E-commerce Product Recommendation System Using Enhanced Product Visibility

Nitin Upadhyay and Dr. Deepika Pathak
Department of Computer Application, Dr. A.P.J. Abdul Kalam University, Indore (M.P.) - 452010, India
Corresponding Author Email : nitin_upadhyay86@yahoo.com

Article Info
Page Number: 775-785
Publication Issue:
Vol. 70 No. 2 (2021)

Article History
Article Received:
05 September 2021
Revised: 09 October 2021
Accepted: 22 November 2021
Publication: 26 December 2021

Abstract

E-commerce is becoming much more popular in recent years. A significant amount of e-commerce consumers increased as well as new product vendors also increased in order to get benefit from this open market. But the new product vendors are suffering from low product sales. Therefore, in order to enhance the product visibility of the new vendors needs some improvements on the existing recommendation system. The current recommendation algorithms consider the products based on ratings and reviews, their existing sales, and brand size. Therefore, the new vendors feel the biased nature of any e-commerce platform. The next option is to enhance the product visibility using advertisements which may be much more expensive for new vendors. Therefore, this paper presents a new e-commerce product recommendation model, which provides a win-win situation for all the parties' consumers, vendors, and e-commerce platform administrators. The proposed model also suggests key points that can be considered during the design of an e-commerce business. Additionally, a weighted recommendation system is proposed and evaluated for recommending a set of products to existing customers. The proposed model utilizes a feedback model for refining and recommending products to consumers.

Keywords— E-commerce, product recommendation, consumer satisfaction, Administrative point of view, visibility enhancement.

I. INTRODUCTION

E-commerce has become popular. It makes it easy to buy products and get products at doorsteps. Additionally, product return and replacement make it more valuable for new buyers also [1]. The consumer can log their compliant and queries about the product using reviews or forums. In addition, it provides a platform for product vendors and small businesses to sell their products to the globe. But, new product vendors always suffer from low conversion rates and sales. Therefore, new sellers may feel biased or may drop out of e-commerce. The main reason behind this negative aspect of e-commerce is recommendation engines which are used for recommending products to consumers.

The existing recommendation models are mainly considering the high quality of products, which are already popular on the platform. On the other hand, another way to increase product visibility is to do advertisements for the product on the website. This may increase the cost of product sales and also reduces the profit of the vendor. Therefore, in this paper, we introduce a machine-learning technique to deal with the product recommendation problem.

The proposed technique utilizes user feedback for improving the next level of product recommendation. The proposed technique is not only enhancing the product visibility to all the product vendors it also helps to deal with the cold start problem of product recommendation. The cold start problem is the problem of the recommendation system where the recommendation system does not have any prior information related to end user behavior for recommending the product.

In this section, we provide an overview of the presented work. Next, the details of the proposed system are discussed, thus first we provide the model of the proposed system and the algorithm is also described. Further, the experiments are conducted on the web access log dataset, and the performance of the proposed technique is measured. Finally, the conclusion of this study is discussed and future directions are planned.

II. PROPOSED WORK

The recommendation system design is an application of the Web mining technique. Web mining is a collection of data mining algorithms utilized for analyzing web data. The aim of web mining is to recover the application-centric patterns available in direct or indirect form on the web. The direct information is visible and accessible to the public openly for example web content on web pages. On the other hand, indirect web data is not directly accessible to all web users. This data is generated on the background of the servers such as web access logs.

In this paper, an e-commerce recommendation system is discussed. The recommendation system was mainly designed for e-commerce to guide or suggest the appropriate products to the e-commerce product buyers. This system utilizes the previous user information of e-commerce access behavior in terms of frequent buying products, brands, cost range, trying new product brands, and other relevant information. This information helps to understand the requirements of the end-user and other historical buying patterns. Behavioral and purchasing patterns are used for recommending the products. But the recommendation techniques are constrained by the business logic. The business logic enables an e-commerce company to make revenue by managing the platform. The business logic is related to marketing, promotions, discounts, and coupons. Additionally, the less information about the client or product the recommendation becomes complex for actualization.

Therefore, we need to design an enhanced recommendation system. The proposed enhancement includes a solution for the cold start problems where no information is available about a new consumer. Additionally, the key points are discussed to employ for enhancing the visibility of new vendors as well as helping manage promotional activities. Therefore, a new strategy is used to design the best condition for product vendors, consumers, and e-commerce administration. The proposed recommendation system for solving the addressed problem is discussed in this section. The modules and the relevant outcomes are also discussed.

A. Initialization of Recommendation System

The aim is to prepare appropriate data for utilizing the proposed recommendation system. In this context, the web access log is used. Figure 1 demonstrates the web server access log generation and processing of the information.

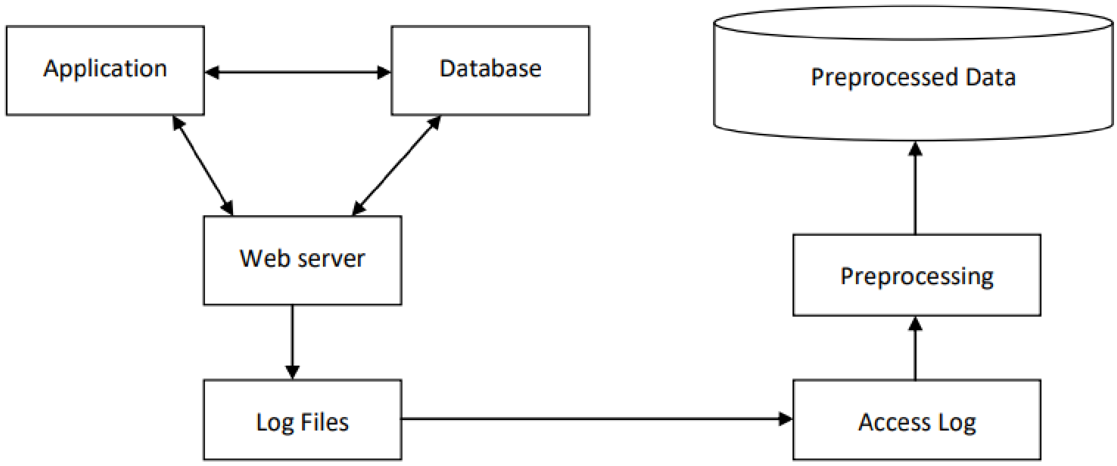


Figure 1: Shows the web server and access log processing

The web access log is generated at the web server, where the e-commerce application is parked. The web servers are maintaining different log files for different server events. Among them, the web access log contains the entry about requests and responses of the application resource. The Web access log contains different attributes such as IP address, timestamp, response, protocol, methods, resource name, and others. Not all web access log information is essential for recommendation system design. Therefore, in preprocessing phase we remove the unwanted web access log attributes. After preprocessing the attributes such as IP address, resource name, and time stamp are preserved and the remaining data is removed. The preserved data of the access log is transformed into a relational data table and kept separately for further processes.

B. Basic Recommendation

The process of performing the initial recommendation without any prior user information on a model is presented in figure 2. Mainly this phase of recommendation utilizes the initial activity of the user. Therefore, the preprocessed database is utilized as given in figure 1. The preprocessed data contains the URLs of the products which are requested by the e-commerce user during product search. The list of products is a combination of promoted, advertised, popular, and highly demanding product brands as well as the new vendor similar products are also considered. This list of combined preprocessed data is further transformed into transaction sets based on different user session activities.

The transactions based on the user's session activities are used for frequent pattern mining. Therefore the apriori algorithm is applied to preprocessed data. The apriori algorithm is a popular frequent pattern mining algorithm. This kind of data analysis returns a subset of the product data which is found frequently in different user sessions. The computed frequent patterns and the user's initial activity of product search are used to find the appropriate product as the first stage product recommendation. This recommendation is based on the initial user's product search data and available frequent products. After this stage of recommendation, the search space is reduced by involving the advanced recommendation.

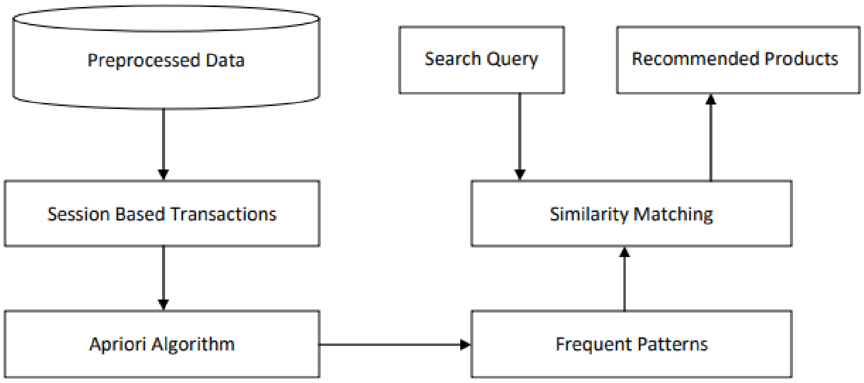


Figure 2 Primary Product Recommendation

Table 1 demonstrates the process involved in the initial product recommendation. The algorithm accepts the web access log file L, business logic-based constraints B, and the initial search query of user Q. First the log file L is read and preprocessed. The preprocessed data is combined with the business logic and generate a set of preprocessed data. The preprocessed data P_n, is used with apriori algorithm to compute the frequent pattern F_n. Next the similarity between F_n and Q is measured to recommend initial list of products P_r

Table 1 Algorithm for Initial 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. $if\left(F_i.contains(Q)\right)$
        i. $P_r.Add(F_i)$
    b. $end\ if$
5. End for
6. Return P_r

C. 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. 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 in the initial product recommendation. 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 for a recommendation.

Table 2 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. $if(P_{r_i}.Clicked)$
        i. $Feedback.Add(P_{r_i})$
    b. $end\ if$
2. End for
3. $S = KNN.Search(Feedback, P_n)$
4. Return S

Table 2 demonstrates the process of feedback involved in the initial product recommendation. 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 3.

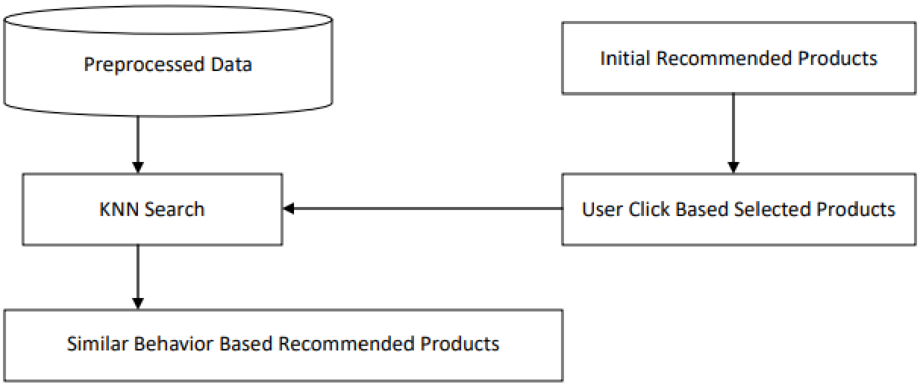


Figure 3 User Feedback based product recommendation

D. Final Recommendation

In this phase, 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 product recommendation. Figure 4 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.

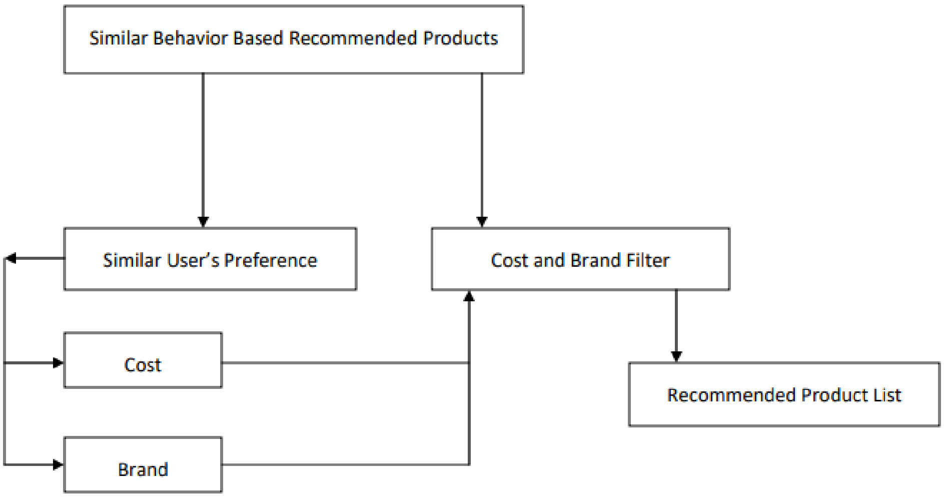


Figure 4 Final stage recommendation

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 range and popular brands among a similar group of users 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. Both the influencing 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 4 and table 3.

Table 3 consists of the steps to follow for performing the final stage recommendation. The algorithm accepts the user feedback based on generated product recommendation S and produces a refined and precise final recommendation R. in this process first we recover two factors first the average cost of similarly behaving user and the list of product brands, which are used or accepted 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 value of all the products we sort the list S. From the sorted product list of S is used to discover the top-ranked list of products is produced as the final recommendation of the products.

Table 3 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

III.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 4 and figure 5.

Table 4 Performance of the proposed recommendation system

Experiment 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 is describing the correctness of the proposed recommendation system. That is measured as the ratio of correctly recognized next products and total samples given to be recognized. The following formula can be used for accuracy.

$$accuracy = \frac{correctly\ identifyied\ samples}{total\ samples} X100$$

The accuracy of the proposed recommendation system is given in figure 5(A) and table 4. 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 recommendation system is measured in terms of percentage (%). According to the experimental results, the accuracy is increasing with the sample size of experimental data for building the model.

The next parameter for evaluation is the Error rate. That is used to find the misclassification or prediction of the recommendation system. 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 5(B) contains the experimental results in terms of error rate (%). The Y axis shows the error rate percentage of the recommendation system and the X axis includes the sample size of data used in experiments. According to the obtained error rate we find that the error rate is not fluctuating highly therefore the performance of the system is much more consistent even when the experimental data samples are increasing in different experiments. Therefore, the developed recommendation system is reliable for solving the addressed issues.

Further for measuring the efficiency of the proposed model we also calculated the time and space utilized by the proposed technique. Thus first we calculated the time required for performing the final prediction. The amount of time consumed for performing this task using the proposed recommendation system is called the time consumption. The time requirement is measured using the following formula:

$$time\ consumed = \ end\ time - start\ time$$

The time consumption of the proposed recommendation model for appropriate product suggestions is demonstrated in figure 5(c). The X-axis contains the sample size of experimental data and the Y-axis contains the time consumed in recommendations. The time is measured in terms of milliseconds (MS). According to the results, the time consumption in

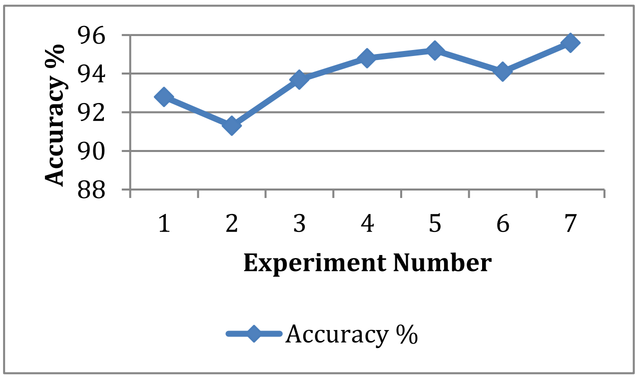
ISSN: 2094-0343

different experiments are varying with the sample size and remains consistently increasing. Time consumption is also reducing with the amount of data increases in the web access log file therefore system is acceptable.

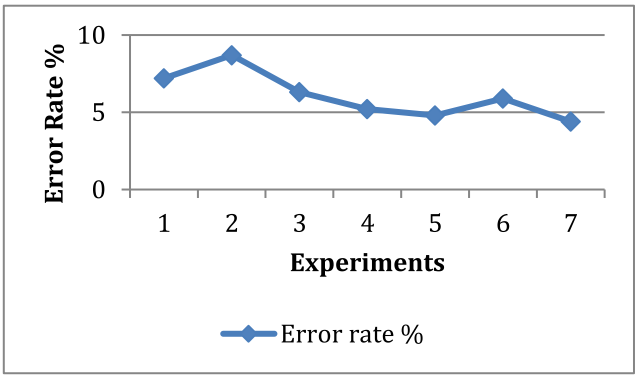
Finally, memory usage is measured in terms of kilobytes (KB). That can be measured based on the process. The execution of a process takes an amount of main memory. This memory size is known as memory usage. Actually, it is the difference between the total memory assigned to the process and the amount of space free to be used. That can be calculated using the following equation:

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

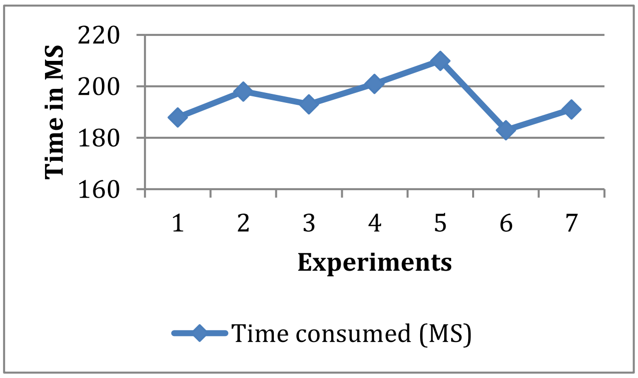
The measured memory usage is given in figure 5(D) and table 4. In this diagram, X-axis contains the sample size used in experiments, and the Y axis shows the used memory by the system. Memory usage is depending on the amount of data to be used. The memory is not varying much rapidly for the product recommendation.



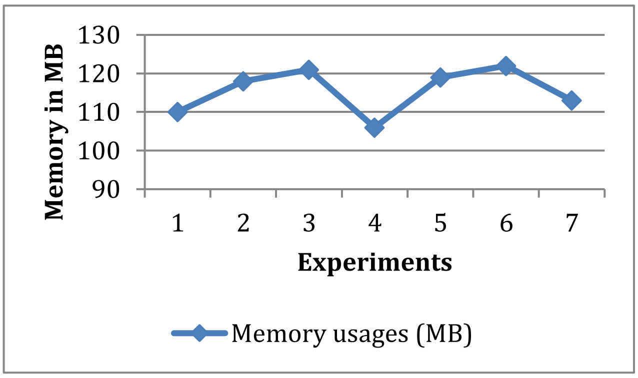
(A)



(B)



(C)



(D)

Figure 5 shows the experimental consequences of the proposed recommendation model in terms of (A) Accuracy (%) (B) Error rate (%) (C) Time utilization and (D) Memory utilization

IV. CONCLUSION & FUTURE WORK

E-commerce is offering products and services online and delivering them to the consumer's door using a website or a mobile application. The e-commerce provides exclusive features to their consumer like easy and timeless return and replacement. However, in literature a number of studies around e-commerce recommendation systems are available. But most of the study is limited to provide the solution centric to the consumers, and there is no study or very fewer studies are available which are dealing with the issue of new product vendors.

In this paper, we first discuss the issues of e-commerce vendors and their struggle to increase visibility of their newly lunched products. In this context, this paper involves the implementation of recommendation system from the administrative point of view. The method is promising to implement business conditions to an ecommerce and also offers the opportunity to enhance the product visibility of new product vendors. This will provide win-win conditions e-commerce administration, product vendors and consumers.

The proposed recommendation system is designed in four steps:

1. **Initialization phase:** in this phase the web access log file is preprocessed and transformed into the session based sequences (transactions) of different users.
2. **Initial recommendation:** After initialization the transactions are used with frequent pattern algorithm. This process used with the linear search for identifying the similar and much frequent products according to user query.
3. **Feedback acceptance:** Initial recommendation is used for collection of user feedback. That feedback is utilized to identify the similar users which follow the similar session transactions. Based on the similar behavior user's preferences the initial recommendation is refined.
4. Final recommendation:

The prototype model of the discussed model is developed and experiments are carried out. The obtained results demonstrate the accurate, reliable, and efficient recommendation model which is not only able to deal with the administrative point of view it also includes the solution for the cold start problem of recommendation.

REFERENCES

[1] Amit Ranjan, Madhvendra Misra, Jitendra Yadav, "Online Shopping Behavior During COVID 19 Pandemic: an Indian Perspective", SSRN Electronic Journal · June 2021

[2] 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, Volume 59, August 2021, 102168

[3] M. Wankhade, A. C. S. Rao, C. Kulkarni, "A survey on sentiment analysis methods, applications, and challenges", Artificial Intelligence Review volume 55, pages5731–5780 (2022)

[4] 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 volume 7, Article number: 33 (2020)

[5] F. O. Isinkaye, Y. O. Folajimi, B. A. Ojokoh, "Recommendation systems: Principles, methods and evaluation", Egyptian Informatics Journal, Volume 16, Issue 3, November 2015, Pages 261-273

[6] A. Iqbal, R. Amin, J. Iqbal, R. Alroobaea, A. Binmahfoudh, M. Hussain, "Sentiment Analysis of Consumer Reviews Using Deep Learning", Sustainability 2022, 14, 10844

[7] X. Zhou, Y. Hu, L. Guo, "Text Categorization based on Clustering Feature Selection", Procedia Computer Science Volume 31, 2014, Pages 398-405

[8] Renchu Guan, Hao Zhang, Yanchun Liang, Fausto Giunchiglia, Lan Huang, and Xiaoyue Feng, "Deep Feature-Based Text Clustering and its Explanation", IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 34, NO. 8, AUGUST 2022

[9] M. Ali Fauzi, "Random Forest Approach for Sentiment Analysis in Indonesian Language", Indonesian Journal of Electrical Engineering and Computer Science, Vol. 12, No. 1, October 2018, pp. 46~50

[10] Anuj Sharma, Shubhamoy Dey, "A Document-Level Sentiment Analysis Approach Using Artificial Neural Network and Sentiment Lexicons", APPLIED COMPUTING REVIEW DEC. 2012, VOL. 12, NO. 4

[11] Alebachew Chiche, Betselot Yitagesu, "Part of speech tagging: a systematic review of deep learning and machine learning approaches", Journal of Big Data volume 9, Article number: 10 (2022)