

Predicting the user navigation pattern from web logs using weighted support approach

Om Prakash P.G¹, Jaya A², Ananthakumaran S³, Ganesh G⁴

^{1,3}Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh, India

²Department of Computer Applications, B.S.A.Crescent Institute of Science and Technology, Chennai, Tamilnadu, India

⁴Department of Computer Science, BlueCrest College, Freetown, Sierra Leone

Article Info

Article history:

Received Mar 13, 2020

Revised Sep 7, 2020

Accepted Sep 18, 2020

Keywords:

Classification

Mining

Prediction

User behavior

User navigation

Web mining

Web traversal

Weblog

ABSTRACT

A weblog contains the history of previous user navigation pattern. If the customer accesses any portal of organization website, the log is generated in web server, based on sequence of user transaction. The weblog stored in the web server as unstructured format, it contains both positive and negative responses i.e. successful and unsuccessful responses, identifying the positive and negative response is not useful for identifying user behavior of individual user. Initially the successful response is taken, from that conversion of unstructured log format to structured log format through data preprocessing technique. The process of data preprocessor contains three step process data cleaning, user identification and session identification. The pattern is discovered by preprocessing technique from that user navigation pattern is generated. From that navigation pattern classifier technique is applied, the conversion of sequence pattern to sub sequence pattern by clustering technique. This research is to identify the user navigation pattern from weblog. The Improved Spanning classification algorithm classifies the frequent, infrequent and semi frequent pattern. To identify the optimal webpage using classificatopn algorithm from that user behavior is identified.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Om Prakash P.G

Department of Computer Science and Engineering

Koneru Lakshmaiah Education Foundation, Vijayawada

Andra Pradesg, Inida

Email: ommail2004@gmail.com

1. INTRODUCTION

Now a day lot of products available in online websites, so the number of users is increased due to attraction of price discounts in online shopping. In the current situation, the percentage of users is increased day by day to access the internet. The users are utilizing the internet lot of time to access the websites; the lot of products offers is available in websites. So more number of user shown interest in online shopping.

To identify the user's optimal page is a difficult task in online shopping. While the users accessing the website, the user searching transactions are stored as log format. Initially, the log as unstructured format, the conversion of unstructured log to structured information by data pre-processing technique. The weblog consists of various entries like Logical address of user, date & time stamp, categories of product details and status code of the webpage. The pre-processing is a three-step process. The data cleaning contains relevant and irrelevant data; it cleans the irrelevant data in log. The user identification is identified by IP address of the specified user that is stored in the weblog. The session identification is classified based on session management of each user accessing the websites.

Every organization identifies the user's demand, so organization will manufacture depends upon users need. If the product is currently not available to the user, the user may purchase some other product that is available in the online websites. If the product is crossed benchmark production, the sale of product to the customer is a biggest challenge, so the organization identifies the need the demand of the user. The calculation of identify the demand is different for each product. So the organization has to concentrate the need of the users demand. With the help of weblog the demand of each of the individual user needs are monitored.

The process of identifying user needs and user interest is a difficult task. With the help of weblog the user needs and interest of every user can identify by user navigation pattern. The user navigation path pattern can be analyzed by history of pattern that is taken from weblog. Using pattern discovery is the process of convert sequence in to sub sequence of similar pattern. By using forward path and backward path technique subsequence pattern can be generated. Then the sub sequence pattern is grouped to form a cluster, that will helpful to identify the number of user needs. The user interest is a biggest task of every organization. By using improved span classification algorithm the user interest will classified.

The paper is organized as follows, Section 2 is deals review of literature, Section 3 deals with problem definitions, Section 4 describes framework for user navigation system, and Section 5 gives the results & discussion. Finally Section 6 concludes the paper.

2. LITERATURE SURVEY

WUM is a mining technique that is to identify the user pattern from that navigational behaviour is analysed that was stored in weblog. Wangshu *et al.* [1] showed that preprocessing has two stages, first it eliminates unwanted and similar datasets and then it removes irrelevant features of ranking. Xin *et al.* [2] showed that (OSN) classified in to extroversive and introversive behavior. First, it identifies how users communicate with online friends then it gathering social information in the following websites.

Guoshuai *et al.* [3] showed that prediction service is based on rating. Ruili *et al.* [4] suggested that user behaviour pattern, it analyzesth of actual usage and anticipated usage. The actual usage is get transaction database based on sessions. The anticipated usage has to get the session from the count of user path is identified. Surbhi *et al.* [5] showed that back navigation approach utilizes forward path and backward path to identify frequent pattern, semi frequent pattern and infrequent pattern. Zliobaite *et al* [6] identified the prediction accuracy by adaptive pre-processing algorithm. Om Prakash *et al.* [7] suggested user pattern is classified by navigation behaviour. User behavior can be analyzed based on the count of similar data. If the count of user is above the benchmark level then the user can be considered as interested user (IU), if it is below the level then it is not interested user (NIU) from that user behavior is analyzed.

Dr. A. R. Patel *et al.* [8] identified initially weblog initially contains raw log, so conversion of raw data to processed information by using pre-processing technique, the data pre-processing has the following process of cleaning the data then identify the user based on ip address and by using session. Hong Cheng *et al.* [9] showed that the traversal pattern, it contains sequence of pattern. To convert the sequence of pattern in to sub sequence pattern based on navigation pattern modeling. The classification algorithm has three sequences of pattern generation, it has frequent, semi-frequent and in-frequent. Kerana *et al.* [10] identified that, the traversal pattern utilizes database transaction, the sequence of similar and dissimilar transactions are converted into sub sequence by reducing number of similar transactions to form clusters, through the clustering technique it classifies frequent, semi frequent and in frequent item data sets, from that user buying behavior is identified.

Mobasher *et al.* [11] identified the personalize of each user. The unprocessed weblogs are converted into processed log by using data pre-processing. After pre-processing the sequence of pattern is generated by using traversal pattern, it uses forward path the sequence of path is breakdown the sequence in to sub sequence path. The clustering algorithm reduces the number of similar path grouped to form a cluster. The classification classifies the user interested pattern and user not interested pattern. Rajesh *et al.* [12] showed the three main classification models are random forest, SVM and K-neighbors classifiers, by using single classifier the prediction accuracy is 30%, while using all the three classifier the prediction accuracy is increased up to 51%. Vinothkanna *et al.* [13] suggested to analysis the future detection RBFN classifier is used for effective classification, RBFN is a good classification and its recognition rate is nearly 99.2%.

Kousar Nikhathet *et al.* [14] discussed text clustering approach along with distance-based approach model helps in optimizing the text documents. Babu *et al.* [15] suggested rapid miner tool is a data mining technique; it helps to cluster the information to get high profit and less risk to customers, that yields profit to the customers. Gummadi *et al.* [16] proposed the algorithm to improve the performance of the system using clustering technique. The clustering technique algorithm uses the LEACH, CLAEER and mean shifted algorithm to improve the accuracy of energy efficient routing in vehicle tracking using wireless sensor networks. Karthikeyan *et al.* [17] suggested neighbor based cluster location aware routing (NCLAR) helps to

achieve more packet delivery rate with high location accuracy. Initially cluster region has greater signal strength with less signal delay, then routing table is constructed is formed with less signal delay value and more signal strength. In second phase, neighbor node routing table is constructed and updated with addition of more fields.

Bommadevara *et al.* [18] identified K-means algorithm is suitable for predicting the route cause for getting the Heart disease. Heart disease prediction is used to determine the root cause of getting heart attack and the probability of getting a heart attack. Akhila *et al.* [19] Suggested frequent pattern mining and cluster analysis, measure and analysis of energy utilization of human activity, to predict human health through smart devices. Angel Prathyusha *et al.* [20] utilized Hybrid soft decision model is a classification techniques, the classifier predict the farmer growth in agriculture development. It helps to cultivate the best suitable crops rather than unsuitable crops, by using this technique it attained 92% of accuracy. Anjali Devi *et al.* [21] utilizes apriori algorithm, it utilizes previous academic information from that to predict the placement of the current student. Arshad *et al.* [22] proposed the machine learning based attack detection model is best than the traditional statistical based learning model and rule based learning models in terms of time, detection rate are concerned. Arvind Selwal *et al.* [23] suggested Machine Learning Model, it classify seven to eleven features, by this model it predict thyroid nearly 99.8%. Chanintorn *et al.* [24] utilized regression model, it will forecast the with optimal solutions. Wahab *et al.* [25] analysed prediction of energy consumption of two laboratories by using artificial neural network algorithm.

Research GAP:

- a) There exists a difficulty in capturing the individual user behaviour.
- b) Visitors face difficult to find the relevant web pages that leads to wastage of time.
- c) Need to reduce the mining time while capturing the user behavior from weblogs.
- d) There is a lack of prediction accuracy to analyze user behavior.
- e) User behavior is difficult to analyze the unstructured data.

3. PROPOSED FRAMEWORK FOR USER NAVIGATION SYSTEM

Figure 1 shows the framework for user navigation system, to identify the user behavior, the following techniques are applied, it consists of preprocessor, navigation pattern modeling and classification technique. The role of preprocessor converts the raw log to processed information. The Navigation pattern modeling converts the above sequence to subsequences. The classification algorithm classifies the above sequence into frequent sequence and infrequent sequence.

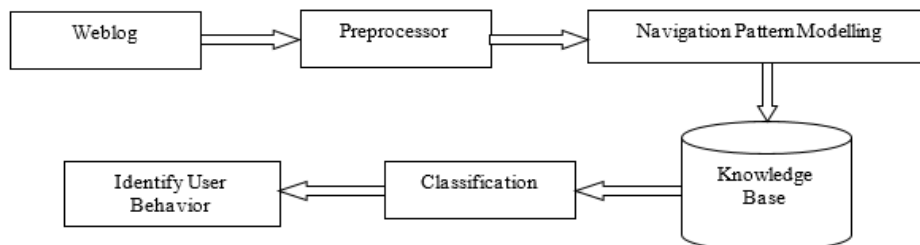


Figure 1. Framework for user navigation system

3.1. Weblogs

Whenever the user accesses the web site, the log is recorded in the form of IP address, categories, date & time, bytes transmission and status code of each webpage. The log consists of following user attributes

- a) IP address
- b) Date and time
- c) Request method
- d) URL of the page
- e) Categories
- f) Bytes transmission
- g) Status code

Figure 2 shows the structure of weblog, the log consists of Ip address, date and time, request method, URL of page, number of bytes transmission and status code for the webpage.

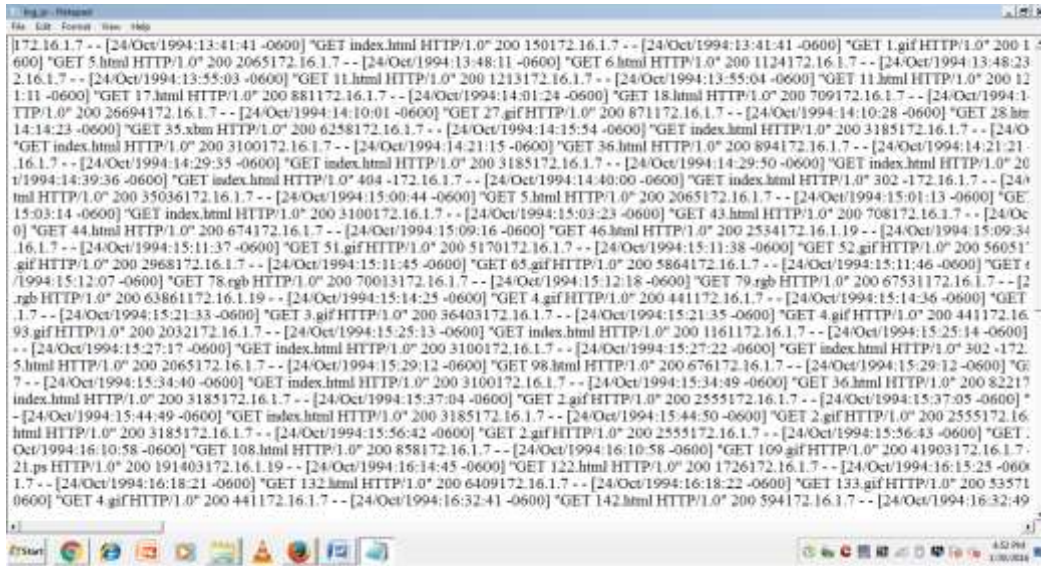


Figure 2. Weblog structure

3.2. Preprocessing

In pre-processing, it has unstructured data format of user transaction. The pre-processing contains three step processes

- a) Data cleaning
- b) User identification
- c) Session identification

3.2.1. Data cleaning

These logs have been taken from the vehicle web server, the logs taken from the period of 23/Oct/2016 to 30/Oct/2016 The log file has 3568 data sets, in that each data set contains the IP address, Categories and status code. The data cleaning will take care of irrelevant information when the user was visited sited earlier. It cleans the incomplete information that is available in websites. The status code is generated whenever user transforms to another page. The status code helps to clean the irrelevant information in log. After the data cleaning process in log 1568 records are obtained.

E.g.: from Figure 2 shows “172.16.2.7- [23/Oct/2016-0600] "GET 9.gif HTTP/1.0" 200 441” the stats code 200-400 is successful webpage of the user.

3.2.2. User identification

The user is identified from cleaning process, it has Users IP address, request time, requested URL, date & time.

E.g.: Figure 2 shows the user IP address 172.16.2.7 is identified from datasets.

3.2.3. Session identification

It classifies the user based on time sessions, if the similar user visits the website two times, so two sessions must be recorded.

E.g.: Figure 2 identified the similar ip address of different time stamp from 13:41:41 to 13:45:46 & 14:05:31 to 14:08:54 is considered as two session.

3.3. Navigation pattern modelling

The Table 1 has the browsing pattern of the users. Figure 3 shows the tree structure of navigation pattern modeling, from the above weblog the tree structure is generated, T1 consider as Home page, T2 as City selection, T3 as Two wheeler, T4 as LMV, T5 as Heavy Vehicle, T6 as Petrol, T7 as Diesel, T8 as TVS, T9 as Honda, T10 as Bajaj, T11 as Enfield, T12 as Maruti suzuki, T13 as Hyundai, T14 as OPEL, T15 as VAN, T16 as Saab, T17 as Bus, T18 as Manual, T19 as Automatic, T20 as Color Selection, T21 as Payment Gateway, T22 as Credit Card, T23 as Debit Card, T24 as Cash, T25 as Loan option.

Table 1. User pattern

Pattern Nos	Pattern
1	< T1,T2,T3,T5,T6,T8,T18,T20,T21,T22>
2	<T1,T2,T3,T5,T6,T10,T18,T21,T22,T23>
3	<T1,T2,T5,T7,T17,T7,T16>
4	<T1,T2,T4,T5,T7,T16,T19,T20,T21,T25>
5	<T1,T2,T4,T5,T7,T15,T19,T20,T21,T24>
6	<T1,T2,T3,T5,T7,T11,T6,T9>
7	<T1,T2,T4,T5,T6,T4,T7>

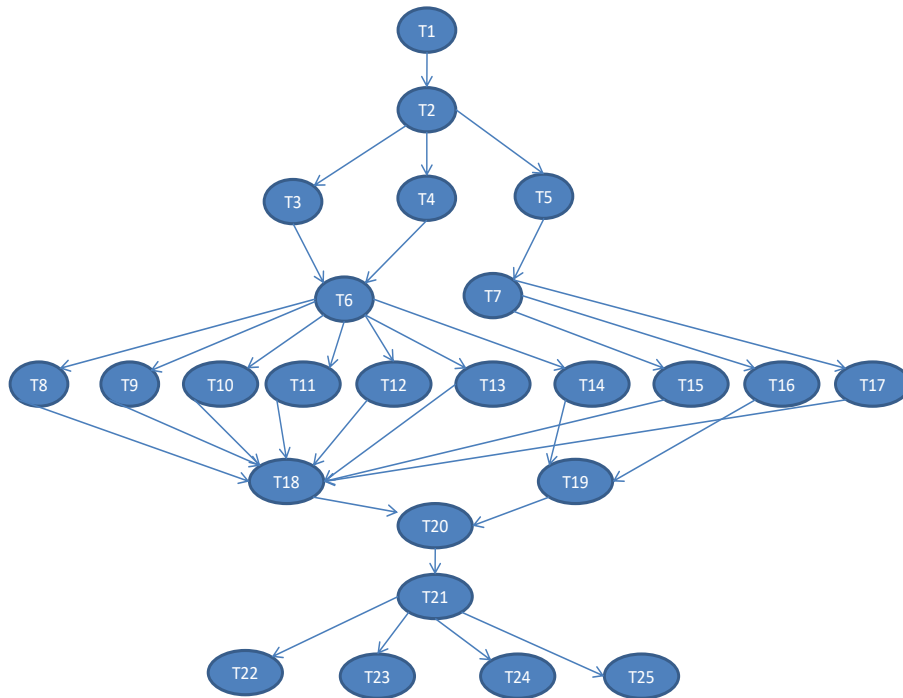


Figure 3. Tree structure of navigation pattern modelling

3.4. Classification

Using Improved Spanning algorithm, the pattern is converted in to FS, SFS and IFS. It classifies the pattern by count. The pattern count is above the threshold then it has IU, so users are interested in purchase, then user as NIU, then the user not interested in purchase. Table 2 shows the count of user pattern. Table 3 shows the buying interest of vehicle for the individual user. If the count of the pattern determines the Buying interest of the vehicle.

Table 2. Pattern count

Pattern Nos	Pattern	Count
1	< T1,T2,T3,T5,T6,T8,T18,T20,T21,T22>	16
2	<T1,T2,T4,T5,T7,T15,T19,T20,T21,T24>	17
3	<T1,T2,T4,T5,T7,T16,T19,T20,T21,T25>	11
4	<T1,T2,T3,T5,T6,T10,T18,T20,T21,T23>	16
5	<T1,T2,T5,T5,T7,T17,T7,T16>	2

Table 3. Interested and not interested users

Pattern Nos	Pattern	Interested/Not Interested
1	< T1,T2,T3,T5,T6,T8,T18,T20,T21,T22>	Interested
2	<T1,T2,T4,T5,T7,T15,T19,T20,T21,T24>	Interested
3	<T1,T2,T4,T5,T7,T16,T19,T20,T21,T25>	Interested
4	<T1,T2,T3,T5,T6,T10,T18,T20,T21,T23>	Interested
5	<T1,T2,T5,T7,T17,T7,T16>	Not Interested

Improved Spanning Algorithm
(Sequential Pattern, FS, SFS, IFS)

```

{
Step 1: Get the Input Sequence Pattern  $\alpha$ 
Step 2: Scan the Sequence item b with  $\alpha$ 
Step 3: for each b with  $\alpha$ 
Step 4: if item b greater than max_sup
Step 5: Push item b in FS "Frequent Sequence" // end of if
Step 6: else if item b greater than min_sup
Step 7: Push item b in SFS "Semi Frequent Sequence" // end of else if
Step 8: else Push item b in IFS "In Frequent Sequence" // end of else
Step 9: End for
} //end of algorithm
    
```

4. RESULTS AND DISCUSSION

The experimentation will be performed using vehicle data. The implementation will be done using Weka tool and the performance will be compared with the existing algorithms based on precision, recall and F-measure. Figure 4 shows the Vehicle compactness, Figure 5 shows Circularity of vehicle, Figure 6. Shows the Categories of individual vehicles and Figure 7. Shows the Different class of vehicles. The proposed method is implemented using improved spanning algorithm with a knowledge base which has 1568 instances in log. The maximal is 119, mean is 93.678, Std.Dev is 8.234 and the minimum compactness is 73. In total 19 attributes were generated. The accuracy is calculated by the number of instances generated by the system and the total number of attributes generated. The IU and NIU can be classified based on the threshold of the system. Figures 4-7 shows the sample output screen for search result.

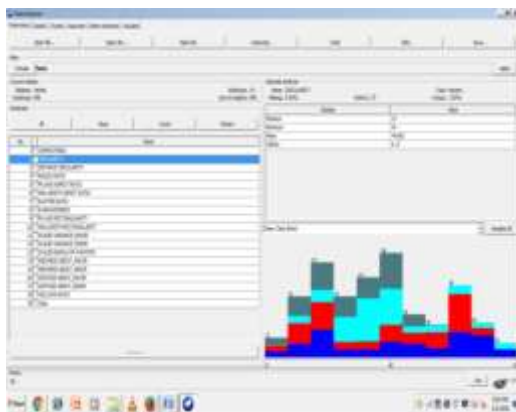


Figure 4. Vehicle compactness

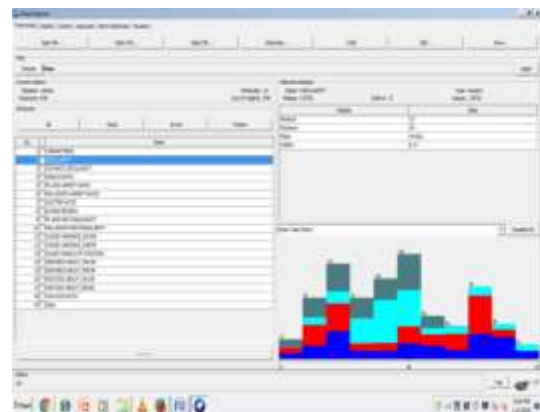


Figure 5. Circularity of vehicle

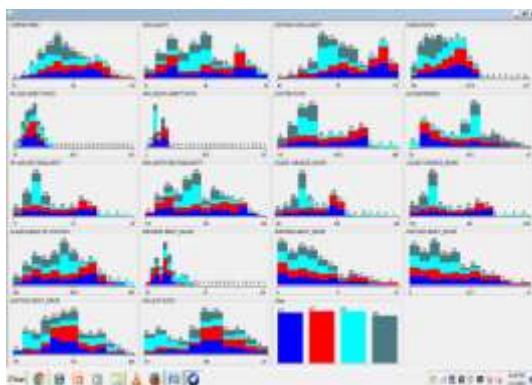


Figure 6. Categories of individual vehicles



Figure 7. Different class of vehicles

The Figure 8 shows the buying vehicle prediction in various classes of people and compare the attributes such as fuel, color, brand, cost, technology and compact, the user has compared with the above maximum attributes from that prediction is identified. The Figure 9 shows the comparison of existing and proposed methods. The graph analysis method has precision, recall, and F-measure, the modified improved span algorithm values are 0.8742, 0.8534, and 0.8775 respectively.

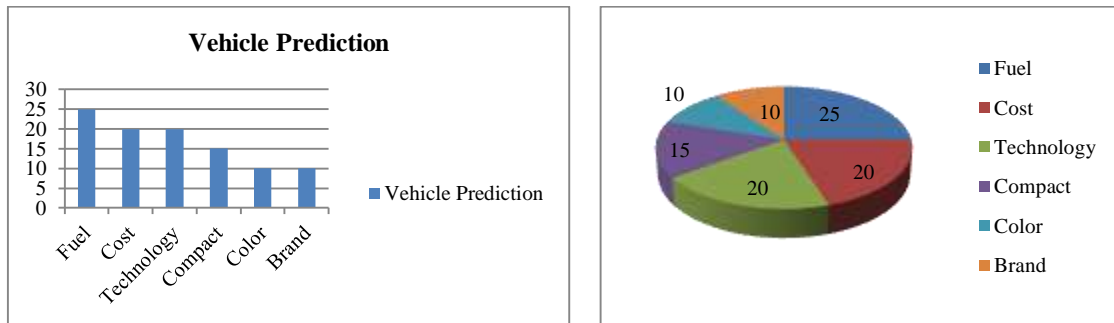


Figure 8. Vehicle prediction

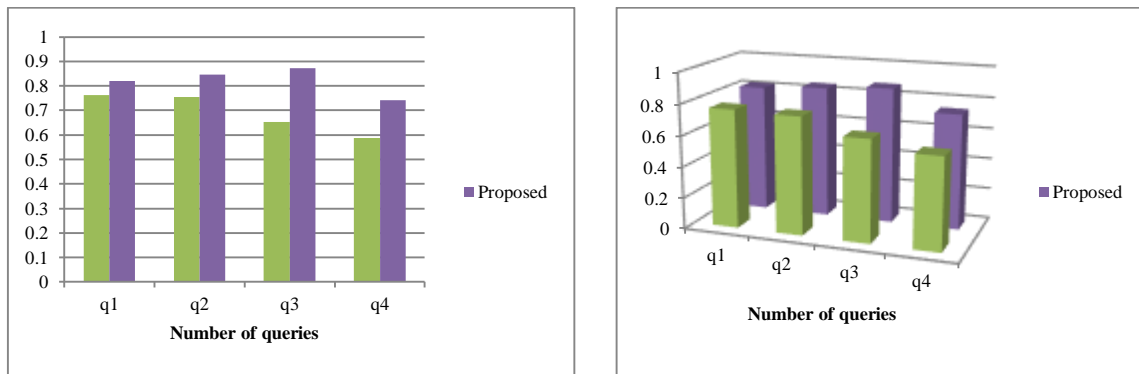


Figure 9. Comparison graphs of existing and proposed algorithm

The comparative discussion Table 4 shows the comparison of existing and proposed methods of predicting the buying behavior of vehicle. The analysis is compared with existing and proposed method. The precision, recall, and F-measure of the modified improved span method are 0.8742, 0.8534, and 0.8775, respectively, that is greater than the existing methods. The precision, recall, and F-measure of existing methods of prefix span are 0.7903, 0.7396, and 0.6784 respectively.

Table 4. Comparative discussion

Datasets	Methods	Precision	Recall	F-measure	Prediction accuracy
Vehicle Dataset	Prefix Span Algorithm	0.7903	0.7396	0.6784	0.5934
	Improved Spanning Algorithm	0.8242	0.8534	0.8775	0.7375

5. CONCLUSION

The proposed method is implemented by using improved spanning classification algorithm, the proposed algorithm is implemented for vehicle prediction for that weka tool is used. The system will show the prediction level of user based on attributes such as fuel, cost, technology, compact, color and brand from that user buying interest is generated and improves prediction accuracy. The prediction accuracy of each user can be identified by the behaviour of user path that is taken from weblog, based on that IU and NIU can be classified by modified improved span classification algorithm. The system will focus to increase the buying prediction behavior from weblog.

REFERENCES

- [1] W. Liu, S. Liu, Q. Gu, J. Chen, X. Chen and D. Chen, "Empirical Studies of a Two-Stage Data Preprocessing Approach for Software Fault Prediction," in *IEEE Transactions on Reliability*, vol. 65, no. 1, pp. 38-53, 2016.
- [2] X. Ruan, Z. Wu, H. Wang and S. Jajodia, "Profiling Online Social Behaviors for Compromised Account Detection," in *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 1, pp. 176-187, 2016.
- [3] G. Zhao, X. Qian and X. Xie, "User-Service Rating Prediction by Exploring Social Users' Rating Behaviors," in *IEEE Transactions on Multimedia*, vol. 18, no. 3, pp. 496-506, 2016, doi: 10.1109/TMM.2016.2515362.
- [4] R. Geng and J. Tian, "Improving Web Navigation Usability by Comparing Actual and Anticipated Usage," in *IEEE Transactions on Human-Machine Systems*, vol. 45, no. 1, pp. 84-94, 2015, doi: 10.1109/THMS.2014.2363125.
- [5] S. Huria and J. Singh, "Implementation of Dynamic Association Rule Mining Using Back Navigation Approach," *2015 Fifth International Conference on Communication Systems and Network Technologies*, Gwalior, pp. 1048-1050, 2015.
- [6] I. Zliobaite and B. Gabrys, "Adaptive Preprocessing for Streaming Data," in *IEEE Transactions on Knowledge and Data Engineering*, vol. 26, no. 2, pp. 309-321, 2014, doi: 10.1109/TKDE.2012.147.
- [7] Om Prakash, P. O. *et al*, "Analyzing and predicting user behavior pattern from weblogs." *International Journal of Applied Engineering Research*, vol. 11, no. 9, pp. 6278-6283, 2016.
- [8] Om Prakash P.G. and A. Jaya, "Analyzing and predicting user behavior pattern from weblogs," *International Journal of Applied Engineering Research*, vol. 11, pp. 6278-6283, 2016.
- [9] Ketul Patel and A. R. Patel, "Process of Web Usage Mining to find Interesting Patterns from Web Usage Data," *International Journal of Computers & Technology*, vol. 3, pp. 144-148, 2004, doi:10.24297/ijct.v3i1c.2767.
- [10] Hong Cheng, Xifeng Yan, Jiawei Han, "IncSpan: Incremental Mining of Sequential Patterns in large database," available in http://hanj.cs.illinois.edu/pdf/kdd04_IncSpan.pdf.
- [11] D. Kerana Hanirex *et al*, "Efficient Algorithm for Mining Frequent Itemsets Using Clustering Technique", *International Journal on Computer Science and Engineering*, vol. 3, no. 3, March 2011.
- [12] Mobasher *et al*, "Automatic personalization based on Web usage mining", *Communications of the ACM*, vol. 43, pp. 142-151, 2000.
- [13] P. Rajesh, N. Srinivas, K. Reddy, G. Vamsipriya, M. Dwija and D. Himaja, "Stock trend prediction using ensemble learning techniques in python," *International Journal of Innovative Technology and Exploring Engineering*, vol. 8, pp. 150-154, 2019.
- [14] R. Vinothkanna and T. Vijayakumar, "Using contourlet transform based RBFN classifier for face detection and recognition," *International Conference on ISMAC in Computational Vision and Bio-Engineering*, vol. 30, pp. 1911-1920, 2019, doi:10.1007/978-3-030-00665-5_176.
- [15] Kousar Nikhath *et al*. "Feature selection, optimization and clustering strategies of text documents". *International Journal of Electrical and Computer Engineering*, vol. 9, no. 2, pp. 1313-1320, 2019.
- [16] Babu, B. S *et al*. "Customer data clustering using density based algorithm". *International Journal of Engineering and Technology(UAE)*, vol. 7, no. 2, pp. 35-38, 2018, doi:10.14419/ijet.v7i2.32.13520.
- [17] Gummedi, A *et al*. "EECLA: Clustering and localization techniques to improve energy efficient routing in vehicle tracking using wireless sensor networks". *International Journal of Engineering and Technology (UAE)*, vol. 7, pp. 926-929, 2018.
- [18] Karthikeyan, T *et al*. "Investigation on maximizing packet delivery rate in WSN using cluster approach". *Wireless Personal Communications*, vol. 103, no. 4, pp. 3025-3039, 2018.
- [19] Bommadevara *et al*. "Heart disease prediction using machine learning algorithms". *International Journal of Innovative Technology and Exploring Engineering*, vol. 85, no. 5, pp. 270-272, 2019.
- [20] Akhila, G *et al*. "A survey on health prediction using human activity patterns through smart devices". *International Journal of Engineering and Technology (UAE)*, vol. 7, no. 1, pp. 226-229. 2018.
- [21] Angel Prathyusha *et al*. "A survey on prediction of suitable crop selection for agriculture development using data mining classification techniques". *International Journal of Engineering and Technology (UAE)*, vol. 7, no. 3.3, pp. 107-109, 2018.
- [22] Anjali Devi *et al*. "Analysis and prediction of student placement for improving the education standards". *International Journal of Engineering and Technology (UAE)*, vol. 7, no. 2.8, pp. 303-306, 2018.
- [23] Arshad, M *et al*. "A real-time LAN/WAN and web attack prediction framework using hybrid machine learning model". *International Journal of Engineering and Technology (UAE)*, vol. 7, no. 3.12, pp. 1128-1136. 2018.
- [24] Ss Arvind Selwal *et al*. "A Multi-layer perceptron based intelligent thyroid disease prediction system", *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 17, no. 1, pp. 524-533, January 2020.
- [25] Chanintorn *et al*. "Proposed algorithm for Regression-based prediction with bulk noise", *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 17, no. 1, pp. 543-550, 2020.
- [26] Wahab *et al*. "Artificial neural network based technique for energy management prediction", *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 17, no. 1, pp. 94-101, 2020.

BIOGRAPHIES OF AUTHORS



P.G. Om Prakash is working as an Assistant Professor in the Dept. of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh, India. He has Pursuing his Ph.D., degree in Computer Science and Engineering from B.S.A. Crescent Institute of Science and Technology, Chennai, Tamilnadu, India. He has more than 9 years of academic and 3 years of Software experience. His recent research interest includes Web Mining, Cloud Computing, Computer Networks.



A. Jaya is working as a Professor in the dept of Computer Applications, B.S.A. Crescent Institute of Science and Technology, Chennai, Tamilnadu, India. She has received her Ph.D degree in Information and Communication Engineering from Anna University, Chennai, Tamilnadu, India in 2011. She is having 23 years of Academic Experience. Area of Research NLP, Knowledge Engineering, Web Mining and Text Mining.



S. Ananthakumaran is working as an Associate Professor in the Dept. of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh, India. He has received his Ph.D., degree in Computer Science and Engineering from Manonmaniam Sundaranar University, Tamilnadu, India in 2016. He has more than 17 years of academic and 9 years of research experience. His recent research interest includes Secure Data Communication, Blockchain Technology, IoT and Deep Learning.



G. Ganesh is working as an Assistant Professor in the Dept. of Computer Science in Blue Crest College, Freetown, Sierra Leone. He has Completed his M.Tech in SRM University, Chennai, Tamilnadu, India. He has more than 11 years of academic and 2 years of Software experience. His recent research interest includes Data Mining, Web Mining, Cloud Computing, Computer Networks.