Applications of Machine Learning in Improving E-Commerce Platform

Dr. Nitin Upadhyay

IPS ACADEMY, Institute of Business Management and Research,Sanwer (M.P)

Author Email-Nitin_upadhyay86@yahoo.com

Abstract

In recent years the popularity of e-commerce has increased. The e-commerce is providing services at user's door step. It is a business ecosystem, which includes three category of users i.e. product buyers, sellers and administrators. The buyers and sellers are the users who are connected using ecommerce platform. The e-commerce is implementing business logic and applications for making the balance between the users, products, sells and other business activities. These activity is generating a significant amount of data. The analysis and utilization of this data will help in improving the functional and operational capability of e-commerce business. The proposed work is motivated to explore opportunity of applying Machine Learning (ML) techniques over the e-commerce data. The aim is to help the e-commerce stakeholders by enhancing the e-commerce services. The work simulate three different applications. First is focused on the spam review classification. The reviews in e-commerce may influence the buyer's decisions. Therefore, authentic reviews are needed for maintaining the healthy e-commerce. The proposed model uses TF-IDF and chi-square test for feature extraction. Then, an Artificial Neural Network (ANN) is trained for identifying the spam reviews. Second module is also focused on the review classification with the prospective of product seller. This include sentiment analysis. That analysis will help the product sellers as a feedback to improve their product or service. Additionally, the model is also recovering the level of user satisfaction. In this context, a ML technique is proposed using the TF-IDF based text features and sentiment score for categorizing the review text according to five satisfaction levels. Third model is a recommendation system as a solution to rare product recommendation and cold start problem. This system is helping to the administrators to make balance between new and old sellers. Thus, a weighted recommendation system is proposed. The model utilizes a feedback for refining and recommending products. The experiments are conducted on Amazon product review dataset. Based on the experiments and comparison, we found the proposed methods are able to provide optimal solutions. Additionally, the employment of ML techniques will help the e-commerce business at different levels.

Keywords: machine learning, e-commerce, spam review classification, sentiment based classification, recommendation system, business logic implementation.

1. Introduction

E-commerce is a business approaching consumers through websites and mobile applications. It also provide opportunity for new sellers for expanding their business. In this paper, we are investigating the applications of Machine Learning (ML) in e-commerce. The primary aim of e-commerce is to providing high quality of product delivery, return and exchange [1]. Additionally, fast resolution of the consumer complains enable reliability and faith. Next, stakeholder of ecommerce is product sellers. The e-commerce is also providing the opportunity to the sellers to reach new consumers [2]. Additionally, management and balance between sellers and consumers is done by e-commerce administration. The activities generate a significant amount of data. This data is large therefore the analysis requires to using ML technique. The ML based data analysis can provide a new research opportunity [3]. In this paper, three different ML applications are studied, which are benefiting the e-commerce business. The study includes: (1) A spam review classification model to deal with spam reviews which influences the buyer's decisions. (2) Performing sentiment based text classification for collecting the product feedback and (3) A recommendation system is proposed for enhancing the new product vendor's product visibility. The aim is to enhance the productivity of e-commerce platforms for all their stakeholders. Additionally, demonstrate how the ML techniques

will help in doing this improvement. E-commerce business is rising and growing day by day. But most of the e-commerce only concentrated on consumer satisfaction but neglecting the supplier's interest. That is a critical business issue. Therefore, we need to maintain healthy relationships and coordination among all the e-commerce stakeholders. The paper is focused on applications to enhance the experience of e-commerce users i.e. consumer, business owner and e-commerce administration. Additionally, empower the users to get positive benefits from the e-commerce.

2. Analyzing Spam Product Review to Help Consumers

The ecommerce is an application for providing the ease in selection of appropriate products. It provides an online showcase of products and services. User are visiting the ecommerce platform and can select appropriate service or product [40]. In order to make buying decision, the product review and rating is providing a significant role. Most of the online buyers are reading the product reviews and evaluate the rating of the product before parches. Therefore, the false or spam reviews are influencing the behavior of buyers [4]. In this context, we need a system, which classify and/or identify the false reviews. In this paper, we are exploring the techniques of spam review classification using ML. The key problem in classifying the review is the limited amount of text and features, which are not sufficient to decide whether a given review is a spam or legitimate. Thus, we proposes a new ML model based on text classification to identify fake or spam product reviews. In this context, the Amazon product review dataset is considered. The preprocessing of the review text has been performed and the feature extraction technique has also been employed. Next, the ML algorithm has been implemented to classify the review text. Further, the performance of the proposed technique is measured in terms of accuracy and time for performing classification. The main aim is to design a spam review detection system for providing correct information about product. In this context, (1) Developing an enhanced spam detection model for ecommerce product review and (2) Evaluation of the proposed technique experimentally and perform comparative study. Here we are motivated to analyze the e-commerce review post text in order to identify the spam or misleading posts. In this context, a spam review classification model has been implemented and overview is presented in figure 1.

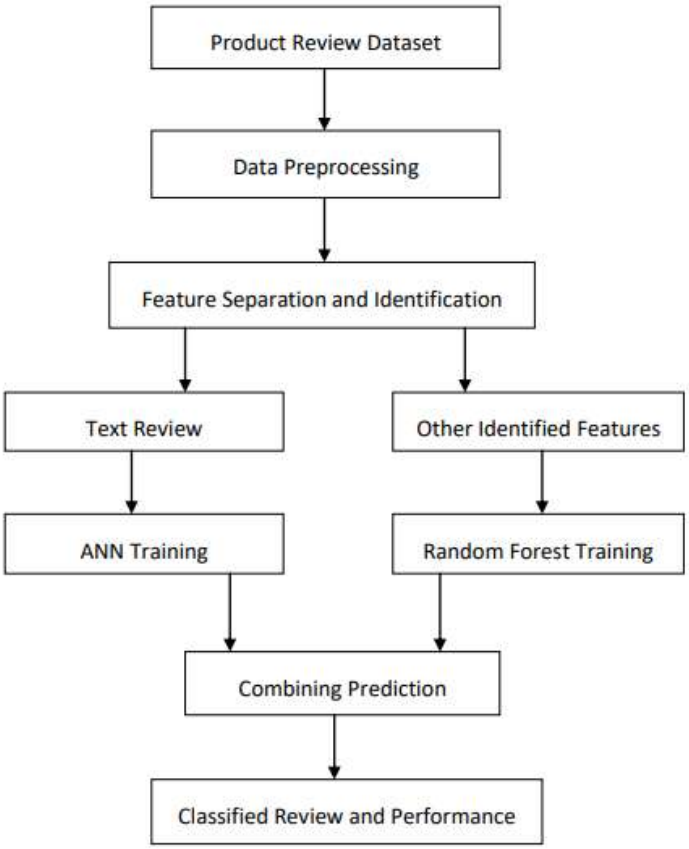

Figure 1 Demonstrate the flow of the spam review classification model

Dataset: The proposed system accepts the product review dataset. This dataset is taken from the Kaggle repository and known as Amazon product review dataset [5]. The dataset consists different categories of product such as electronics, home and kitchen, toys and games, and many others. The dataset is available in the format of JSON. Among them, we have selected the Toys and Games category. Therefore, we first read and parse the JSON files for obtaining the attributes. The obtained attributes from the review file are: Id, Reviewer Name, votes-down/up, Review Text, Rating, Summary, Review Time, Category, and Class. Now we need to preprocess the data. Preprocessing is used to remove the non-essential data. In this context, we remove some non-essential attributes like ID, Reviewer Name, review time, and category. After the elimination, we prepare two subsets of the dataset. This data processing is demonstrated as Feature separation and identification. The two subsets are: (1) **Set 1:** In this set of information the following attributes are included: Votes Down, Votes Up, Rating, and Class. (2) **Set 2:** In this set of information the review text and summary are combined with the class attribute for classification.

Training of set 1 attributes: Now, we have implement two ML algorithms for processing both the set of data. Thus, the random forest algorithm has been applied on information set 1. Random forest is an ensemble learning technique for classification, and regression. That is constructed by a number of decision trees during training. During classification the output is calculated by most trees outputs. The random forest can also be used for regression or prediction. In order to predict a continuous values, the mean prediction of the all the trees are used. It overcome the issues of over fitting. Random forests are normally more accurate than the individual decision trees.

Training of set 2 attributes: On the other hand, for processing of the data, we combine the review summary and review text. So, we utilize the Term Frequency and Inverse Document Frequency (TF-IDF) method for extracting the text features. TF-IDF is one of the popular and frequently used text feature selection technique. It is computed in two parts, first part is measuring the frequency of keywords additionally the second part include the measurement of importance of the keyword in a document. Based on both the parts are combined as a weight, which is used to select essential keywords. The term frequency (TF) is measured using:

$$TF = \frac{N}{T}$$

Where, N is the count of a keyword in a document and T is the total word count in a document.

Additionally the Inverse Document Frequency (IDF) is denoted using:

$$IDF = \log\left(\frac{N_d}{df}\right)$$

Where, N_d is number of document and df is the number of documents contains the target keyword.

Additionally, the combined weight or TF-IDF is denoted by:

$$w = TF * IDF$$

Using the calculated weight w, top k keywords from the list of keyword has selected as the text feature. This vector of text is utilized with Artificial Neural Network (ANN) algorithm for performing the training. The ANN is an artificially created neural network, used for solving various problems based on classification, recognition and prediction. This technique is effectively works on all kinds of data formats. The ANN is made with of three main layers: Input layer, Hidden layer, and Output layers. The input layer consist of similar number of neurons as the number of inputs, hidden layer may consist of a number of layers and huge number of neurons for performing the calculation. Finally, at the output layer the similar number of neurons is available as the output variables. For improving the learning the error is back propagated for optimizing the weights of neurons.

After training both random forest and ANN the model is become able to predict the spam reviews. In this context, with the similar ID the test dataset has been prepared. This test dataset is used with the trained algorithms using the trained attributes. Here both the models are predicting the nature of review according to their learning. After getting the prediction from both the algorithms the predictions are combined using the following equation.

$$F_p = 0.4 * R_p + 0.6 * A_p$$

Where, F_p is the final prediction value for the given review, R_p is the prediction made by the random forest and A_p is prediction done by the ANN algorithm.

Here, the weight of ANN-based prediction is considered higher as compared to the random forest because the ANN provides the prediction based on the content of the review and summary of the review given by the reviewer. The random forest-based prediction considers the attributes of the opinion of others and the reliability of the reviewer. The implementation has been done on the basis of python and performance analysis was performed.

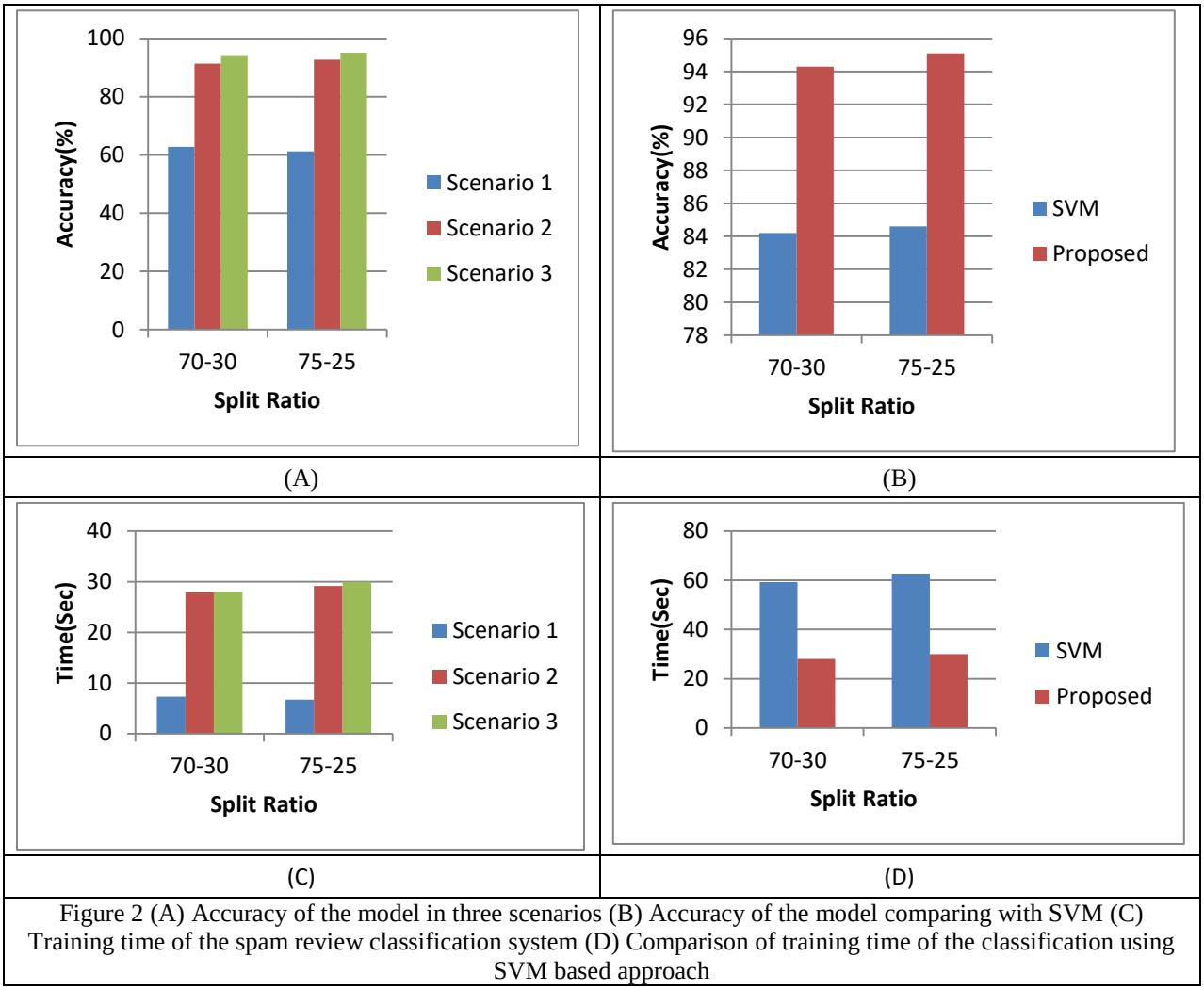

Figure 2 (A) Accuracy of the model in three scenarios (B) Accuracy of the model comparing with SVM (C) Training time of the spam review classification system (D) Comparison of training time of the classification using SVM based approach

2.1 Results Analysis

Spam reviews are very influential facts by which the decision of buyers is frequently changed. Therefore, accurate product review is essential for a healthy e-commerce system. In this section, we have tested the spam product review detection technique and compared it with a traditional approach of SVM-based spam detection. In this context, we have considered the two key performance parameters namely accuracy and training time. The performance analysis is performed based on the following experimental scenarios: (1) Providing the performance of random forest classifier with the additional information (2) Providing the performance of ANN algorithm based on the review content and (3) Providing the performance of combined prediction and compared with the SVM based classification approach

The accuracy describes the correctness of spam identification by the trained classifiers. In our case the accuracy can be defined as a ratio of total correctly identified spam reviews over total reviews provided for classification. That is given by the following equation:

$$accuracy = \frac{Correctly\ classified\ spam\ review}{total\ reviews} X100$$

The accuracy of the experimental scenarios is given in figure 2(A) and comparison with the SVM is given in figure 2(B). According to the accuracy as given in figure 2(A), we found that the accuracy for classifying the additional review parameters provides lower accuracy as compared to review content classification. Moreover, the combination of additional review features with the review content enhances the accuracy of the spam identification. On the other hand, the performance reported on figure 2(B) shows the proposed technique is far superior than the traditional SVM based classification approach. Additionally, the training time is also measured, which is an essential parameter of measuring resource consumption. That indicates the amount of time required to train the algorithm using the given amount of data. The time consumption can be measured using following equation:

$$time\ consumed = End\ time - Start\ Time$$

Table 1 Accuracy of three different scenarios

Accuracy			
Sample ratio	**Scenario 1**	**Scenario 2**	**Scenario 3**
70-30	62.8	91.4	94.3
75-25	61.2	92.7	95.1
Training time			
70-30	7.3	27.89	28.01
75-25	6.7	29.14	29.97

The performance of experimental scenarios in terms of training time is demonstrated in figure 2(C). According to the performance, additional feature based classification requires very fewer time as compared to the content based approach and the combined feature based approach. Here, the combined feature based approach and content based approach consumes similar amount of time. Similarly, in figure 2(D) shows comparative training time of the proposed and traditional technique. The traditional approach includes SVM classifier. Based on the performance, the proposed technique found efficient as compared to SVM based spam review classification. The proposed methodology is efficient as well as accurate for spam review classification on e-commerce platform.

Table 2 comparing accuracy of the proposed model with SVM

Accuracy		
Sample ratio	**SVM**	**Proposed**
70-30	84.2	94.3
75-25	84.6	95.1
Training time		
70-30	59.26	28.01
75-25	62.68	29.97

3. Sentiment based Review Analysis for Feedback and consumer satisfaction

When some consumer has any compliant about the product, then the consumer utilize the reviews or forums to discuss the complaints about the product. In both the cases, the user feedback is obtained is valuable for the product

designers, vendors and manufactures to improve the quality of the product [6]. In this context, the automatic review classification based on the product quality and consumer sentiment evaluation is essential [7]. However, a number of efforts are done for analyzing the e-commerce product reviews based on sentiments. But most of the work is not considering the product quality [8]. Additionally, very limited work is focused on product quality feedback, which are either not much accurate or not much effective due to partial data exploration [9]. The proposed work is motivated to design a ML technique, which will deal this problem in two steps. First, the technique of information retrieval is utilized for extracting the review post related to product quality. And then utilize the sentiment based text classification and grading to describe the consumer satisfaction level in five grade scale. The satisfaction grade is defined as highly satisfied, satisfied, neutral, not satisfied and disappointed. Next, the details of proposed system is discussed, thus first we provide the algorithm to identify the review post related to product quality, then the sentiment analysis process and review grading technique is described. Further, the experiments are conducted on Amazon product review dataset and performance is measured. The proposed work is providing benefit to the product vendors or manufactures by providing the consumer feedback and quality improvement suggestions. In this context, the system design is divided into main parts: (1) Extracting the review post which are discussing the product quality and (2) Analyzing the review post to indicate the consumer satisfaction level

3.1 Review Post Extraction

The aim is to understand the review of user and collect the feedback, which may help the product vendors to improve their quality. In this context, it is required to find an appropriate dataset, which can use for analysis. In this context, a product review dataset available on Tensorflow library is utilized. This dataset is known as the Amazon product review dataset [10]. This dataset may contain a significant amount of product categories, among them we just utilize a subset of review dataset which is termed as "Mobile_Electronics". The attributes of the dataset is listed as: customer_id, helpful_votes, Marketplace, product_category, product_id, product_parent, product_title, review_body, review_date, review_headline, review_id, star_rating, total_votes, verified_purchase and Vine. From these set of attributes, we need only limited attributes thus we just selected "review body", "review headline" and "star rating". The dataset has 104975 instances and a total 3 attributes. Now, before the utilization of data samples, we have prepared three subsets. Among 10% of random samples (10497) are selected for vocabulary development. Next 20% of samples (20995) are used for testing and remaining 70% of samples (73483) samples are preserved for training. In addition, the star rating is encoded. The rating is available between 1 and 5. Disappointed is given as (-1), Not satisfied as (-0.5), Neutral is given by (0), Satisfied is denoted by (+0.5) and Highly satisfied is given by (+1).

3.2 Vocabulary development

In this phase we consider the 10% of samples and then performed cleaning of the samples using the following steps: (1) Stop word removal, (2) Special character removal and (3) Also removed the abbreviations. Then, the remaining keywords are organized according to the rating associated with the considered samples. In this manner, we produced five individual vectors to process the remaining samples. Let these vector are defined as $V = \{v_1, v_2, v_3, v_4, v_5\}$.

3.3 Extracting Quality Relevant Reviews

In order to select relevant post, the concept of information retrieval is being utilize. Therefore, the k-means clustering is used for finding the related review post. The k-means algorithm is slightly modified and prepared two versions. The k-means clustering algorithm accepts N reviews $(x_1, x_2, \dots, x_n)$ to be cluster [11]. Algorithm select k samples as initial cluster centers, but in this modified clustering algorithm we have k=5 and our initial cluster centers are $V = \{v_1, v_2, v_3, v_4, v_5\}$. Next, we calculate the distance between each review x_i and each cluster center V, then assign each object to the nearest cluster [12]. For calculating distance, we use the following equation:

$$d(x_i, v_j) = \sqrt{\sum_{j=1}^{d} (x_i - v_{j1})^2}, i = 1 \dots N, j = 1 \dots k \dots \dots . (1)$$

Additionally, for making another variant we utilize the cosine similarity as:

$$\cos(x_i, v_j) = \frac{x_i * v_{ij}}{\|x_i\|\|v_{ij}\|} \dots\dots\dots (2)$$

Additionally, for converting it to distance we use the following equation:

$$d(x_i, v_j) = 1 - \cos(x_i, v_{ij}) \dots\dots\dots (3)$$

Where, $d(x_i, v_j)$is the distance between data i and cluster j.

Additionally, to calculate the new cluster centers the mean of reviews are calculated as:

$$v_i = \frac{1}{N}\sum_{j-1}^{n_i} x_{ij}\,, i = 1,2, \dots, K \dots\dots\dots (4)$$

Where, N_i is the number of reviews in current cluster i.

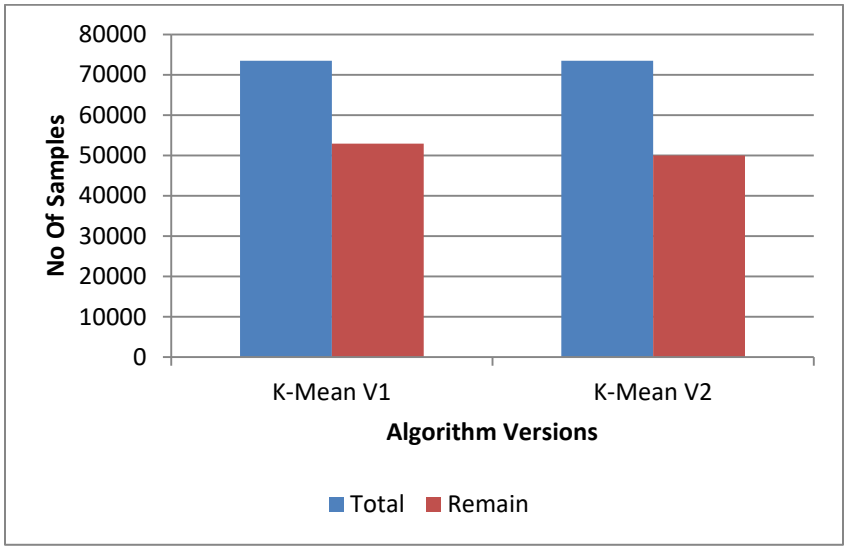

Figure 3 Filtered Reviews based on both the k-means versions

We repeat these steps till the stopping criteria is not reached, and finally calculated cluster centers are returned as the outcome. Now, let the resultant centroids are $U = \{u_1, u_2, u_3, u_4, u_5\}$. To filter the less relevant reviews, we create a threshold by using:

$$T = \frac{|U - V|}{2} \dots\dots\dots\dots\dots (5)$$

Using this threshold we refine the reviews and obtain relevant reviews. The number of samples are given in figure 3. The figure demonstrate the number of samples, before and after applying the threshold for filtering the review content. Here, we found reduced number of reviews which are more relevant. According to the results, by using k-means algorithm using Euclidean distance the 28% of samples (20575) samples are removed thus we get a total of 52908 relevant samples. On the other hand, when we used the cosine similarity based k-means then 32% of samples (23515) are reduced, and finally we get the 49968 review samples for training.

3.4 Sentiment Analysis

In this section, we utilized refined samples for performing the sentiment analysis. In this context, we first apply the preprocessing on the review data. Further, we extract the relevant text features using the TF-IDF. Here, we have selected only 2000 features for utilizing with the classification algorithm. But, we need to include the sentiment features also, therefore, we have utilized the Natural Language Toolkit (NLTK) for parsing the reviews text.

Table 3 Filtered Reviews of the k-means versions

Algorithms	Total	Remain
K-Mean V1	73483	52908
K-Mean V2	73483	49968

Using this toolkit, we extract the part of speech tagging based features and we obtained 40 features [13]. Then, a combined feature vector is prepared. The total feature size of the review data is becomes 2040. Both the kinds of training samples and testing samples are utilized with the feature selection techniques and the feature vectors are prepared. Finally, two classification algorithms are used namely ANN [14] and random forest [15] for classification and validation. Here, we utilize the star rating as the class label. After training, we performed the classification of the test review. The predicted class is in form of star rating and further we denote this predicted rating using C_p. After classification, we need to grade the review according to user satisfaction level. Therefore, we utilize a sentiment scoring API **V**alence **A**ware **D**ictionary and s**E**ntiment **R**easoner (VADER). In order to calculate the sentiment score the library provides a function which is known as "polarity_scores". The polarity score results four different scores and the value range. The negative, positive and neutral are having range between 0 and 1. Additionally the compound score is varying between -1 and +1. Here, we utilize the compound sentiment score for review headlines and review body. This compound sentiment score is denoted as C_h and C_b. Finally, to calculate the satisfaction level of the consumer, we calculate a score W using:

$$W = C_p * w_1 + C_h * w_2 + C_b * w_3 \dots \dots \dots (6)$$

Where, w_1, w_2 and w_3 are weighting factor and decided by the constraint of $w_1 + w_2 + w_3 = 1$. In this work we have utilized $w_1 = w_2 = w_3 = 0.33$.

Finally, for grading the following decision function will be used:

$$f(w) = \begin{cases} if\ w = 0 \quad Then\ neutral \\ if\ w > 0 \leq 0.5 \quad Then\ \text{satisfied} \\ if\ w > 0.5 \ \leq 1\ Then\ highly\ \text{satisfied} \dots \dots \dots (7) \\ \text{if w} < 0 \leq -0.5\ then\ not\ \text{satisfied} \\ \text{if w} < -0.5 \leq -1\ Disappointed \end{cases}$$

3.5 Results Analysis

The proposed work involves the evaluation of two different experimental scenarios: (1) **Evaluation of clustering algorithms:** Both the prepared versions of k-means are evaluated in terms of accuracy and training time. and (2) **Evaluation of sentiment based text classification:** The classification performance of random forest and ANN algorithm is provided for both kinds of review samples. Here, we also calculated the accuracy and training time of both the techniques. Figure 4(A) demonstrate the performance of clustering approach for filtering the less relevant review posts. In figure 4(A) the performance in terms of accuracy is reported. The X-axis of the diagram shows the variants of algorithms, where K-Mean V1 is developed on the basis of Euclidean distance and K-means V2 shows the performance of cosine based k-means algorithm. Additionally, Y axis shows accuracy in terms of percentage (%).Similarly, figure 4(B) shows the performance in terms of training time.

Table 4 clustering algorithm's Performance

Algorithms	Accuracy	Training Time
K-Mean V1	71.2	189
K-Mean V2	79.5	204

The training time is measured in terms of seconds (Sec). In this diagram, X axis shows the versions of the clustering algorithms and Y axis shows the training time. According to the performance, modified k-means is providing high accurate results but the method consumes additional time.

Table 5 Sentiment based review classification Accuracy (%)

	Accuracy (%)		Training Time	
Algorithms	**Random Forest**	**ANN**	**Random Forest**	**ANN**
K-Mean V1	89.7	88.9	59.26	128.88
K-Mean V2	94.5	94.1	62.68	152.21

Figure 4(C) shows the performance of the proposed sentiment based review classification system. The classification performance for both kinds of samples is evaluated and the performance of both the classifiers (i.e. Random forest and ANN) is also described. Figure 4(D) describes the performance of sentiment based review classification in terms of accuracy. The accuracy is demonstrated in terms of percentage (%). In X-axis, the sample generations algorithms are given and in Y axis the accuracy is given.

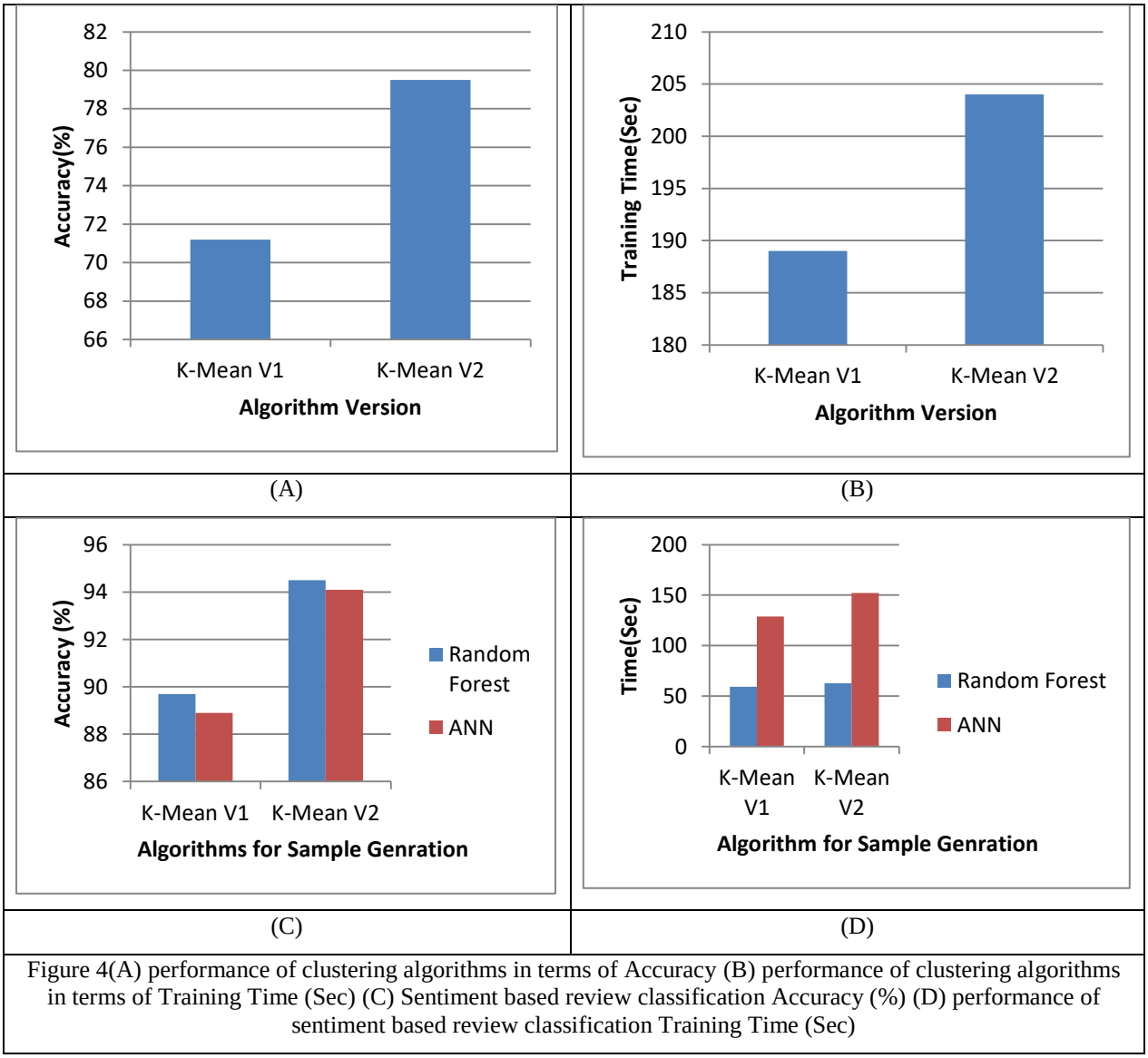

Figure 4(A) performance of clustering algorithms in terms of Accuracy (B) performance of clustering algorithms in terms of Training Time (Sec) (C) Sentiment based review classification Accuracy (%) (D) performance of sentiment based review classification Training Time (Sec)

Similarly, in figure 4(D) the training time of the algorithms are discussed where the X axis is remains same and Y axis shows the training time. Based on the performance in terms of accuracy, we found the random forest is providing more accurate classification as compared to ANN, additionally, it is efficient also as compared to ANN.

4. Product Recommendation Using Enhanced Product Visibility

New product sellers are not making profit in ecommerce, due to low product visibility and low conversion rates. In this situation, new sellers may drop e-commerce. Additionally, the promotions and advertisements are the expensive way of increasing the product visibility. This problem is known as rare product recommendation problem. In addition, the recommendation systems are not able to recommend products for new costumers. This problem is known as the cold start problem. In this presented study, both the problem has considered to solve using redesign of a recommendation system. The available recommendation systems are utilizing already popular products or those products which are already have higher sales. Therefore, we introduce a ML technique to enhance the recommendation problem. This technique is utilizing user feedback and progressive method to improve the recommendation. This technique is useful for enhancing the rare product visibility and also can handle the cold start problem. Next section provides the details of the proposed recommendation system. Based on the implemented technique the experiments are conducted using web access log file, and the performance is measured.

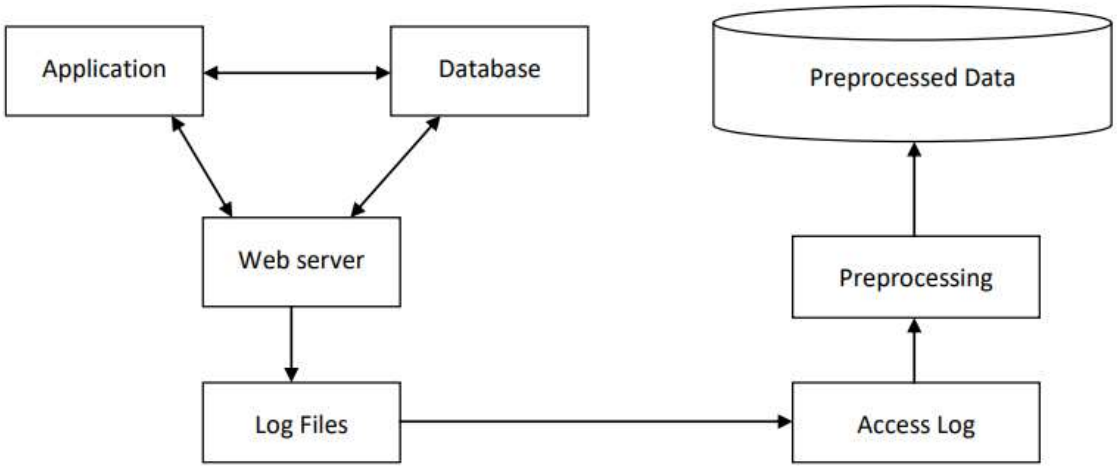

Figure 5 The web server and access log

4.1 Proposed Recommendation System

The recommendation systems are designed for e-commerce for supporting the consumer's need and suggesting the appropriate products. It is the representation of user web access behavior. Therefore, recommendation system is an application of the Web mining. This data is generated at the servers during the different activities involved in server and application execution on server such as web access logs. The aim of this system is to guide for suitable products to the consumer. This system is a data driven application and consumes historical user information. The user information is used with the data analytics techniques of recovering the user access behavior. This recovered information can be relevant to the user's intrest based on frequency, brands, cost, new products/brands, and others. This information used to establish relationship among user and product, and also help to understand the requirements of the user. Additionally, highlighted features are utilized with ML techniques to suggest relevant products. The recommendation systems are also used to employ business logic. The business logic is an essential component of a successful e-commerce and help to create funds. These business logics are marketing, promotions, discounts, and coupons. Therefore, the proposed enhancement incorporates the solution for rare product recommendation and the cold start problems using a consumer feedback system.

4.2 Initialization of Recommendation System

The proposed system is a feedback oriented system. The recommendation of products has been done in steps. At initial phase of recommendation, the system utilizes the web access log. The access log generation process is demonstrated in Figure 5. According to the diagram, web access log is prepared on the web server, which is maintaining different log files for different events. Among them, the web access log contains the entry about requests and responses of the web resource. The generated access log is utilized with the proposed system for

extracting the essential insights. Web access log can include different properties such as timestamp, IP address, protocol, response, methods, resource name, and others. However, in this work all access log attributes is essential. Therefore, we eliminates unwanted attributes. Only limited attributes i.e. IP address, resource name, and time stamp are preserved. These attributes are transformed and kept into a database table for further utilization.

4.3 Basic Recommendation

The recommendation system provide suggestions without any prior user information as given in figure 6. This system performs recommendation based on initial user activities. Therefore, previously preprocessed web log is utilized. The preprocessed data contains the products URLs, which are requested by the user. The aim is to produce a list of products, which is combination of promoted, advertised, popular, highly demanding product and brands, as well as the new seller's relevant products are also used. This product list is needed to use with frequent pattern analysis therefore this list of product has transformed into transaction sets based on user sessions. Among different frequent pattern mining algorithms the apriori algorithm is one of the popular association rule mining algorithm.

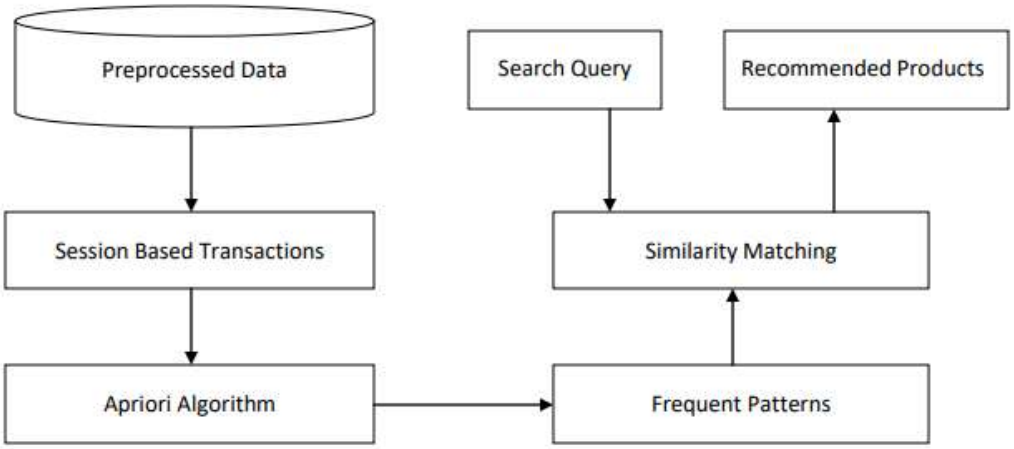

Figure 6 Primary Product Recommendation

That algorithm is accurate than other association rule mining techniques therefore the apriori algorithm is used. The apriori calculates the subset of the product, which are frequently occurred in user sessions. The apriori algorithm calculate frequent set of products and user's product search keywords are used to find the appropriate product from the product list. The relevant product from this list is recommended as first stage recommendation. The user's product search keywords and frequent products obtained from apriori algorithm is used as final outcome of this stage. Next, the user feedback is involved to optimizing the product search results. Table 6 highlights the steps of the product recommendation based on user initial search activity. This algorithm works as a search algorithm, which accepts the web access log file L, business logic B, and user search query Q. In first step the algorithm read the log file L and then preprocessed. The preprocessed data and the business logic is used to generate list of products. The product list P_n, and the frequency based on access log file is used with apriori algorithm. The apriori algorithm prepares frequent pattern F_n and works as database for making the search using user query. Then relevancy between F_n and Q is calculated to recommend the products P_r.

Table 6 Algorithm for search based Recommendation

Input: Web access log L, Business constraints B, Initial Search query Q
Output: Products recommended P_r
Process:
1. $R_n = readAccessFile(L)$
2. $P_n = preProcessData(R_n) + B$

3. $F_n = Apriori.FreqSet(P_n)$
4. $for(i = 1; i \leq n; i++)$
    a. $if\big(F_i.contains(Q)\big)$
        i. $P_r.Add(F_i)$
    b. $end\ if$
5. End for
6. Return P_r

4.4 Involving User Feedback in Recommendation

The initial recommendation generates a list of products that are relevant to the user search product and the frequency of purchased products of different users. It is a large list of products. Now, we identify the user behavior and requirement based on the user feedback by using product clicked.

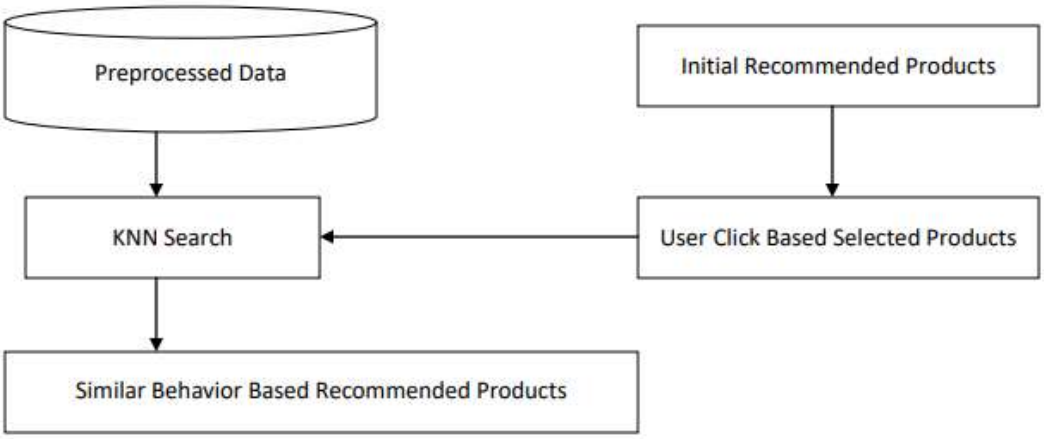

Figure 7 User Feedback based product recommendation

The generated list of products P_r is demonstrated in front of the user, and the user can select some products among the results according to their interest. The list of products in which the user is interested is recognized here as the user feedback. The selected product by the user is used with the k-Nearest Neighbor (k-NN) algorithm for finding similar kinds of patterns. Here, the user interest-based behavior is matched with the other user's product interest behavior. Similar behavior-based products are recommended by including user feedback. Table 7 demonstrates the process of feedback involved in the initial product recommendation.

Table 7 second stage of prediction

Input: Initial Recommendation P_r, Entire Preprocessed data P_n

Output: Refined recommendation S

Process:

1. $for(i = 1; i \leq P_r.length; i++)$
    a. $if\big(P_{r_i}.Clicked\big)$
        i. $Feedback.Add\big(P_{r_i}\big)$
    b. $end\ if$
2. End for
3. $S = KNN.Search(Feedback, P_n)$
4. Return S

Therefore, the initially recommended product P_r is used for obtaining user feedback about the initially recommended products. The feedback-based selected products are then utilized as queries for searching similar products in the entire preprocessed product database. In this experiment, the KNN algorithm is used for searching and the outcomes of the KNN algorithm are produced as the feedback-centric recommendation. The process of feedback-centric recommendation is also demonstrated in figure 7.

4.5 Final Recommendation

The aim is to provide precise product recommendations by using previous-stage recommendations. This phase of recommendation is working as a filter to remove additional products obtained from the previous phase of recommendation. Figure 8 shows the limited search space for making precise product recommendations to the end user. This recommendation is made based on the recommendation made in the feedback-based recommendation, which contains less data as compared to the entire preprocessed data. Using the feedback-based product recommendation we recover two key facts namely the product cost range and brands popular among a similar set of users. Therefore, based on the cost and popular brands we tried to filter the list of products. In order to find the mean cost of the product and bands, all the brands and costs of the product are considered, which are found in likely hood of the feedback-based recommendation.

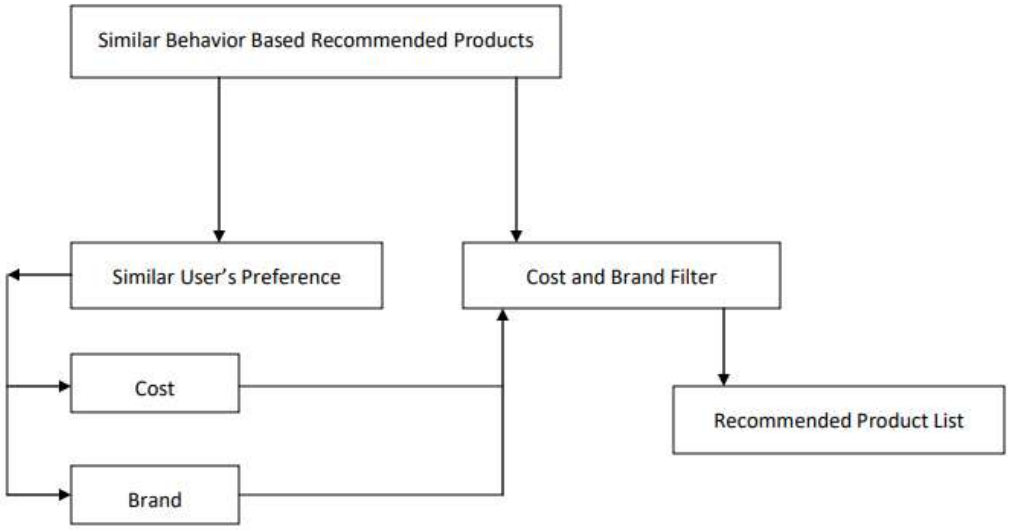

Figure 8 Final stage recommendation

Both the factors cost and brand are combined to rank the list of recommended products from the behavior based recommended products. The entire process of product recommendation is demonstrated in figure 8 and table 8. Table 8 consists of the steps to follow for final stage recommendation. The algorithm accepts the user feedback based on generated product recommendation S and produces a refined and precise final recommendation R. First, we recover two factors average cost of similarly behaving user and the list of product brands, which are used by the similar group of user. Then using both factors, we calculated a rank for each product obtained in the recommended product list S. using the calculated rank of all the products we sort the list S. From the sorted product list of S is used to return the top-ranked list of products as the final recommendation.

Table 8 Final Stage of product recommendation

Input : Feedback based product recommendation S
Output : Final recommendation R
Process:
1. $for(i = 1; i \leq S.length; i++)$
a. $C = GetCost(S_i)$

b. $B = getBrand(S_i)$
c. $Rank_i = C * B$

2. $end\ for$
3. $R = Sort(S, Rank)$
4. Return R

4.6 Results Analysis

In this section, we evaluated the performance of the proposed recommendation system. Therefore, a list of products from the Amazon e-commerce platform is created, and using a previously available web access log file is modified using the collected product URLs. In this context, accuracy, error rate, time used in prediction, and memory utilization of the proposed system is calculated and reported in table 9 and figure 9.

Table 9 Performance of the proposed recommendation system

Exp. No	Accuracy %	Error rate %	Time consumed (MS)	Memory usages (MB)
1	92.8	7.2	188	110
2	91.3	8.7	198	118
3	93.7	6.3	193	121
4	94.8	5.2	201	106
5	95.2	4.8	210	119
6	94.1	5.9	183	122
7	95.6	4.4	191	113

The accuracy of the proposed recommendation system is given in figure 9(A) and table 9. In this figure, dataset sample size used in experiments is given in the X axis and the obtained accuracy of experiments is included in the Y axis. The accuracy of the system is measured in terms of percentage (%). According to the results, the accuracy is increasing with the sample size for building the model. Next parameter is the Error rate. That is used to find the misclassification of prediction. That is measured based on outcomes and total samples for the recommendation. The error rate can be computed using:

$$error\ rate = \frac{misclassified\ samples}{total\ samples} X100$$

Figure 9(B) contains the results in terms of error rate (%). According to the error rate, we find that the error rate is not fluctuating highly, thus the performance is consistent even when the data samples are increasing. So, the developed recommendation system is reliable for solving the addressed issues. Further, we also calculated the time and space utilized. Thus, we calculated the time required for performing the prediction. The amount of time consumed is called the time consumption. The time consumption of the proposed model is demonstrated in figure 5.7. The X-axis contains the sample size and the Y-axis contains the time consumed. The time is measured in milliseconds (MS).

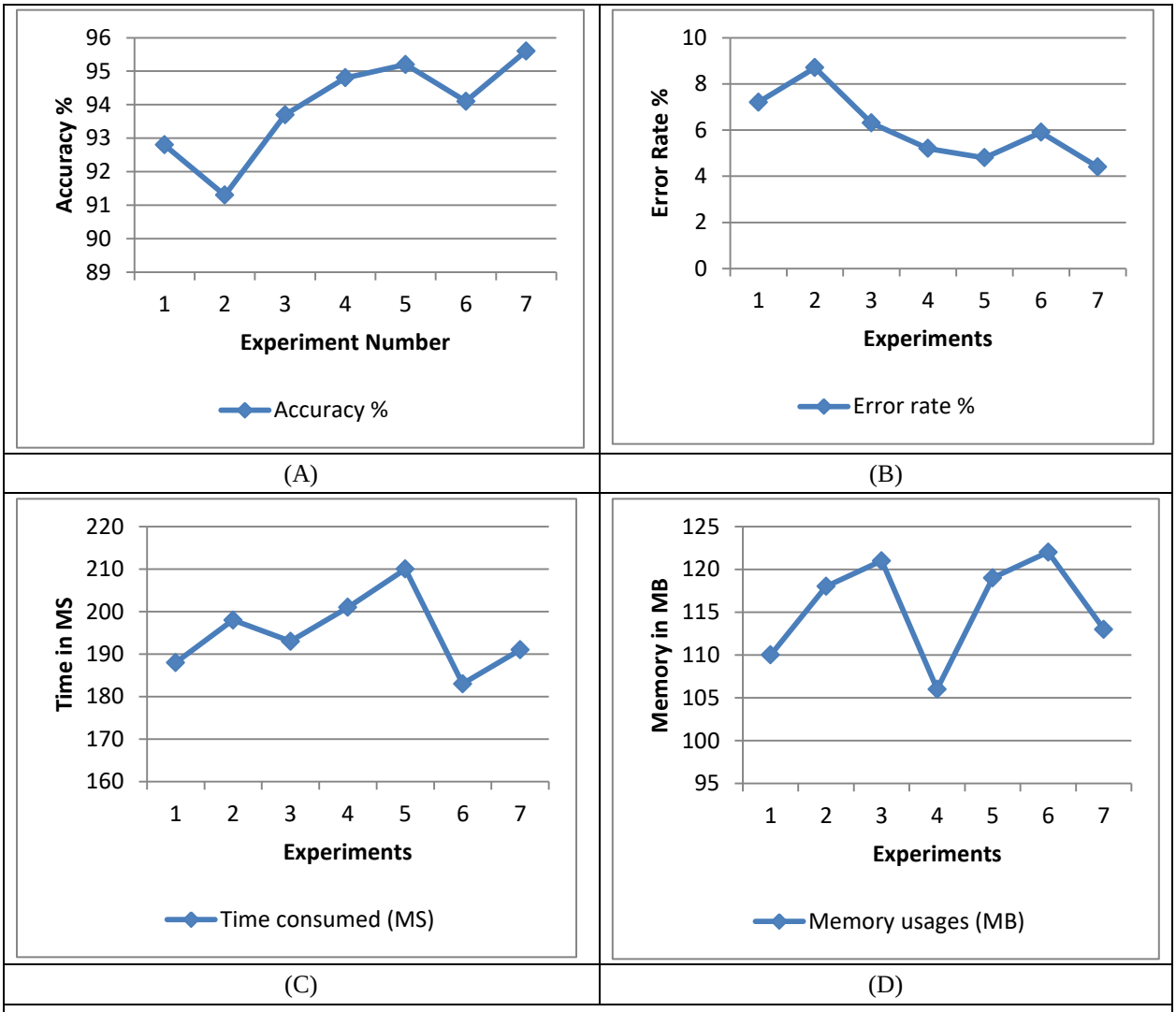

Figure 9 (A) Accuracy (%) of the proposed recommendation model (B) Error rate (%) of the recommendation model (C) Time utilization (D) Memory utilization of the proposed model

According to the results, the time consumption are varying with the sample size and remains consistently increasing. Time consumption is also reducing with the amount of data in web access log file. Therefore, system is acceptable. Finally, memory usage is measured in terms of kilobytes (KB). That can be measured based on the process. The execution of a process takes an amount of main memory. This memory size is known as memory usage. It is the difference between the total memory assigned and the free space. That can be calculated using:

$$memory\ usage = total\ memory - free\ memory$$

The measured memory usage is given in figure 5.8 and table 4. In this diagram, X-axis contains the sample size used, and Y axis shows the used memory. Memory usage is depending on the amount of data. The memory is not varying much rapidly for the product recommendation.

5. Conclusion and future work

The e-commerce applications are including three main users i.e. consumer, seller and administrator. The consumer is the product buyer. The consumer browses the products and decides the appropriate product. In deciding the appropriate product, the product reviews are playing essential role. Therefore, for providing the legitimate

information to the consumers a spam review classification technique is proposed. Second, a sentiment based review classification technique was introduced to classify the product review for identifying the satisfaction level of consumer. This helps the product seller to understand the product feedback. Next, the recommendation system is studied with the aim to solve the administration issue for implementing the business logic and improve the new product visibility. That system will also help to make balance between new and old seller's product recommendation.

The ecommerce clients are utilizing product review and rating for making product purchasing decisions. Thus, a spam review identification model for helping the users is introduced. This model usage TF-IDF and chi-square test for text review feature extraction. Then, an ANN is used for spam review identification. The experiments on Amazon product review dataset have been carried out. Additionally, compared with SVM based method. The e-commerce review classification can be analyzed for finding consumer feedback. The business companies are utilizing these reviews for identifying consumer's expectations. Thus, a sentiment analysis technique for product review classification was proposed, to understand consumer satisfaction level. This model utilizes the text features and sentiment score for grading the product review in to five satisfaction levels. The Amazon product review dataset used. The new product vendors are suffering from low product sales. To enhance the product visibility needs some improvements on the existing recommendation system. The current algorithms are considering the products based on ratings and reviews, existing sales, and brand size. Therefore, new vendors feel the biased. Another option to enhance product visibility is using advertisements, which is expensive. Therefore, a new recommendation model is proposed for obtaining a win-win situation for all the ecommerce parties' i.e. consumers, vendors, and administrators. A weighted recommendation system is proposed and evaluated for recommending a set of products to existing customers. The model utilizes a feedback for refining products to consumers.

The ecommerce and application of ML is studied. In this work, different solutions at the different user prospective is discussed and relevant ML techniques are implemented. Based on the results, we suggest the following extension: (1) Most of the work in research is focused on consumer centric approach. Now, we need to explore the domain of administration and business logic implementation on e-commerce using ML. (2) The weighted classification approach is a promising for adopting and identifying properties of data. Therefore, it is needed to explore the performance influence of algorithm by extending attribute selection approaches. (3) The recommendation module need more parameters for identifying appropriateness of product, which will enhance the criteria of product visibility. Therefore, in near future it is required to consider.

References

[1] G. Taher, "E-Commerce: Advantages and Limitations", International Journal of Academic Research in Accounting, Finance and Management Sciences, Vol. 11, No. 1, pages 153-165, 2021.

[2] J. Oláh, N. Kitukutha, H. Haddad, M. Pakurár, D. Máté, J. Popp, "Achieving Sustainable E-Commerce in Environmental, Social and Economic Dimensions by Taking Possible Trade-Offs", Sustainability, Vol. 11, Issue 89, pages 89, 2019

[3] N. Chen, "Research on E-Commerce Database Marketing Based on Machine Learning Algorithm", Hindawi Computational Intelligence and Neuroscience Vol. 2022, Article ID 7973446, pages 13, 2022

[4] T. Chen, P. Samaranayake, X. Y. Cen, M. Qi, Y. C. Lan, "The Impact of Online Reviews on Consumers' Purchasing Decisions: Evidence From an Eye-Tracking Study", Front. Psychol., Sec. Decision Neuroscience, Vol. 13, pages 1-13, 2022

[5] https://www.kaggle.com/datasets/naveedhn/amazon-product-review-spam-and-non-spam

[6] A. Ranjan, M. Misra, J. Yadav, "Online Shopping Behavior During COVID 19 Pandemic: an Indian Perspective", SSRN Electronic Journal, pages 1-24, June 2021, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3874348

[7] Y. K. Dwivedi, E. Ismagilova, D. L. Hughes, J. Carlson, R. Filieri, J. Jacobson, V. Jain, H. Karjaluoto, H. Kefi, A. S. Krishen, V. Kumar, M. M. Rahman, R. Raman, P. A. Rauschnabe, J. Rowley, J. Salo, G. A. Tran, Y. Wang, "Setting the future of digital and social media marketing research: Perspectives and research propositions", International Journal of Information Management, Vol. 59, Id 102168, pages 1-37, Aug 2021

[8] M. Wankhade, A. C. S. Rao, C. Kulkarni, “A survey on sentiment analysis methods, applications, and challenges”, Artificial Intelligence Review, Vol. 55, pages5731–5780, 2022
[9] P. Sasikala, L. M. I. Sheela, “Sentiment analysis of online product reviews using DLMNN and future prediction of online product using IANFIS”, Journal of Big Data, Vol. 7, Article number 33, pages 1-20, 2020
[10] F. O. Isinkaye, Y. O. Folajimi, B. A. Ojokoh, “Recommendation systems: Principles, methods and evaluation”, Egyptian Informatics Journal, Vol. 16, Issue 3, Pages 261-273, Nov 2015
[11] X. Zhou, Y. Hu, L. Guo, “Text Categorization based on Clustering Feature Selection”, Procedia Computer Science, Vol. 31, Pages 398-405, 2014
[12] R. Guan, H. Zhang, Y. Liang, F. Giunchiglia, L. Huang, X. Feng, “Deep Feature-Based Text Clustering and its Explanation”, IEEE Transactions on Knowledge and Data Engineering, Vol. 34, No. 8, pages 3669-3680, Aug 2022
[13] A. Chiche, B. Yitagesu, “Part of speech tagging: a systematic review of deep learning and machine learning approaches”, Journal of Big Data, Vol. 9, Article number 10, pages 1-25, 2022
[14] A. Sharma, S. Dey, “A Document-Level Sentiment Analysis Approach Using Artificial Neural Network and Sentiment Lexicons”, Applied Computing Reviews, Vol. 12, No. 4, pages 67–75, Dec. 2012
[15] M. A. Fauzi, “Random Forest Approach for Sentiment Analysis in Indonesian Language”, Indonesian Journal of Electrical Engineering and Computer Science, Vol. 12, No. 1, pages 46~50, Oct 2018