers. Each layer in DenseNet121 connects to every other layer in a feed-forward manner. This improves feature reuse and makes the use of parameters more effective, which in turn improves the speed of the model with smaller datasets. ResNet50 is based on the idea of residual learning. The ResNet50 network consists of residual blocks which allow information to flow directly from one layer to another without any modification. This approach has achieved state-of-the-art performance on various image recognition tasks as it helps to avoid the vanishing gradient problem which occurs in deep neural networks. Overall, these CNN architectures have shown excellent performance on image recognition tasks and are also widely used in various fields like computer vision, medical imaging, and natural language processing.

An E-Commerce recommendation system may benefit businesses by providing tailored and relevant recommendations that cater to the user's preferences and behaviours. It can increase consumer engagement, improve conversion rates, and enhance customer happiness. To achieve this, we focus on a product (basically a query) and based on its context, propose further related items. The trade of goods and services over online platforms like the Internet is known as electronic commerce, or E-Commerce. It is essential for the daily operations of modern customers and businesses alike. According to estimates, these user-driven recommendation systems account for about 35% of the revenue that E-Commerce platforms generate. Overall, an E-Commerce Recommendation System is a highly sophisticated and technologically advanced tool vital for online retail success. Businesses may boost customer engagement, conversion rates, and customer satisfaction by delivering highly personalised product recommendations and insights into client behavior.

UNDERSTANDING THE ARCHITECTURE OF E-COMMERCE RECOMMENDATION SYSTEM

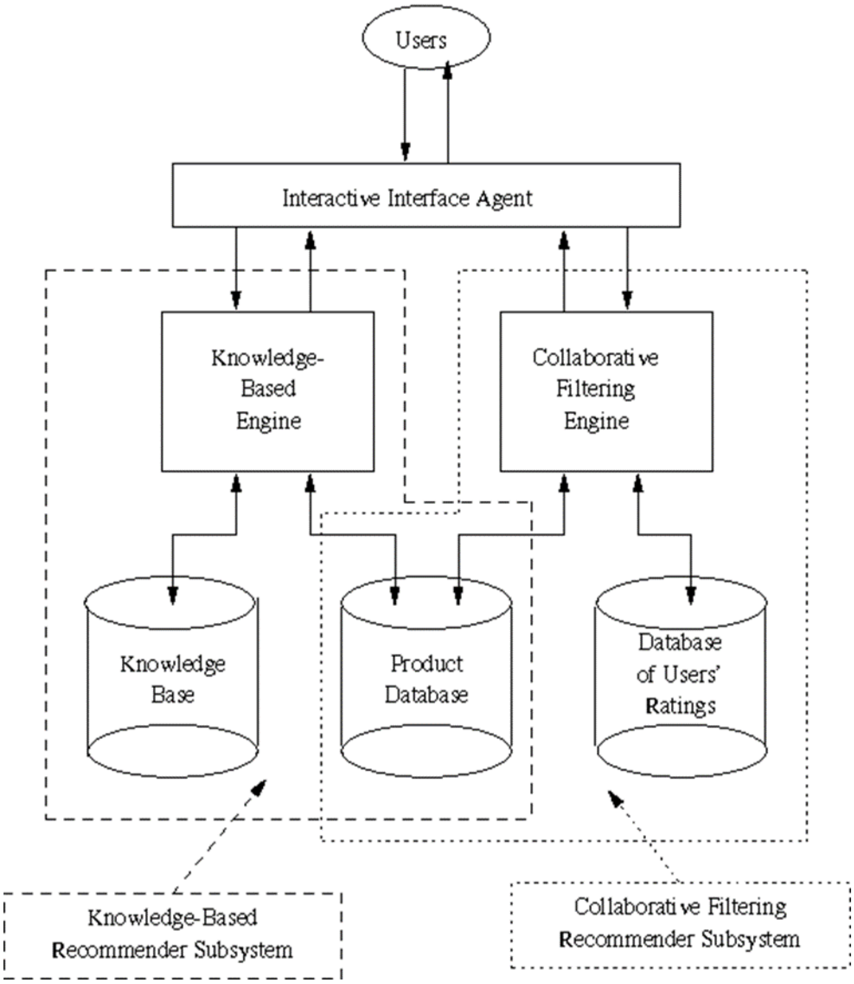


Fig 1.1: Architecture of Recommendation system [11]

On the basis of comprehending and learning about customers' wants and preferences, e-commerce recommendation systems offer users product information and suggestions, suggest products that may be of interest to users, and assist users in completing the purchase process. E-commerce recommendation systems use collaborative filtering, content filtering, knowledge discovery, interactive, and other recommendation technologies based on the unique characteristics and needs of users to suggest information to the appropriate users in the right scene, at the right time, and through the right channel, resulting in a customised shopping experience for customers. Today, e-commerce is an inseparable part of the business for a variety of reasons, including the ease of use, universal accessibility, wide variety, and manageable compassion of products from different vendors, trusted payment ways, and the convenience of home shopping with the least waste of time. E-commerce websites, such as eBay, Alibaba, Amazon, and Netflix, provide great opportunities for their users. These e-commerce platforms tend to utilize the social network features, expert cloud technology, big data, and using mobile devices to provide user satisfaction for their users.

The Recommendation Systems (RSs) gather data and perspectives from the client actions before recommending the goods or services that they believe are the best appropriate among the available choices. They also enable the capabilities required to better customised programmes for each user. Recommendation systems have been utilised in a wide range of industries, including marketing, business, e-learning, music, food and nutritional information systems, social networks for health, and e-commerce. They automate personalization in the context of e-commerce using both traditional and modern techniques, such machine learning. By supplying each customer with the appropriate products, recommending new goods and products to encourage cross-selling, and enhancing customer loyalty, recommender systems in e-commerce increase turnover. One of the aspects that RSs offer to e-commerce platforms is the huge significance of Internet marketing activities, especially e-commerce that leads to personalised recommendation strategies. Compared with the traditional search engine, personalized recommendation can meet the needs of people with diverse backgrounds, purposes, and interests in different periods. Moreover, this recommendation mechanism has changed the original traditional mode of "information seeking", created a new mode of "information seeking" and active push, and provided users with a new way of e-commerce website experience.

Thus, the recommender system is presented as an intelligent system, which identifies the user category having a basis on the user information and then user interest analysis. Once such information is obtained, in the second stage, the analysis is performed to obtain the similarity group respective to necessity products and services. When we search for a product and choose it, we are taken to the product page, which is mostly made up of information like the brand name, product image, title, price, and description. We have proposed a content-based strategy employing text semantics and visual feature-based product similarity in this paper using the product image and its description. We have obtained the data by acquiring approximately 44k samples of apparel from Kaggle.

TYPES OF RECOMMENDATION SYSTEMS:

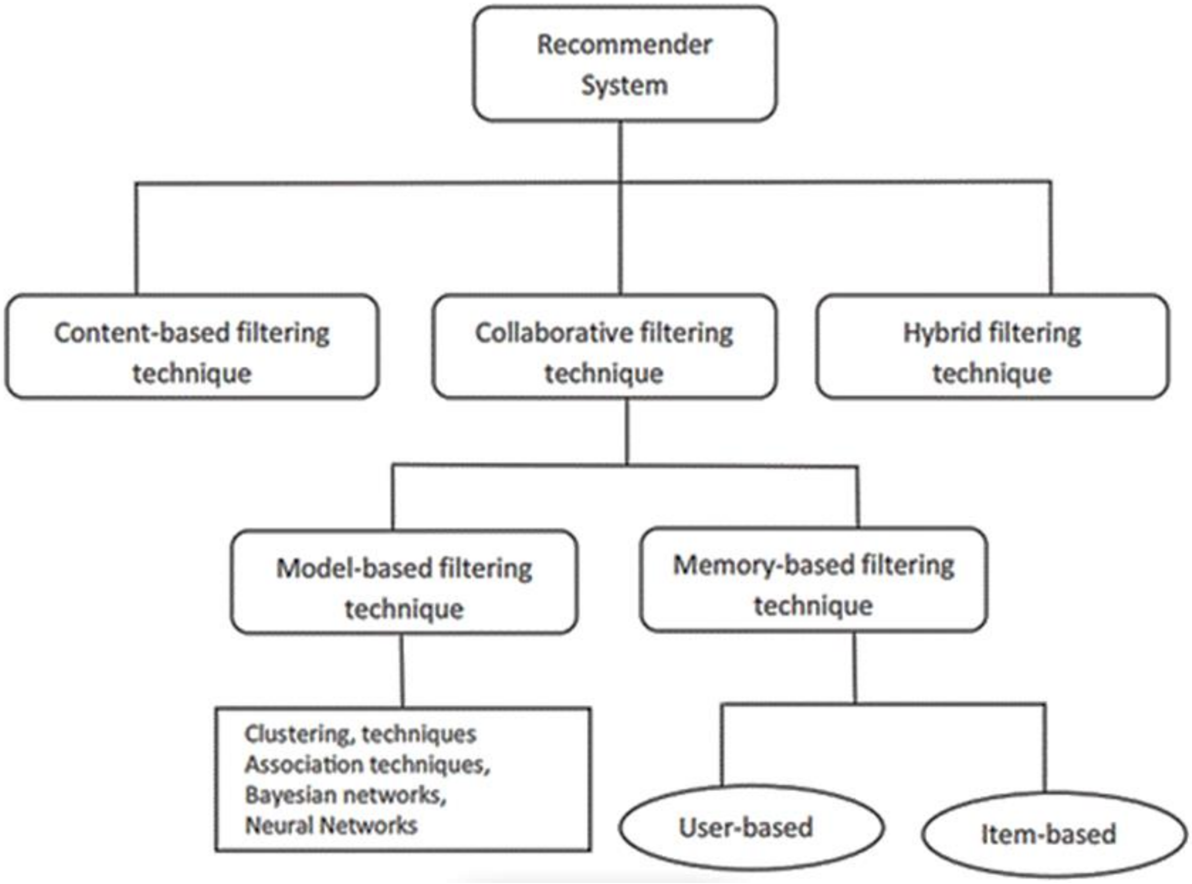


Fig 1.2: Types of Recommendation System

Recommender systems are a potent information-filtering tool that may be used to either anticipate whether a certain user would like a specific item or to identify a set N of items that will be of interest to a specific user. Three suggestion systems—Collaborative-based Filtering, Content-based Filtering, and Hybrid Filtering System—are discussed in this section. The collaborative filtering method is based on collecting and analysing user behaviour data. There are two types of collaborative filtering approaches used: user-to-user collaborative filtering and item-to-item collaborative filtering. Content-based filtering methods rely on a product description and a profile of the user's preferred choices. To offer a greater choice of products to clients, hybrid recommendation systems use both content-based and collaborative filtering at the same time.

Collaborative-based Filtering System

The most popular and effective recommendation technology in RSs is collaborative filtering. It is a technique used for connecting a customer's data to the data of other customers who share similar purchasing habits to provide the user with directions for prospective shopping. Amazon uses CF techniques for making its recommendations depending on a client's previous purchases and the purchase patterns of those that bought the same products. This technique is used by E-commerce platforms to create market segmentation depending on client behavior compared to the psychographics and demographics metrics. A database of user preferences is processed optimally for the preparation of recommendations is the primary performance of this method. The two different critical issues associated with the CF technique are scalability and accuracy. The problems of scalability and efficiency on a global scale require modern processing space and speed optimization to support customer satisfaction. Thus, CF is the most popular and successful technique that utilizes client ratings, details, and reviews gathered from the total clients to construct recommendations.

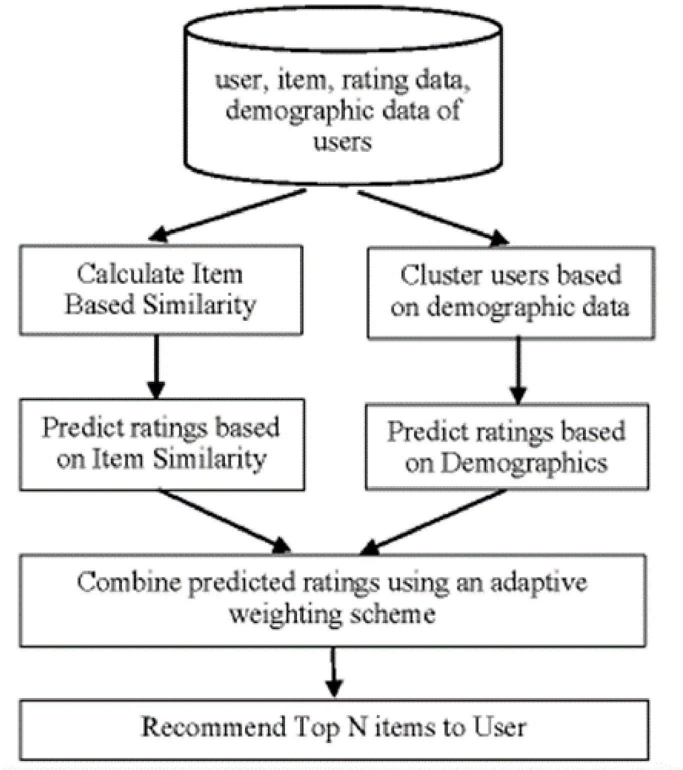


Fig 1.3: Collaborative Filtering Process [12]

There have been two principal modes of implementing collaborative learning: a) memory-based approach and b) model-based approach. The memory-based approach computes the resemblance among the customers by computing the similarity function like the cosine formula. The model-based approach uses some sophisticated methods like machine learning techniques to and patterns in the dataset and learn from them to employ the new data and also some approaches like matrix factorization.

Content-based Filtering System

Content-based recommendation systems suggest an item to a user based on a representation of the things and items common to those that users have previously bought or evaluated. The customer will receive recommendations that are similar to those they previously liked. These algorithms are used to suggest online sites, TV shows, news stories, advertising, etc. All content-based recommender systems share a few common elements, such as item descriptions, user profiles, and methods for comparing profiles to products to determine which recommendation is most appropriate for a given user. A user's preferences and personal information are contained in their profile. Taste is influenced by how the products are evaluated in the previous log. In order to prevent the new-user dilemma, recommender systems put up a survey as part of the profile-building process to gather basic information about a user. In order to make recommendations, an engine analyses products with positive user ratings to unrated items and looks for similarities. The user will be given the unrated items that are most like the positively ranked ones. Content-based systems base their recommendations on user interest profiles and item factors. Similar to personalised profiles, which indicate the kinds of products a person loves and are generated automatically based on user feedback. The acquired user data is compared to the content characteristics of the items to be examined in order to estimate which ones to recommend. For instance, in the case of movie recommendations, the content-based recommender system looks for similarities among the movies the user has rated highly in the past (specific actors, actress, directors, genres, subject matter, etc.). Finally, recommendations would be made for movies that closely resemble the user's choices. The system has a sizable database called the "Item Profile" that contains the things to be recommended and their features. The recommender system receives some kind of information from the users about their preferences. The system creates a profile of the users using the item information and user preferences.

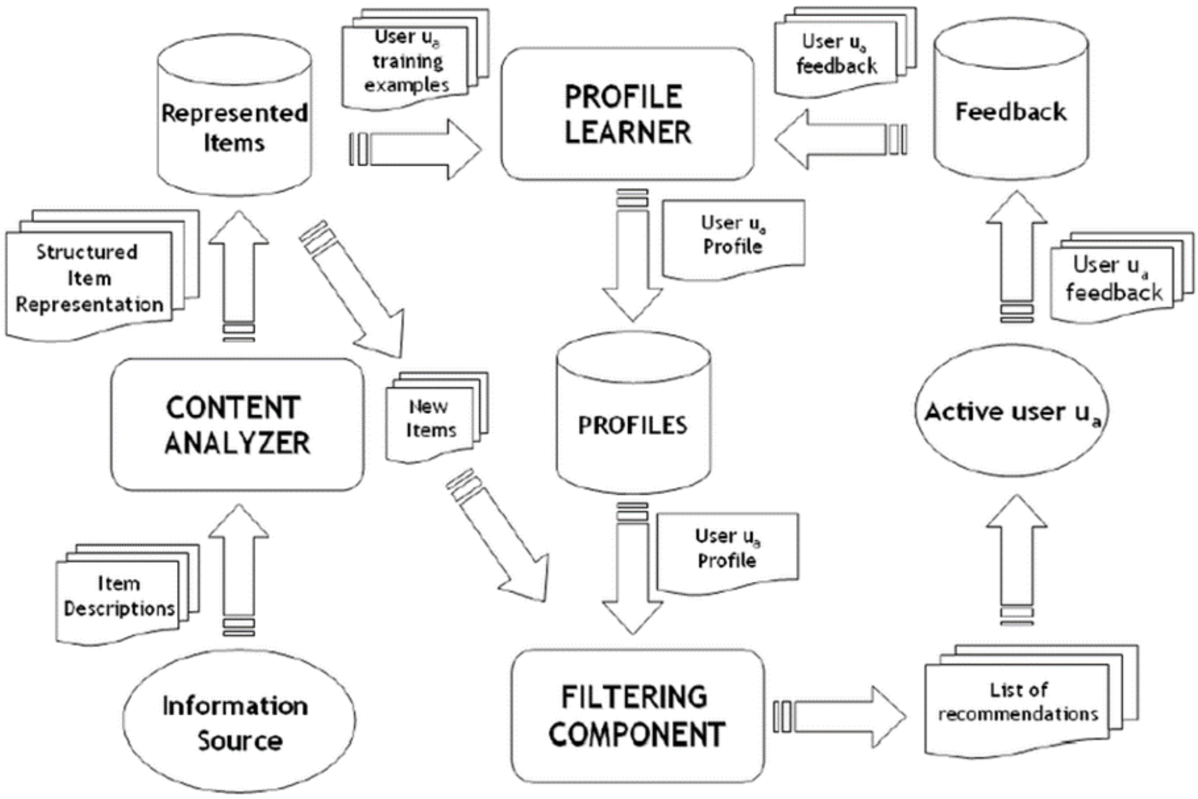


Fig 1.4: Content-based Recommendation System [13]

Hybrid Filtering System

Hybrid Recommendation is a combination of both collaborative filtering and content based approach .With Hybrid Recommendation many types of problems such as Cold-Start problem can be handled using the hybrid recommendations .Different ways of hybridization are implementing CF and CB separately and joining their predictions, Incorporating some content based methods into collaborative filtering, Incorporating some of the collaborative characteristics into content based approach, Constructing a unifying model that possess both content-based and collaborative characteristics. Many combination approaches that are used for building hybrid recommendation systems are as follows:

- Mixed: This method points to the suggestions and recommendations from a set of various recommendation systems that are presented simultaneously.
- Weighted: Produces only single recommendation by using the votes and rates that are created by some recommendation approaches.

- Feature combination: The characteristics that relate to various recommendation data resources are assembled into a single recommendation system algorithm.
- Cascade: One of the recommendation systems corrects the prompts and recommendations that are presented by another recommendation system.
- Feature Augmentation: Output results from one approach are employed as input data and characteristics for another recommendation method.
- Meta level: The approach that is established by one recommendation system is utilized as an input for another approach.
- Switching: In this method, the recommendation system shifts among different recommendation approaches acceding to the current situation.

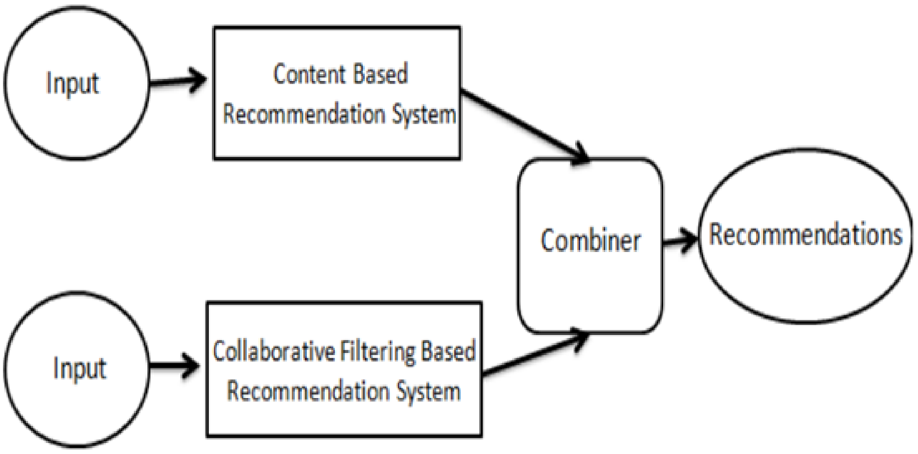


Fig 1.5: Hybrid Filtering System [14]

TABLE I
COMPARISON AMONG THE RECOMMENDATION SYSTEM

Types	Content Based	Collaborative	Hybrid
Advantages	Transparency Quality is improved over time due to user independence.	Easy to create and use. It is simple to add new data. More suitable.	Improves user preference. Prevent cold start problems. No problem with sparsity.
Disadvantages	Limited content analysis. Over-specialization and serendipity.	It is entirely dependent on user ratings. Insufficient Data due to sparity and cold start problem.	Building a network of trust is challenging.

BACKGROUND OF THE PROBLEM

Customers have difficulty finding the things they like in the present E-Commerce recommendation methods. Some of the concerns that have not been adequately taken into account by existing E-Commerce recommendation systems are lack of customer satisfaction, the absence of customised recommendations, the inability to resolve cold start issues, improper management of the confined resource situation, improper handling of the data validity period, and decreased efficiency are only a few of the constraints. Collaborative filtering's key drawback is that it needs certain types of data in order to be effective. Sparsity and scalability are two of its main downsides [1].

One of the biggest challenges in building an effective recommendation system for an E-Commerce platform is the "cold start problem". This refers to the difficulty of making personalized recommendations for new users who haven't yet provided any explicit feedback or interacted with the platform enough to establish a profile. Without any data to work with, it's difficult to make accurate recommendations that are tailored to a user's interests and preferences.

Another challenge is the "sparsity problem". This occurs when there are many items available for purchase on the platform, but any given user has only interacted with a small subset of them. This leads to a sparse matrix of user-item interactions, which can make it difficult to accurately estimate the similarity between users or items. Additionally, the "data quality problem" can arise if the data used to train the recommendation system is incomplete, inaccurate, or biased. This can lead to poor recommendations that don't reflect users' actual preferences. Finally, the "privacy problem" is a growing concern, as users become more aware of the risks associated with sharing personal data. To address this, E-Commerce platforms need to find ways to protect user privacy while still collecting enough data to train effective recommendation models.

Customers often find it challenging to find the products they need quickly and easily, resulting in a poor shopping experience. E-Commerce businesses are aware of this problem and are looking for ways to provide a more personalized and efficient shopping experience to their customers. One of the ways to address this problem is by developing an effective recommendation system that can provide accurate and personalized product recommendations to customers.

However, developing an effective recommendation system poses several challenges, including dealing with data sparsity, handling the cold-start problem, ensuring algorithmic fairness, and protecting user privacy. Therefore, there is a need for an E-Commerce recommendation system that can overcome these challenges to provide customers with personalized and relevant product recommendations while maintaining user privacy and algorithmic fairness. The success of such a system will depend on its ability to provide accurate and diverse recommendations, improve customer engagement, and enhance the overall shopping experience on the E-Commerce platform.

STATEMENT OF PROBLEM

The E-Commerce industry has seen significant growth over the years, with more customers opting to shop online. With an ever-increasing number of products available on E-Commerce platforms, it has become challenging for customers to find the products they need quickly. As a result, there is a growing need for an effective recommendation system that can provide personalized and relevant product recommendations to customers. The primary objective of this project is to design and develop an E-Commerce recommendation system that can help customers find the products they need quickly and easily. The system should leverage machine learning algorithms such as content-based filtering to analyze customer data such as purchase history, browsing history, and search queries to provide accurate and personalized product recommendations. The system should be scalable, efficient, and easy to use for both customers and administrators. The success of this project will be measured by its ability to increase customer engagement, improve sales, and enhance the overall customer experience on the E-Commerce platform.

LITERATURE SURVEY

By virtue of the ever-increasing prominence of mobile internet, an increasing number of businesses in various industries are switching to E-Commerce from conventional business practices. In this context, E-Commerce Recommendation Systems have become indispensable as a means of providing consumers with individualised product information and suggestions based on their interests and preferences. These recommendations can be found in a variety of formats, including friend recommendations on Facebook, queries with similar topics on Quora, and purchase suggestions on E-Commerce websites. The suggestion system facilitates the user's ability to efficiently complete their purchases. Utilising data garnered from the user's previous search activity; these systems automatically generate personalised recommendations for the user. To gain a deeper understanding of the various methodologies used in the development of these recommendation systems, we have read a large number of publications with care.

A Survey of E-Commerce Recommender Systems [2]

In 2016, Farida Karimova analysed more than sixty distinct recommendation system techniques, culminating in the publication of her paper titled "A Survey of E-Commerce Recommender Systems." Recommendation systems have grown in prominence in recent years due to their capacity to provide personalised recommendations and increase productivity. As a result, researchers have conducted literature reviews to scrutinise the most recent advancements in this field and assess the challenges these systems currently face. In the field of recommendation systems, collaborative-based filtering (CBF) and content-based filtering (CBF) are two predominant methods. However, it has been discovered that both methods have limitations. CBF systems are known to experience frigid starts and sparsity issues, whereas CB systems are susceptible to overspecialization. Multiple research projects have identified and discussed these issues. In the sphere of recommendation systems, hybrid recommendation techniques have become increasingly prevalent. These techniques integrate multiple recommendation algorithms to enhance the precision and efficacy of user recommendations. Consequently, hybrid recommendation techniques have garnered considerable interest from researchers and practitioners. To address the limitations of single-recommendation methods, hybrid recommendation strategies have been proposed. These strategies integrate two or more techniques, such as content-based (CB) and collaborative filtering (CF), to enhance system performance and address issues such as the cold-start problem. The researchers investigated the effectiveness of the social network-based recommender approach in their study. This strategy utilises data from social networks, such as user preferences and social buddy influence, to improve the accuracy of recommendations and surmount significant obstacles, such as cold-start and data sparsity issues. It has been discovered that in the domain of E-Commerce recommendation systems, conventional approaches such as Collaborative Filtering (CF) and Hybrid methods continue to reign supreme. Recent studies in the discipline have uncovered this observation as a key takeaway. Scholars have actively sought solutions to the limitations of collaborative filtering (CF) algorithms, according to the available literature. In particular, researchers have investigated various approaches to reduce computational complexity and improve recommendation precision. Collaborative filtering (CF) has been identified in the literature as a technique with a number of essential characteristics. These characteristics include precision, acceptability, the ability to offer a broader selection of products, and the potential to increase the loyalty of existing users. The development of highly accurate personalised recommendation systems (RS) has been the subject of extensive research. Despite the advancements made in this field, there are still significant limitations to consider. The review has identified unresolved concerns regarding the RS, leading to the conclusion that further research is required on this topic.

Research on the Application of Collaborative Filtering Algorithm in Mobile E-Commerce Recommendation System [3]

Xiangpo Li elaborates on the notion that the recommendation system is based on the user's personal information such as features, historical behaviour, and items by utilising collaborative filtering, content filtering, knowledge discovery, interactive and recommended recommendation technology. A recommendation algorithm known as collaborative filtering is currently one of the most widely used recommendation algorithms that may be found in a variety of recommendation algorithms. This research paper evaluates the collaborative filtering algorithm in the context of the application of a mobile electronic commerce recommendation system experiment. The findings led them to arrive at the conclusion that collaborative filtering technology is the most successful technology in the use of personalized recommendation. This technology in particular plays an essential role in increasing the sales revenue of E-Commerce websites.

Recommendation Systems for E-Commerce: A Review [4]

Priya S. and Mansoor Hussain D. assess the term "recommendation system," which refers to an intelligent technology that generates a categorised list of products that a consumer may find appealing. Mansoor Hussain D. and Priya S. examine the term "recommendation system." There are numerous methodologies and algorithms in use for data filtering and recommendation generation. This paper contrasts, elaborates on, and investigates the limits of various techniques by presenting an overview of the challenges faced by recommendation strategies. They reached this conclusion by comparing and contrasting content-based, collaborative-based, and hybrid filtering systems. Based on their findings, they concluded that these recommendation algorithms have provided a vast array of search and filtering techniques. Recommendation systems are becoming an increasingly vital component of online E-Commerce platforms and video streaming websites. Contributing to the improvement are the enhanced modelling of users and items, the addition of contextual information to the process of making recommendations, support for multicriteria ratings, and the availability of a flexible and less disruptive approach to making recommendations. To improve the quality of online commerce predictions in the future, enhanced clustering algorithms and improved prediction development methodologies will be established.

Content Based Apparel Recommendation for E-Commerce Stores [5]

Through the development of a content-based recommendation system for women's apparel, in which the system generates similar products for consumers based on the requested item, we can provide superior customer service. The 2022 publication "Content-Based Apparel Recommendation for E-Commerce Stores" was written by Utpal Chandra De, Shobhan Banerjee, Manas Kumar Rath, Tanmaya Swain, and Tapaswini Samant. In the paper, the authors utilise numerous text-based strategies to retrieve information from the product page using the product image and description. They created the algorithm using approximately 180k samples of women's blouses obtained through the Amazon Product Advertising API. They identified seven main categories for the system's construction. Their apparel suggestion task was performed using a text-based semantic search (Weighted IDF-W2V) and a convolutional neural network (VGG-16). The Bag of Words technique, the TF-IDF method, or just simple IDF can be used to determine whether or not text-based product comparisons in the product title pass muster. Due to the large number of incomplete values, the data had to be cleaned, resulting in the loss of approximately 167 thousand samples. As a result, they concluded that in order to obtain more accurate and superior recommendations based on a particular product that was sought for, one must train the model on a large number of samples.

A Systematic Study on the Recommender Systems in the E-Commerce [6]

Pegah Malekpour Alamdari, Nima Jafari Navimipour, Mehdi Hosseinadeh, Ali Asghar Safaei, and Aso Darwesh conducted a comprehensive literature review in the field of E-Commerce recommender systems. Their survey generated a substantial quantity of data that can inform future research in this field. The authors examine five distinct categories of recommendation system algorithms in their research: knowledge-based filtering (KBF), hybrid filtering, collaborative filtering (CF), demographic-based filtering (DBF), and content-based filtering (CBF). In addition, the key themes of the selected papers are briefly summarised. The specified literature consists of articles published between 2008 through 2019. In their research, the authors investigated the field of recommender systems in the context of E-Commerce, with a particular emphasis on examining the fundamental models of recommendation approaches. In the field of Recommender Systems (RS), the most fundamental models rely on two distinct kinds of input data: user-item interactions and quality information about the objects and users. User-item interactions refer to the feedback provided by users, such as ratings or purchase behaviour, whereas quality information refers to additional information about objects and users, such as associated keywords or textual profiles. These two categories of data are the basis for fundamental RS models. In contrast to content-based filtering (CBF), collaborative filtering (CF) relies on user-item interactions. This dichotomy is widely acknowledged as a fundamental distinction between these two classes of recommendation algorithms in the literature. Utilising the collective behaviour of users, CF methods identify patterns of similarity and make recommendations based on these patterns. CBF approaches, on the other hand, concentrate on the intrinsic properties of products and users to generate recommendations that match their preferences. Despite their distinctions, both CF and CBF methods have been demonstrated to be effective in a variety of recommendation scenarios and have been extensively researched in the literature. Knowledge-Based Framework (KBF) approaches necessitate the explicit articulation of user requirements, constraints, and external knowledge foundations. This information is essential for the formulation and implementation of KBF strategies. Demographic recommendation systems (RSs) employ user demographic information to generate personalized recommendations based on unique mapping demographics for ratings or purchase preferences. As a means of attaining improved performance across diverse environments, hybrid systems that integrate multiple types of recommendation systems have been proposed. These hybrid systems seek to provide users with more effective and accurate recommendations by combining the strengths of various recommendation techniques. In their research, the authors provided a comprehensive analysis of the relevant literature. They provided a comprehensive comparison of the chosen works, highlighting their strengths and limitations. Security, response time, scalability, accuracy, operation cost, novelty, implicit or explicit data source, and independence were evaluated by the authors. The results were presented as tables so that readers could readily compare and contrast the various papers. The authors of the study followed a three-step process when conducting a literature review. They began by formally establishing the research queries. Second, they conducted a search for relevant documents. Finally, they applied specific constraints to select the most relevant papers for review. The researchers presented a comparison chart of metrics and review issues for the selected papers in their study. The findings of a study can provide useful information for future research endeavours.

Algorithm in E-Commerce Recommendation [7]

In a sea of information, the E-Commerce sector faces a challenge in recommending high-quality products to users. Through this paper, Zezhou Fan, Dan Chang and Jinhong Cui propose using support vector machines (SVM) to classify products and obtain both positive and negative feedback for each product. For positive feedback, the exhaustive score of product ratings and reviews is computed, and a collaborative filtering recommendation algorithm is developed using this information to generate a final recommendation list based on the preference score. The proposed algorithm is evaluated on Taobao's online data and demonstrates accurate and rapid recommendation generation. The experiment encompasses eight product categories with approximately 25,000 items, and the hardware configuration used is a ThinkPad E445 with 3.3 GHz and 4GB of memory. The proposed algorithm uses SVM for classification and negative item feedback information to filter out items that the user may not like, thereby enhancing the efficiency and precision of recommendation systems. Indirectly, the comprehensive score derived from the weighted average of positive feedback scores and reviews improves the algorithm's precision. This algorithm has practical value for the E-Commerce industry because it improves the accuracy and efficiency of product recommendations while reducing the number of items that the user may not like. The proposed solution addresses the issue of recommending high-quality products to consumers in the E-Commerce industry effectively.

An Enhanced Recommendation Algorithm Based on Modified User-Based Collaborative Filtering [8]

Since there is so much data available online, recommendation algorithms have gained a lot of traction. Typically, collaborative filtering is used by traditional recommendation algorithms to determine user and item similarity. However, data sparsity and overfitting can negatively impact the accuracy of these systems, resulting in subpar-quality recommendations. Ramil Lumauag, Ariel Sison, and Ruji Medina propose an improved recommendation algorithm based on modified user-based collaborative filtering to improve the quality of recommendations by addressing the aforementioned problems. Using the MovieLens dataset, the algorithm's precision and performance were compared to the traditional algorithm using Root Mean Square Error (RMSE), Precision, and Recall. The results demonstrated that the improved algorithm outperforms the conventional algorithm and enhances the precision of recommendations. The purpose of this paper was to modify the conventional user-based collaborative filtering to address the data sparsity and overfitting problems, and to compare the performance of the modified algorithm to that of the conventional algorithm. Through experimental evaluation, the paper analyses the shortcomings of user-based collaborative filtering in the conventional recommendation algorithm and proposes an improved algorithm to address these problems. The paper suggests verifying the enhanced algorithm's coverage, determining the neighbour’s quality, and comparing the proposed algorithm to other enhanced methods for future research. By modifying the collaborative filtering based on user similarity, the improved algorithm improves prediction accuracy, improves the performance of recommendations, and resolves data sparsity and overfitting issues. This paper presents an efficient method for enhancing the quality of recommendations in the presence of voluminous data.

E-Commerce Intelligent Recommendation System Based on Deep Learning [9]

Gang Huang examines the significance of intelligent recommendation systems (IRS) in E-Commerce and how deep learning algorithms can improve the performance of such systems. He proposes a novel convolutional neural network (CNN)-based E-Commerce IRS architecture and evaluates its performance on the Alibaba dataset.Following a discussion of the CNN-based recommendation algorithm, the paper outlines the functional modules and architecture of the proposed E-Commerce IRS. The proposed algorithm is contrasted with the Item method and the Bayesian personalized ranking (BPR) method, two infamous recommendation algorithms. The experimental results demonstrate that the CNN-based recommendation algorithm outperforms the other two algorithms in terms of prediction accuracy and efficiency and thus has both practical and promotional value.The paper emphasises the significance of developing effective recommendation algorithms that take into account both the co-location of various products on E-Commerce websites and the users' overall preferences. It is argued that deep learning algorithms can help mine user session sequences for interest and behaviour preferences and can considerably enhance the efficiency of algorithm model calculations. However, the accuracy of the proposed algorithm is still inadequate, and the authors suggest that structural modifications are required to enhance its performance.This paper contributes to the increasing body of research on intelligent recommendation systems in E-Commerce and demonstrates the potential of deep learning algorithms to enhance the efficiency and precision of such systems. The proposed architecture and algorithm can be further refined and adapted to various E-Commerce scenarios, and future research can investigate methods to increase the diversity and originality of the recommended results.

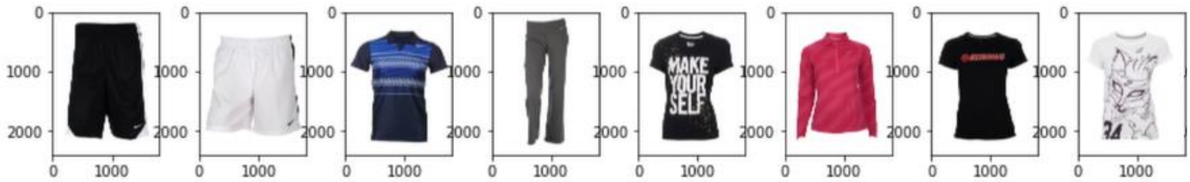
Research on Recommendation Algorithm Based on E-Commerce User Behavior Sequence [10]

A dynamic and promising area of the economy is E-Commerce, which the growth of the internet has enabled. Personalised recommendation systems have been devised to enhance the efficiency and effectiveness of E-Commerce platforms. Bai Li, Allam Maalla, & Mingbiao Liang propose an architecture for a personalized recommendation system based on offline mining, real-time mining, and deep learning. Flume, Kafka, and Spark Streaming are utilised by the architecture to capture and pre-process user behaviour data. Offline mining is conducted with MapReduce on Hadoop and Spark, while real-time mining is performed with clusters of Kafka and Spark. On a Spark distributed cluster with multiple GPUs, the DeepLearning4j framework is used for deep learning to generate personalized product recommendations based on user profiles and preferences. The architecture can increase the conversion rate of E-Commerce products and enhance the quality and efficiency of users' purchasing decisions. Deep learning provides a greater understanding of user requirements, historical interactions, and product characteristics than conventional recommendation systems. The proposed architecture employs Hadoop and Spark technology platforms, such as the HDFS distributed file system, the MapReduce programming paradigm, and the Spark calculation engine. By integrating big data technology and deep learning technology, a personalized E-Commerce recommendation system is created that accurately recommends products that users need, thereby reducing the time it takes for users to locate products and increasing the conversion rate of E-Commerce products. This architecture is a valuable instrument for E-Commerce platforms seeking to stand out in a competitive market, as it benefits both users and merchants.

CHAPTER 2

DATASET

We used the fashion dataset from Kaggle which had more than 44,000 sample high resolution images. These images were sourced from the fashion e-commerce website, myntra.com. These images have been allotted multiple labels and they allow to describe the features of each image. These said features are extracted by the algorithm and then utilized to provide recommendations. The dataset includes images with the main categories such as Apparel, Accessories, Footwear, Personal Care, Free Items and Sporting Goods. They also include subcategories such as Topwear, Bottomwear, Watches, Shoes, Belts, Flip Flops, Bags, Innerwear, Sandal, Shoe Accessories, Fragrance, Jewellery, Lips, Saree, Eyewear, Nails, Scarves, Dress, Loungewear and Nightwear, Wallets, Apparel Set, Headwear, Mufflers, Skin Care, Makeup, Free Gifts, Ties, Accessories, Skin, Beauty Accessories, Water Bottle, Bath and Body, etc.



	id	gender	masterCategory	subCategory	articleType	baseColour	season	year	usage	productDisplayName	image
0	15970	Men	Apparel	Topwear	Shirts	Navy Blue	Fall	2011	Casual	Turtle Check Men Navy Blue Shirt	15970.jpg
1	39386	Men	Apparel	Bottomwear	Jeans	Blue	Summer	2012	Casual	Peter England Men Party Blue Jeans	39386.jpg
2	59263	Women	Accessories	Watches	Watches	Silver	Winter	2016	Casual	Titan Women Silver Watch	59263.jpg
3	21379	Men	Apparel	Bottomwear	Track Pants	Black	Fall	2011	Casual	Manchester United Men Solid Black Track Pants	21379.jpg
4	53759	Men	Apparel	Topwear	Tshirts	Grey	Summer	2012	Casual	Puma Men Grey T-shirt	53759.jpg

Fig 2.1: Sample Images and the categories of their features

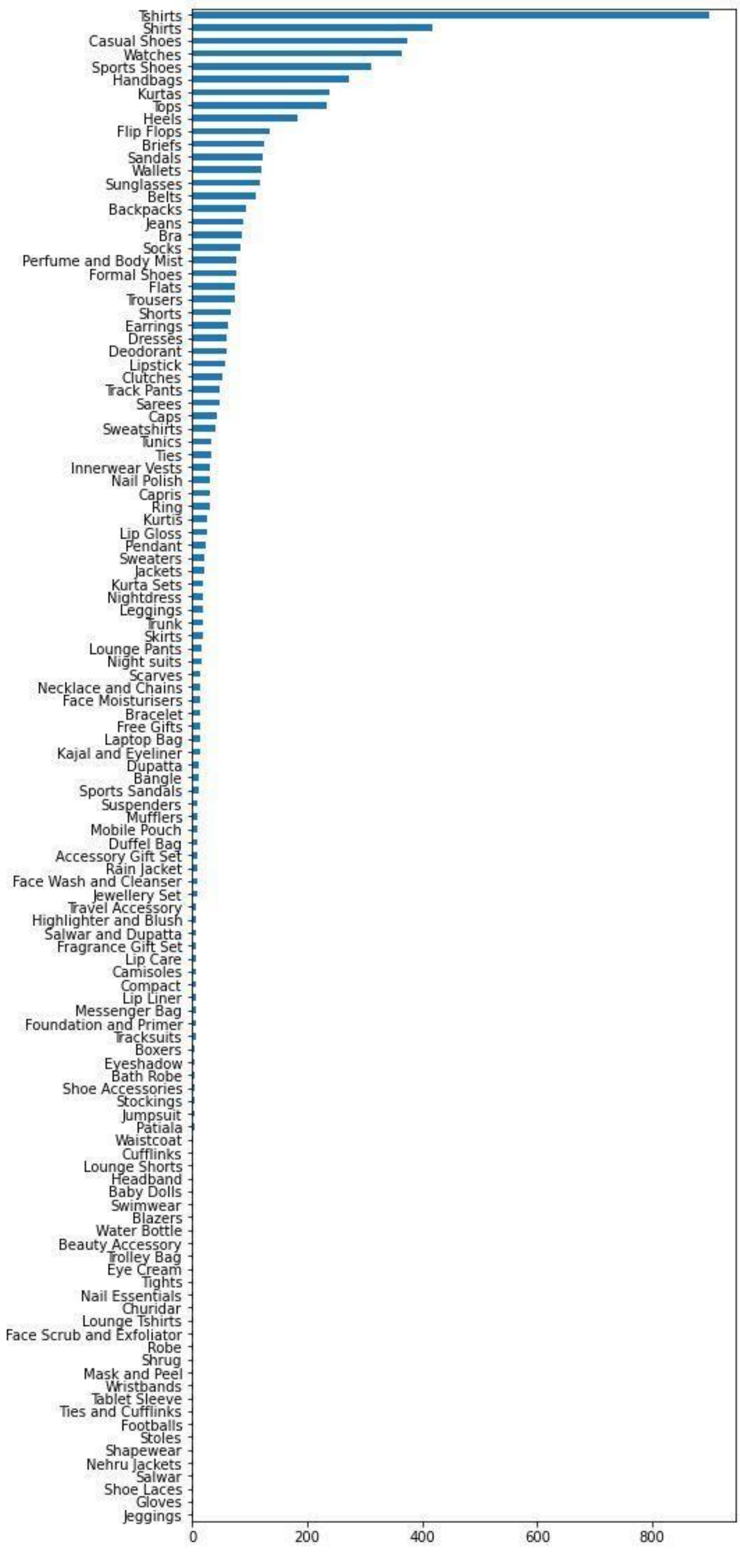


Fig 2.2: Dataset Bifurcation of the 44k samples

MODULES

- *ResNet50*

ResNet-50 is a deep convolutional neural network architecture that belongs to the ResNet (Residual Network) family. The "50" in ResNet-50 represents the number of layers in the network, including both convolutional layers and fully connected layers. The key innovation of ResNet is the introduction of residual connections, or skip connections, which allow information to bypass several layers and directly propagate from one layer to another. This enables the network to learn residual mappings instead of explicitly trying to learn the desired underlying mappings, which helps alleviate the degradation problem that arises in very deep networks.

ResNet-50 was pre-trained on the ImageNet dataset, which contains millions of labeled images from various object categories. This pre-training allows the model to learn meaningful representations of images and can be fine-tuned or used as a feature extractor for various computer vision tasks, such as object detection, image segmentation, and image recognition.

In our project we use ResNet50 to send our 44k images through the 50 layers of ResNet and be able to extract the features of those images and then dump that information in a pickle file to be used to find recommendations using nearest neighbours afterwards.

- *TensorFlow*

TensorFlow is an open-source, end-to-end ML platform. It has an extensive, adaptable ecosystem of community resources, libraries, and tools that allows researchers to advance the state-of-the-art in machine learning and developers to create and deploy ML-powered applications with ease. TensorFlow was initially developed by engineers and developers working on the Google Brain team from Google's Machine Intelligence Research organisation to undertake machine learning and deep neural networks research. The system is sufficiently general to be applicable in numerous additional domains. We utilize Tensorflow with the keras model for image pre-processing and algorithm implementation.

- *Streamlit*

  Streamlit is an open-source Python library that facilitates the creation and sharing of aesthetically pleasing, custom web applications for ML and data science. We developed our website's interface using this. It enables us to create an interactive webpage that allows users to browse for recommendations by uploading an image as a reference.

- *keras*

  Keras is the high-level API of TensorFlow 2; it is an approachable, highly-productive interface for solving machine learning problems, with an emphasis on contemporary deep learning. It provides essential abstractions and building elements for rapidly developing and deploying machine learning solutions. We used it for image processing, downsizing the input images and for implementation of our algorithm.

- *pickle*

  Binary protocols for serialising and deserializing a Python object structure are implemented in the pickle module. Python's "pickling" function converts an object hierarchy into a byte stream, whereas "unpickling" does the opposite, transforming a byte stream back into an object hierarchy. Pickle data employs a concise binary encoding by default. Pickled data can be efficiently compressed if appropriate size characteristics are required. The features and names of our samples are stored in two pickle files, embeddings.pkl and filenames.pkl, which we create in this phase. At the time of the recommendation's actual implementation, these files are used.

- *tqdm*

  tqdm is a python module that is used to create a progress bar in the python script. Progress bar is basically a graphical control element and we use it to visualize the progression of an extended computer operation. Here in this experiment we need the tqdm module during the extraction of features from the images while creating the pickle file. Since the process is long and extensive, we use tqdm to be updated on whether the extraction is working correctly or not.

- *sklearrn.neighbors*

sklearn.neighbors is a module in the library scikit.learn, a machine learning library in python. This module is used for performing various tasks related to nearest neighbor algorithms in our code.

The module includes the KNeighborsClassifier and KNeighborsRegressor classes, which are used for classification and regression tasks, respectively, using the k-nearest neighbors algorithm. These classes allow you to train a model based on a labeled dataset and make predictions for new, unseen data points based on their proximity to the training examples.

The KNeighborsClassifier and KNeighborsRegressor classes support different distance metrics, such as Euclidean distance and Manhattan distance, to measure the proximity between data points. They also allow you to specify the number of neighbors to consider when making predictions.

For our code we used sklearn.neighbors to perform nearest neighbor classification using euclidean distance formulas to find the closest images to the ones we upload.

- *cv2*

cv2 is a widely used Python library for computer vision and image processing tasks. It is a part of the OpenCV (Open Source Computer Vision) library, which is a collection of computer vision algorithms and tools. cv2 provides a convenient interface to access and manipulate images and videos.

We use cv2 in our project to perform the image processing tasks of being able to read the images and features and to show the output image during the test run.

- *os*

The os module is a built-in module in Python that provides a way to interact with the operating system. It offers a range of functions for working with files, directories, and other operating system-related functionalities.

We use it in our project to interact with the directories for the pickle files and to extract images.

- *PIL*

  PIL (Python Imaging Library) is a popular Python library for image processing and manipulation. It provides a wide range of functions and methods for working with images. We use it in our project to open images and extract features and append them into a pickle file.

- *numpy*

  NumPy (Numerical Python) is a python library which is used for numerical computations and mathematical operations on multi-dimensional arrays. It provides efficient data structures, functions, and tools for working with large, multi-dimensional arrays and matrices. We use this library in our project to create image arrays and work with them

WORKING

We designed the system model for this ecommerce recommendation project by dividing its working system into three different parts. Those three parts are:

1. Image Pre-processing

In this step we use computer algorithms to manipulate digital images. We need to refine the image and remove the noise and enhance the pixels of the image (224 x 224). Now an image is a two dimensional array with numbers between 0 to 255. An image can be expressed in the form of f(x,y) where x, y stands for the horizontal and vertical components. For pre-processing the steps we follow are

- **Read image**: Take provided image as input and store it somewhere.
- **Resize image**: Then we resize the image in accordance with input size of model.
- **Segmentation**: The saved image is converted from RGB to BGV which helps in getting better extraction of features
- **Flatten**: The saved 2D matrix of the image is converted into Vector.

The algorithm that we used for our system was **ResNet50**.

2. Recommendation Engine

The goal of the recommendation engine is to find a product that is similar based on the image that the user uploads. It uses different algorithms to find and recommend a product. The working of this engine goes as:

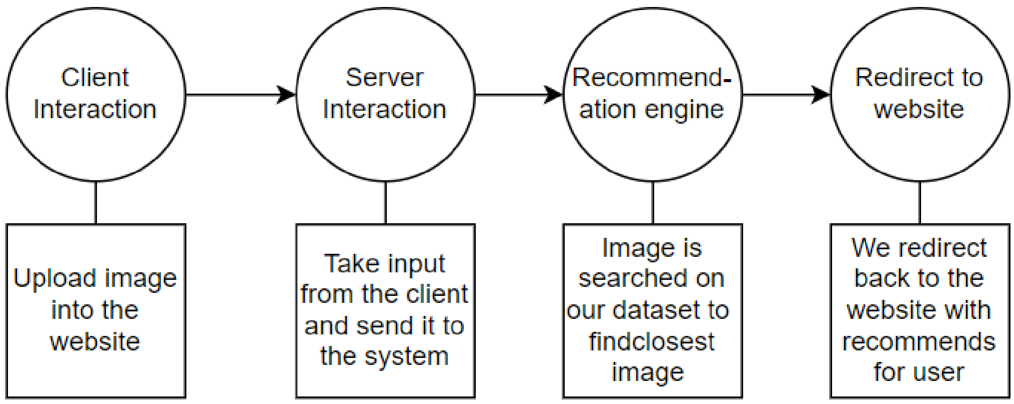


Fig 2.3: Recommendation Engine Flowchart

We use Keras for tensorflow in our system and we apply transfer learning through the model of Resnet50 as the base of the network and then compare it. Transfer networks can be explained to be where we take features learned on one problem and then we leverage them to find new similar problems. In our system we use a simple sequential model as it is and then use a euclidean algorithm to find the similarity. Through this similarity we find the products which are most similar to each other and hence give the user 5 different products which are similar to the uploaded image.

3. Web App

For creating the web application for our recommendation system we use the python library Streamlit to create it where the user first uploads the necessary items needed which then go through the recommendation system on the backend and gives back 5 images on the website which are very close to the uploaded image.

OUTPUT

On the Google COLAB, we ran tests on around 44k samples using the VGG16, DenseNet121, and ResNet50 algorithms. For an input of a shoe photo, each algorithm provided us with five recommendations. From all of them we decided to apply ResNet50 for our final system because it is the **quickest algorithm** giving accurate recommendations.

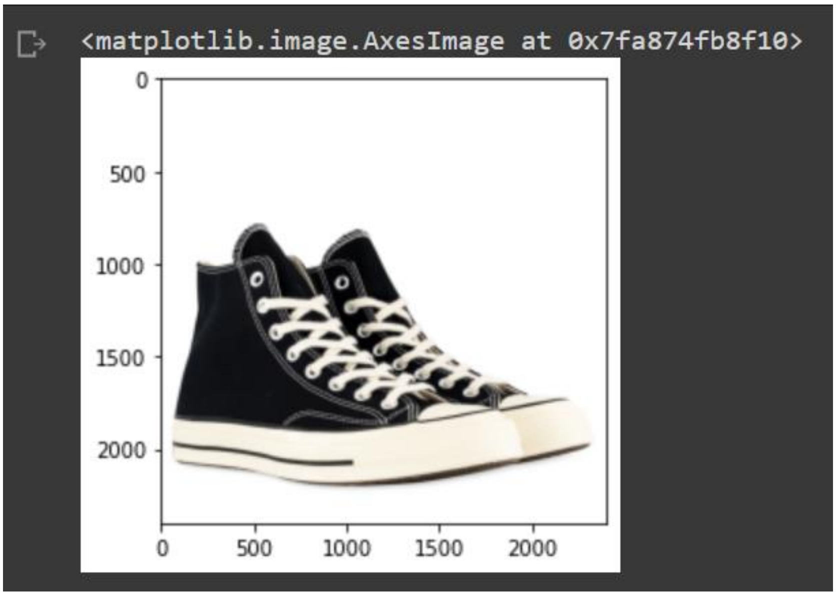


Fig 2.4: Input image for each algorithm

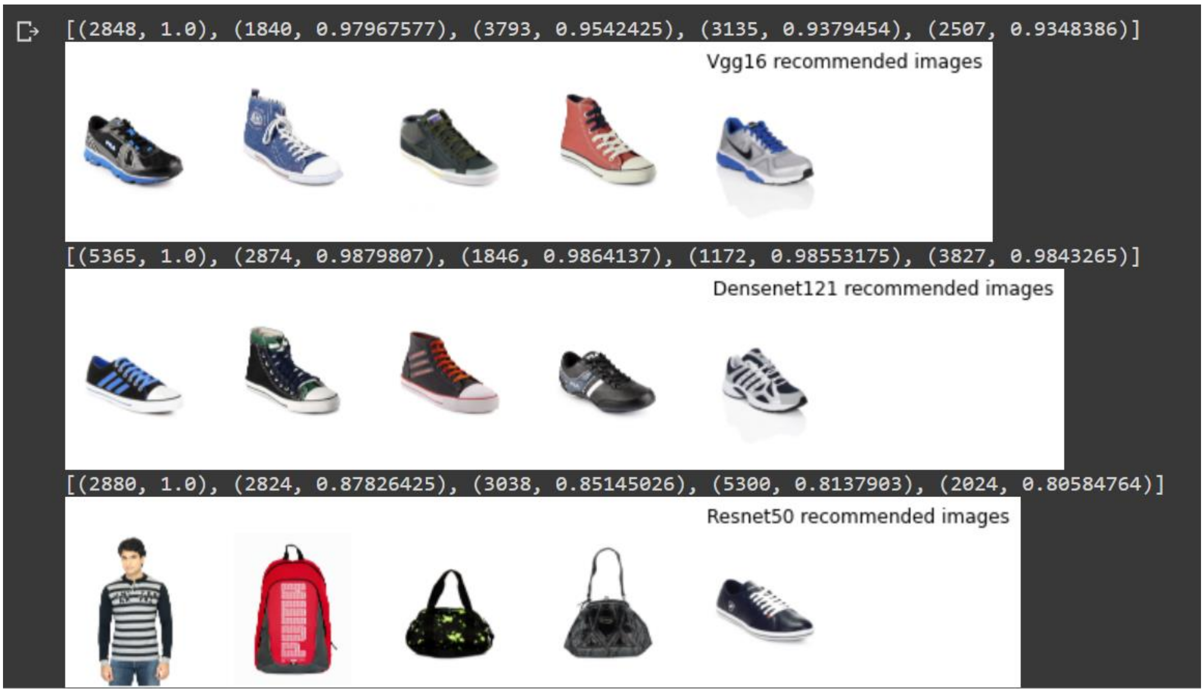


Fig 2.5: Output recommendations using the three algorithms

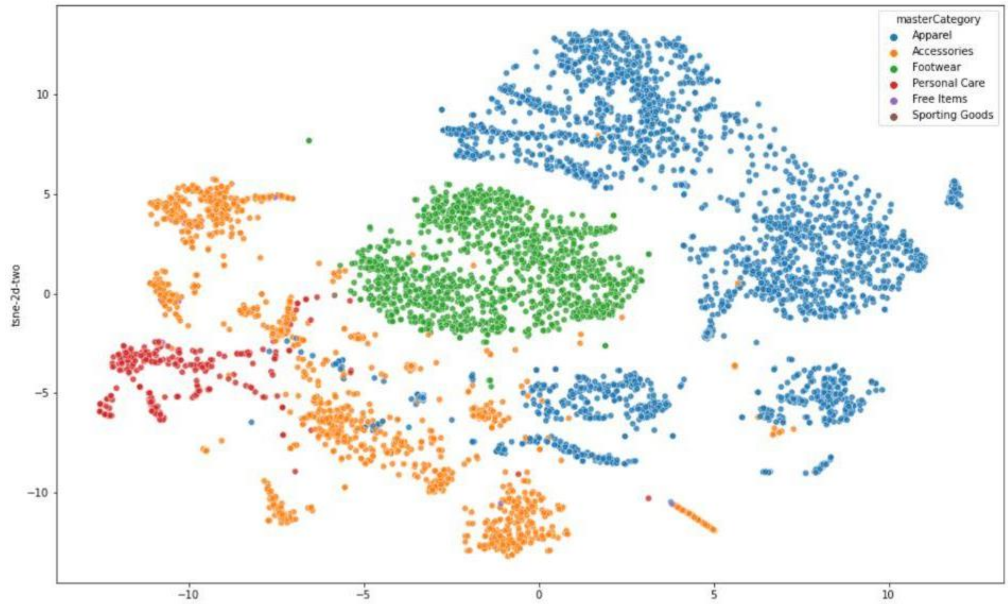


Fig 2.6: Graphical representation of dataset

We also checked the graphical representation of the files of features of the images to check and see how we can use the nearest neighbor algorithm while maintaining the accuracy of the recommendation model. As you can see from the image above the style of images are divided into apparels, accessories, footwear, personal care, free items, sporting goods, etc.Also, every similar item makes a cluster and through this cluster the nearest neighbor is found which is why we get the most accurate recommendation.

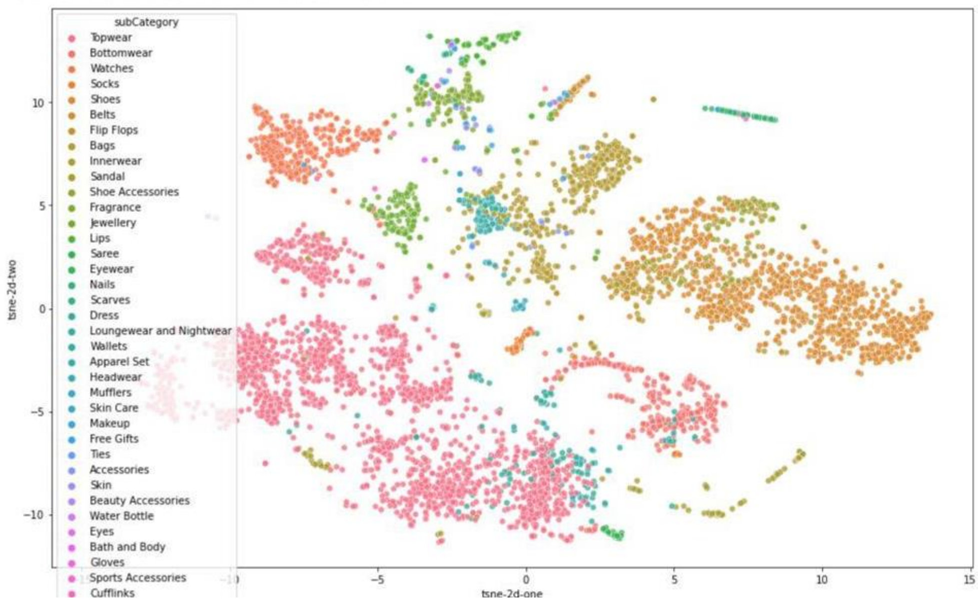


Fig 2.7: In depth Graphical representation of dataset

In the image above we have just divided the categories into further subcategories to notice and see how similar products make similar clusters which we use to get the recommendations.

To make our system more adaptable, we employed PyCharm to construct a three-stage process consisting of testing, displaying, and preprocessing. The sample photos are first pre-processed via app.py, which sorts them by style and other distinctive characteristics before exporting them as compressed pickle files labelled embeddings.pkl and filenames.pkl. We extract these features using the ResNET50 algorithm, which combines imagenet with the nearest neighbour approach.

The app.py created pickle files are then used in the testing phase via test.py, where a sample image is provided and 5 suggestions are returned. The user-uploaded image was scaled down using GlobalMaxPooling2D so that it could be processed more quickly. We make use of the cv2 module for demonstrating the generated samples.

The final step is to run main.py and put the algorithm and system into action. The webpage for our recommendation system was developed using the Python module StreamLit. Multiple picture recommendations were made using StreamLit as well. After the user has uploaded the necessary goods, our recommendation engine will show products that are comparable to those uploaded. A few of our recommendation output examples are shown in the figures below.

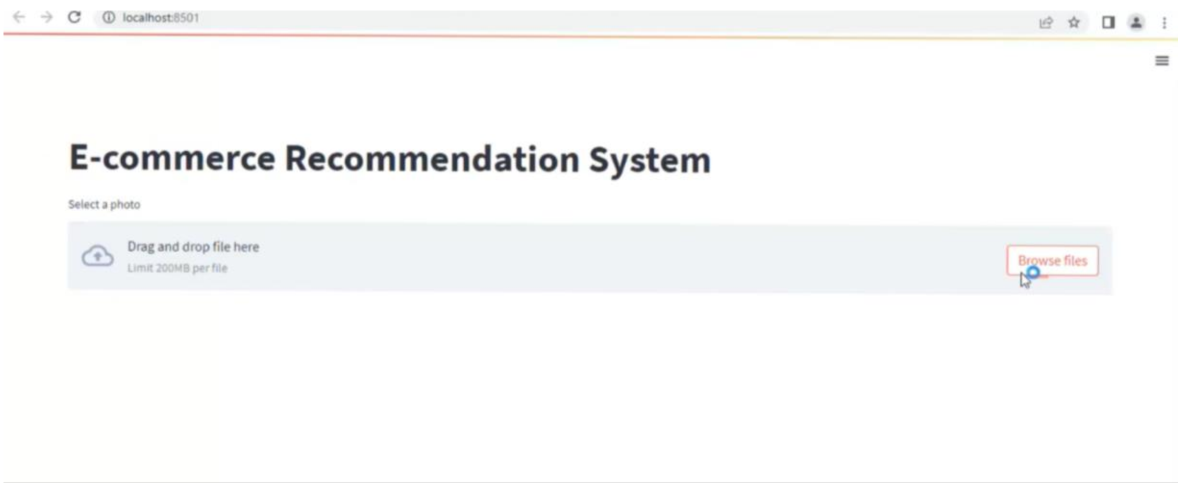


Fig 2.8: Streamlit website

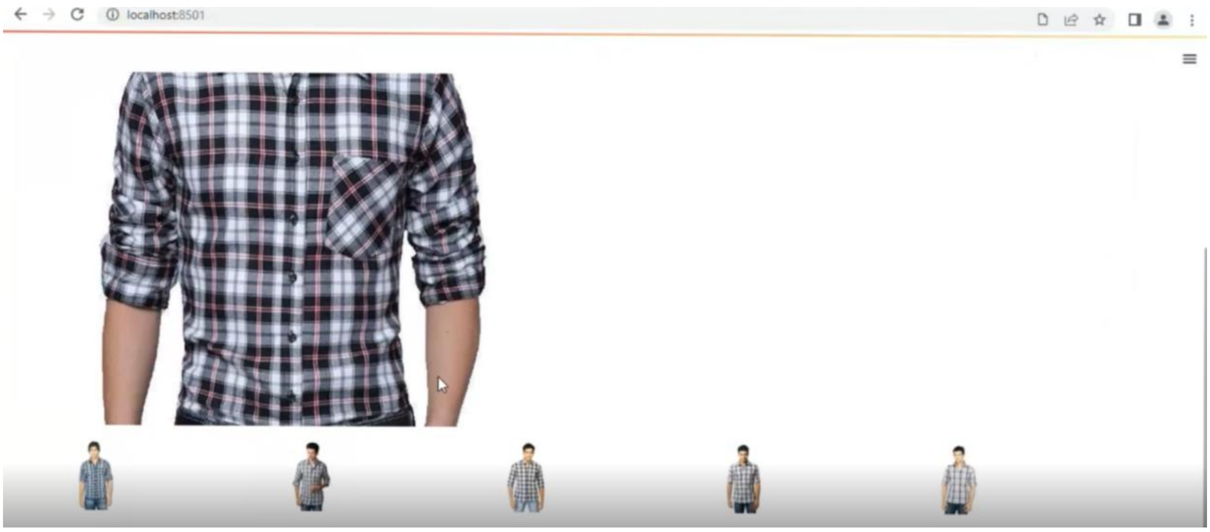


Fig 2.9: Recommendation Output 1 – Checkered Shirt

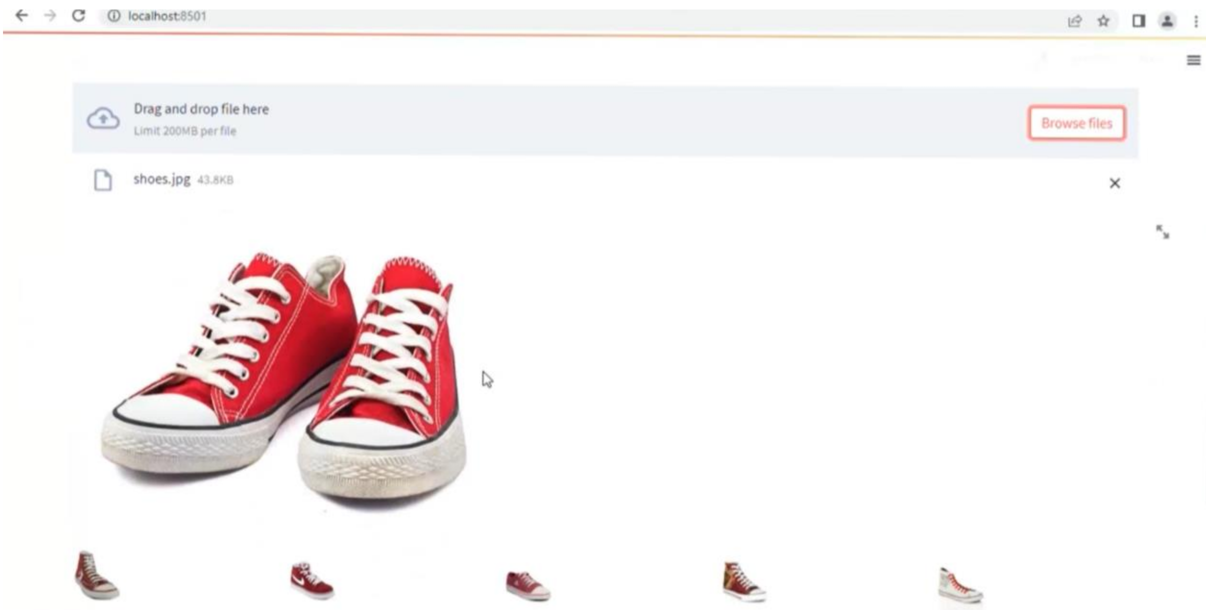


Fig 2.10: Recommendation Output 2 – Red Converse Shoes

CHAPTER 3

APPLICATIONS

E-commerce sites employ recommender systems to make product recommendations to customers and to give them information that will enable them to make an informed decision about what things to buy. The products may be suggested based on a site's best-selling items generally, the consumer's demographics, or an analysis of prior purchasing patterns that predicts future behaviour. Consumers can receive recommendations for products, tailored product information, opinions of the community as a whole, and community critiques. These strategies for making recommendations are, in general, a part of site personalization because they enable the site to adjust to each user. In a way, recommender systems make it possible to build a new store that is specifically tailored to each customer.

Essentially, a recommendation system does the following:

- **Turning Browsers into Buyers**: Most website visitors browse the site without making a purchase. Consumers can find the things they want to buy with the use of recommender systems.
- **Increasing Loyalty**: By fostering a relationship between the website and the user that adds value, recommendation systems increase client loyalty.
- **Cross-selling**: By influencing customers to buy additional products, recommender systems enhance cross-selling.

We have seen similar technologies being used in practice on e-commerce platforms like Amazon, Flipkart, and others.

An amazingly complex yet very successful algorithm is used by the e-commerce giant Amazon. The three primary relationship dependencies in the algorithm are User-Product, Product-Product, and User-User. In order to generate recommendations that are more accurate, it also makes use of user behaviour, demographic information, and product attribute information. It is best and most efficient to combine collaborative and content-based approaches into a single, all-encompassing recommendation engine. This is how the Amazon algorithm operates; it gives its customers the highest-quality recommendations by fusing the best aspects of both approaches. Based on how customers interact with various product offers, its hybrid, Bandit-based and Casual Inference algorithms choose from a number of recommender models in real-time. Thus, using artificial intelligence and machine learning, Amazon's tailored recommendation engine enhances the shopping experience for customers and displays products in real time based on their browsing history.

To create "Similar Product Recommendations," Flipkart, another industry giant, combines content and collaborative filtering processes. The collaborative filtering method is applied over user browse data like product page views, wishlist, add to cart, etc. to determine the most commonly co-browsed goods for a particular product. The content matching is done over product attributes and photos in the catalogue. Combining these several sources yields the ordered list of comparable goods based on relevance. By highlighting the "correct" products inside the appropriate set, they aim to further assist the user in making a selection and increase their level of confidence in making a purchase. To boost the "trust factor," enable popular products to sell out more quickly, and optimize the product based on engagement and conversion, ranking functions including Product Quality, Performance, and Diversity are applied.

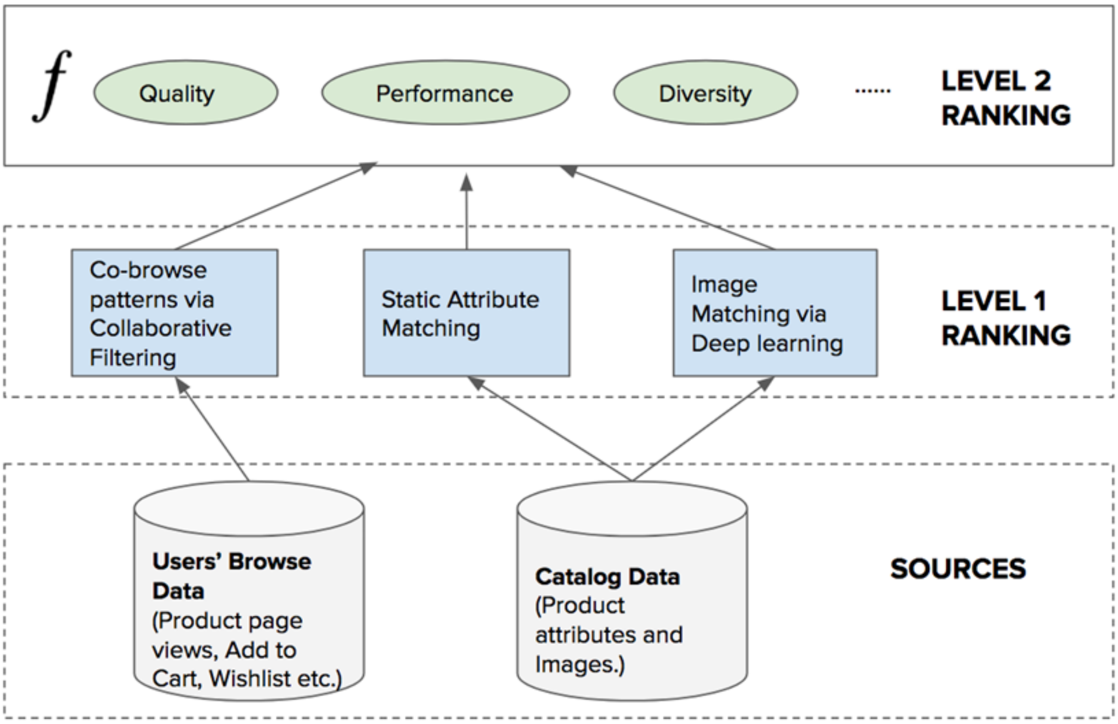


Fig 3.1: Flipkart Recommendation Flow (source: www.medium.com)

FUTURE SCOPE

E-Commerce recommendation systems have already come a long way in providing personalized recommendations to online shoppers. However, there is still significant potential for improvement and innovation in this field. Here are some potential future directions for E-Commerce recommendation systems:

1. **Contextual recommendations**: E-Commerce recommendation systems could use more contextual information to make better recommendations. For example, they could consider the user's location, weather, time of day, and other relevant factors when making recommendations.

2. **Multimodal recommendations**: With the increasing availability of multimedia content, E-Commerce recommendation systems could start to incorporate images, videos, and audio into their recommendations. This could provide a more engaging and interactive shopping experience for users.

3. **Explainability**: One challenge with E-Commerce recommendation systems is that they can sometimes be seen as "black boxes" that make recommendations without providing any explanation for why they are making them. In the future, E-Commerce recommendation systems could become more transparent and provide explanations for their recommendations.

4. **Personalization**: Personalization is already a key feature of E-Commerce recommendation systems, but there is still potential for improvement. In the future, E-Commerce recommendation systems could use more advanced machine learning algorithms to better understand each user's unique preferences and make more accurate recommendations.

CONCLUSIONS

Recommendation systems is an exponentially growing field that has attracted the attention of various researchers and academicians. It is not easy to make a choice among numerous options and gigantic data, it is always going to be a tough and confusing task. Recommendation systems help us to overcome this. To do the job competently and accurately, the recommendation systems use efficient information retrieval and filtering mechanisms. Over the past years, immense research work has been devoted to meet these ends, and several recommendation approaches and techniques are proposed.

In the recommendation system that we ended up designing we created an ecommerce content based recommendation system that used a dataset with 44k fashion related images. We then sent the images through a ResNet50 convolution neural network to extract the features of the images. Now those features can be represented graphically and through those features we can get recommendations that we need by using a nearest neighbor algorithm. In our system we used Euclidean distance. This gives us the 5 images from the dataset which are closest to the input image we put in and recommend that to us.

We also made the frontend of this system by using a python module streamlit which creates a local host to create a website and we are able to browse and upload the input image and get recommendations online.

PAPERS PUBLISHED

1. Khushi Gandhi, Jill Karia, Aarya Shah, Hetankshi Vora, and Sanjay Deshmukh, "E-Commerce Recommendation System" was published in DJ SPARK Magazine with the ISBN: 978-93-5777-300-3.

REFERENCES

[1] Sivapalan, Sanjeevan, Alireza Sadeghian, Hossein Rahnama, and Asad M. Madni. "Recommender systems in ecommerce", 2014.

[2] Farida Karimova (2016), "A Survey of E-Commerce Recommender Systems".

[3] Xiangpo Li (2021), “Research on the Application of Collaborative Filtering Algorithm in Mobile E-Commerce Recommendation System”, IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC).

[4] Priya S, & Mansoor Hussain D (2017), “Recommendation Systems for E-Commerce: A Review”.

[5] Utpal Chandra De, Shobhan Banerjee, Manas Kumar Rath, Tanmaya Swain, & Tapaswini Samant (2022), “Content Based Apparel Recommendation for E-Commerce Stores”, 3rd International Conference for Emerging Technology (INCET).

[6] Pegah Malekpour Alamdari, Nima Jafari Navimipour, Mehdi Hosseinzadeh, Ali Asghar Safaei, & Aso Darwesh, “A Systematic Study on the Recommender Systems in the E-Commerce”.

[7] Zezhou Fan, Dan Chang & Jinhong Cui, “Algorithm in E-Commerce Recommendation”.

[8] Ramil G. Lumauag, Ariel M. Sison, & Ruji P. Medina, “An Enhanced Recommendation Algorithm Based on Modified User-Based Collaborative Filtering”, 2019 IEEE 4th International Conference on Computer and Communication Systems.

[9] Gang Huang, “E-Commerce Intelligent Recommendation System Based on Deep Learning”, 2022 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC).

[10] Bai Li, Allam Maalla, & Mingbiao Liang, “Research on Recommendation Algorithm Based on E-Commerce User Behavior Sequence”, 2021 IEEE 2nd International Conference on Information Technology,Big Data and Artificial Intelligence (ICIBA 2021).

[11] T. Tran (2006), "Designing recommender systems for e-commerce: an integration approach".

[12] J. Ben Schafer, Joseph A. Konstan, & John Riedl, "E-Commerce Recommendation Applications"

[13] Koene, Ansgar & Perez, Elvira & Carter, Chris & Statache, Ra mona & Adolphs, Svenja & O'Malley, Claire & Rodden, Tom & McAuley, Derek. (2015). "Ethics of Personalized Information Filtering".

[14] M Sharma, Smann (2013), "A survey of Recommender systems: Approaches and Limitations"