International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

An Intelligent Weighted Recommendation Technique utilizing Ensemble System for Enhanced Prediction Accuracy for better Consumer Decision

Analp Pathak[1, 2], Dr. B. K. Sharma[3]

Submitted: 06/02/2024 **Revised**: 14/03/2024 **Accepted**: 20/03/2024

Abstract: A recommendation system can intelligently employ machine learning algorithms to suggest diverse options tailored to user interests based on multiple sources of information. Most recommendation systems heavily rely on the collaborative filtering (CF) approach, where user preference data is amalgamated with that of other users to predict additional items of potential interest to the consumer. In this study, an innovative weighted recommendation system is developed to enhance consumer decision-making using CF. Equations to calculate the weight of both the product and review, as well as the similarity between consumer reviews, are devised in the methodology. The methodology employs machine learning techniques such as Multi-nomial Naïve Bayes (MNB), Multi-Layer Perceptron (MLP), and Logistic Regression (LR) as intelligent ensemble models. Ensemble Classifiers (MNB+MLP+LR) are utilized to implement the methodology's results, aiming for superior outcomes compared to previous studies. The proposed model is trained and tested using an open-source dataset. Numerical analysis of the proposed model demonstrates its superior performance over conventional methods in terms of accuracy (0.952), precision (0.908), recall (0.897), F-measure (0.941), error rate (0.087), and other metrics.

Keywords: Intelligent Recommendation System, Machine Learning, Multi-nomial Naïve Bayes, Multi-Layer Perceptron, Logistic Regression, and Ensemble Classifiers

1 Introduction

Information affecting to online business is advancing Information affecting to 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 systems [1][2][3].

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 more [4].

[1]Research Scholar, Dr. A. P. J. Abdul Kalam technical University, Lucknow
[2]Department of Information Technology, KIET Group of Institutions, Delhi-NCR, Ghaziabad, India
[3]Principal Scientific Officer and Head Computer Science and Engineering Department, NITRA Technical Campus, Ghaziabad
[1]analp.pathak@gmail.com
[2]drbksharma@nitratextile.org

1.1 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. These preferences or tastes may be gleaned from the customers' purchase histories. Actually, as the significance of online business continues to grow, the recommender system will become an increasingly important instrument for the execution of personalized marketing. 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 services [5]. When employing a parallel strategy to generate recommendations are anticipate getting several benefits. 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 processing [6]. Figure 1 shows three stages of the recommended system as given below.

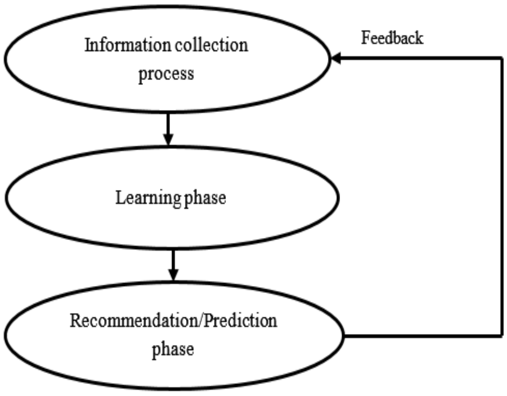

Fig1. Recommendation System.

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, which requires individual information that is difficult to terminate; 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 future [5].

1.2 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 users [7]. CF uses user ratings and comments to determine whether or not to display material to a certain user. It looks through a massive user base to locate folks who have similar preferences as the current user. There are two primary classes of recommendation strategies for recommending data: content-based recommendation and CF recommendation approaches. 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 U= $\{u_1, u_2, \dots, u_m\}$and n item lists. $I = \{i_1, i_2, \dots, i_n\}$. The set of things on which each user u i has commented is designated by I_{u_i}. 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 write that $I_{u_i} \subseteq I$, then I_{u_i} might be a collection of zero elements. For a key user, $U_a \subseteq I$ 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 forms [8].

Recommendation: It provides a list of the top N hotels according to the user's overall satisfaction with each establishment $I_r \subset I$. $I_r \subseteq I_{u_a} = \phi$ 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 interface [9].

Prediction: $P_{a,i}$ is a numeric number that reflects the anticipated possibility of the item $i_j \notin I_{u_a}$ for the active user u_a, 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 u_a indicates that this forecasted number is somewhere within the same range of 1 to 5 [8].

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.

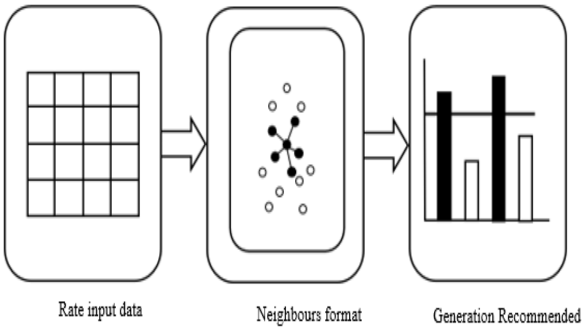

Fig2. The CF process [9].

2 Related Work

filter is discussed below:

Mohd Sabri and Nurul (2022) [10] 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 effectiveness of the CF approach used in the book recommendation system. CF is one of the strategies that has been the most extensively modified and used for the recommendation system. The dataset that was used in this initial investigation was retrieved from the Kaggle website, where it had already been put through the rigors of the 10-fold cross-validation method. A random selection of one thousand data had been made, with nine hundreds of those data designated for use in the training phase of the process and one hundred reserved for use in the testing phase. 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 %.

Han et. al., (2022) [11] 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. A selection of exemplary cutting-edge similarity techniques in a nonlinear manner is improved to give them the ability to understand human emotions. 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 due to the effectiveness, explicability, and consistency of the suggestions it provided.

Sharma et. al., (2022) [12] predicted book suggestions using a hybrid system's prediction algorithm The proposed system's three stages may be broken down into the two types of filtering they combine: content-based filtering and CF. First, it compares the current user's profile to a database of users to find others with similar characteristics. Using the user's profile and the item's contents as inputs, the system then selects the candidate's item for each comparable user in the second phase. It is time to provide recommendations to the end user when the Resnick prediction equation has been used to determine an item's prediction value. Current state-of-the-art recommendation techniques, including CF and content-based filtering, were used to evaluate our proposed system. Experiments reveal that the suggested hybrid filtering strategy is superior to both traditional CF and content-based filtering.

Zhao (2022) [13] combined the traditional CF recommendation system with a subsystem based on cluster analysis using a genetic algorithm. Both the idea and execution of the system are shown. In addition, only users who are part of the user base's clusters are looked for when determining who is the nearest neighbor. Because of this, the system is now a genetic clustering-based CF recommender system. 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.

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. Before making any suggestions, the algorithm analyses the product to determine which aspects are most important to the consumer. Following that, it does an aspect-based sentiment classification to locate those elements in customer evaluations and assigns a sentiment score to each of those features. The credibility of each customer review determines how much weight it is given. 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.

Chen et. al., (2021) [15] intended a CF and recommendation system based on dynamic clustering and a double-layer network (DCCF-DN). To begin, a model of user and object attributions in the form of a double-layer network is built. To further describe the interaction between users and objects, the rounding-forgetting function is used to simulate interest. Second, clustering equations are used to group users and goods into distinct communities, much like how the evolutionary process does. Measures of similarity are used in each group to identify nearby nodes that have common interests. At last, expected scores are obtained, and a list of the top N recommendations for the intended audience. The effectiveness of the DCCF-DN algorithm has been shown throughout a range of studies. When compared to other algorithms, DCCF-simulation DN's results show that it improves recommendation performance. On the other hand, suitable parameters are necessary for better outcomes. K1 − K8.

Lai et. al., (2021) [16] presented an innovative approach to rating prediction using a deep learning model with semantic components based on attention-based gated recurrent units (GRUs). Here, a two-step process is suggested for extracting feature aspects from user preferences, one that integrates the word attention technique with review semantics. To begin, a bidirectional GRU neural network is used to extract key phrases from user evaluations of their attention. In the second step, 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.

Bi et. al., (2020) [17] intended a recommendation system based on deep neural networks, whereby item average rating, user basic data (user gender, user age, user profession, user ID), and use basic data are all employed, along with item basic data (name, category, ID), user ID, and 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

and the item data. It is put through three tests using data taken directly from the Movie Lens website to prove that the suggested method works. In addition to outperforming state-of-the-art CF recommendation algorithms, experiments show that the proposed method also addresses the data sparsity issue and the cold-start problem that would otherwise emerge.

There is a wide range of authors who used the technique and presented their discoveries, as can be found in table 1.

S. No.	Authors and Years	Techniques/ Methods	Outcomes
1	**Mohd Sabri and Nurul (2022) [10]**	Item Based Collaborative Filtering technique	The experimental investigation finds that the book recommender achieves an adequate performance, based on an F-measure value of 80.38 percent.
2	**Han et. al., (2022) [11]**	MFPS	The results of experiments demonstrated that MFPS's effective, explanatory, and steady suggestions set it apart from its rivals.
3	**Sharma et. al., (2022) [12]**	collaborative filtering and content-based filtering	The experimental results demonstrate the superiority of the suggested hybrid filtering strategy over the state-of-the-art collaborative filtering and content-based filtering methods.
4	**Zhao (2022) [13]**	genetic algorithm	The experiments show that as the number of users increases, the response time of the traditional CF recommendation system increases linearly, but the response time of the genetic clustering-based system does not change.
5	**Tahira et. al., (2022) [4]**	aspect-based sentiment analysis	The research indicates that 91% of the target audience was satisfied with the presented model.

6	**Chen et. al., (2021) [15]**	DCCF-DN	The simulation results show that DCCF-DN improves upon the recommendation performance of existing algorithms in several ways. The best outcomes, however, need more realistic starting points. K1 − K8.
7	**Lai et. al., (2021) [16]**	Attention based GRU	The experimental results confirm the effectiveness of the suggested strategy in raising prediction accuracy above and above that of conventional methods.
8	**Bi et. al., (2020) [17]**	Deep neural networks-based recommendation algorithm	Experimental results show that the proposed approach not only outperforms existing CF recommendation systems, but also addresses the problems of data sparsity and cold start that emerge when using such algorithms.

3 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 latter [18].

4 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.

5 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.

6 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.

6.1 Technique Used

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

6.1.1 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 own 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, 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|c_i) of document d gave class c in the MNB model, the formula is given below [19]:

$$Multi-nomial\ P(|c_i) = P\left((f_1, f_2, \dots, f_n)|c_i\right) = \prod_{1\le j\le n} P(f_j|c_i) \quad (1)$$

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

$$C_{map} = \underset{c_i \in C}{argmax}\ \hat{P}(c_i) \prod_{1\le j\le n} \hat{P}(f_j|c_i) \quad (2)$$

The next step is to calculate a probability score. As soon as a term meets the criteria of $\hat{P}(f_j|c_i)$, it is added to the document's vocabulary. Therefore, to get $\hat{P}(w_j|c_i)$, divide the total number of keywords in class c_i by N_{jr}, where N_{jr} is the number of times word w_j appears in document d_r from class c_i. Then evaluate the likelihood of a document based on its classification as follows:

$$P(d_r|c_i) = \left(\sum_j^{|v|} N_{jr}\right)! \prod_{j=1}^{|v|} \frac{\hat{P}(w_j|c_i)^{N_{jr}}}{N_{jr}!} \quad (3)$$

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

$$\hat{P}(w_j|c_i) = \frac{count(w_j|c_i)}{\sum_{w\in V} count(w, c_i)} \quad (4)$$

- **Multi-Layer Perceptron**

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 possible [20].

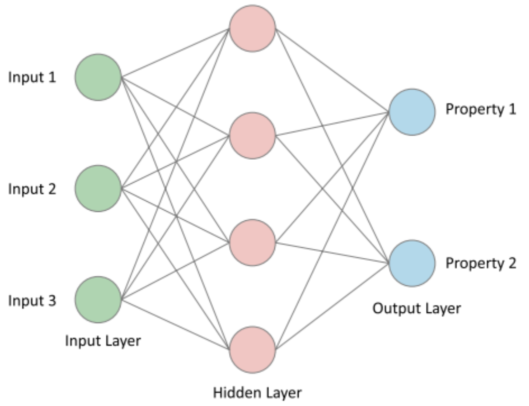

Fig3. MLP [21].

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 produces [22]:

$$f(x) = max(0, x) \quad (5)$$

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. 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 contexts [23].

$$\text{Log} = \left[\frac{p}{1-p}\right] = \beta_0 + \beta(\text{Age}) \tag{6}$$

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 4 [23].

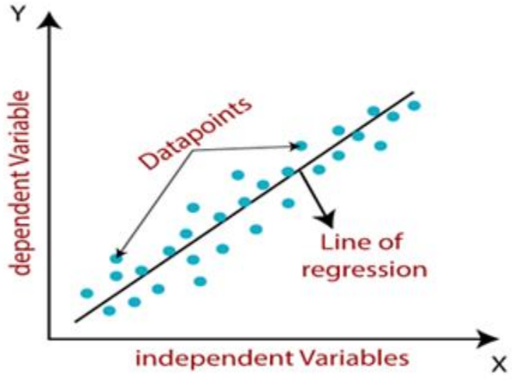

Fig 4. Logistics Regression [23].

6.2 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

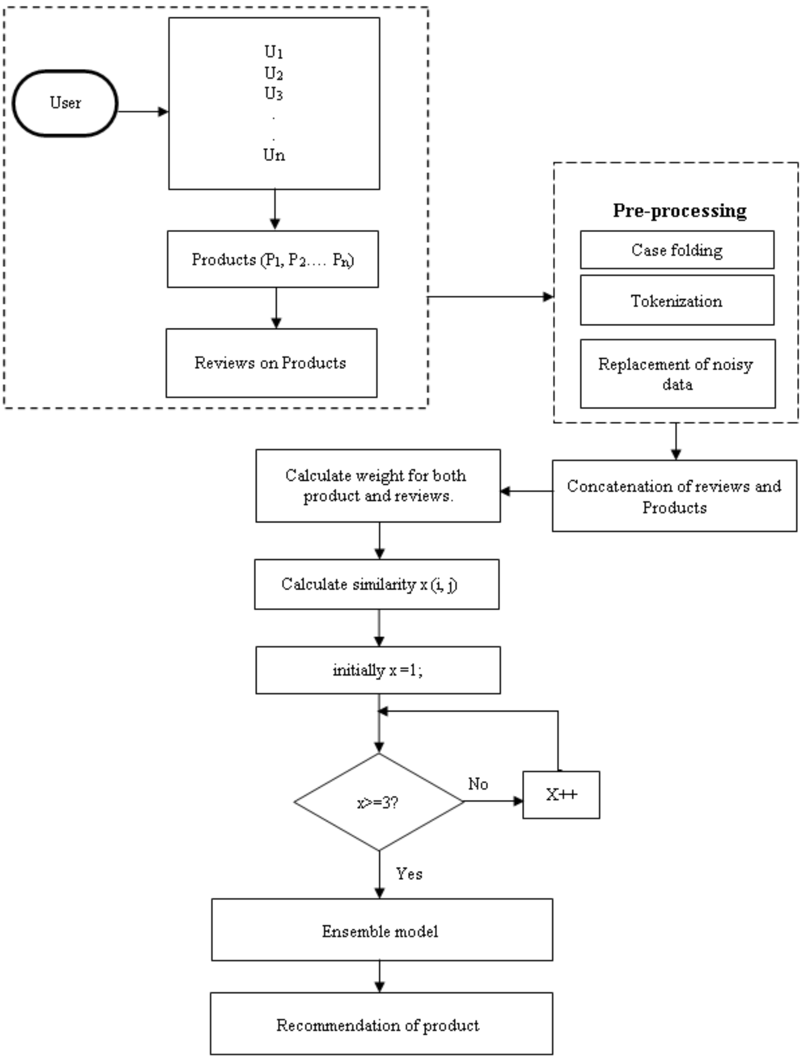

Fig 5. Block Diagram of Proposed Methodology

Step 1: Data Collection

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 is 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 of the data, 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 Weight for both Product and Review

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

$$W^{P}_{(u_p)} = R_{P_{U_i}} / \sum^{n}_{U_i \in U} R_P \tag{1}$$

Where,

$W^{P}_{(u_p)}$= Denotes the total weight of product (P) and review by users

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

Step 5: Calculate Similarity of users' reviews.

In this step 5, similarity between users' reviews is calculated by using a formula as given below:

$$Sim\ U(i,j) = (1-d) + d \sum_{U_j \in U} Sim\ (U_j) \times R_{P_{U_i}} \tag{2}$$

Where,

$Sim\ U(i,j)$= Denotes the similarity of two consumers' review on same products

d = Dampening factor

Step 6: Ensemble Model

Ensemble models are applied on 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 modelling 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 process products are recommended to the consumers as per their needs.

6.3 Proposed Algorithm

ALGORITHM: CONSUMER DECISION USING COLLABORATIVE FILTERING

Start

INPUT: → PRODUCT: $P_{i \epsilon 1,2,\ldots,n}$
CONSUMER: $U_{j \epsilon 1,2,\ldots,n}$ User per review

CO-RELATION: → [PRODUCT ↔ CONSUMER] $P_{1,2,3,\ldots U_{i \neq j}}$

OUTPUT: → PRODUCT ADVOCACY based on TOP-RATED REVIEWS.

Phase – I: Data Collection

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

Phase – II: Data Pre-processing

Step 3: The review for each product and its conciseness will be in Text format. So, it is a Text classification problem.
Step 4: Each review will be cleaned with punctuation, escape sequence, stop words, emoji, unwanted spaces, and digits, then apply WordNet Lemmatization.
Step 5: Each review's conciseness will be converted into a numerical value with either OneHotEncoding.
Step 6: The rating will be visualized with a histogram plot.

Phase – III: Concatenated Review with weight Calculation

Step 7: Each product's threshold will be a minimum with three ratings.
Step 8: Each product's review's conciseness will be considered the top two most frequent ratings.
Step 9: The product will be concatenated with its topmost reviews.
Step 10: 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:

$$W_{(u_R)}^{P} = \frac{R_{P_{U_i}}}{\sum_{U_i \in U}^{n} R_P}$$

where,

$W_{(u_R)}^{P}$ → denotes total Weight of the product (P_i) per user's review.

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

Phase – IV: Interlinked between more than two webpages

Step 11: The maximum occurrence of n-gram tokens for each product's review will be considered a dominant specification.
Step 12: For each product [P_i] minimum of three reviews ($R_{P_{U_j}}$) will be extracted to find the occurrence of 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.

$$Sim\ (i, j) = (1 - d) + d \sum_{U_j \in U} Sim\ (U_j) \times W_{(u_p)}^{P}$$

where,

$Sim\ U(i, j)$ → similarity of two consumers' product review

d = Dampening factor

Step 13: 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 best algorithm

Step 14: Ensemble Methods build sentiment analysis models. Ensembles combine classifiers to improve outcomes. MULTINOMIAL NAÏVE BAYES (**MNB**), MULTI-LAYER PERCEPTRON (**MLP**), and LOGISTIC REGRESSION(**LR**) are used.
Step 15: Provide classification report.

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

End

7 Results and Discussion

In this section, the results 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 two brands such as Amazon and Amazon digital services [24].

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;

$$Precision = \frac{TP}{(TP + FP)} \quad (1)$$

Where,

TP = True positive value,

FP = False Positive value.

Recall obtained by using.

$$Recall = \frac{TP}{(TP + FN)} \quad (2)$$

Where,

FN = False Negative

The F1-score measured by using formula is;

$$F - Measure = \frac{(2 * Precision * Recall)}{(Precision + Recall)}$$

(3)

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

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

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. The values of the specified parameters that are computed for the MNB in this work are shown in table 2 below. A graph representation of this result is shown in figure 6 which is given below table 2.

Table 1. Calculated Values of parameters for MNB

Technique	Precision	Recall	F1-Measure	Accuracy
MNB	0.819	0.812	0.823	0.8523

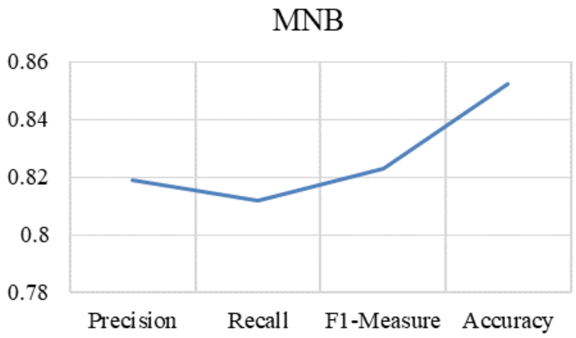

Fig 1. Graph of MNB Results

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 table shows the values of the parameters which are calculated for the MLP in the work are shown in table 3 and the graph representation of this result is shown in figure 7 as given below.

Table 3. Calculated Values of parameters for MLP

Technique	Precision	Recall	F1-Measure	Accuracy
MLP	0.865	0.831	0.865	0.8298

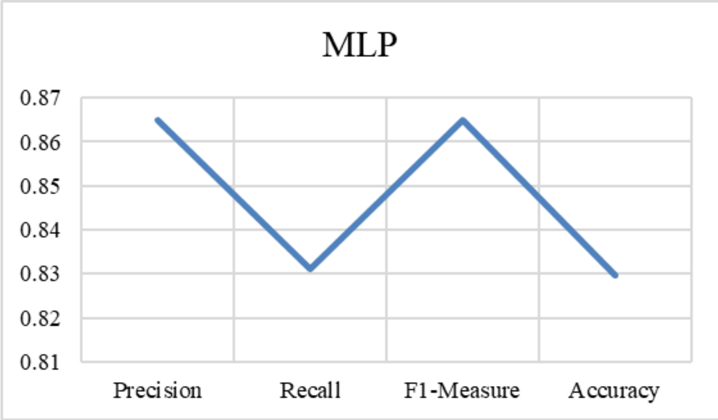

Fig 2. Graph of MLP results

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 values of the parameters that are computed for the LR in this work are displayed in table 4, and a graph representation of this result is shown in figure 8 which is presented below.

Table 2. Calculated Values of parameters for LR

Technique	Precision	Recall	F1-Measure	Accuracy
LR	0.798	0.791	0.821	0.925

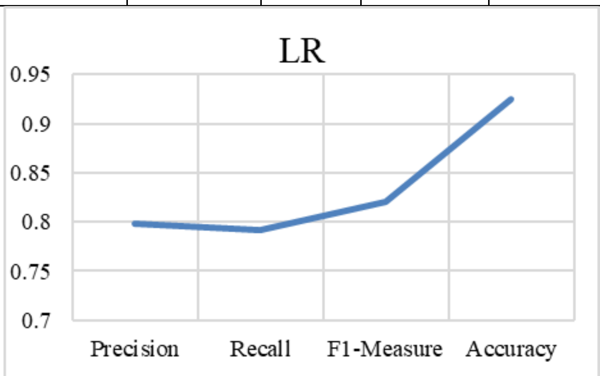

Fig3. Graph of LR results

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. The work presents table 5 that provides the values of the parameters that are computed for the ensemble model, and it presents a figure that shows a graph representation of the result.

Table 3. Calculated Values of parameters for Ensemble Classifiers.

	Parameters			
Technique	**Precision**	**Recall**	**F1-Measure**	**Accuracy**
Ensemble Classifiers	0.908	0.897	0.941	0.952

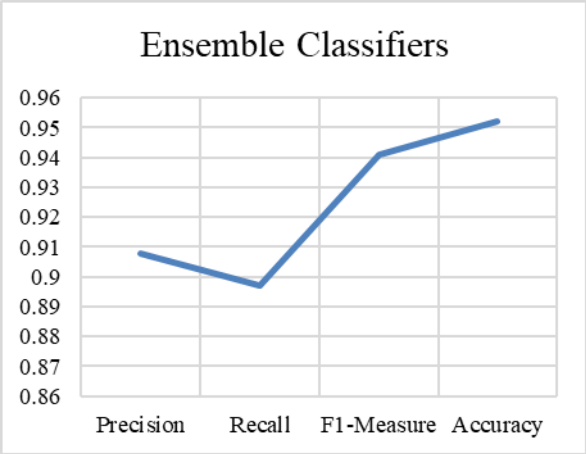

Fig 4. Graph of Ensemble Classifier results

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 all these techniques is depicted in table 6 and a graph representation of this result is shown in figure 10 as given below.

Table 4. Calculated Values of Error Rate of all the techniques

Technique	Error Rate
MNB	0.108
MLP	0.153
LR	0.1097
Ensemble Classifiers (MNB+MLP+LR)	0.087

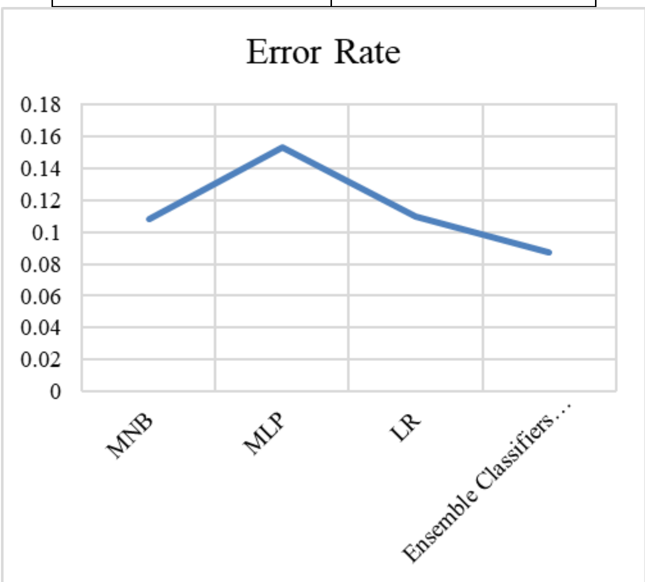

Fig 5. Graph of Error rate results for all techniques

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 modelling.

8 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. Table 7 shows the overall comparison of the proposed model with other conventional techniques in terms of precision, recall, f1-measure, and accuracy, and table 8 shows the overall comparison of the proposed model with other conventional techniques in terms of error rate.

Table 5. Comparison of Results

Parameters	Models						
	MNB [18]	**MLP [18]**	**LR [18]**	**MNB**	**MLP**	**LR**	**Proposed Ensemble Classifiers (MNB+MLP+LR)**
Precision	0.783	0.753	0.783	0.819	0.865	0.798	0.908
Recall	0.774	0.748	0.774	0.812	0.831	0.791	0.897
F1-measure	0.777	0.747	0.777	0.823	0.865	0.821	0.941
Accuracy	0.7907	0.7599	0.7911	0.8523	0.8298	0.925	0.952

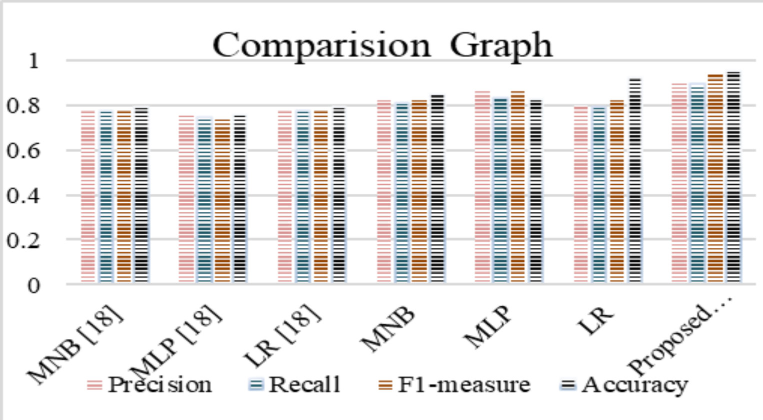

Fig 6. Comparison of the proposed work's performance to that of similar current schemes in terms of Precision, Recall, F1-measure, and Accuracy

Table 8. Comparison of Results (Error Rate)

Parameter	Models						
	MNB [18]	**MLP [18]**	**LR [18]**	**MNB**	**MLP**	**LR**	**Proposed Ensemble Classifiers (MNB+MLP+LR)**
Error Rate	0.207	0.239	0.208	0.108	0.153	0.1097	0.087

Fig 7. Comparison graph of Error rate among all techniques

9 Conclusion and Future Work

This work focuses mainly on proposing novel and group trust models to enhance consumer decisions. Recommendation systems have been and continue to be applied in various applications to support item (e.g., movies or music) recommendation and to solve the information overload problem by suggesting items of possible interest to consumers. In the work, CF is used to create a novel weighted recommendation system to better consumer decisions. Most recommendation systems mainly depend on a method known as CF, in which the customer's preference data is joined with that of other users to produce predictions about further items that the consumer may be interested in. 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 the approach, MNB, MLP, and LR are the components that make up the ensemble model. Ensemble Classifiers, consisting of MNB, MLP, and LR, are considered when putting the findings of the approach into practice to get superior outcomes compared to those obtained from earlier research. The suggested model is trained and evaluated with the use of a publicly accessible open-source dataset that can be found on the Kaggle website. In addition to that, a comparison of the outcomes is included in the study. According to the results of a numerical study of the suggested model, it outperformed other traditional approaches in many different respects, including accuracy (0.952), precision (0.908), recall (0.0897), F-measure (0.941), and error rate (0.087). In further work, other algorithms would be evaluated for performance comparison to determine the recommendation system that is the most effective, particularly for online shopping websites.

References

[1] Cui, Laizhong, Linyong Dong, Xianghua Fu, Zhenkun Wen, Nan Lu, and Guanjing Zhang. "A video recommendation algorithm based on the combination of video content and social network." Concurrency and Computation: Practice and Experience 29, no. 14 (2017): e3900.

[2] Liu, Mengsi, Weike Pan, Miao Liu, Yaofeng Chen, Xiaogang Peng, and Zhong Ming. "Mixed similarity learning for recommendation with implicit feedback." Knowledge-Based Systems 119 (2017): 178-185.

[3] 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.

[4] Bouras, Christos, and Vassilis Tsogkas. "Improving news articles recommendations via user clustering." International Journal of Machine Learning and Cybernetics 8 (2017): 223-237.

[5] Kim, Byung-Do, and Sun-Ok Kim. "A new recommender system to combine content-based and collaborative filtering systems." Journal of Database Marketing & Customer Strategy Management 8 (2001): 244-252.

[6] 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

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

[8] He, Xiangnan, Lizi Liao, Hanwang Zhang, Liqiang Nie, Xia Hu, and Tat-Seng Chua. "Neural collaborative filtering." In Proceedings of the 26th international conference on world wide web, pp. 173-182. 2017.

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

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

[11] 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.

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

[13] Zhao, Yan. "Design of Garment Style Recommendation System Based on Interactive

Genetic Algorithm." Computational Intelligence and Neuroscience 2022 (2022).

[14] 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.

[15] 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.

[16] 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.

[17] 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

[18] 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.

[19] 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.

[20] 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 popular convolutional neural network and potential visual transformer." *Artificial Intelligence Review* (2022): 1-58.

[21] https://towardsdatascience.com/recsys-series-part-5-neural-matrix-factorization-for-collaborative-filtering-a0aebfe15883

[22] 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.

[23] Ribokaitė, Lina. "Outlier detection in multidimensional streaming data." PhD diss., Vilniaus universitetas, 2021.

[24] https://www.kaggle.com/code/haojie98/amazon-product-recommendations/data