COMPUTERS & SECURITY 98 (2020) 102006

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/cose

Computers & Security

Efficient classification model of web news documents using machine learning algorithms for accurate information

Aos Mulahuwaish[a], Kevin Gyorick[a], Kayhan Zrar Ghafoor[b,c], Halgurd S. Maghdid[d], Danda B. Rawat[e,*]

[a] Department of Computer Science and Information Systems, Saginaw Valley State University MI 48710, USA
[b] Department of Software Engineering, Salahaddin University-Erbil, Iraq
[c] School of Mathematics and Computer Science, University of Wolverhampton, Wulfruna Street, Wolverhampton, WV1 1LY, UK
[d] Department of Software Engineering, Faculty of Engineering, Koya University, Kurdistan Region-F.R. Iraq
[e] Department of Electrical Engineering and Computer Science, Howard University, Washington, DC 20059, USA

ARTICLE INFO

Article history:
Received 17 August 2020
Accepted 18 August 2020
Available online 26 August 2020

Classification:
SVM
kNN
DT
LSTM

Keywords:
Web News Applications
Data mining

ABSTRACT

Web applications are regarded as a popular platform to exchange information with users. These applications have to be able to process Big-Data quickly and to serve users in a timely manner with accurate information posted in news portals which can be a huge challenge to overcome. Huge computation power is needed to crawl the web and process big-data and the methods are needed to be developed to reduce space and time complexity of this process. Data mining is considered to be a solution to mitigate the aforementioned challenges by extracting specific information based on explicit features. This paper proposes an efficient model for web that extracts news information and sorts news documents into four different categories business, technology & science, health and entertainment. Four different machine learning classifiers Support Vector Machine (SVM), K-Nearest Neighbors (kNN), Decision Tree (DT) and Long Short-Term Memory (LSTM) are compared. These classifiers are implemented separately and are then compared using accuracy and receiver operating characteristic curves. The attained results show that the accuracy of kNN was the worst at 88.72% and SVM was the best at 95.04%.

1. Introduction

We have witnessed that more applications are moving to mobile or web page-based applications instead of desktop applications. Services like shopping, billing, communication, and transportation can all be done with an internet browser or web page application instead of a sizeable application that is installed onto the desktop computers. Web page applications are commonly used to exchange information between users. However, there is a common issue with this approach, the processing of huge amounts of information on the web. Web pages for serving news have issues for providing accurate information in a timely manner since they have to serve thousands of users in short periods of time and have many articles to be processed for accuracy for easy filtering.

* Corresponding author.
E-mail addresses: amulahuw@svsu.edu (A. Mulahuwaish), kpgyoric@svsu.edu (K. Gyorick), kayhan@ieee.org (K.Z. Ghafoor), First.Last@koyauniversity.org (H.S. Maghdid), db.rawat@ieee.org, danda.rawat@howard.edu (D.B. Rawat).
https://doi.org/10.1016/j.cose.2020.102006
0167-4048/

Despite having to handle Big-Data, web applications are the most accessible application for users to get updated information. Being able to load an application in seconds without installation attracts many users resulting in the need for huge amounts of computational power to serve users in seconds with the results they were expecting. This network traffic can be mitigated by categorizing data more affectively and classifiers can be used for data mining. The process of data mining is to find features in huge amounts of data records to minimize how much data needs to be evaluated when a user is filtering through it. These features depend on the application, with web news the features could be news content, users' behavior, visiting frequency, etc. There are different branches of data mining as described in (Smita, 2014), one is the predictive branch which includes classification, regression, time series analysis, and prediction. Another is the descriptive branch which includes clustering, summarization, association rules, and sequence discovery. The applications for data mining are to improve search engines by classifying web documents and identifying usage patterns allowing for the understanding of user behavior on a web application. The understanding of what users want to see will keep them using the application for longer and consistently.

The contribution of this article is to demonstrate how features like article content using classifiers such as Support Vector Machine (SVM), Decision Tree (DT), Long-Short Term Memory (LSTM), and K-Nearest Neighbor (k-NN) can reduce computation strain on time and space complexity. In particular, we conducted a comparative study among several machine learning algorithms for filtering web-based news contents.

The rest of this article is structured as following: section II presents related work to the study. Section III presents the proposed method using the four classifiers as mentioned above. Section IV demonstrates the obtained results after a set of experiments. Section V concludes the article with key points.

2. Related work

Other types of research using machine learning classifiers have been performed to strengthen web mining accuracy and to decrease time and space complexity. This section is to identify what more research has been done and what can be done on web-data mining.

Blogs and social networks are important, web applications used by many internet users to share their thoughts, daily activities, and photographs. In (Gharehchopogh et al., 2015), a hybrid method was proposed combining k-nearest neighbors (k-NN) and artificial neural network (ANN) to classify bloggers. A study was conducted on the Kohkiloye and Boyer Ahmad Province bloggers' dataset with this hybrid classifier and achieved an accuracy of 89.4%. A user's interest in different types of web content can be measured using view count, rating, comments, feedback, and sharing history. For example, in (Broxton et al., 2013), probability rules were used on the site YouTube to measure the rank of sharable videos. The study noticed that social videos were more likely to go viral than non-social videos. Another popular use of web application is for e-commerce. E-commerce is how many modern companies conduct most of their business. Systems have been developed to handle the big-data generated from e-commerce sites however, they run into issues such as needing to recommend items to a growing user base, determining whether or not highly clicked items are actually being purchased and retargeting items that are already stored in a user's cart. The authors in (Lopes and Roy, 2015) help solve these issues with a new dynamic recommendation system using techniques such as the product-based technique for anonymous users, the user-based technique for users with an account, and the action-based technique for using user behavior. The system uses three different metrics user interest, frequency of visits and similarity and demonstrates an accuracy of 82%. In (Devi et al., 2012), the Naïve Bayesian technique is used to mine user's interests for e-commerce sites. It was found to be a time-consuming approach. In (Lopes and Roy, 2014), an agglomerative hierarchical clustering approach is used that uses the Euclidean distance Eq. (11) to calculate similarity and clusters users together based on similar browsing features. News web applications use machine learning classifiers to pick out interesting and trending news to show an interested user. The better the classifiers are at showing users articles they want to read the longer the user will stay on the site. In (Kaur and Rashid, 2016), a backpropagation neural network was developed for web mining news applications such as BBC. The classes are entertainment, health, business, etc. The time complexity will reduce over time as it learns a user's interests.

Data and text mining are used to provide features for the classifiers used in this article. Data mining is the process of finding meaningful relationships in data (Bora, 2013). While text mining is the process of extracting interesting complex patterns or knowledge from unstructured text documents (Akilan, 2015). A small introduction on data mining can be found in (Bora, 2013) and on text mining (Akilan, 2015). Parallel corpora are required by statistical machine translation, cross-lingual information retrieval, word sense disambiguation and lexical acquisition (Tomás et al., 2005). In (Tomás et al., 2008, Tomás et al., 2005, Tomás et al., 2005), a web spider and machine learning models were used to create large multilingual corpora for use in such tasks. The experiment resulted in 652 parallel webpages being found on the web.

After reviewing the above research works, the following was established. Accuracy varies according to features used to implement the classifier, precision and accuracy need further improvement, and time complexity needs improvement as the data applications have to deal with will only grow day by day. The proposed solution will remedy the situation by providing high and stable classification accuracy for web news applications using various categorical classes to further reduce space and time complexity.

3. The proposed method

This section presents the proposed approach and the steps for preprocessing as well as the classification techniques (Fig. 1).

3.1. *Preprocessing*

Preprocessing step is significant in order to make the dataset ready for classification. Documents are processed using the

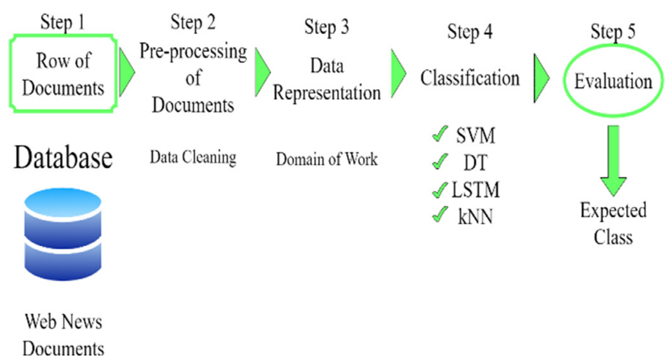

Fig. 1 – Summary of the proposed approach.

following steps. First, the documents are stripped of HTML and XML tags. Then, words that are considered noise are removed such as “is”, “the” and “it”. Afterword, punctuation, and special symbols are removed including “.”, “%” and “@”. The documents are then converted into lower case letters changing words like NEWS or News to news. Next, the documents are tokenized turning the articles into arrays of meaningful words so the frequency of each word can be calculated. The infrequence words that appear two or fewer times are then removed along with short words of two or fewer letters and long words. The words are then processed by Porter-Stemmer normalization algorithm (Chen et al., Apr. 2018), which removes common morphological and inflexional endings from English words leaving the stem of the word for example the words “running” and “runner” would become the word “run”. After the above process, some documents will become empty, so they are removed from the array. The prepressing steps are shown in order in Fig. 2.

3.2. Classifiers

We used four best suited classification techniques: Support Vector Machine (SVM), Decision Tree (DT), Long Short-Term Memory (LSTM) and K-Nearest Neighbors (kNN). These techniques require the training dataset that is used when training the classifier and the test dataset used to test the accuracy of the classifier after training.

3.2.1. *Support Vector Machine*

SVMs work by graphing the dataset where the number of features equals the dimensions of the graph. The dataset points are then separated by hyperplanes which in the 2D senses with two classes consist of three parallel lines, two of these lines are known as the support vectors they boarder the closest data point(s) between classes. The last line marks the midpoint between the two support vectors. The hyperplane with the greatest distance between the two support vectors is the best fit. If there isn’t one the dataset must be scaled to the next dimension using a kernel function. Popular kernel functions include polynomial as in Eq. (1), gaussian as in Eq. (2), radial basis function (RBF), Laplace RBF, sigmoid, etc.

$$k\left(x_i, x_j\right) = \left(x_i x_j + 1\right)^d \tag{1}$$

$$k(x, y) = \exp\left(-\frac{\|x - y\|^2}{2\sigma^2}\right) \tag{2}$$

3.2.2. *Decision Tree*

A DT builds a flow chart diagram that looks like a tree and branches off at every decision or variable. The topmost node of the tree is known as the root node, the tree’s creation originates at the root node and is built in a top-down fashion. Entropy shown in Eq. (3), is used to measure how unpredictable a decision is within the tree. 1 being very unpredictable, a 50% chance and 0 being a guarantee, either a 0% or 100% chance were p_i is the probability of a class i and c is the total classes. Information gain is used to measure the reduction of uncertainly when additional nodes are used before the given node, Eq. (4). The set of nodes that provide the most information gain are used. Decision trees are prone to overfitting since this above process is repeated on every branch of the tree.

$$E(s) = \sum_{i=1}^{c} -p_i log_2 p_i \tag{3}$$

$$IG(X, Y) = E(Y) - E(Y|X) \tag{4}$$

3.2.3. *Long Short-Term Memory*

LSTM specializes in text classification since the classifier can learn long-term dependencies between the text. The LSTM classifier is a form of recurrent neural network or RNN, which is a layered network that uses the previous outputs for the inputs of the next layer. LSTM has feedback connections allowing it to work with sequences of data instead of just single data points. An LSTM node consists of a cell, input gate, output gate, and forget gate. The cell is what remembers values over a time interval and the three gates regulate how the information will flow through the cell. The following Eqs. (5)–(10) are used in the creation of an LSTM with a forget gate. Where W and U are matrices containing the weights for the inputs and recurrent connections. x_t is the input vector unit, f_t is the forget activation vector, i_t is the input activation vector, o_t is the output activation vector, h_t is the output vector unit, $\tilde{c}_t$ is the cell input activation vector and c_t is the cell state vector. σ_g and σ_h are the activation functions sigmoid and hyperbolic tangent, respectively.

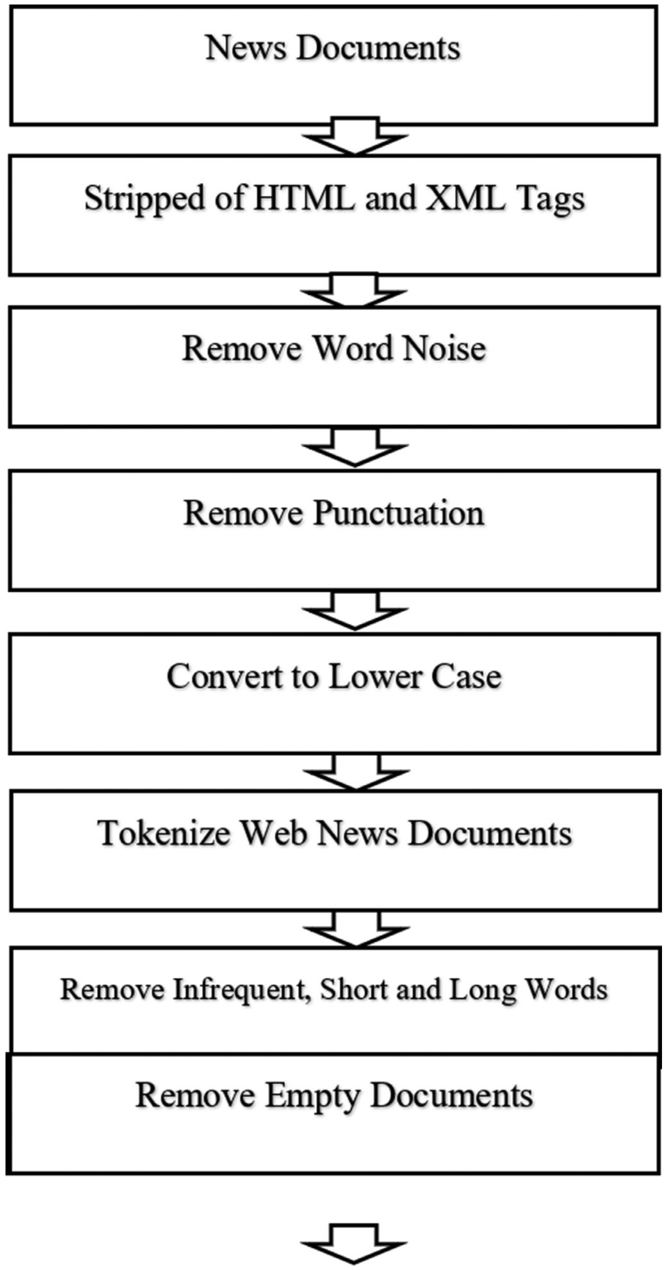

Fig. 2 – The preprocessing step.

$$f_t = \sigma_g\left(W_f x_t + U_f h_{t-1} + b_f\right) \tag{5}$$

$$i_t = \sigma_g(W_i x_t + U_i h_{t-1} + b_i) \tag{6}$$

$$o_t = \sigma_g(W_o x_t + U_o h_{t-1} + b_o) \tag{7}$$

$$\tilde{c}_t = \sigma_h(W_c x_t + U_c h_{t-1} + b_c) \tag{8}$$

$$c_t = f_t \circ c_{i-1} + i_t \circ \tilde{c}_t \tag{9}$$

$$h_t = o_t \circ \sigma_h(c_t) \tag{10}$$

3.2.4. *k-Nearest Neighbors*

The kNN technique works by using documents from the training dataset after they are preprocessed to build its model. The model uses the Euclidian distance equation shown in Eq. (11). The Euclidian distance is calculated between a new document from the test dataset that was also preprocessed, and all of the training data points.

$$d_j = \sqrt{\sum_{i=1}^{n} (WN_c - WN_o)^2} \tag{11}$$

where d_j is the Euclidian distance between the current document (WN_c) and another document (WN_o). After all the Euclidian distances are calculated for the new news document they are sorted from the smallest distances to the largest. The

Fig. 3 – Class distribution in the dataset

k nearest data points are selected based on the distance and the mode class in the k points is used as the prediction. The model is recreated many times with different values for k to see which k will produce the fewest errors when new data is introduced to the model.

4. Implementation and testing of the proposed method

The dataset being utilized to train the classification algorithms is the news aggregator dataset provided by "Center for Machine Learning and Intelligent Systems" – "Information system and computer science" – "University of California, Irvine". The dataset spans the period from October 4, 2014, to October 8, 2014. The news articles are in four different clusters that represent the content of the articles. The Artificial Intelligence Lab at the Faculty of Engineering, Roma Tre University – Italy originally provided the dataset (Gasparetti, 2017).

The dataset provides 422,937 news articles in four different categories, entertainment, science and technology, business, and health. For the experiment 25,000 news articles were randomly chosen from all categories and for classification purposes divided into two subsets. The training dataset which includes 90% or 95% of the random dataset and the test dataset containing the remaining data in the random dataset. Fig. 3 shows the distribution of the four different categories in the dataset used in the experiment. "B" denotes the business category, "e" denotes entertainment, "m" denotes health and "t" denotes science and technology.

Measures used in this article are accuracy, sensitivity, and specificity. Accuracy is how close the classification algorithm is to the desired results. To measure accuracy of the different classification algorithms, Eq. (12) is used where Y_{pre} is the classification's predictions for the test-dataset documents and Y_{test} is the correct predictions for the test-dataset documents. Sensitivity is a percentage that represents the true positive rate the classification algorithm achieved. A true positive is a correctly identified member of a class that is labeled as positive and a false positive is an incorrectly identified member of a class that is labeled as positive. Sensitivity is calculated using Eq. (13), where TP is the total true positives and FN is the total false negatives.

Specificity is a percentage that represents the true negative rate the classification algorithm achieved. A true negative is a correctly identified member of a class that is labeled as negative and a false negative is an incorrectly identified member of a class that is labeled as negative. Specificity is measured using Eq. (14), where TN is the total true negatives and FP is the total false positives.

$$accuracy = \frac{\sum (Y_{pre} = Y_{test})}{Y_{test}} * 100 \tag{12}$$

$$sensitivity = \frac{TP}{TP + FN} \tag{13}$$

$$specificity = \frac{TN}{TN + FP} \tag{14}$$

In this article, the confusion matrix is used to demonstrate the preferred result against the predicted result. Each column of the matrix represents samples in the predicted class while each row represents samples in the actual class. Receiver operating characteristics (ROC) curves are also used throughout the article to demonstrate a true positive rate against the false positive rate at various thresholds. The dataset is being organized into four classes instead of two so in order to use ROC curves the one verse all approach is used. The area under a ROC curve (AUC) and accuracy of a classification technique are good indicators for comparison.

1) Support Vector Machine Technique

Support vector machine was implemented using MATLAB R2020a's fitcecoc function. The SVM achieved an accuracy of 95.04% making SVM the leading classification technique for the dataset. Since SVMs have no trouble handling large feature datasets. The ROC curve as seen in Fig. 5 has an average area under the curve (AUC) of 99.4% for the SVM technique. Fig. 4 shows the SVM is classifying most of the dataset correctly with a high average sensitivity of 94.97% and an average specificity of 5.03%. Health has the highest false negative rate at 6.4%.

2) Decision Tree Technique

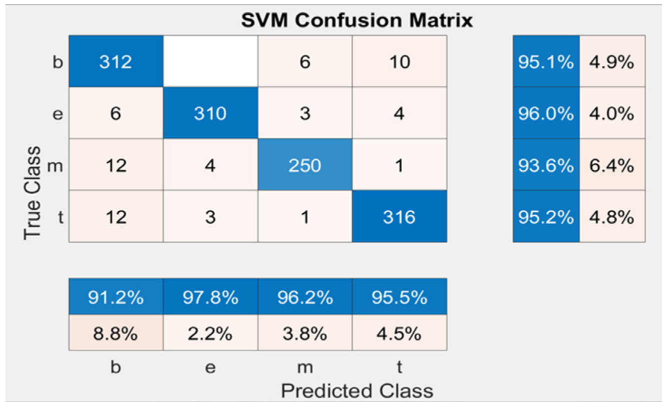

Fig. 4 – Confusion matrix for SVM.

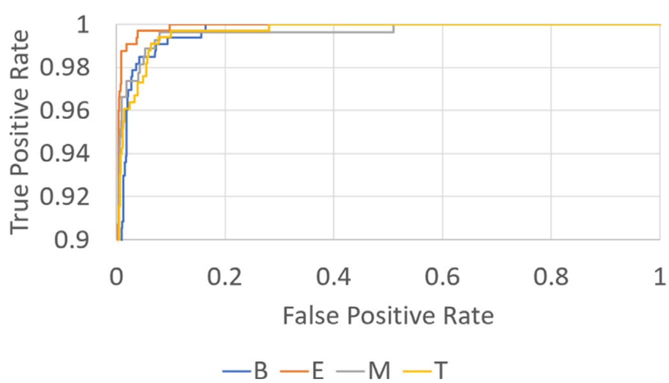

Fig. 5 – ROC curve for SVM.

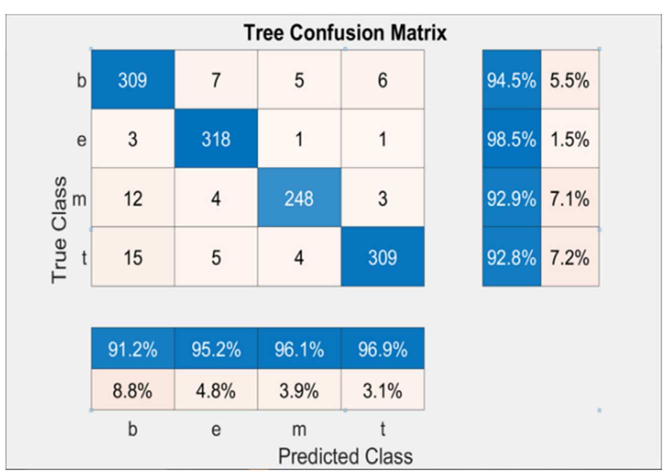

Fig. 6 – Confusion matrix for DT.

Decision Tree was implemented using the MATLAB *fitctree* function. DT achieved an accuracy of 94.72% making DT the second-best solution. The ROC curve in Fig. 7 has an average AUC of 97.78% for the decision tree technique. Fig. 6 shows the DT is classifying the dataset at an average sensitivity of 94.68% and an average specificity of 5.32%. Technology has the highest false negative rate at 7.2% just 0.1% higher than health.

3) Long Short-Term Memory Technique

The long short-term memory classifier was implemented with the MATLAB trainNetwork function. LSTM achieved an accuracy of 93.52% making LSTM the third best solution. The ROC curve as seen in Fig. 9 has an average AUC of 99.3% for the LSTM technique. Fig. 8 shows the LSTM classifier achieved

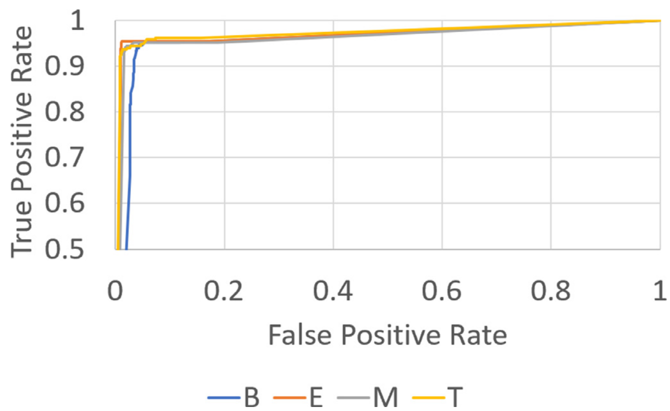

Fig. 7 – ROC curve for DT.

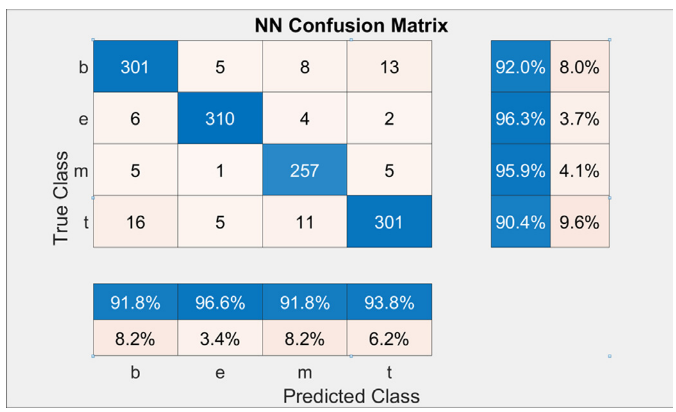

Fig. 8 – Confusion matrix for LSTM.

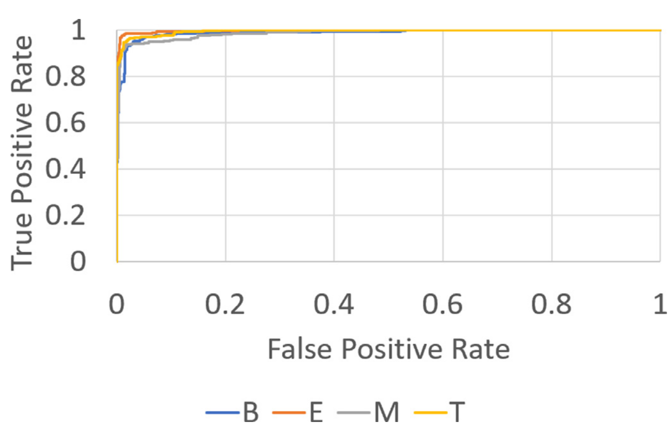

Fig. 9 – ROC curve for LSTM.

an average sensitivity of 93.65% and an average specificity of 6.35%. Technology has the highest false negative rate at 9.6%.

4) K-nearest Neighbors Technique

K-nearest neighbors was implemented using the MATLAB fitcknn function. kNN achieved an accuracy of 88.72% making kNN the worst solution. The ROC curve in Fig. 11 has an average AUC of 97.99% for the kNN technique. Fig. 10 shows the underperformance of the kNN classifier with only an average sensitivity of 88.73% and an average specificity of 11.27%. Technology was misidentified the most with a false negative rate of 13.3%.

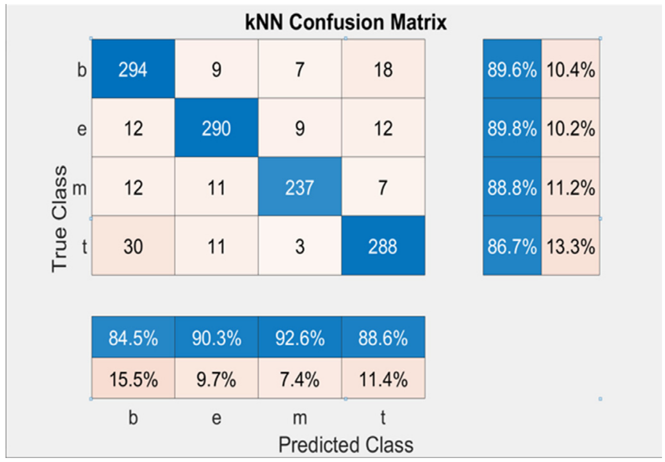

Fig. 10 – Confusion matrix kNN.

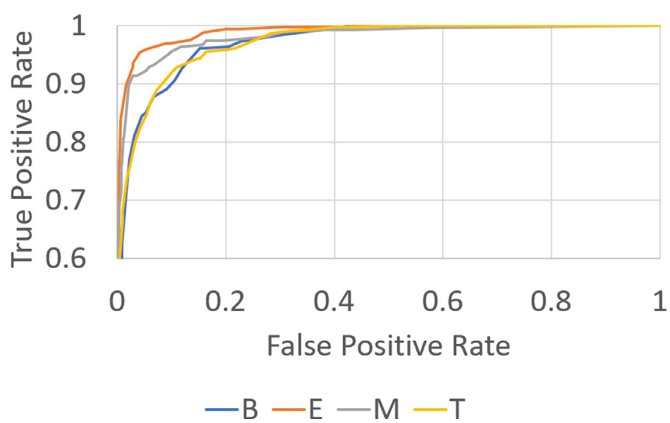

Fig. 11 – ROC curve for kNN.

5. Conclusion and future work

After recognizing the daily impact, the web applications have on users and their popularity, related studies were discussed to find the research gap and different suitable machine learning classifiers are studied. We studied the machine learning classifiers that demonstrated a more accurate results with less time and space complexity for web applications based big data. Specifically, we have compared four classifiers K-Nearest Neighbors (kNN), Support Vector Machine (SVM), Decision Tree (DT), and Long Short-Term Memory (LSTM). SVM coming out on top with the best accuracy of 95.04% and kNN with the worst at 88.72%. It is important to note that the web mining in combination with classification has become a valued research topic over recent years. Existing solutions fulfill the needs of most web applications. However, existing solutions require a huge amount of computation power, so the proposed approach works on reducing space and time complexity. The results of the top classifiers demonstrated good accuracy, a reduction in time needed for the training and testing phases of classification (time complexity), and reduction of total documents (space complexity).

The experiments along with other web mining research provides proof that classifiers are effective at sorting web news for accuracy. However, there is more research in web mining to be desired. For example, time and user-visiting information can be used in product mining for a web application to recommend products the customer is more likely to buy. An admin panel can be created for the classifiers created in the to show the application admin the results of the classifier at a glance right on the site they're managing. Other classification techniques can be tested for classifying web news to check if they can further reduce space and time complexity along with other advantages over the classifiers tested here.

Declaration of Competing Interest

None.

REFERENCES

Akilan A. Text Mining: Challenges and Future Directions. IEEE 2015:1679–84.
Bora S. Data mining and ware housing. IEEE 2013:1–5.

Broxton T, Interian Y, Vaver J, Wattenhofer M. ‘Catching a viral video,. J. Intell. Inf. Syst. 2013;40(2):241–59.
Chen P-H, Zafar H, Galperin-Aizenberg M, Cook T. ‘Integrating natural language processing and machine learning algorithms to categorize oncologic response in radiology reports. J. Digit. Imag. Apr. 2018;31(2):178–84.
Devi B, Devi Y, Rani B, Rao R. Design and implementation of web usage mining intelligent system in the field of e-commerce. Procedia Eng. 2012;30:20–7.
Gasparetti F. ‘Modeling user interests from Web browsing activities. Data Mining Knowl. Discovery 2017;31(2):502–47.
Gharehchopogh FS, Khaze SR, Maleki I. ‘A new approach in bloggers classification with hybrid of k-Nearest neighbor and artificial neural network algorithms. Indian J. Sci. Technol. 2015;8(3):237–46.
Kaur S, Rashid EM. ‘Web news mining using back propagation neural network and clustering using K-means algorithm in big data. Indian J. Sci. Technol. 2016;9(41):1–8.
Lopes P, Roy B. ‘Dynamic recommendation system using Web usage mining for e-commerce users. Procedia Comput. Sci. 2015;45(Mar.):60–9.
Lopes P, Roy B. Dynamic recommendation system using web usage mining for e-commerce users. Int. J. Eng. Res. Technol. 2014;3(7).
Smita Sharma P. Use of data mining in various field: a survey paper,. IOSR J. Comput. Eng. 2014;16(3):18–21.
Tomás J, Bataller J, Casacuberta F, Lloret J. Mining Wikipedia as a parallel and comparable corpus. Lang. Forum 2008;34:123–36.
Tomás J, Sánchez-Villamil E, Casacuberta F, Lloret J. WebMining: an unsupervised parallel corpora web retrieval system. Proc. Corpus Linguist. Conf. 2005.
Tomás J, Casacuberta F, Lloret J. WebMining: Non supervised system to obtain parallel corpus from the Web. Corpus Linguist. Conf. Ser. 2005.

Aos Mulahuwaish is with the Department of Computer Science and Information Systems, Saginaw Valley State University, USA.

Kevin Gyorick is with the Department of Computer Science and Information Systems, Saginaw Valley State University, USA.

Dr. Kayhan Zrar Ghafoor is with School of Mathematics and Computer Science, University of Wolverhampton, Wulfruna Street, Wolverhampton, WV11LY, UK. He was ay a post doctorate researcher with Department of Computer Science and Engineering, Shanghai Jiao Tong University. Before that, he was visiting researcher at University Technology Malaysia for 6 months. He received the BSc degree in Electrical Engineering from Salahaddin University, the MSc degree in Remote Weather Monitoring from Koya University and the PhD degree in Wireless Networks from University Technology Malaysia in 2003, 2006, and 2011, respectively. He has published over 60 scientific/research papers in ISI/Scopus indexed international journals and conferences. He is the receipt of the UTM Chancellor Award at 48th UTM convocation in 2012.

Halgurd S. Maghdid is with the Department of Software Engineering, Faculty of Engineering, Koya University, Kurdistan Region-F.R. Iraq.

Dr. Danda B. Rawat is a Full Professor in the Department of Electrical Engineering & Computer Science (EECS), Director of the *Howard Data Science and Cybersecurity Center, Director of Cyber-security* and Wireless Networking Innovations (CWiNs) Research Lab, Graduate Program Director of Howard CS Graduate Programs and Director of Graduate Cybersecurity Certificate Program at Howard University, Washington, DC, USA. Dr. Rawat is engaged in research and teaching in the areas of cybersecurity, machine learning, big data analytics and wireless networking for emerging networked systems including cyber-physical systems, Internet-ofThings, multi battle domain, smart cities, software defined systems and vehicular networks. His professional career comprises more than 15 years in academia, government, and industry. He has secured over $6 million in research funding from the US National Science Foundation (NSF), US Department of Homeland Security (DHS), US National Security Agency (NSA), US Department of Energy, National Nuclear Security Administration (NNSA), DoD Research Labs, Industry (Microsoft, Intel, etc.) and private Foundations. Dr. Rawat is the recipient of NSF CAREER Award in 2016, Department of Homeland Security (DHS) Scientific Leadership Award in 2017, Researcher Exemplar Award 2019 and Graduate Faculty Exemplar Award 2019 from Howard University, the US Air Force Research Laboratory (AFRL) Summer Faculty Visiting Fellowship in 2017, Outstanding Research Faculty Award (Award for Excellence in Scholarly Activity) at GSU in 2015, the Best Paper Awards (IEEE CCNC, IEEE ICII, BWCA) and Outstanding PhD Researcher Award in 2009. He has delivered over 20 Keynotes and invited speeches at international conferences and workshops. Dr. Rawat has published over 200 scientific/technical articles and 10 books. He has been serving as an Editor/Guest Editor for over 50 international journals including the Associate Editor of IEEE Transactions of Service Computing, IEEE Transactions of Network Science and Engineering, IEEE Internet of Things Journal and IEEE Network. He has been in Organizing Committees for several IEEE flagship conferences such as IEEE INFOCOM, IEEE CNS, IEEE ICC, IEEE GLOBECOM and so on. He served as a technical program committee (TPC) member for several international conferences including IEEE INFOCOM, IEEE GLOBECOM, IEEE CCNC, IEEE GreenCom, IEEE ICC, IEEE WCNC and IEEE VTC conferences. He served as a Vice Chair of the Executive Committee of the IEEE Savannah Section from 2013 to 2017. Dr. Rawat received the Ph.D. degree from Old Dominion University, Norfolk, Virginia. Dr. Rawat is a Senior Member of IEEE and ACM, a member of ASEE and AAAS, and a Fellow of the Institution of Engineering and Technology (IET).