

An Optimal Feature Selection with Wavelet Kernel Extreme Learning Machine for Big Data Analysis of Product Reviews

R.Sathya^a, and L. R.Aravind Babu^b

^a

Research Scholar, Department of Computer and Information Science, Annamalai University, Annamalainagar-608002, Tamilnadu, India.

^bAssistant Professor, Department of Computer and Information Science, Annamalai University, Annamalainagar-608002, Tamilnadu, India.

Article History: Received: 11 January 2021; Accepted: 27 February 2021; Published online: 5 April 2021

Abstract: In recent times, generation of big data takes place in an exponential way from diverse textual data sources like review sites, media, blogs, etc. Sentiment analysis (SA) finds it useful to classify the opinions of the big data to different kinds of sentiments. Therefore, SA on big data helps a business to take beneficial commercial understandings from text based content. Though several SA approaches have been presented, yet, there is a need to improve the performance of SA to interpret the customer's feedback and increase the product quality. This paper introduces a novel social spider optimization based feature selection based wavelet kernel extreme learning machine (SSO-WKELM) model. The proposed model initially undergoes pre-processing to remove the unwanted word removal. Then, Term Frequency-Inverse Document Frequency (TF-IDF) is utilized as a feature extraction technique to extract the set of feature vectors. Besides, a social spider optimization (SSO) algorithm is utilized for feature selection process and thereby achieves improved classification performance. Subsequently, WKELM is employed as a classifier to classify the incidence of positive or negative user reviews. For experimental validation, a Product review dataset derived from Amazon along with synthetic data is used. The experimental results stated the superior classification performance of the SSO-WKELM model.

Keywords: Big data, Product reviews, Sentiment analysis, Parameter tuning, TF-IDF, KELM

1. Introduction

Nowadays, electronic-commerce is extensively developed as a simulation of using Internet [1]. Independent unprejudiced user feedbacks are examined as scalable base for service data. People with inclination in making decision about purchasing goods from stores. The reviews are assumed to the part of making decision through trading model of a provider. Similarly, learning and using the feedbacks are applicable in gaining the trust of a consumer that supports trade development. Followed by, data deployment over social networking webpages are progressive to greater extent [2]. Thus, reduced interest in users for mining the wider resources of data to examine the sentiment or emotions. Finally, social media investigation provides the hints for best increment and smart principles for goods of upcoming generation. Additionally, the above-defined statement is applicable in businesses regarding user's perception of brand, and make decisions regarding the image. Generally, Big data is referred as massive amount of data which is derived from web pages, social media, remote sensing data, and clinical archives, etc, in a structured, semi-structured, or unstructured manner [3]. The big data is derived from various sources like television, mobile, and many other sources like industry data storage. For example, big data is usually unstructured and demands real-world application. This deployment urges for making a new model for data acquisition, transmission, memory space, and huge-scale data computation approaches [4]. Thus, the firms require organized base for initializing big data and details for executing the central domains. At this point, newly developed feedbacks and polls are important in estimating the assessments, which represents the impact of degree of services [5]. In recent times, in Artificial Intelligence (AI) applications, text classification is significant according to the size of text data.

Sentiment identification and analysis desire to find the basic attitudes, emotions, and subjectivity to certain entities like learners to understand the resources. Because of the massive advantages for smart education, sentiment analysis is adopted as significant mechanism to find and classify the sentiments from multimodal as well as multisource data in complete learning process. Also, Sentiment Analysis (SA), predicts the subjectivity of documents which is a required process in text classification and prediction [6]. Numerous works in SA have been presented and recently it has been applied in businesses as significant tool for learning the market and user division, and for predicting stock price in tactical market scheduling as well as maneuvering. Besides, human life is full of emotions and suggestions which play a significant role most of the applications. Fig. 1 shows the data complexity in big data.

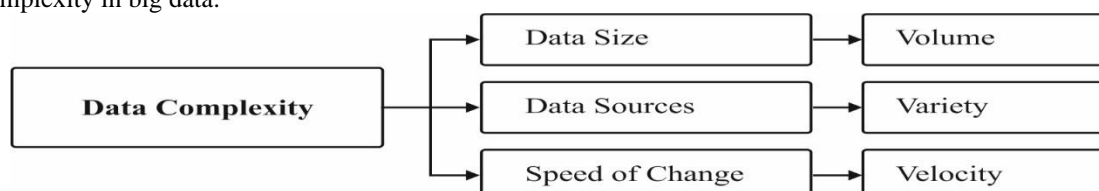


Fig. 1.Data complexity in big data

Prediction of sentiments in decisions making is a major concern used for preventing the wrong solution in human-based actions of businesses. Moreover, SA is depicted as a principle of examining the emotions expressed in a text [7]. Hence, in this competing world, learning the user demand and market-related manufacturing is one of the promising objectives in businesses. In line with this, text and SA can be applicable for a business to make possible decisions from text-based content such as word documents, email, and posts on social networks like Twitter, Facebook (FB), and LinkedIn. Consequently, it results in enhanced attention for presenting views and replies for pervasive SA in academic searching as well as business.

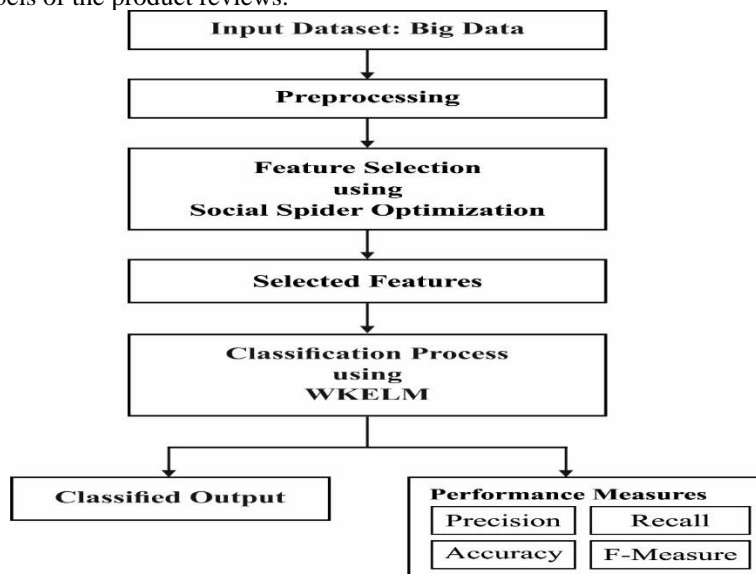
Tsai et al [8] reviewed the studies developed on data analytics from classical data analysis to present big data examination. From the system perception, knowledge discovery is carried out and consolidated as 3 portions namely, input, prediction, and result. Sharef et al [9] outlined the advanced models in SA, such as sentiment polarity forecasting, SA features, sentiment classification models, and utilities of SA. Moreover, Graham et al [10] examined the classification of “big data logistics.” It has provided a novel view for accurate words and predict “positive” and “negative” emotions. Diverse businesses like public and private, industrial, healthcare, and retail are defined for combining big data and logistics. The deployment of KNIME text modeling is learned and converted as mathematical objects to apply normal KNIME data mining models.

Besides, Sharma et al [11] originated learning on SA for big data. It has recent updates in opinion mining. Also, developers have identified that sentiment mining is one of the well-known and important objectives. As a result, numerous works have been processed, unfortunately, specific challenges to sentiment mining are relevant to unstructured data. Lastly, Balaji et al. [12] performed a study on stock prediction according to emotion and big data analysis. This work has implied the review of each framework with stock prediction models and big data analytics. It demonstrates that numerous approaches are used in various processes and consequently predictive analytics is extremely optimal in-stock prediction.

This paper introduces a novel optimal social spider optimization based feature selection (FS) with wavelet kernel extreme learning machine (SSO-WKELM) method. The proposed model initially undergoes pre-processing to remove the unwanted word removal. Then, Term Frequency-Inverse Document Frequency (TF-IDF) is utilized as a feature extraction method to extract the set of feature vectors. Besides, a social spider optimization (SSO) algorithm is utilized for FS process thereby achieves improved classification performance. At last, the WKELM model is applied as a classification technique to allocate the class labels of the product reviews. To ensure the effective performance of the SSO-WKELM model, a series of simulations were performed on the product review dataset from amazon and synthetic dataset.

2. The Proposed SSO-WKELM Model

The workflow involved in the SSO-WKELM model is illustrated in Fig. 2. The input product review dataset is initially preprocessed to remove the unwanted words that exist in it. Afterward, a set of feature vectors are effectively extracted by the use of TF-IDF technique. At the same time, the SSO algorithm properly chooses the appropriate set of features. At last, the WKELM model is applied as a classification technique to allocate the class labels of the product reviews.

**Fig. 2.** Working process of SSO-WKELM model

2.1. TF-IDF Model

The prominently applied metrics in data retrieval is “TD-IDF. The data weighting approaches are employed in measuring the probability-weighted amount of data in the applied document. In traditional data theory, IDF is interpreted as ‘volume of data’ provides the log of inverse probability. Based on the definition, TF-IDF is defined as a value that increases 2 two quantities namely TF and IDF. Followed by, the term frequency offers evaluation of probability incidence where it is normalized by TF in a document which depends upon the scope of estimation. According to the fundamental expression of data theory [13], a document is regarded as unordered set of terms.

Assume $D = \{d_j, \dots, d_n\}$ as a collection of documents and $W = \{w_i, \dots, w_M\}$ be the set of diverse system in D . Here, documents D is implied by corpus of data gained from tweeter are induced while W implies the query term. Hence, parameters N is overall number of documents whereas M signifies the count of terms. In data theory, selecting the term w_i from W and election of document d_j from D has been assumed.

2.2. SSO based Feature Selection Process

In general, SSO considers the search space as communal spider web and a candidate solution in this population implies a spider. A spider takes a weight regarding the fitness measures of a solution which is symbolized. In this model, 2 diverse search set of evolutionary operators are used in simulating the variations among cooperative behaviors adopted in the colony.

It is modeled for resolving the nonlinear global optimization issues with box limitation as:

$$\begin{aligned} & \text{minimize: } f(x) \quad x = (x^1, x^2, \dots, x^d) \in \mathbb{R}^d; \\ & \text{subject to } x \in X \end{aligned} \tag{1}$$

where $f: \mathbb{R}^d \rightarrow \mathbb{R}$ means a nonlinear function and $X = \{x \in \mathbb{R}^d \mid l_h \leq x \leq u_h, h = 1, \dots, d\}$ defines a limited and possible space restricted by low (l_h) and upper (u_h) limits.

Here, SSO applies a population S of N candidate solutions for solving optimization issues. A solution indicates the new position while the typical web implies a search space X . Here, population S is classified as 2 search agents namely, Male (Ms) and Female (Fs). By this aid of simulating an actual spider colony, number N_f is of females, and N_f is selected randomly from the overall population S , while the rest N_m is assumed as male individuals ($N_m = S - N_f$). During this scenario, group Fs develops female individuals ($Fs = \{f_{s_1}, f_{s_2}, \dots, f_{s_{N_f}}\}$), thus Ms group, male individuals, ($Ms = \{ms_1, ms_2, \dots, ms_{n_m}\}$), where $S = Fs \cup Ms$ ($S = \{s_1, s_2, \dots, s_N\}$). In this algorithm, a spider i is composed of weight we_i which upon the solution fitness, and weight is estimated using the given expression:

$$we_i = \frac{fit_i - worst}{best - worst} \tag{2}$$

where fit_i depicts the fitness of i -th spider position, $i \in 1, \dots, N$, and $best$ and $worst$ illustrates the best and worst fitness value of complete population S . The data exchange is a basic action of SSO in optimization model [14]. As a result, it can be simulated by the vibrations generated in a web. Hence, vibration of a spider i receives from a spider j is developed as given in the following:

$$V_{i,j} = we_j e^{d_{i,j}^2} \tag{3}$$

where we_j denotes the weight of j -th spider and d depicts the distance from 2 spiders. A spider i is capable of perceiving 3 class of vibration, $v_{i,n}$, $v_{i,b}$, and $v_{i,f}$. $v_{i,n}$ which are the vibration originated by adjacent spider n with maximum weight regarding ($w_n > w_i$). $v_{i,f}$ has generated by closest female spider and suitable if i is a male spider and consequently $v_{i,b}$ is generated by optimal spider in population S . Here, a population N of spiders are computed as initial stage $k = 0$ for determining value it of iterations ($k = it$). Based on gender all spiders are estimated by a diverse set of evolutionary functions. For female spiders, novel position fs_i^{k+1} is attained by changing the recent spider position fs_i^k . Hence, the modification is managed by a probability factor Pf and movement is generated in relation to alternate spiders and the vibrations forwarded by the search space:

$$fs_i^{k+1} = \begin{cases} fs_i^k + \alpha \cdot v_{i,n} + \beta \cdot v_{i,b} \cdot (s_b - fs_i^k) + \delta \cdot \left(\text{rand} - \frac{1}{2} \right) & \text{with probability } Pf \\ fs_i^k - \alpha \cdot v_{i,n} - \beta \cdot v_{i,b} \cdot (s_b - fs_i^k) + \delta \cdot \left(\text{rand} - \frac{1}{2} \right) & \text{with probability } 1 - Pf \end{cases} \tag{4}$$

where α, β, δ , and rand implies the random values from $[0, 1]$ and k denotes the iteration value and individuals s_n and s_b implies the closer spider with maximum weight when compared with fs_i^k and optimal spider in a communal web, correspondingly.

Conversely, the male spiders are categorized into 2 classes namely, Dominant (D) and non-dominant (ND). The D group is combined with male spider which has the fitness values about complete male set. Therefore, ND group is developed by remaining male spiders. In optimization method, male spiders ms_i^k are processed by the given expression:

$$ms_i^{k+1} = \begin{cases} ms_i^k + \alpha \cdot v_{i,f} \cdot (s_f - ms_i^k) + \delta \cdot \left(\text{rand} - \frac{1}{2} \right) & \text{if } ms_i^k \in D \\ ms_i^k + \alpha \cdot \left(\frac{\sum_{h \in ND} ms_h^k \cdot w_h}{\sum_{h \in ND} w_h} - ms_i^k \right) & \text{if } ms_i^k \in ND \end{cases} \tag{5}$$

where α, δ , and rand defines the random values from $[0, 1]$ and s_f indicates the closest female spider to male individual i .

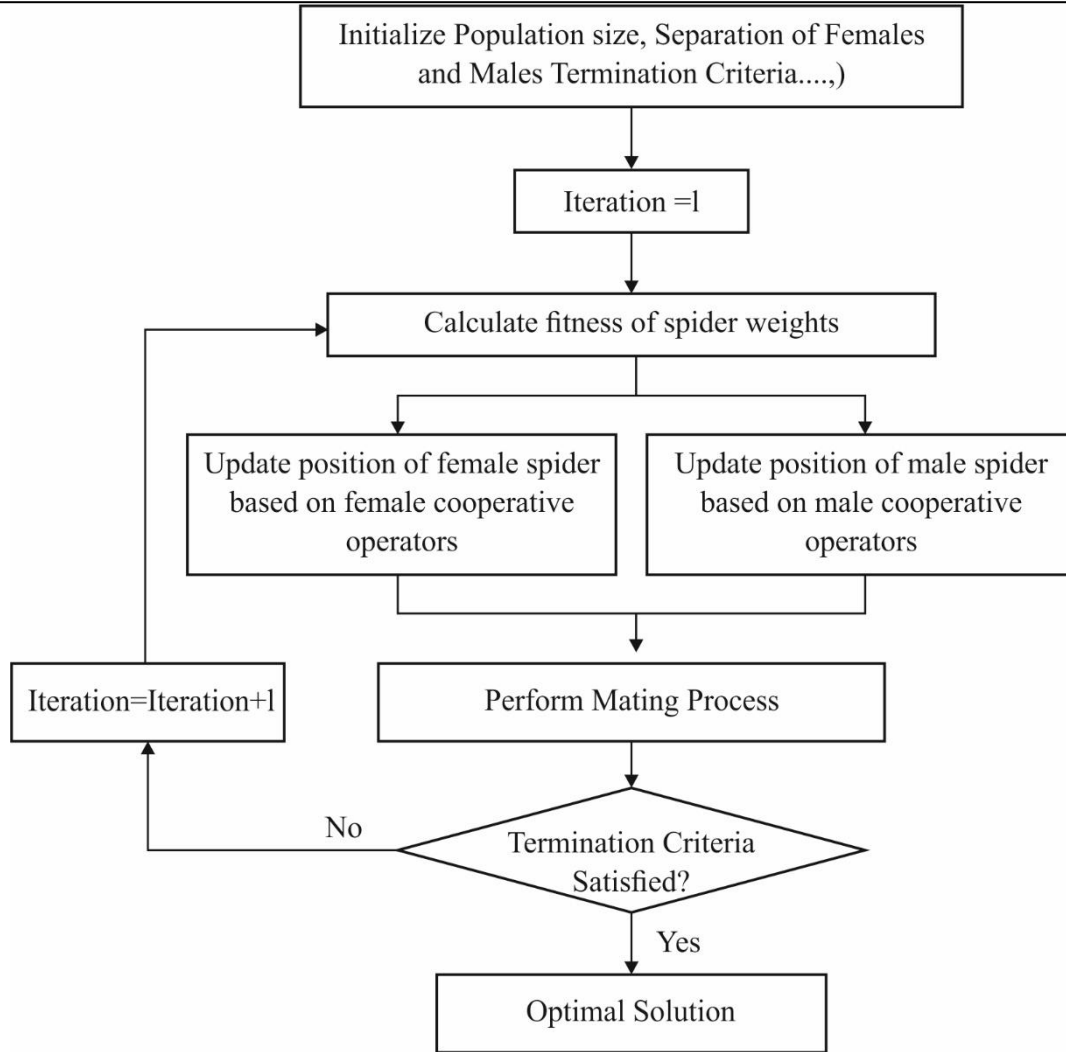


Fig. 3. Flowchart of SSO algorithm

In SSO, mating process is carried out among dominant male m_d and female individuals within a limited range r ; and creates the new offspring s_{new} . The weight of spider describes the possibilities of control of a spider on s_{new} ; the heavier component has maximum probability for influencing new individual s_{new} . When a novel spider is produced, it can be related to remaining population, and when the new spider contains optimal fitness, then worst spider is changed by s_{new} ; else s_{new} is removed. Fig. 3 illustrates the flowchart of SSO technique.

The SSO algorithm for FS process involves the following operations [15].

Initial stage

Here, the newly developed approach is initialized by generating random population X with collection of solutions (spiders' location) $x_i, i = 1, \dots, N$. To estimate the solutions, fitness function (FF) has been applied; but, a solution $(x_i(t))$ at iteration t has to be transformed as binary vector with the help of given expression:

$$x_i^j(t + 1) = \begin{cases} 1 & \text{if } \frac{1}{1 + e^{-x_i^j(t)}} > \epsilon \\ 0 & \text{otherwise} \end{cases}, \quad (6)$$

where $\epsilon \in [0, 1]$ indicates a random value. The boolean vector x_i depicts the decided features corresponding to 1 and remaining features are 0 which are neglected finally.

The FF value applied in the newly projected approach is illustrated as:

$$fit_{x_i} = \rho \times \frac{\text{No of in correctly samples}}{\text{Total Number of samples}} + \rho_1 \times \left(\frac{|x_i|}{d} \right) \quad (7)$$

where $| \cdot |$ means the count of selected features and parameters ρ_1 and ρ are applied in balancing classification error and $| \cdot |$. Eq. (7) is comprised of 2 portions, initially, the first part is employed to estimate the classification error of using the WKELM classification which depends upon the selected features whereas the second portion implies the count of decided features. The WKELM classifier is learned according to the training set and forecasts the labels of testing set; followed by, error is evaluated by relating to the simulation outcome with actual labels.

Updated stage

It is simulated by dividing X 2 classes namely, Female and Male; followed by, a female spider x_{fi} and a male x_{mj} . Afterward, mating is processed among female and male spiders which depend upon the mating values. The FF value for a spider is processed and the values are related to offspring FF, in which offspring interchanges the

poor male when it can be female spider. Once the spider locations are upgraded, the FF values are estimated $fit_{\bar{x}_i}$ for \bar{x}_i . The optimal N solutions are decided from 2 populations for developing new population and optimal solution x_{best} is performed. Therefore, the predefined steps are followed still the termination condition is accomplished and result from the FS process is optimal which illustrates the subset of chosen features $Feat_{Sel}$.

2.3. WKELM based Classification m

Here, the wavelet-mix kernel function for KELM method and develop the theme of weighted model in KELM has presented weighted WKELM framework. In sigmoid kernel function is defined as a global kernel function as well as translation-invariant kernel function. In addition, wavelet kernel functions are considered as local kernel functions. Thus, the integration of 2 kernel functions manages the learning as well as generalization capability and createto complete sense of features in rotation invariant as well as translation invariant. The weighted ELM approach is presented for dealing with the samples that are imbalanced in probability distribution, and it carries out quite well when compared with classical frameworks. Hence, weighted WKELM scheme develops the weighted mechanism for cost function and accomplishes identical effect of weighted ELM [16]. Therefore, KELM method has existed from ELM scheme, and weighted cost function is represented as shown below:

$$\min L_{ELM} = \frac{C}{2} \|\beta\|^2 + \frac{W}{2} \|T - HV\|^2 \tag{8}$$

$$\beta = \begin{cases} H^T (CI + WHH^T)^{-1} WT, & N < L \\ (CI + H^T WH)^{-1} H^T WT, & N \geq L \end{cases} \tag{9}$$

In KELM approach, the simulation outcome is illustrated in the following:

$$y = \begin{bmatrix} k(\vec{x}, \vec{x}_1) \\ k(\vec{x}, \vec{x}_2) \\ \dots \\ k(\vec{x}, \vec{x}_N) \end{bmatrix} \cdot (CI + WK^T)^{-1} WT \tag{10}$$

where K denotes the kernel matrix, W refers the weighted matrix, and C defines regularization variable. Next, the wavelet-mix kernel functions are introduced for KELM and examine the weighted WKELM framework.

3. Experimental Analysis

To validate the performance of the presented SSO-WKELM technique, Amazon product review dataset with synthetic data. The dataset includes a set of 235,000 positive and 147,000 negative review files in the dataset. In experimental is performed in 3 Intel Xeon E3-1220 processors running at 3.1 GHz speed with 64 Gb RAM and Hadoop Framework.

Table 1 and Fig. 4 investigates the accuracy analysis of the presented SSO-WKELM model on the applied dataset. From the experimental outcomes, it is apparent that the MI-CART method has resulted in an insignificant classification outcome with an accuracy of 0.7945 whereas the MI-RF model has resulted in a slightly increased accuracy of 0.8285. At the same time, the FSO-CART and FSO-RF models have demonstrated moderate and closer accuracy values of 0.8599 and 0.8914. But the SSO-WKELM model has accomplished superior performance by offering a maximum accuracy of 0.9456.

Table 1 Result Analysis of Existing with Proposed SSO-WKELM Method in terms of Accuracy

Methods	Accuracy
MI-CART	0.7945
MI-RF	0.8285
FSO-CART	0.8599
FSO-RF	0.8914
SSO-WKELM	0.9456

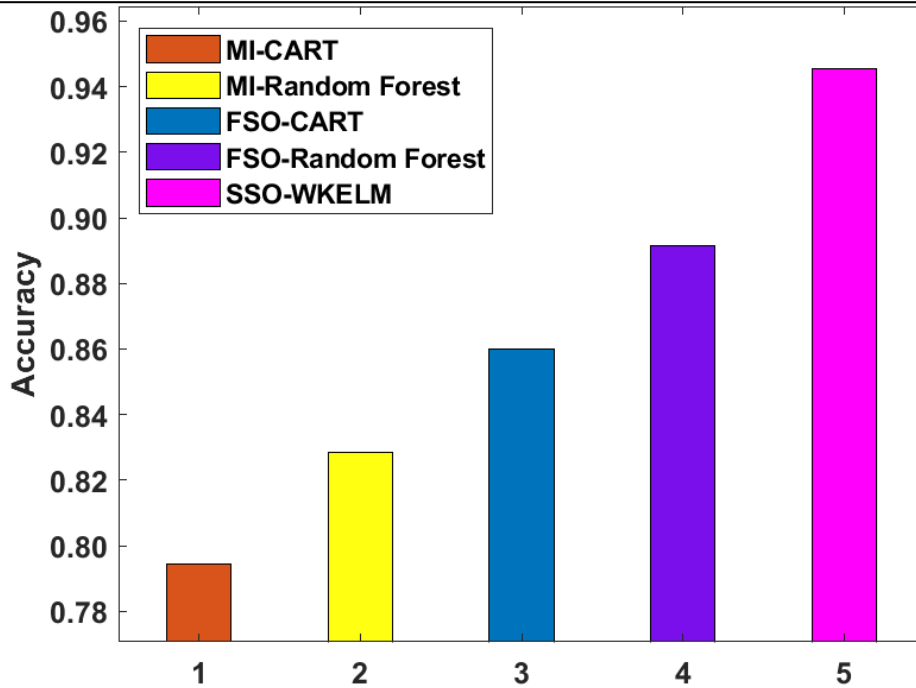


Fig. 4. Accuracy analysis of SSO-WKELM model

Precision analysis of the SSO-WKELM model is made with the existing techniques on the classification of positive and negative sentiments as shown in Fig. 5 and Table 2. The simulation outcome ensured that the MI-CART technique has resulted in the least outcome by offering a minimum precision of 0.6971 and 0.8757 on the classification of positive and negative sentiments respectively. Besides, the MI-RF model has showcased certainly improved results by achieving a precision of 0.7444 and 0.8937 on the classification of positive and negative sentiments respectively.

Table 2 Result Analysis of Existing with Proposed SSO-WKELM Method in terms of Precision

Methods	Positive Sentiment	Negative Sentiment
MI-CART	0.6971	0.8757
MI-random forest	0.7444	0.8937
FSO-CART	0.7945	0.9067
FSO-random forest	0.8499	0.9184
SSO-WKELM	0.9267	0.9675

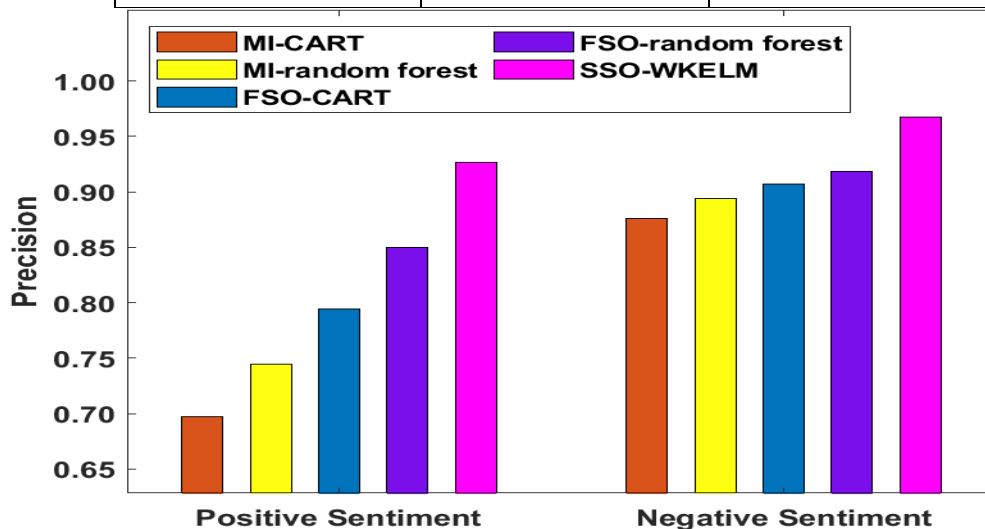


Fig. 5. Precision analysis of SSO-WKELM model

Moreover, the FSO-CART technique has offered a moderate performance by achieving a precision of 0.7945 and 0.9067 on the classification of positive and negative sentiments respectively. Furthermore, the FSO-RF model

has accomplished manageable results with the precision of 0.8499 and 0.9184 on the classification of positive and negative sentiments respectively. But the SSO-WKELM model has surpassed the other methods by providing a maximum precision of 0.9267 and 0.9675 on the classification of positive and negative sentiments respectively.

Table 3 Result Analysis of Existing with Proposed SSO-WKELM Method in terms of Recall

Methods	Positive Sentiment	Negative Sentiment
MI-CART	0.8240	0.7761
MI-random forest	0.8444	0.8816
FSO-CART	0.8580	0.8612
FSO-random forest	0.8716	0.9037
SSO-WKELM	0.9345	0.9568

A recall analysis of the SSO-WKELM methodology is made with the existing models on the classification of positive and negative sentiments as depicted in Fig. 6 and Table 3. The simulation result makes sure that the MI-CART model has resulted in the worst result by offering a minimum recall of 0.8240 and 0.7761 on the classification of positive and negative sentiments respectively. Likewise, the MI-RF model has outperformed certainly enhanced outcomes by obtaining a recall of 0.8444 and 0.8816 on the classification of positive and negative sentiments respectively. Furthermore, the FSO-CART approach has offered a moderate performance by attaining a recall of 0.8580 and 0.8612 on the classification of positive and negative sentiments respectively. At the same time, the FSO-RF model has accomplished manageable outcomes with the recall of 0.8716 and 0.9037 on the classification of positive and negative sentiments correspondingly. But the SSO-WKELM model has surpassed the other techniques by providing a superior recall of 0.9345 and 0.9568 on the classification of positive and negative sentiments respectively.

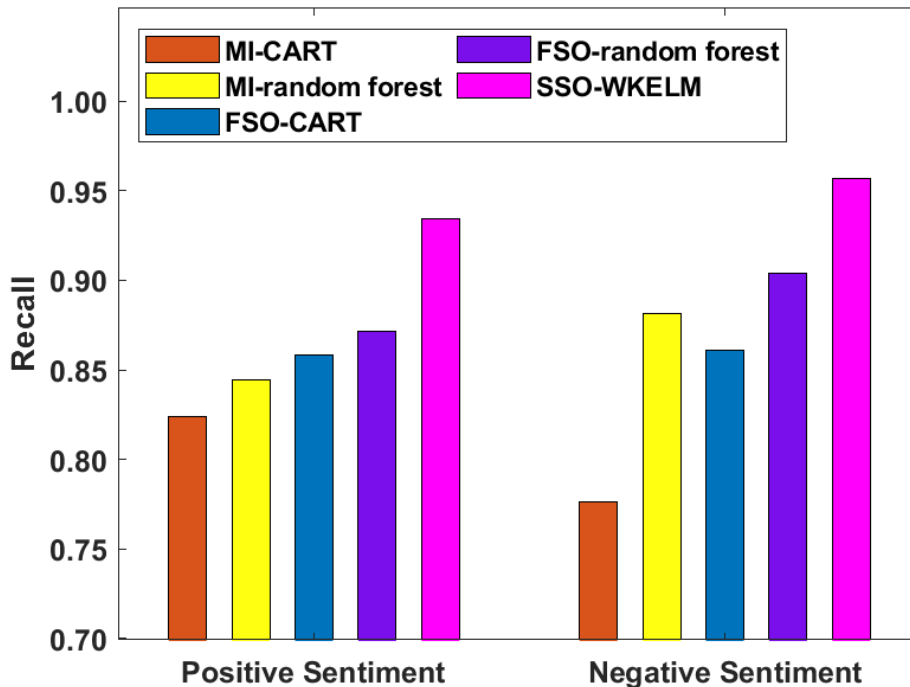


Fig. 6. Recall analysis of SSO-WKELM model

An F-measure analysis of the SSO-WKELM model is made with the existing methods on the classification of positive and negative sentiments as demonstrated in Fig. 7 and Table 4. The simulation result ensured that the MI-CART approach has resulted in the least result by offering a lesser F-measure of 0.8229 and 0.7553 on the classification of positive and negative sentiments correspondingly. Likewise, the MI-RF methodology has exhibited certainly improved results by reaching an F-measure of 0.8545 and 0.7913 on the classification of positive and negative sentiments respectively. Along with that, the FSO-CART manner has offered a moderate performance by attaining an F-measure of 0.8833 and 0.8250 on the classification of positive and negative sentiments correspondingly.

Table 4 Result Analysis of Existing with Proposed SSO-WKELM Method in terms of F-Measure

Methods	Positive Sentiment	Negative Sentiment
MI-CART	0.8229	0.7553
MI-random forest	0.8545	0.7913
FSO-CART	0.8833	0.8250
FSO-random forest	0.9110	0.8606
SSO-WKELM	0.9621	0.9273

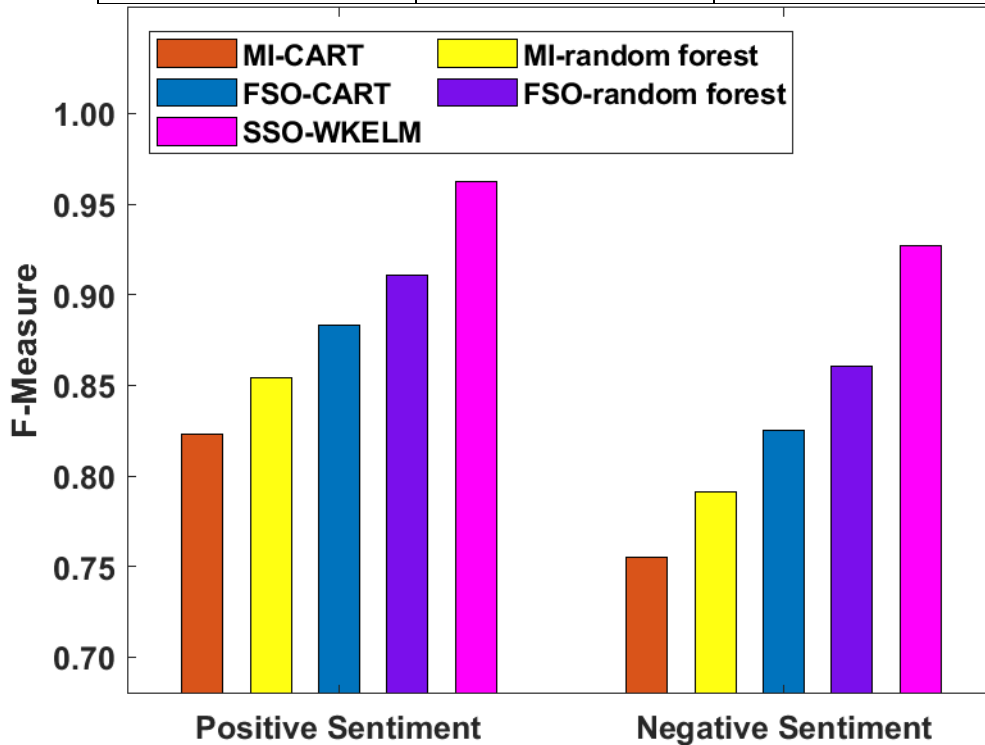


Fig. 7. F-measure analysis of SSO-WKELM model

Moreover, the FSO-RF model has accomplished manageable outcomes with the F-measure of 0.9110 and 0.8606 on the classification of positive and negative sentiments respectively. However, the SSO-WKELM model has surpassed the other techniques by providing a higher F-measure of 0.9621 and 0.9273 on the classification of positive and negative sentiments correspondingly.

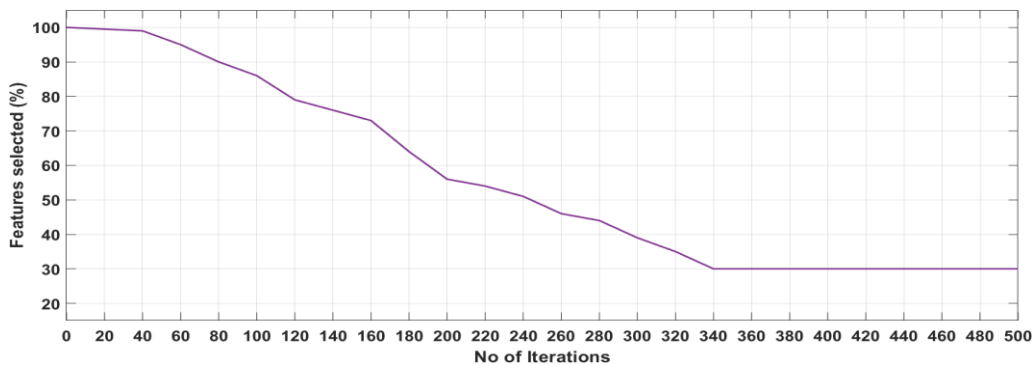


Fig. 8. Analysis of Feature Selection Results

Fig. 8 investigates the feature selection analysis of the SSO algorithm under varying number of iterations. The figure depicted that the SSO algorithm has effectively chosen a subset of features under different set of iterations.

4. Conclusion

This paper has presented an efficient SSO-WKELM model for the SA of big data for online product reviews. The SSO-WKELM model performs SA on big data through different processes such as preprocessing for

unwanted noise removal, TF-IDF for feature vector generation, SSO algorithm for FS and WKELM for classification. The usage of SSO algorithm for FS processed to enhanced classifier results. In order to ensure the effective performance of the SSO-WKELM model, a series of simulations were performed on the product review dataset from amazon and synthetic dataset. The resultant experimentation outcome verified the proficient results analysis of the SSO-WKELM model on the applied dataset. As a part of future scope, the classification performance can be improvised by the use of FS methodologies.

References

- Ramanujam RS, Nancyamala R, Nivedha J, Kokila J. Sentiment analysis using big data. Paper presented at: 2015 International Conference on Computation of Power, Energy, Information and Communication (ICCPEIC); 2015:0480-0484.
- Sehgal D, Agarwal AK. Real-time sentiment analysis of big data applications using Twitter data with Hadoop framework. *Soft Computing: Theories and Applications*. Singapore: Springer; 2018:765-772.
- Kurian D, Vishnupriya S, Ramesh R, et al. Big data sentiment analysis using hadoop. *Int J Innov Res Sci Technol*. 2015;1:92-96.
- Wang H, Xu Z, Fujita H, Liu S. Towards felicitous decision making: an overview on challenges and trends of big data. *Inform Sci*. 2016;367: 747-765.
- Chandana R, Harshitha D, Ramachandra A. Big data migration and sentiment analysis of real time events using hadoop ecosystem. Paper presented at: International Conference on Intelligent Data Communication Technologies and Internet of Things; 2018:764-770.
- Roshanfekar B, Khadivi S, Rahmati M. Sentiment analysis using deep learning on Persian texts. Paper presented at: 2017 Iranian Conference on Electrical Engineering (ICEE); 2017:1503-1508.
- Keshavarz H, AbadehMS, AlmasiM. A new lexicon learning algorithm for sentiment analysis of big data. Paper presented at: 2017 IEEE 15th International Symposium on Intelligent Systems and Informatics (SISY); 2017:000249-000254.
- Tsai C-W, Lai C-F, Chao H-C, Vasilakos AV. Big data analytics: a survey. *J Big Data*. 2015;2:21.
- Sharef NM, Zin HM, Nadali S. Overview and future opportunities of sentiment analysis approaches for big data. *JCS*. 2016;12:153-168.
- Graham G, MeritonRF, Hennelly P. Sentiment Analysis Using KNIME: a Systematic Literature Review of Big Data Logistics. Leeds: The University of Leeds; 2016
- Sharma S, Bansal M, Kaushik A. A survey on sentiment analysis for big data. *Int J Adv Res Sci Eng*. 2017;6:412-416.
- Balaji SN, Paul PV, Saravanan R. Survey on sentiment analysis based stock prediction using big data analytics. Paper presented at: 2017 Innovations in Power and Advanced Computing Technologies (i-PACT); 2017:1-5.
- Caballero, A., Niguidula, J.D. and Caballero, J.M., 2017, November. Analysis and Visualization of University Twitter Feeds Sentiment. In *International Conference on Big Data Technologies and Applications* (pp. 132-145). Springer, Cham.
- Luque-Chang, A., Cuevas, E., Fausto, F., Zaldivar, D. and Pérez, M., 2018. Social spider optimization algorithm: modifications, applications, and perspectives. *Mathematical Problems in Engineering*, 2018.
- Ibrahim, R.A., AbdElaziz, M., Oliva, D., Cuevas, E. and Lu, S., 2019. An opposition-based social spider optimization for feature selection. *Soft Computing*, 23(24), pp.13547-13567.
- Ding, S., Zhang, J., Xu, X. and Zhang, Y., 2016. A wavelet extreme learning machine. *Neural Computing and Applications*, 27(4), pp.1033-1040.