Novel Weighted Recommendation Technique for Consumer Decision using Collaborative Filtering

### Analp Pathak\*, Dr. B. K. Sharma

### \*Research Scholar, Dr. A. P. J. Abdul Kalam Technical University, Lucknow, 0000-0002-3555-3217 (analp.pathak@gmail.com)

### Principal Scientific Officer and Head Computer Science and Engineering Department, NITRA Technical Campus, Ghaziabad,

### Abstract

Customers are sometimes overwhelmed and make poor decisions due to the rapid growth and variety of online information and e-commerce products. Improving the precision and variety of these methods to meet user needs remains a formidable obstacle. To mitigate the above challenges in this research, a novel weighted recommendation system is created for better consumer decisions using Collaborative Filtering (CF). Initially, an equation is derived to calculate the weight of both the product and review, and then an equation is derived to calculate the similarity between the consumer’s review. Multi-nomial Naïve Bayes (MNB), Multi-Layer Perceptron (MLP), and Logistic Regression (LR) are used as ensemble models for product recommendation. The proposed model is trained and tested using an open-source dataset that is available on the website of Kaggle. Numerical analysis of the proposed model shows that it performed better than other conventional methods in terms of accuracy (0.952), precision (0.908), recall (0.897), F-measure (0.941), and error rate (0.087).

**Keywords** Recommendation System, CF, Multi-nomial Naïve Bayes, Multi-Layer Perceptron, Logistic Regression, and Ensemble Classifiers

# Introduction

Information affecting online business is advancing Information affecting online business is advancing at a dizzying rate because of the arrival of the age of big data. For instance, Amazon receives an average of 900 million customers every day. Users of e-commerce platforms are now confronted with the issue of information explosion. The problems caused by an unnecessary amount of information are addressed by a recommendation system, which can determine the requirements and pursuits of users through an analysis of the collected data from their previous interactions. And then assist those users in making decisions regarding appropriate options. In recent years, researchers have been able to get the great majority of the research efforts that have been done on recommendation systems123.

One of the most effective ways to filter through this information is to rely on the suggestions of other people who are faced with a large amount of data that the normal internet user encounters daily. Recommendations might have come in a variety of formats, including spoken words, letters of recommendation, reports from the news media, public surveys, travel guides, website evaluations, and so on. Over the last 15 years, several large electronic sites have included recommendation systems to facilitate this natural social process. The major goal of these systems is to help consumers find the most relevant and useful information among the vast amount of material accessible online, including but not limited to news articles, web pages, pictures, and more4.

## Recommendation System

A recommendation system is an electronic operator that helps consumers identify the most useful items or services based on the customers' past preferences or tastes. A well-designed recommender system would analyze the preferences that are either inferred or explicitly expressed by each consumer and will automatically offer a selection of items or services5. Delivering outcomes promptly is one of the advantages of a recommendation system. The parallel execution of algorithms can make it easier to generate output without sacrificing performance. These benefits are:

* Large volumes of data can be handled quickly, resulting in increased efficiency.
* A wide range of item types can be suggested.
* It is simple to convert an existing method to parallel processing6. Figure 1 shows three stages of the recommended system as given below.

Diagram

Description automatically generated

**Fig1.** Recommendation System.

There are two primary points of view when it comes to making recommendations: (a) content-based, in which users are recognized by attempting to identify their key characteristics; and (b) CF, in which one can take advantage of the fact that people who had common interests in the past might even agree on one‘s tastes in the future5.

## Collaborative Filtering

CF has emerged as one of the most important methods for developing tailored recommender systems because of its precision and scalability. The core function of CF is to deduce the preferences of users based on the activity data of both the users themselves and other users7. CF uses user ratings and comments to determine whether or not to display material to a certain user. There are two primary classes of recommendation strategies for recommending data: content-based recommendation and CF recommendation approaches. This layer processes data on the production logistics sent from the intelligent gateway layer, including data verification, filtering, fusion, categorization, analysis, and other processing activities. To accomplish the value-added use of mass production logistics data8. The most common approach to recommending is CF. There might be two main categories of CF, namely, those based on the items themselves and those based on the people who use them. Discovering a group of people who have similar interests to the current users is the goal of user-based CF. Item-based CF, which is focused on item similarity, is often estimated by analyzing user behavior. However, there are significant problems with cold starting and data sparsity. problems with informational "cold starts" and conventional methods' lack of detail. As more information is gathered from the internet, such as text, images, and tag data, which includes specialized demand data and detailed project information, this risk might be reduced. Using a person's past choices and the opinions of other individuals with similar tastes, a CF method might offer new items to the user or assess the items/usefulness hotels for the user. In a typical CF setup, there will be m user lists Uand n item lists. . The set of things on which each user u I have commented is designated by . An open rating, usually a numerical scale, is one way in which users might express their thoughts, but ratings can also be derived inferentially from user behavior such as purchasing behavior, time spent on site, and even connecting behavior. Note that if anyone writes that , then might be a collection of zero elements. For a key user, also referred to as the active agent, the goal of a CF process is to determine an item's probability, and this likelihood may take one of two forms9.

**Recommendation:** It provides a list of the top N hotels according to the user's overall satisfaction with each establishment . must include on the suggested list both things and hotels that the present consumer has not previously bought. In certain circles, it is also referred to as a Top-N recommendation method interface10.

**Prediction:** is a numeric number that reflects the anticipated possibility of the item for the active user , depending on the user's behavior. This value is expressed as a percentage, and it may range from 0 to 1. The information that was supplied by indicates that this forecasted number is somewhere within the same range of 1 to 59.

The CF procedure is shown in Figure 2. It begins with the inputting of data, which is followed by the setting of parameters according to the format of the neighbors in the third stage. In the third stage, it makes suggestions for the production of brand-new things.

Rate input data

Neighbors format

Generation Recommended

**Fig 2.** The CF process10.

Bi et. al., (2020)11 intended a recommendation system based on deep neural networks, whereby item average rating, user basic data, and use basic data are all employed, along with item basic data user ID. The algorithm's basic concept is built on using deep neural networks to construct a regression model that can anticipate users' evaluations. Four separate types of neural network layers are used to construct a user feature matrix and an item feature matrix, respectively, using the user data. After this Lai et. al., (2021)12 presented an innovative approach to rating prediction using a deep learning model with semantic components based on attention-based gated recurrent units (GRUs). User reviews are analyzed into individual words and utilize Latent Dirichlet Allocation and the attention weights of the selected words to produce the aspect-based attention semantic vectors from these reviews. Next, the aspect-based attention semantic vectors are used in conjunction with the XGBoost technique to predict user preference ratings. The experimental results confirm the effectiveness of the suggested strategy in raising prediction accuracy above and above that of conventional methods. In the same context, Chen et. al., (2021)13 intended a CF and recommendation system based on dynamic clustering and a double-layer network (DCCF-DN). When compared to other algorithms, DCCF-simulation DN's results show that it improves recommendation performance. However, these machine learning techniques have some limitations such as popularity bias, lack of transparency, limited data, and many more. Therefore, Tahira et. al., (2022)14 developed a recommender system that makes use of online customer evaluations within the context of the Internet of Things to match the characteristics of a product that are significant to the buyer. Experimental research will be carried out to investigate how the impact of the suggested algorithm changes with hedonistic and utilitarian items. This is because the influence of recommender systems is dependent on the nature of the product being recommended. Followed by Zhao (2022)15 combined the traditional CF recommendation system with a subsystem based on cluster analysis using a genetic algorithm. From various experiments, the response time of a classic CF recommendation system increases linearly with the number of consumers, while the response time of a genetic clustering-based CF recommendation system remains constant regardless of the size of the user base. As Hybrid recommendation systems combine multiple recommendation algorithms to provide more accurate and diverse recommendations therefore, Sharma et. al., (2022)16 predicted book suggestions using a hybrid system's prediction algorithm. CF and content-based filtering were used to evaluate the proposed system. Experiments reveal that the suggested hybrid filtering strategy is superior to both traditional CF and content-based filtering. But these hybrid systems have some drawbacks such as complexity, increase computational environment, and integration challenges therefore, these are not suitable for recommendation systems further to mitigate these issues Han et. al., (2022)17 revealed a multilayer fuzzy perception similarity algorithm (MFPS) to perceive and interpret user similarities, ultimately leading to an improvement in the suggestion quality in a way that is subjective. This is the first time that triangular fuzzy numbers have been used for RSs, and it was done in this work. In addition, a layered structure is developed to enhance a particular RS's capacity to subjectively perceive similarities across user qualities. The results of the studies described above demonstrated that MFPS stood out above other competitor baselines. Followed by Mohd Sabri and Nurul (2022)18 designed and tested a book suggestion using a method called item-based CF. The book suggestions were accurately forecasted by the recommendation system, which received an F-measure % that was acceptable at 80.38 %. The Precision, the Recall, and the F-measure were the metrics that were used in the assessment of the book suggestion prototype. Based on this first investigation, the book recommender has effectively suggested reading materials that have satisfactory performance, as measured by their F-measure value of 80.38 %. Fuzzy systems can be useful in recommendation systems as they allow for the handling of imprecise or uncertain information. But there are some drawbacks to using fuzzy systems in recommendation systems such as scalability, overfitting, and interpretability.

Current research in recommendation systems is focused on improving the accuracy and relevance of recommendations, as well as addressing issues related to data privacy, fairness, and transparency. Some of the latest trends in recommendation research include:

Multi-objective optimization: This approach involves optimizing recommendation systems for multiple objectives, such as accuracy, diversity, novelty, and serendipity.

* Explainability: Researchers are developing new methods to provide users with explanations for why they are being recommended certain items or products.
* Context-aware recommendation: This involves incorporating contextual information, such as location, time, and user behavior, into the recommendation process to provide more relevant and personalized recommendations.
* Hybrid recommendation: This approach combines multiple recommendation techniques to provide more accurate and diverse recommendations.

Despite the advancements in recommendation systems, there are still some limitations that need to be addressed. Some of these limitations include:

* Cold start problem: This refers to the difficulty of making recommendations for new users or items with limited data.
* Data sparsity: This is a common problem in recommendation systems where there is a large number of items and users, but only a small fraction of them have ratings or interactions.
* Algorithmic bias: This refers to the potential for recommendation algorithms to produce biased recommendations based on factors such as race, gender, or socioeconomic status.
* Privacy concerns: Recommendation systems typically require access to users' data, which raises concerns about data privacy and security.

In this analysis, a novel weighted recommendation system is created for better consumer decisions using CF. Machine learning models MNB, MLP, and LR are combined to form an ensemble model for product recommendation. The dataset that is used to train and test the proposed ensembled model is an open-source dataset and is easily available on the website of Kaggle. Section 1 provides an introduction related to the recommender system and CF, then displays the literature review that discusses the latest research that is performed on the recommender system section 2 displays the background study, section 3 depicts the problem formulation, section 4 discusses the objectives that are obtained after evaluating the previous work, section 5 discuss the methodology section, section 6 displays the results section and finally section 7 displays the conclusion and future scope of the research.

# Background Study

More people are interested in review-based recommender systems now than ever before due to the proliferation of social networking sites. The development of such systems is motivated by a desire to put to good use the insightful data included in users' written critiques. With the use of sentiment analysis, this study introduces a CF recommendation system. Sentiment analysis is used on a data set consisting of 7210 reviews of 221 novels culled from the Amazon website to achieve this end. To get user feedback, an ensemble of models is used. For ensemble modeling, an approach based on weighted vote classifiers is also used. Java Web Crawlers were used to get the necessary information from Amazon.com. Comments made by Amazon customers on individual book titles, such as Business Intelligence, were the only source of information. Sentiment analysis is done using several approaches, including text normalization and ensemble techniques. Users were more likely to suggest popular items and recommender system performance was improved when sentiment analysis of customer reviews was included. As a result of incorporating sentiment analysis into recommender systems, this research demonstrates significant gains in the effectiveness of the latter19.

# Problem Formulation

Customers with an Internet connection may shop for necessities whenever and wherever they choose. People prefer to buy books, for instance, from online retailers like Amazon. In addition, users have the option to provide written feedback on a product, which may ultimately influence the purchasing decisions of other consumers. It will be impossible to create relevant queries to retrieve meaningful data from such a massive volume of data if the substance of documents is unknown. Users need aid in comparing papers, categorizing them according to relevance, and uncovering patterns. Customers are often overburdened and make bad judgments because of the rapid increase and diversity of information accessible on the internet and the rapid creation of new e-commerce products (purchasing items, product comparison, various auctions, etc.). That's a bad thing since it cuts into profits.

# Research Objectives

* To study and evaluate previous studies on recommendation systems for consumer decisions.
* To create a novel recommendation technique by implementing a new formula for better consumer decisions by using CF.
* To prove the robustness of the proposed model by comparing it with another conventional model in terms of accuracy and other performance evaluation parameters.

# Research Methodology

The concept of designed architecture is examined in the context of research methodology. The term "research methodology" refers to the process through which authors outline the specifics of how they plan to conduct their studies, and the name "research methodology" itself refers to this process. It's a way of approaching a study problem that is reasonable and systematic. It is common practice for authors to provide a brief explanation of their technique to guarantee that their research produces accurate and trustworthy results and achieves the stated goals and objectives. The process considers, not just the data itself but also its origins, potential uses, and methods of acquisition. After it, ensembling models are used to ensemble the methodology, and then a recommendation system is modeled for better consumer decisions.

## Technique Used

In this section, a brief description of all the techniques which are taken into consideration is given below:

* **Ensemble Classifiers:**

The Ensemble Methods are used in the construction of the sentiment analysis model. The classification of comments is accomplished using an ensemble technique of modeling based on many classification methods. Meta-algorithmically, ensemble approaches combine the insights of many different intelligent models into a single prediction algorithm. The goal of most ensemble algorithms is to improve the overall performance of the algorithm by combining the efforts of several weak learners. Each ensemble approach has its unique focus, with bagging looking to reduce variance, boosting to reduce bias, and stacking to raise prediction accuracy. The ensemble technique relies on merging several classifiers to get better results than anyone classifier could achieve alone. Multiple models, including Multinominal Nave Bayes (MNB), Multi-layer Perceptron (MLP), and Logistic Regression (LR), are used in this study. Such algorithms are used in the predicting phase of supervised learning.

* **Multi-nomial Naïve Bayes**

For precise tallies, the MNB model is the only option. For text classification tasks, the MNB classifier is considered, where each document d is characterized by a characteristic vector (f1, f2, … fn) containing the integer number of the occurrence of each word in the document. To calculate the conditional probability P (d|) of document d gave class c in the MNB model, the formula is given below19:

(1)

A Naive Bayes classification may provide the following final equation for the most likely categories:

(2)

The next step is to calculate a probability score. As soon as a term meets the criteria of , it is added to the document's vocabulary. Therefore, to get , divide the total number of keywords in class by , where is the number of times the word appears in document from class . Then evaluate the likelihood of a document based on its classification as follows:

(3)

where v is an intersection of all possible word categories. From the training dataset, the possibility of in can be calculated as follows [19]:

(4)

* **Multi-Layer Perceptron**

If the hidden layers utilize a nonlinear activation function, such as the logistic function and the model incorporates estimated weights between the inputs and hidden layers, the model becomes nonlinear. A multilayer perceptron is a name given to the final model (MLP)20. In a typical MLP architecture, there are three layers: input, hidden, and output as shown in Figure 3. It is a kind of artificial neural network (ANN) known as an early MLP. The early MLP, like the perceptron, uses error corrective learning to set the weight of the connections between its layers, adjusting the weight as needed to minimize the difference between the predicted and actual output. Multi-layer error-correcting learning, however, is not possible21.

Input 1

Input 2

Input 3

Input layer

Hidden layer

Output layer

**Fig3.** MLP22.

Because of this, the early MLP employs the usage of a random integer to decide the connected weight between the hidden layer and the input layer. On the other hand, error corrective learning is used to alter the connection weight that exists between the hidden layer and the output layer. MLP now has the capability, thanks to the BP method, to modify the connection weight on a layer-by-layer basis. If an MLP is constructed with only a single hidden layer, which is comprised of three neurons, The activation function of the network is the rectified linear unit function, which produces23:

(5)

Where f(x) is an activation function.

* **Logistic Regression**

The LR method is widely used for linear classification. It enables the formation of a multivariate regression by allowing a connection to arise between a variable that is independent and dependent variables. apply the binary logistic regression technique to investigate the P2P platform risk features. According to the findings, platform stability, profitability, risk management, liquidity, and transparency may all be used to estimate the likelihood that a platform would have issues24. LR is the framework of multivariate analysis that may be used to forecast the existence or absence of a function or consequences based on the values of several different predictor variables in a series. This can be helpful in several different contexts25.

(6)

Where p represents the probability, and β\_0 indicates the value of the intercept. With the assistance of a line of regression, the LR just splits the data into two distinct categories, as seen in Figure 425.

Dependent variable

Independent variable

Line of regression

Datapoints

**Fig 4**. Logistics Regression26.

## Proposed Methodology

The architecture of the proposed research work has been shown in Figure 5 and the work process of this methodology has been given in the below steps.

initially x=1;

x>=3?

Ensemble model

Recommendation of product

Concatenation of reviews and Products

Calculate weight for both product and reviews.

Calculate similarity x (i, j)

X++

Yes

No

Reviews on Products

User

U1

U2

U3

.

.

Un

**Pre-processing**

Case folding

Tokenization

Replacement of noisy data

Products (P1, P2…. Pn)

**Fig 5.** Block Diagram of Proposed Methodology

**Step 1: Data Collection**

Data collection is the first step of this work. In this step, data is collected from the websites in terms of users (U1, U2, U3, … Un), products (P1, P2, P3, … Pn), and reviews of the users. The reviews which are taken into consideration are given for the same products by all the users or customers.

**Step 2: Pre-processing of the Data**

After collecting data, pre-processing of this data is done in this step. Pre-processing of the data is the most important step. In the methodology, pre-processing is done by using case folding, tokenization, and replacement of noisy data to enhance the performance of the methodology.

**Step 3: Concatenation of the Reviews and Products**

In this step 3, after pre-processing the data, the concatenation of the reviews and products is done. Products and the reviews of the products which are given by several users are linked to each other.

**Step 4: Calculate the Weight for both Product and Review**

After done concatenation of the product and reviews in the previous step, here weight is calculated for products and reviews of the user by using a newly created formula as given below:

(7)

Where,

= Denotes the total weight of the product (P) and review by users

= Review (R) of the Product given by the user ) where i = 1, 2, 3, … n

**Step 5: Calculate the Similarity of users’ reviews.**

In this step 5, the similarity of users’ reviews is calculated by using a formula as given below:

(8)

Where,

= Denotes the similarity of two consumers’ reviews on the same products

d = Dampening factor

**Step 6: Ensemble Model**

Ensemble models are applied to the data which is obtained in step 5 by calculating similarities of users’ reviews. Multi-nomial Naïve Bayes (MNB), Multi-Layer Perceptron (MLP), and Logistic Regression are employed for modeling with the help of a few classifier algorithms, which collaborate to assign labels to the feedback.

**Step 7: Recommendation of products**

This is the last step of the whole process. After completing all the processes products are recommended to the consumers as per their needs.

## Proposed Algorithm

ALGORITHM: CONSUMER DECISION USING COLLABORATIVE FILTERING

**Start**

Text

Description automatically generated

**Phase – I:** Data Collection

1. Data collection should be accurate based on consumer reviews [Uj] based on each product [Pi].
2. Data contains each product's [Pi] review, available on every e-commerce website.

**Phase – II:** Data Pre-processing

1. The review for each product and its conciseness will be in Text format. So, it is a Text classification problem.
2. Each review will be cleaned with punctuation, escape sequence, stop words, emoji, unwanted spaces, and digits, then apply WordNet Lemmatization.
3. Each review's conciseness will be converted into a numerical value with either OneHotEncoding.
4. The rating will be visualized with a histogram plot.

**Phase – III:** Concatenated Review with Weight Calculation

1. Each product's threshold will be a minimum with three ratings.
2. Each product's review's conciseness will be considered the top two most frequent ratings.
3. The product will be concatenated with its topmost reviews.
4. To determine each product's review-dominancy, Term Frequency and Inverse Document Frequency (Tf-Idf) will be calculated to understand the context of the corpus, which can be calculated as:

*where,*

→ denotes the total Weight of the product (P*i*) per user's review.

→ Review (R) of the Product given by the user )

**Phase – IV:** Interlinked between more than two webpages

1. The maximum occurrence of n-gram tokens for each product's review will be considered a dominant specification.
2. For each product [Pi] minimum of three reviews ( will be extracted to find the occurrence of the most frequent tokens. The damping factor, d, is the likelihood that a user will click on a link, and (1-d) is for non-direct connections to any webpage.

***where***,

→ similarity of two consumers' product review

d = Dampening factor

1. The Page Ranking Algorithm is helpful since it ranks webpages by significance, but it is limited because it ranks at indexing time, not retrieval time, and if a webpage with no outlinks is detected, the user goes to a random bookmark.

**Phase – V:** Ensemble algorithms to find the best algorithm

1. Ensemble Methods build sentiment analysis models. Ensembles combine classifiers to improve outcomes. Multi-nomial Naïve Bayes (**MNB**), Multi-layer Perceptron (**MLP**), and Logistic Regression(**LR**) are used.
2. Provide classification report.

Calculate the best **A**ccuracy, **P**recision, **F**1 – **S**core, and **R**ecall for the best algorithm

**End**

# Results and Discussion

In this section, the result demonstrated that are generated based on the proposed methodology. Also, there is a brief explanation of the dataset that is used for the training and testing of the model. Finally, the proposed model is compared with another conventional model to investigate its efficiency of it.

* **Dataset**

The dataset that is used in the proposed methodology is known as Amazon Products Recommendation 2016-2017. It is an open-source dataset that is easily available on the website of Kaggle. It is a vast dataset that Amazon published in 2017 that can be used for computer vision applications such as instance segmentation, object identification, key point recognition, and semantic segmentation. It is a collection of 34661 reviews by the same number of consumers as reviews. In this data, these reviews given on various electronic devices which are sold by Amazon are given by the consumers. These devices are from two brands such as Amazon and Amazon digital services27.

Evaluating the performance of the recommender system is an essential step that must be taken to ensure that it can be expanded successfully. There are a few well-known measurements that might be used to analyze the precision or performance of recommender systems. These measurements include accuracy, precision, recall, error rate, and F-measure (to balance the two measures). These measurements can accurately represent the recommendation system's performance.

Precision is calculated by;

(9)

Where,

TP = True positive value,

FP = False Positive value.

Recall obtained by using.

(10)

Where,

FN = False Negative

The F1-score measured by using the formula is;

(11)

To calculate the accuracy of the model a formula is used which is given below;

(12)

The results which are calculated for the proposed methods are given in R1, R2, R3, R4, and R5 as shown below;

**R1: MNB**

This result presents the results of calculating the precision, recall, F1-measure, and accuracy of the proposed MNB for the products and reviews that are taken from the dataset. A graphical representation of this result is shown in Figure 6.



**Fig 6.** Graph of MNB Results

**R2: MLP**

The value of the precision, recall, f 1-measure, and accuracy are calculated for the proposed MLP in this result in the context of reviews and products. The graphical representation of this result is shown in Figure 7 as given below.



**Fig 7.** Graph of MLP results

**R3: LR**

In the context of reviews and products, the value of the precision, recall, f 1-measure, and accuracy are determined for the suggested LR in this result. The graphical representation of the obtained result is shown in Figure 8 which is presented below.



**Fig8.** Graph of LR results

**R4: Ensemble Classifiers**

This result is obtained for ensemble classifiers. MNB, MLP, and LR are combined in an ensemble model to obtain better results in comparison to the previous research. In the context of reviews and products, the value of the precision, recall, f1-measure, and accuracy are determined for the suggested MLP in this result. Figure 9 shows the results of the ensemble classifier graphically as shown below.



**Fig 9.** Graph of Ensemble Classifier results

**R5: Error rate of all techniques**

In this result, firstly the error rate of all techniques is calculated separately. Then, the error rate of the ensemble model (MNB+MLP+LR) is calculated to get a lower error than the previous work. The error rate of MNB, MLP, LR, and ensemble classifiers are 0.108, 0.153, 0.1097, and 0.087 respectively. Figure 10 shows the error rate of all classifiers in a graphical manner.



**Fig 10.** Graph of Error rate results for all techniques

These results were obtained by repeating the model for each classifier 3 times. It was revealed that the ensemble model with an accuracy rate of 95.2% had higher accuracy than the other algorithms for modeling.

# Comparison Results

In this section, the proposed model is compared with other conventional methods such as MNB, MLP, and LR. It is compared based on positive metrics parameters such as accuracy, precision, recall, f1-measure, and error rate. Figure 11 shows the comparison of the conventional technique with the proposed model based on precision, recall, f1-measure, and accuracy and it is seen the proposed model is higher among all the methods. Figure 12 shows the comparison of the conventional technique with the proposed model based on error rate and it is seen the proposed model is a lower error rate among all the methods.



**Fig 11.** Comparison of the proposed work's performance to that of similar current schemes in terms of Precision, Recall, F1-measure, and Accuracy



**Fig 12.** Comparison graph of Error rate among all techniques

# Conclusion and Future Work

This work focuses mainly on proposing novel and group trust models to enhance consumer decisions. In the work, CF is used to create a novel weighted recommendation system to better consumer decisions. Using CF, a unique weighted recommendation system has been developed in this study to improve consumer choice-making. The technique includes the creation of a formula to compute the weight of both the product and the review, as well as a calculation to assess the similarity between different consumers' reviews. In addition to that, a comparison of the outcomes is included in the study. According to the results the suggested model outperformed other traditional approaches in many different aspects, including accuracy (0.952), precision (0.908), recall (0.0897), F-measure (0.941), and error rate (0.087). In future work, combining multiple recommendation techniques, such as content-based and CF, with weighted collaborative filtering can improve the accuracy and diversity of recommendations.

# References

Yadav, Vikash, Rati Shukla, Aprna Tripathi, and Anamika Maurya. "A new approach for movie recommender system using K-means Clustering and PCA." *Journal of Scientific & Industrial Research 80, no. 02* (2021): 159-165.

Shen, Jian, Tianqi Zhou, and Lina Chen. "Collaborative filtering-based recommendation system for big data." *International Journal of Computational Science and Engineering 21, no. 2* (2020): 219-225.

Ding, Linlin, Baishuo Han, Shu Wang, Xiaoguang Li, and Baoyan Song. "User-centered recommendation using us-elm based on dynamic graph model in e-commerce." *International Journal of Machine Learning and Cybernetics 10* (2019): 693-703.

Raza, Shaina, and Chen Ding. "News recommender system: a review of recent progress, challenges, and opportunities." *Artificial Intelligence Review* (2022): 1-52.

Geetha, G., M. Safa, C. Fancy, and D. Saranya. "A hybrid approach using collaborative filtering and content-based filtering for recommender system." *In Journal of Physics: Conference Series, vol. 1000, no. 1, p. 012101*. IOP Publishing, 2018.

Sivaramakrishnan, N., and V. Subramaniyaswamy. "GPU-based Collaborative Filtering Recommendation System using Task parallelism approach." 2018 *2nd International Conference on I-SMAC* (IoT in Social, Mobile, Analytics, and Cloud) (I-SMAC) I-SMAC (IoT in Social, Mobile, Analytics, and Cloud) (I-SMAC), 2018 2nd *International Conference on. IEEE*, 2018

Deng, Xiaoyi, Fuzhen Zhuang, and Zhiguo Zhu. "Neural variational collaborative filtering with side information for the top-K recommendation." *International Journal of Machine Learning and Cybernetics* 10 (2019): 3273-3284.

Kunamalla, Chiranjeevi, and K. Shahu Chatrapati. "Network Path Optimization Strategy using Collaborative Cache for Delay Tolerant Networks." *Journal of Scientific & Industrial Research 82*, no. 02 (2023): 210-218.

Valarmathi, P., R. Dhanalakshmi, and Narendran Rajagopalan. "Boosting a Hybrid Model Recommendation System for Sparse Data using Collaborative Filtering and Deep Learning." *Journal of Scientific & Industrial Research* *Vol. 79,* (2020).

Wang, Wei, et al. "Trust-enhanced collaborative filtering for personalized point of interests recommendation." *IEEE Transactions on Industrial Informatics 16.9* (2019): 6124-6132.

Bi, Jian-Wu, Yang Liu, and Zhi-Ping Fan. "A deep neural networks-based recommendation algorithm using user and item basic data." International Journal of Machine Learning and Cybernetics 11 (2020): 763-777.

Lai, Chin-Hui, Duen-Ren Liu, and Kun-Sin Lien. "A hybrid of XGBoost and aspect-based review mining with attention neural network for user preference prediction." International Journal of Machine Learning and Cybernetics 12 (2021): 1203-1217.

Chen, Jianrui, Bo Wang, Zhiping Ouyang, and Zhihui Wang. "Dynamic clustering collaborative filtering recommendation algorithm based on double-layer network." International Journal of Machine Learning and Cybernetics 12 (2021): 1097-1113.

Tahira, Anum, Walayat Hussain, and Arif Ali. "Based Recommender System for Hedonic and Utilitarian Products in IoT Framework." In IoT as a Service: 7th EAI International Conference, IoTaaS 2021, Sydney, Australia, December 13–14, 2021, Proceedings, pp. 221-232. Cham: Springer International Publishing, 2022.

Zhao, Yan. "Design of Garment Style Recommendation System Based on Interactive Genetic Algorithm." Computational Intelligence and Neuroscience 2022 (2022).

Sharma, Sunny, Vijay Rana, and Manisha Malhotra. "Automatic recommendation system based on hybrid filtering algorithm." Education and Information Technologies (2022): 1-16.

Han, Di, Yijun Chen, and Shuya Zhang. "Implicit social recommendation algorithm based on multilayer fuzzy perception similarity." International Journal of Machine Learning and Cybernetics 13, no. 2 (2022): 357-369.

Mohd Sabri, Norlina, and Nurul Azeymasnita Jaffar. "Book recommendation based on collaborative filtering technique." ESTEEM Academic Journal 18 (2022): 92-103.

Abbasi, Fatemeh, and Ameneh Khadivar. "Collaborative Filtering Recommendation System through Sentiment Analysis." *Turkish Journal of Computer and Mathematics Education (TURCOMAT)* 12, no. 14 (2021): 1843-1853.

Stanković, N. Lj, Marija Dragovan Blagojević, M. Ž. Papić, and D. Karuović. "Artificial Neural Network Model for Prediction of Students’ Success in Learning Programming." *Journal of Scientific & Industrial Research 80, no. 03* (2021): 249-254.

Ghosh, Monalisa, and Goutam Sanyal. "An ensemble approach to stabilize the features for multi-domain sentiment analysis using supervised machine learning." *Journal of Big Data* 5 (2018): 1-25.

Zhang, Jinghua, Chen Li, Yimin Yin, Jiawei Zhang, and Marcin Grzegorzek. "Applications of artificial neural networks in microorganism image analysis: a comprehensive review from conventional multilayer perceptron to the popular convolutional neural network and potential visual transformer." *Artificial Intelligence Review* (2022): 1-58.

<https://towardsdatascience.com/recsys-series-part-5-neural-matrix-factorization-for-collaborative-filtering-a0aebfe15883>

Pandian, R., A. Sabarivani, R. Ramadevi, and N. R. Krishnamoorthy. "Effect of data preprocessing in the detection of epilepsy using machine learning techniques." *Journal of Scientific & Industrial Research 80, no. 12* (2022): 1066-1077.

Bakshi, Garima, Rati Shukla, Vikash Yadav, Aman Dahiya, Rohit Anand, Nidhi Sindhwani, and Harinder Singh. "An optimized approach for feature extraction in multi-relational statistical learning." *Journal of Scientific & Industrial Research* *80*, no. 6 (2021): 537-542.

Greco, Claudia, Pasquale Pace, Stefano Basagni, and Giancarlo Fortino. "Jamming detection at the edge of drone networks using Multi-layer Perceptrons and Decision Trees." *Applied Soft Computing* 111 (2021): 107806.

Ribokaitė, Lina. "Outlier detection in multidimensional streaming data." Ph.D. diss., Vilniaus universitetas, 2021.

<https://www.kaggle.com/code/haojie98/amazon-product-recommendations/data>