An Extensive Analysis of Product Suggestion Systems

Ms. Shadna Yadav
sadhna.yadav02@rediffmail.com

Abstract: *- In modern times, the Internet has become a package of parts and everyday life. During the Covid 19 pandemic, all sectors range from food delivery, e-commerce, education, learning, health services and more. With lighter product selection and wider accessibility, e-commerce gains popularity at a faster pace. In it, consumers are exposed to a variety of options for the products of their interest. It is a major challenge for users to select high quality products at the right cost. Product recommendation systems are used for this purpose and support users in their decisions to select the right product. Product Recommendations System analyzes product tests and provides recommendations for the product. Based on previous reviews, this recommendation may be positive or negative. A comprehensive study of recommendation systems provides the work presented here. It covers a variety of aspects of the various methods, ranging from content-based filtering methods to hybrid filtering methods used in current recommended systems. The current study shows the impact of recommendation systems on e-commerce businesses*

Keywords:*E-commerce, Quality product, decision making, Recommendation, Content-based filtering, collaborative base filter, hybrid filtering.*

1.1 History of Recommendation Systems

The 1990s laid the foundation for many of today's recommendation systems. Early information calling systems were used before the 1990s. In the 1950s and 1960s, early information calling systems were also known as simple information calling systems. For example, researchers have developed a system based on simple, key collaboration coordination at the moment. His goal was to help users find relevant documents or content based on search queries. The joint filter system recommends the elements based on previous interactions and reviews. A content-based recommendation system was also developed in the 1990s. These systems recommend elements based on the characteristics and properties of the elements. The hybrid system was developed to improve the recommended accuracy (2000s). The Netflix Awards (2006 2009) is one of the most innovative and well-known content-based recommendation initiatives. The aim was to predict film reviews.

1.2 Definition of recommendation system

It is also called a recommended system, or a recommended application or algorithm that simply contains a user's personalized suggestions or recommendations. The idea behind these recommendations is to help users find the right product or content that you may find interesting or useful. As a rule, these recommendations are based on user preferences, user preferences, historical behavior, or explicit feedback. The main goal of the recommended system is to improve the user experience by addressing the issue of information overload in environments where many options are available. Online content platforms, social media, streaming services, e-commerce and other industries create many systems of recommendation.

1.3 Need of recommendation system

Systems that provide recommendations are essential for many internet platforms and apps. Their main goal is to provide personalized suggestions or recommendations by analyzing user behavior and interests. Below are the main reasons why the recommended system is important.

1.3.1 An overload of information

Many of the options can be overloaded in the context of e-commerce websites, streaming services, or news websites with an excessive number of content or elements. The recommendation engine helps consumers decide on the right factors or content from a variety of options.

1.3.2 Customization

Many of the options can be overloaded in the context of e-commerce websites, streaming services, or news websites with an excessive number of content or elements. The recommendation engine helps consumers decide on the right factors or content from a variety of options.

1.3.3 User Involvement

Users interact with the platform rather when they provide appropriate, individual recommendations. Further user sessions, higher conversion rates, and more respectful user base can arise from this improved commitment.

1.3.4 Maintaining Customers:

Personalized experience using recommendations improves customer satisfaction and loyalty. Users who feel like a platform and accept their preferences tend to use them again.

1.3.5 Time Efficiency

Perhaps by presenting interesting options, the recommendation system saves time by eliminating the need for comprehensive search or browsing.

1.3.6 Optimization of Advertisements

The recommendation system has the ability to provide specific ads based on the user's interests and behavior. Showing important ads improves the user experience and improves the effectiveness of your ads.

1.4 Advantage of Recommendation system in E-commerce

For many reasons, the recommendation system is useful in e-commerce. The main benefits of the recommended system are:

1.4.1 Enhanced Income and Sales

The recommendation system includes suggestions for relatives or complementary products that will act to interest customers in order to bring previous purchases, users, and more products to your cart and increase the total amount.

1.4.2 Improved Experience for Users

The recommendation system provides specific product recommendations based on browser courses, user preferences and connections. Related and adaptive content are provided through this adjustment to improve the overall user experience.

1.4.3 More Effective Product Discovery

The recommendation system helps consumers find products they may not have found through manual search queries. This knowledge of gratitude can end with greater interactions and conversion rates.

1.4.4 Extended Time on Platform

Users will remain more if the platform for online purchases provides relevant product recommendations. Intense relationships between increasing conversion rates and organizations.

1.4.5 Market Differentiation

In overcrowded markets, e-commerce websites that use the recommended system effectively have the opportunity to distinguish themselves. Delivering a personalized, effective shopping experience attracts and retains customers and provides a great competitive advantage for your business.

2 Literature Review

Leung W. C et all [1], In text mining and computational linguistics, sentiment analysis a type of text classification has grown in popularity, especially for the analysis of product reviews. Evaluative terms are extracted, sentiment polarity and strength are assessed, and reviews are divided into sentiment classes. This paper discusses two approaches: the machine learning approach and the sentiment orientation approach. Whereas the machine learning approach makes use of classifiers like Naive Bayes and Support Vector Machines, the sentiment orientation approach uses seed adjectives and semantic relationships to determine sentiment. Three steps include a typical sentiment analysis model presented in this paper: Sentiment Classification, Review Analysis, and Data Preparation.

The paper points out how sentiment analysis is constantly changing and how advanced techniques and models are required for opinion extraction and classification. Sentiment analysis's relevance in user preferences and decision-making is highlighted by its possible uses in collaborative filtering. In summary, the article highlights sentiment analysis as an emerging field with promising applications and continuous research.

F Fang X. et all [2], This research study addresses the fundamental issue

of sentiment polarity categorization with a focus on sentiment analysis, often known as opinion mining. The authors provide a sentiment analysis procedure that consists of stages for feature vector creation, sentiment score calculation, and negation phrase detection using product evaluations from Amazon as their dataset.

The challenges associated with sentiment analysis are discussed in the paper, especially when handling the shortcomings of internet data, like the nature of opinions and the absence of ground truth labels. The authors get around these problems by using product reviews that they get from Amazon. These reviews are verified before being posted and include a star rating that serves as ground truth.

The suggested method for sentiment analysis entails determining negation phrases, calculating token sentiment scores, and creating feature vectors for classification. The authors use three classification models (Naive Bayesian, Random Forest, and Support Vector Machine) in their studies for categorising sentences and reviews.

The outcomes demonstrate good competence in classification at the sentence and review levels. With high F1 scores in both categorization tests, the sentiment score exhibits strength as a feature. Out of the three classification models, the Random Forest model consistently exhibits the best performance.

Rao Y. K. et all [3], This study suggests a recommendation algorithm that

makes use of emotion data collected from user evaluations on social media. To predict product ratings, the model combines item reputation similarity, interpersonal sentiment influence, and user sentiment similarity into a single matrix factorization framework. One of the main contributions is the creation of a brand-new connection termed interpersonal sentiment influence, which illustrates the emotional influence that users' friends have on them. Using user reviews, the model calculates sentiment quantitatively and concludes item reputation based on user sentiment distribution.

The experimental results highlight the usefulness of the suggested emotional elements in rating predictions and show significant improvements over current methods. The study highlights how crucial it is for recommendation algorithms to take user preferences and social sentiment dynamics into account, especially when it comes to online reviews. For improved sentiment analysis, the authors propose extending existing language rules, improving sentiment dictionaries, and modifying current hybrid factorization models.

Overall, the suggested approach improves the accuracy of rating predictions in recommender systems by addressing the issues of information overload and including user sentiment, interpersonal relationships, and item reputation.

Haque U. T. et all [4], This research paper suggests a supervised learning model for polarising a large amount of unlabelled product review dataset. As e-commerce has grown in importance, so too have customer reviews in the digital age. The model attempts to classify customer feedback into positive and negative sentiments using a combination of active learning, feature extraction techniques (bag of words, TF-IDF, Chi-square), and different classifiers. The methodology involves obtaining a dataset from Amazon product reviews, including the categories for Electronics, Cell Phone and Accessories, and Musical Instruments.

For every product category, experiments are carried out using several classifiers such as Logistic Regression, Random Forest, Stochastic Gradient Descent, Multinomial Naive Bayes, Linear Support Vector Machine, and Decision Tree.

The findings show that the suggested model performs well in terms of accuracy, with Support Vector Machine consistently yielding the best outcomes. A comparative study of the

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 04 | April - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

suggested method with comparable works shows how well it works to get more accuracy.

Devi U. [5], The research study that is being presented focuses on sentiment analysis of online product reviews, with a specific emphasis on Amazon. The study looks at the growing popularity of online shopping and the growing amount of product reviews, which has led to the need for effective ways to analyse client opinions.

An overview of sentiment analysis, its uses, and the difficulties in manually analysing such a large volume of user-generated information are given in the paper's literature review. Part-of-speech tagging, semantic linkages, and the Bag-of-Words model are among the sentiment analysis models covered by the writers. In order to categorise texts according to their emotional orientation, the review highlights the significance of sentiment polarity and sentiment score.

Additionally, word level, sentence level, document level, and entity level are the four primary areas into which the study divides sentiment analysis methodologies. It highlights the difficulties in document-level categorization and presents the usage of lexicons and corpora in sentiment analysis.

They talk about how consumer trust has changed over time based on online sentiment ratings, highlighting the crucial part sentiment analysis plays in the decision-making process. After that, the study article looks at a variety of sentiment analysis methods, breaking them down into supervised and unsupervised categories. The paper covers topics such as term presence or frequency, part-of-speech tags, grammar, negation, and a brief description of each technique, as well as the significance of feature engineering.

The authors provide the accuracy values of three machine learning algorithms Naive Bayes, SVM, and MLP applied to product reviews in the section on experimentation findings. The results show that all three approaches work well, with MLP showing the best accuracy. The study paper concludes by addressing the growing significance of sentiment analysis in the digital era and offering a thorough overview of numerous methods and strategies. According to the reported experimental data, MLP performs better in the particular context of Amazon.com product reviews than SVM and Naive Bayes. The study offers important information about machine learning applications for text classification and sentiment analysis.

Sultana N. et all [6], This study examines the impact of various adjective, adverb, and verb combinations on the precision of sentiment prediction with a focus on sentiment analysis. A range of classifiers, such as Naive Bayes, Logistic Regression, Linear SVC, and Decision Tree, are employed in the study, which performs experiments using a benchmark dataset of 50,000 movie reviews from Stanford.

The results indicate that, out of all the possible combinations of parts of speech, the combination of verb, adverb, and adjective produces the highest accuracy. Naive Bayes regularly performs better than other classifiers, identifying the ideal combination with an accuracy of 89.855%.

Additionally, the suggested method adds efficiency in terms of execution time, with Linear SVC executing testing datasets quickly and Logistic Regression needing the least amount of training time.

All things considered, this study offers insightful information about the importance of taking into account various speech parts in sentiment analysis and establishes Naïve Bayes as a reliable classifier for this purpose. The results hold significance for automated sentiment analysis across diverse fields, including social media, e-commerce, and film reviews.

Ni P. et all [7] In the context of beer products, this paper focuses on combining user reviews and ratings to improve the effectiveness of recommendation systems. It tackles the issue of rating data's inability to accurately reflect particular product attributes and investigates the extensive application of reviews and ratings for sophisticated recommendation systems. The research examines data from beer reviews, contrasts ten sentiment analysis classification models, and suggests a recommendation model based on the Spark-ALS collaborative filtering algorithm. A thorough literature review, effective outcomes on beer product recommendation, and the creation of an LSTM and Spark-based recommendation model for improved accuracy are among the contributions. Multiple deep learning models are used for sentiment analysis, and Spark-ALS is used for collaborative filtering. Based on user reviews and ratings, the study explores sentiment analysis and beer recommendation. The research compares deep learning models for sentiment analysis with traditional machine learning models using a Craft Beer dataset from Kaggle. Pyspark is used for data experiment rating. With a 62% accuracy rate, the results reveal that the ALS recommendation model performs more effectively than the others. With an accuracy of 0.75, the LSTM model works best in sentiment analysis. Sentiment analysis is incorporated into the recommendation algorithm to improve beer recommendations, which improves user acceptance and increases online beer sales.

Taparia A. et all [8] in an attempt to understand the connection betweentextual content and star ratings, the study investigates the use of natural language processing and text analytics to analyze customer reviews on Amazon.com. A dataset is gathered, pertinent features are chosen, data is preprocessed, word clouds are used to visualize sentiments, and features like polarity and review length are extracted as part of the suggested framework. The authors use a 5-core Amazon review dataset for smartphones and accessories to test different classifiers, such as logistic regression and multinomial naïve bayes. With an accuracy of 54.1%, the results demonstrate that the Logistic Regression classifier outperforms the others, emphasizing the impact of sentiments on ratings. The study comes to the conclusion that understanding customer opinions and forecasting product ratings depend heavily on sentiment analysis. It is suggested that future research incorporate review dates, examine how reviews affect sales, and expand the framework to incorporate reviews from additional e-commerce platforms. To sum up, the study highlights the importance of sentiment analysis in interpreting consumer attitudes and predicting product evaluations, providing useful information for companies.

3. Proposed Methodology

3.1 User Interface:- The user interface is a system in which the user interacts with the proposed system. The user interface

(UI) contains all the elements and components that allow users to provide input, receive feedback, and navigate the capabilities of the model.

3.2 Data Collection:- Sentiment analysis begins with textual data that can be obtained from a variety of technologies from a variety of sources. Text data can usually be created or collected through learning purposes, external parties, or via web scraping or crawling.

3.2.1 Web scraping:- Web scraping is a method of extracting information from a website. This includes accessing data from the website by submitting an HTTP inquiry, parsing the website's HTML or other markup language, and then submitting the information of your interest.

3.3 Data Preprocessing:- Most data from many sources is unstructured. This data is very large in raw form and contains some spelling and grammatical errors. As a result, text must be cleaned and processed in advance before analysis. In addition to improving analysis, the preprocessing steps are intended to reduce the dimensions of the input data. This is because many words like articles, prepositions, characterizations, special characters, etc. are not required and need to be removed as they do not affect the polarity of the text.
Data Preprocessing is done using the following steps:-

3.3.1 Removing links:

This process involves repetition of each line, distribution of text to individual words, and filtering words that contain the substring "HTTP". The resulting cleaned evaluations without hyperlinks are then saved again in the data frame.

3.3.2 Removing Reviews with empty text:

Filter reviews with empty or zero text.

3.3.3 Dropping duplicate rows:

Remove double lines from data records to ensure unique data.

3.3.4 Resetting index:

Returns the index of the dataframe after dropping the replica.

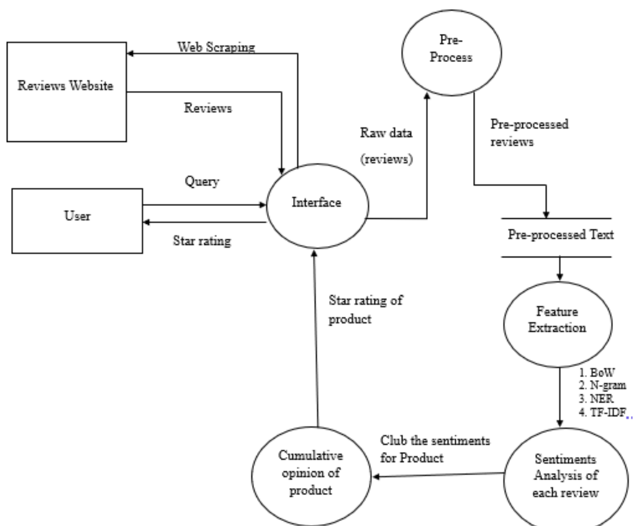


Figure 3.1 Proposed methodology

3.3.5 Removing Punctuations, Numbers, and Special Characters:
Use regular expressions to remove punctuation, digits, and special characters.
3.3.6 Function to remove emoji:
Define a function to remove emojis from the text using regular expressions.
3.3.7 Removing Stop Words:
Remove common stop words (e.g., "the," "and," "is") that do not contribute much to the sentiment.
3.3.8 Tokenize Clean_Reviews:
Tokenize the cleaned reviews into individual words or tokens.
3.3.9 Converting words to Stemmer:
Apply stemming to reduce words to their root form (e.g., "running" to "run").
3.3.10 Converting words to Lemma:
Apply lemmatization to reduce words to their base or dictionary form (e.g., "better" to "good").
3.4 Feature Extraction:- Functional extraction (FE) or functional engineering is a key operation in the mood analysis process as it directly affects mood classification performance. The purpose of this assignment is to extract wise data that characterizes important aspects of text such as: However, there is even more problem with dealing with texts from social media. Some important features used in mood analysis are:
3.4.1 Bag of words:- The method of extracting text features is called bag-of-words (sheets). The tokens of the document, the vocabulary structures of different terms, and the expressions of each document as vectors with word frequencies are all part of the process. The deposits are recorded in the generated sparse document matrix (DTM). On the lines "I like programming" and "I'm enjoying programming," the sheet matrix shows how often words like "I", "How", "programming" and "" appear in all documents, for example. Nevertheless, it is complex and efficient for basic text analysis.

3.4.2 N-gram:- N-grams are coherent sequences of n objects (words, letters, or symbols) from a particular text or language. It is used in a variety of NLP activities (natural language processing) such as text analysis, speech modeling, and machine learning. Unigrams (individual words), Bigrams (couple of consecutive words), Trigrams (three consecutive words) and others are examples of common n-gramm species. Compared to words, n-gram is a more refined representation of text by capturing context and local patterns. However, as the rise increases, the characteristic area expands exponentially, making it even more difficult to achieve data savings. N-Gram is used in speech recognition, mechanical translation and mood analysis, among other things.

3.4.3 Named Entity Recognition (NER):-

Named Entity Recognition (NER) is a natural language processing (NLP) technique that uses text to identify and categorise named entities (people, places, and objects). Tokenization, entity classification, and part-of-speech tagging are all involved. By extracting structured data from unstructured text, NER makes it easier to comprehend connections and context. NER is essential for applications like sentiment analysis and information retrieval because it helps create meaningful representations of textual input. The strategies used range from conventional ones like conditional random fields to more recent ones that use deep learning, including transformers and recurrent neural networks. For a variety of applications, NER is essential to improving language semantic understanding.

3.4.4 TF-IDF (Term Frequency-Inverse Document Frequency):-

The TF-IDF (Term Frequency-Inverse Document Frequency) measure in information retrieval evaluates the relevance of a word in a document in comparison to a database. It combines Inverse Document Frequency—the inverse proportion to the number of documents containing the word—with Term Frequency, which measures the frequency of words in a document. This produces a score that indicates how important a term is in both the particular document and the larger corpus. Greater relevance is indicated by higher ratings. Text mining, information retrieval, and document ranking all make extensive use of TF-IDF, which serves as a foundation for feature extraction and facilitates the comprehension of the significance of phrases in a particular context.

3.5 Sentiments Analysis of each review:

To determine the polarity of user sentiments, rule-based sentiment analysis algorithms automatically categorise input information based on a set of predetermined rules. NLP approaches are used to carry out these guidelines. These methods include tokenization, parsing, lexicons, stemming, and part-of-speech tagging.

The following scenario shows an example of a rule-based sentiment analysis methodology. Word lists are used to express various emotions. those that are negative include "horrible," "worst," and "bad," whereas those that are favourable include "great," "love," and "amazing." After then, a sentence input is examined and its polarised word count is recorded. Following that, this count is divided into buckets of various groups, such as "positive," "negative," or "neutral," depending on the greatest number of polarised terms connected to the supplied sentiment. Although rule-based algorithms are simple to use, they frequently ignore the nuances of word pairings and text.

3.6 Cumulative opinion of product:

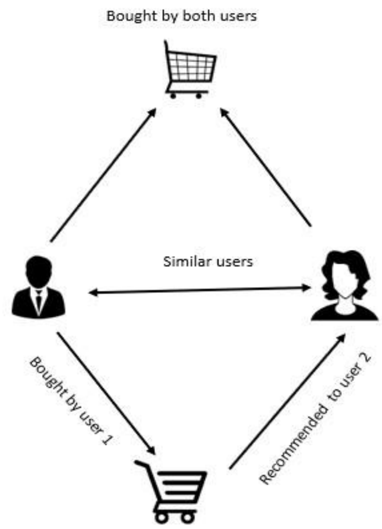
To determine the cumulative sentiment for the complete dataset, add up or average the sentiment scores. Based on the opinions gathered, you are given an overall sentiment score for the product.

4 Classification of Recommendation Systems

Different types of recommendation systems can be defined according to their methods and approaches. The primary kinds of recommendation systems are listed below.

4.1 Collaborative Filtering

Collaborative filtering is an increasingly popular method used by recommendation systems to provide users with personalized recommendations based on the preferences and behavior of other users who are similar to them.

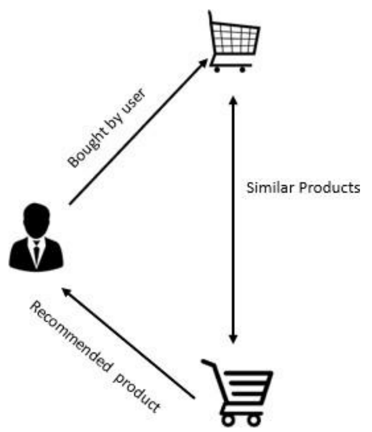


The basic idea of this method is that if users A and B share preferences on some items, then it is likely that they will share preferences on other items as well. Collaborative filtering can be broadly categorized into two categories.

- **User-Based Collaborative Filtering.**
- **Item-Based Collaborative Filtering.**

4.2 Content-Based Recommendation Systems

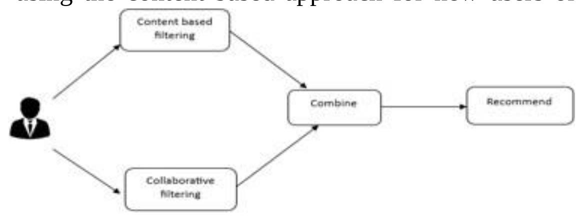
Content-based recommendation systems provide users with recommendations based not only on the user's declared preferences but also on the features and characteristics of the items. These systems aim to suggest items having similar qualities to previously liked or interacted with content, providing priority to the content of the item.



Content-based filtering has the advantage that it's able to solve the cold start problem, which arises when new items or users have little to no data available. But it may be an over-specialization issue, suggesting only products that are overly similar to the user's previous actions, which reduces variety and randomness.

4.3 Hybrid Models
Multiple recommendation techniques are combined by hybrid models to improve their scope and accuracy.

By using the content-based approach for new users or items



and the collaborative filtering approach for users or items with sufficient information, a hybrid model, for instance, can combine content-based and collaborative filtering.

5 Collaborative Filtering:
A popular recommendation method that uses user interactions with things to generate individualised recommendations is called collaborative filtering. It depends on the concept that people who are familiar with similar things.

5.1 User-Item Interactions:
Collaborative filtering is based on interactions between users and items. These interactions can take many different forms, including as clicks, purchases, reviews, ratings, and other types of involvement with goods within a recommendation system by users. Since it serves as the foundation for interpreting user preferences and behavior, this data is essential. Recommendation systems can make suggestions for products that consumers are likely to find interesting by examining these interactions, improving the user experience as a whole.

5.2 User-Based Collaborative Filtering:
One of the standard collaborative filtering strategies is user-based collaborative filtering, sometimes referred to as user-user collaborative filtering. Based on users' previous interactions with things, it finds users who share similar tastes and preferences. The system then suggests products that people with similar tastes to the target user preferred after identifying these similar users. As the user based grows, this strategy may become computationally costly even if it can offer precise recommendation.

5.3 Item-Based Collaborative Filtering:
The similarity between items is the main focus of item-based collaborative filtering as opposed to user-based filtering. It finds products that users are involved with in similar manners and suggests such products to the targeted audience. As the item space is often smaller than the user space, this approach tends to be more scalable. This is because it requires less processing. Because of its efficiency and effectiveness, item-based filtering is frequently employed in recommendation systems.

5.4 Matrix Factorization:

A mathematical method called matrix factorization is used in collaborative filtering to improve the level of recommendations. It involves splitting the matrix of user-item interactions into hidden components. These hidden components, which include preferences, tastes, and attributes, are the hidden qualities or aspects of both users and things. Highly customized and accurate suggestions are made possible by filling in missing data and predicting how a user may interact with items they haven't seen before by approximating the original interaction matrix through matrix factorization.

6 Content-Based Filtering:
Content-Based Filtering is a recommendation system technique that looks at the fundamental characteristics and attributes of items within a system to provide users with customized item recommendations. This strategy depends on your ability to understand and make use of the item's information, content, and features. The basic idea behind content-based filtering is that the system can suggest other items that are similar to those the user has previously interacted with, if the user has expressed interest in particular items. When it comes to recommending products movies, or articles - domains where item attributes strongly influence user preferences in content-based filtering. It works especially well.

6.1 Item Representation:
The essential step in content-based filtering is item representation, which includes containing and providing a detailed description of an item's attributes. In order for content-based filtering to be applied successfully, every item needs to be meaningfully and systematically represented. This representation includes written descriptions, tags, categories, attributes, and other relevant information that help define the item, among other aspects. In order for the recommendation system to understand an item and properly match it with a user's preferences and profile, a strong item representation is essential.
Feature design is essential for accurate item representation. This process involves identifying relevant attributes and converting them into a format that the recommendation system can use. For example, an item's attributes in a movie recommendation system might include its category, the director, the actors, and user ratings. In order to enable relevant comparisons and suggestions, these attributes must be measured and organized.

6.2 User Profile Creation:
The foundational step in content-based filtering is called User Profile Creation, which aims to create a detailed profile that describes a user's interests and preferences. Based on a user's previous interactions, activities, and behavior within the system, this profile is created. It includes analyzing the characteristics of the things that the user has interacted with, rated, or liked. A user's profile might indicate a preference for action and adventure films, for example, if they watch these types of films a lot. Similarly, if a user regularly shows a preference for horror literature, that category will be highlighted in their profile.
User profiles are dynamic, changing as a result of interactions with new content and preferences modifications. It is essential that an exact user profile be created in order for the system to

understand and adapt to evolving user preferences and make recommendations that are appropriate to those preferences.

6.3 Similarity Measures:

An essential part of content-based filtering is Similarity Measures, which provide a way to evaluate how well an item's characteristics match a user's profile. These metrics express the degree of similarity or relationship between an item's attributes and those in the user profile.

7 Hybrid Recommendation Models

7.1 Combining Collaborative and Content-Based Approaches

Hybrid recommendation system is made by combing the algorithms according to need like combine the collaborative and content-based approach. In the hybrid recommendation system first choose the approach which can be used and then combining the output of their result or prediction. The result can be combine based on weighted approach or voting mechanism. Hybrid approaches are used to find the accurate and efficient predictions. In hybrid systems Bayesian networks, clustering is also used.

7.2 Model Weighting

Model weight techniques is the technique that use the weights of the two or more-recommendation algorithm for improving the performance.

Weighting technique predicts the result as the sum of the two or more-recommendation approach weighted result.

According to weighted strategy, c recommendation techniques use to predict the result of user to the item then it computed as follows: -

$$p_{u,i} = \sum_{f}^{c} \sigma_f p_{u,i}^{(f)}$$

Where σ_f denotes the weight.

If there are two recommendation techniques then c=2

$$p_{u,i} = \sigma_1 \cdot p_{u,i}^{(1)} + (1-\sigma_1) \cdot p_{u,i}^{(2)}$$

A weighted hybrid recommendation system that combines the content-based and knowledge-based recommendation techniques. It takes into account user sessions, user profiles, product data, product ratings, user attributes, areas of interest, and user preferences to generate a list of recommended products for each user session.

First, we create the empty list to store the products of user session.

For creating semantic clusters, Semantic clusters group products or users with similar characteristics or attributes, and this information is used to generate personalized recommendations. then we apply the content- based recommendation technique, based on user preferences and user profile, and knowledge-based recommendation technique user's areas of interest and the key features of products. We iterate this for creating semantic clusters and after processing all the semantic clusters this gives us the list which now contains the combined recommendations from both the content-based and knowledge-based approaches. These recommendations are intended to be more personalized by considering user preferences, user attributes, and product features.

7.3 Fusion Techniques

Fusion techniques are used to combine the output result of two or more recommendation algorithm to predict the more accurate recommendation. These techniques used in hybrid recommendation systems that are combining such algorithms like collaborative filtering, content-based filtering, matrix factorization. There are some types of fusion techniques are: -

1 Weighted Fusion: - it is used to combining the result of the algorithms by assigning weights to the outputs.

2 Switching Fusion: - It is used to choose the best performing algorithms for each user or item based on their history or characteristics.

3 Feature Fusion: - It is used to combine the feature extracted from the different algorithms.

4. Cascading Fusion: - It is used to create a cascade recommendation model where output of the one recommendation algorithm serves as the input for the next. For example, a hybrid model could first use collaborative filtering to generate initial recommendations and then refine them using content-based filtering.

5. Contextual Fusion: - It is use to take contextual information like location, time or device to weight the different recommendation algorithms.

7.4 Evaluation Metrics

Evaluation metrics are used to evaluate the performance, accuracy and effectiveness of the recommendation algorithms. In this Recall, Precision, Mean Average Precision (MAP), F-score and metrics are evaluated.

Recall: -

Recall measures the proportion of relevant items that were successfully recommended. It is calculated as the number of correct predictions divided by the total number of relevant items.

Recall = (Number of correct predictions) / (number of total relevant items)

Precision: - Precision measures the proportion of recommended items that are relevant to the user. It is calculated as the number of relevant items recommended divided by the total number of recommended items.

Precision = (Number of correct predictions) / (number of total predictions)

F- Score: The F- Score is the harmonic mean of precision and recall. It provides a balance between precision and recall, which is useful when both are important.

F- Score = (2 * Precision * Recall) / (Precision + Recall)

Mean Average Precision (MAP): MAP considers the average precision at different points of a recommendation list and then computes the mean. It is particularly useful when you have ranked recommendation lists.

8 Challenges and Limitations

8.1 Data sparsity: - As the work shows that user and item interaction play an important role in recommendation system. But the less interaction between the user and item gener1ates the data sparsity problem. Data sparsity problem is that most of the users do not give the rating to the items so that finding the set to users of similar ratings is difficult.

8.2 Cold Start Problem: - Cold start problem is an issue in recommendation system when a new item adds then it has not rated and reviews. So, the system is unable to form any relation between users and items. It makes difficulties to the recommendation system to predict the items.

8.3 Scalability Issues: - Recommendation system works on user and item data. At present a recommendation system works where a large number of user and item dataset. So, to handle the large number of dataset and to predict in real time is an issue to recommendation system. Design the efficient algorithm to handle the large dataset is the challenge in recommendation system.

8.4 Privacy and Securities: - To predict the efficient and appropriate result the recommendation system needs the user's information as much as possible. Because recommendation system works on information. This creates an issue in privacy of users because recommendation system impact's the user choices and behaviour.

As the work shows that recommendation system works on large dataset of users and their choices. So, the security of their personal data and information is issue of concern. Data breach, vulnerable system and cyber-attacks create the security issue of user personal information and behaviour.

9 Implementation and case studies

9.1 Building a Recommendation System

recommendation systems are used to enhanced the user experience by providing personalized content suggestions. For building the recommendation system we should follow the following steps.

9.1.1 Define the objective: - To build a recommendation system, it is crucial to define the objectives clearly. The recommendation system offers readers personalized suggestions based on their reading history, preferences, or specific topics of interest.

9.1.2 Data collection and preprocessing: - An effective recommendation system lies in the quality of data which is used to processing.preprocessing is cleaning the data, handling missing values, and formatting it for analysis. To ensure that the recommendation model is accurate, the dataset must be carefully selected.

9.1.3 Choosing a Recommendation Algorithm: - Recommendation systems typically use collaborative filtering, content-based filtering, or a combination of both. The choice of algorithm depends on the goals and characteristics of the application.

9.1.4 Feature Engineering: - Effective feature engineering is essential for raw data to be transformed into relevant input for the recommendation system.

9.1.5 Model Implementation: - Building the recommendation model involves translating the chosen algorithm into a functional system. Using programming languages such as Python and libraries like scikit-learn or TensorFlow, developers can implement the model.

9.1.6 Evaluation Metrics: - It is essential to measure the recommendation system's performance in order to improve it. Metrics that reveal how well the system is working include precision, recall, and Mean Squared Error (MSE).

9.2 Case study

Amazon:

Amazon is a US-based technology organization that provides cloud computing, streaming, and e-commerce services. Jeff Bezos established the company in 1994, and it has since been recognized as the biggest internet organization globally in terms of revenue as well as the largest online marketplace globally in terms of market value and revenue. According to Alexa's rankings, the company is now ranked fourteenth globally for internet involvement.37.7% of all American e-commerce sales in 2019 were made on Amazon.com.

Recommendation systems are implemented by Amazon in several of its services, such as its streaming and e-commerce platforms. Product sales, targeted advertising, subscriptions (like Amazon Prime), and the fees suppliers pay to be able to sell on the platform are how Amazon makes money from its e-commerce platform. A user will see more products and maybe buy them as well as more ads the longer they stay on the platform. The company's recommendation system is an effective tool because it is able to predict and understand user behaviour and interests, and it generates recommendations based on these observations.

A Technical overview of amazon recommendation system:

Nearly twenty years ago, Amazon implemented a recommendation algorithm for its online store. Before the implementation of this algorithm, Amazon used best-seller lists and human selection to suggest products to users. But according to Amazon, it was discovered that this strategy was fundamentally biased and did not properly offer suggestions to individuals with specialized interests. After that, the business created and implemented an algorithmic system that compares items that a user has bought and evaluated with similar items, and then combining the results into a list of suggested products for the user. The term "item-based collaborative filtering" or "item-to-item collaborative filtering" refers to this novel method of generating recommendations.

The item-to-item collaborative-filtering model's introduction has changed the e-commerce industry and functioned as a key targeted marketing tool. According to Amazon, offering a customized experience will improve users' overall platform experience. But in the process, the business also aims to raise the average order value—the amount spent on the platform each time a user places an order—the amount of time a user spends on the platform, and the total revenue the business makes from each user. Because the company also sells ads, users who use the platform for extended periods of time are likely to see and click on more ads, which generates revenue through ad impressions (views) and clicks.

On its e-commerce platform, Amazon employs recommendation algorithms to present users with various categories of recommendations at various points in time. However, the business is not very transparent about these use cases for the recommendation system. This restricts users' comprehension and agency regarding these tools. However, a number of researchers, journalists, and engineers have recognized a few of the product recommendation categories produced by the company's recommendation system.

10 Applications

The application of recommendation system is very wide in now a day. Recommendation system helps users to discover new content, products, and services according to preferences and need.

This chapter will explore a wide range of applications of recommendation system.

E-commerce: - E- commerce platforms are using the recommendation system at wide range. The goal of using the recommendation system is this domain is to provide the personalized product to users, to increase sales, higher customer engagement and improved user experienced. It is helpful in Product Recommendation recommends products based on the behaviour and preferences of similar users

Cross-selling involves recommending complementary products if a customer buys a camera, the system might suggest camera accessories like a tripod or a camera bag. and Upselling involves recommending higher-priced alternatives to what the user is currently viewing. For example, if a user is looking at a smartphone, the system might suggest a more advanced model.

User personalization create a unique shopping experience for each user by considering their browsing history, past purchases, and even demographic information.

Content Streaming: - Content Streaming is very used in the field of entertainment by offering the TV shows, movies, music. In the

Movies platforms like amazon Prime Video, Netflix, YouTube give the suggestions with the help of recommendation system.

Music platform like Spotify, Apple Music, Savana etc are used the recommendation system.

Social media: - In the social media app recommendation system plays a crucial role in that field. It is used in Friends recommendation on the Facebook, Instagram, and twitter etc.,

Content like post, articles, photos, videos based on like, share and post.

Advertising: - Recommendation system plays an important role in ad campaigns, digital advertising. It is used to target the ads users. It is used to target ads to user based on their interest, online behaviour and past history. It is also used in personalized ads based on user preference.

Job portals: - Recommendation system helps to assist and find the suitable job to the users at the career website and job portals. Recommendation systems provide personalized job suggestions to job seekers based on their skills, experience, and career preferences.

Recommendation systems can help job seekers improve their resumes by suggesting keywords, skills, or certifications to increase their chances of being noticed by employers.

Users can set up job alerts based on their preferences, and recommendation systems send notifications when new job listings match their criteria.

Travel: - Recommendation systems are used very widely in travel and hospitality industry to help plan trips, discover destinations according to their preference.

Traveling websites like booking.com etc are use the recommendation system to suggest hotels, transport options, travel dates etc. It also used to explore the new destinations on their travel history, interests, local culture and weather.

Search Engines: - Recommendation system used in search engines to help users discovering relevant information, product and service.

Search Engines like Google, Bing and Yahoo use the recommendation system to offer search query as search in the type bar.

It also recommends based on the user's recommendation, local business.

REFERENCES

[1] Leung, Cane & Chan, Stephen. (2008). Sentiment Analysis of Product Reviews.

[2] Xing Fang and Justin Zhan.(2015).Sentiment analysis using product review data.

[3] AdinarayanaSalina,KYogeswara Rao Yogi Rao,G. S. N. Murthy.(2017).System from Users Reviews using Sentiment Analysis

[4] TanjimUlHaque,Nudrat Nawal Saber,Faisal Muhammad Shah.(2018).Sentiment analysis on large scale Amazon product reviews.

[5] Uma Devi.(2019) Product review sentiment analysis - a survey

[6] Najma Sultana,PintuKumar,Monika Rani Patra,Sourabh Chandra,(2019), Sentiment analysis for product review.

[7] Pin Ni,YumingLi,Victor Chang.(2020).Recommendation and Sentiment Analysis Based on Consumer Review and Rating.

[8] Ankit Taparia,Tanmay Bagla.(2020).Sentiment Analysis: Predicting Product Reviews' Ratings using Online Customer Reviews.