Bayesian Deep Learning Enabled Sentiment Analysis on Web Intelligence Applications

Abeer D. Algarni*

Department of Information Technology, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, Riyadh, 11671, Saudi Arabia
*Corresponding Author: Abeer D. Algarni. Email: adalqarni@pnu.edu.sa

Abstract: In recent times, web intelligence (WI) has become a hot research topic, which utilizes Artificial Intelligence (AI) and advanced information technologies on the Web and Internet. The users post reviews on social media and are employed for sentiment analysis (SA), which acts as feedback to business people and government. Proper SA on the reviews helps to enhance the quality of the services and products, however, web intelligence techniques are needed to raise the company profit and user fulfillment. With this motivation, this article introduces a new modified pigeon inspired optimization based feature selection (MPIO-FS) with Bayesian deep learning (BDL), named MPIO-BDL model for SA on WI applications. The presented MPIO-BDL model initially involved preprocessing and feature extraction take place using Term Frequency—Inverse Document Frequency (TF-IDF) technique to derive a useful set of information from the user reviews. Besides, the MPIO-FS model is applied for the selection of optimal feature subsets, which helps to enhance classification accuracy and reduce computation complexity. Moreover, the BDL model is employed to allocate the proper class labels of the applied user review data. A comprehensive experimental results analysis highlighted the improved classification efficiency of the presented model.

Keywords: Social media; data classification; Bayesian deep learning; artificial intelligence; web intelligence; feature selection

1 Introduction

With the tremendous growth of Internet and World Wide Web (WWW), people have entered the digital era. The Web offers an entire novel media for communication, which exceeds the conventional transmission media, like television, radio, and telephone. Web Intelligence (WI) utilizes Artificial Intelligence (AI) and modern Information Technology (IT) on the Internet and Web [1]. The WI technique includes the application of AI in the area of web based application. The upcoming web network and growth of outstanding machines and devices are fully dependent upon AI method. This would be one of the milestone techniques attained by human intelligence. The upcoming applications of AI in web technology consist of web mining semantic web, web farming, web spiders, web information retrieval, social intelligence design, and so on, the application was declared that is

unlimited to the above-mentioned and creation of higher end web security system, business markets, anti-hacking security, etc. Organizations utilize WI applications for obtaining real-world perception to the public insight of their products. It is determined by effective social media postings, news media coverage and real-time activities, insight of a product could alter quickly and the mood of consumers [2]. Consumers can deliberate the products via digital channels and reply to message, however, it also played a significant part in shaping product reputation, for instance, while commenting or repeating a story. The affective and sentiment classes are the major indicator acquired from this client's behaviors. They assist organizations to understand better the public discussion and trajectory evolve insights of their products. For measuring effective transmission, though, common emotional classes always don't serve. The domain specific affective methods integrate specific emotional classes that are not made in common methods. Their interpreting can differ from standard insights. An additional benefit of domain specific affective method is possible to involve emotional classes and other undesired/desired semantic relations certain to the conditional context.

Sentiment Analysis (SA) [3,4] provides several tools and techniques that are utilized for extracting and recognizing subjective data that denotes client's opinions. These tools and techniques are classified in the natural language processing (NLP) [5]. The mining of user option from client review is named opinion mining as well as it is accomplished with text mining. SA consist of opinion Learning (ML) based SA, hybrid methods, and lexicon SA met [6]. The Lexicon SA mostly process the polarity of token (i.e., words) in a sentence. A lexicon is a container or dictionary, which comprises a huge collection of regular words that are classified depending upon the polarity score. But more people utilize informal words in analyses, which is not existing in lexicons. Thus, scientists highlight that employing equivalent method for sentiment recognition in the text. Henceforth, the secondary class used an ML method for SA [7,8]. The method could be trained on an instance dataset and utilized to accomplish prediction on distinct datasets. The challenge is equated as classification process, for instance, a document is denoted by a collection of features [9]. Later, this document is labelled depending upon the polarity (like positive, neutral, and negative), and transformed to a feature matrix. In this manner, ML method provides a higher efficiency compared to lexicon-based technique, which is utilized for detecting sentiment [10].

This paper has presented a new modified pigeon stimulated optimization-based feature selection (MPIO-FS) with Bayesian deep learning (BDL), named MPIO-BDL model for SA on WI applications. The presented MPIO-BDL model initially involves preprocessing and feature extraction for deriving a suitable set of information from the user reviews. Moreover, the MPIO-FS approach is employed for the choice of optimal feature subsets, which aids to enhance classification accuracy and decrease computation complexity. Furthermore, the BDL model is utilized for allocating the appropriate class labels of the applied user review data. A comprehensive experimental result analysis takes place to ensure the effective performance of the suggested MPIO-BDL models.

2 Literature Review

Hakh et al. [11] accomplished SA on travelers' opinions regarding airlines. The researcher has made the FS and oversampling methods are also essential for achieving advanced outcomes. Feature analyses are executed for selecting the optimum feature that not only enhances the entire efficiency of the method however it decreases the trained time. Moreover, the skewed distribution of the class determined in most of the small data sets is decreased without affecting overfitting. The outcome of the study displays convincing proof that the presented method contains high classification accuracy while forecasting the 3 types of neutral positive, and negative. Liu et al. [12] presented an equivalent

method and executed a multi-class sentiment classification. An FS method is utilized for extracting the significant feature to train the ML based technique. The efficiency of k-nearest neighbor (KNN), Support Vector Machine (SVM), Naive Bayes (NB), and Radial Basis Function (RBF)-neural network (NN) is verified with ten-fold cross-validation [13].

Hasan et al. [14] accomplished SA by utilizing ML method. The polarity is made by utilizing word sense disambiguation (WSD), TextBlob, and SentiWordNet SA. The TextBlob contains fundamental features of NLP basis that are utilized for subjectivity and polarity estimation of twitter posts. The SentiWordNet is a widely accessible analysis for English language, which comprises ideas extract from a wordnet database. Additionally, W-WSD can able to identify the accurate word sense in a certain context. Pandey et al. [15] proposed a Metaheuristic technique termed CSK that is depending upon k-means (K) and Cuckoo Search (CS). As clustering perform a major part in investigating the perspectives and opinions in client tweets, the study presented a technique, which is utilized for finding the best cluster head from the Twitter datasets. The research demonstrates the possible results. Catal et al. [16] examined the effect of several classification methods on Turkish sentiment classification. The voting method utilizes SVM, bagging, and ND for evaluating the efficiency. The outcome shows that the usage of several classifications raises the efficiency of separate classification. The study endorses that the several classification methods have high significance for sentiment classification. Additionally, the utilization of several classification and preprocessing methods assist to enhance the classification.

Eler et al. [17] proposed that the choice of suitable preprocessing method might generate improved classification efficiency. They studied several preprocessing methods involving stop word elimination, frequency cut, weighting, and stemming for analyzing the effect of ML based classification techniques. The study demonstrates that the integration of several preprocessing techniques has played a significant part in detecting the optimum classification rate. They investigated the preprocessing method and related effect on the featured space by visualization. Similarly, the utilization of several feature extraction methods has proved for improving classification accuracy [18]. Dzisevič et al. [19] examined the usage of 3 feature extraction methods with NN for text analyses. The Term Frequency—Inverse Document Frequency (TF-IDF) is accompanied by 2 alterations as Latent Semantic Analysis (LSA) and Linear Discriminant Analysis (LDA) are employed for evaluating the efficiency of every feature analysis method. The research demonstrates that TF-IDF assists the method for attaining high accuracy with larger dataset. For small dataset, the integration of LSA and TF-IDF is suitable for achieving equivalent accuracy.

3 Proposed Method

The complete system architecture of the presented MPIO-BDL technique has been shown in Fig. 1. The input user reviews are primarily preprocessed in two stages for removing the unwanted data that exist in them. Afterward, the group of features are extracted from the user reviews. Then, the optimal feature subset selection process is carried out using the MPIO-FS model. Finally, the classification of user reviews takes place using BDL model and proper class labels are allocated to the applied user reviews.

3.1 Preprocessing

The presented technique utilizes the subsequent pre-processing approach in 2 stages beforehand extract the feature:
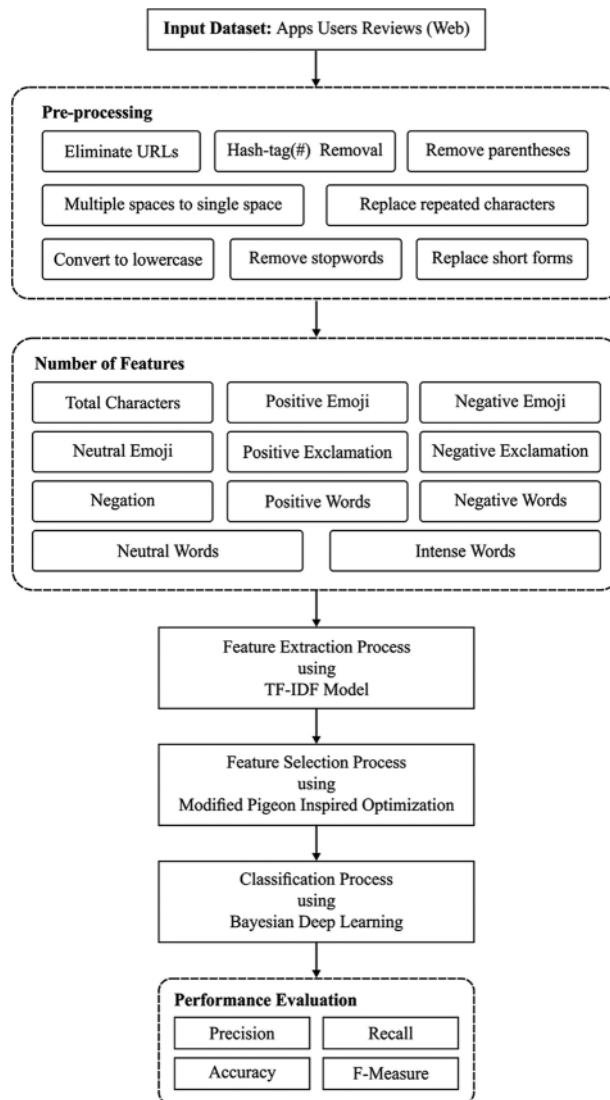
Figure 1: Working process of MPIO-BDL model

Phase 1

This stage removes unnecessary noise components from the input dataset by utilizing the subsequent processes:

- Remove entire URLs by regular expression equivalent and also it represents textual pattern, which determines a searching pattern for text or string. It is utilized for searching the email address, URL, and so on.
- Exchange "@Username" with "user" by utilizing regular expression equivalent.
- As "hash-tag(#)" give beneficial data, eliminate # and keep the word accordingly. Especially "#Lee" is exchanged with "Lee".
- Eliminate dash, backward slash (\) , parenthesis, and forward slash (/) from user review.
- Exchange several white spaces with individual white spaces.

Phase 2

This stage consists of 2 dictionaries such as acronym and stops word, which are utilized for improving the accuracy of resulting input data dataset of Stage 1. The steps involved in Stage 2 is given by:

- Transform each word of user reviews to lowercase.
- Eliminate entire stop words like the, as, is, a, and so on. In comparison with stop word dictionary.
- Exchange series of repetitive characters (like 3 and so on) in a word by 1 character especially "hellooooo" is transformed to "Hello".
- Remove words that don't begin with an alphabet.
- Exchange short-form in the corresponding full-forms by utilizing acronym dictionary.

3.2 Feature Extraction

Next to employing the pre-processing [20], user reviews are transformed to the featured vector using the TF-IDF technique by computing the succeeding eleven features from the Input dataset.

- Overall Features: It denotes the whole number of words accessible in the user reviews.
- Negative Emoji: The special symbols utilized for expressing negative or sad feelings, like : (, : '(, > : (, and so on so-called negative emoji. To acquire the overall numbers of negative emoji in user reviews the negative emoticon dictionary is utilized
- Positive Emoji: Positive emoji's, like: D, ;), :), and so on are the symbols utilized for expressing happy moments. This feature utilizes a positive emoticon dictionary for counting the whole count of positive emojis in the user reviews.
- Neutral Emoji: Neutral emojis (i.e., straight faced emoji) don't give any specific emotion. Overall neutral emoji is count by relating user reviews with neutral emoticon dictionary.
- Positive Exclamation: Exclamatory words, like wow! hurrah! and so on, is utilized for conveying a robust opinion or feeling regarding the subject. For similar, positive exclamation dictionary is utilized for counting the positive exclamation.
- Negative Exclamation: Negative exclamations are count by relating the user review with negative exclamation dictionary.
- Negative Words: This feature denotes the overall counts of negative words like as lost, bad, and so on, in the user reviews.
- Positive Words: This feature counts the amount of positive words such as confidence, achieve, and so on, utilizing positive word dictionary. When there are 2 negative words (i.e., double negation) then these words are counted as individual positive words.
- Neutral Words: Neutral words (i.e., rarely, okay) don't offer any specific feeling or emotions. Overall numbers of neutral word are attained by relating the user reviews with neutral word dictionary.

3.3 Feature Selection

At this stage, the MPIO based FS technique gets executed to choose an optimal subset of features.

The PIO method is one of the bioinspired SI techniques [21]. Pigeons are highly conventional type of bird, which has the capability to fly longer distances for searching the food. Pigeon acquired their homing nature from 2 major operators: landmark, map, and compass. Several studies of pigeon homing knowledge simplified that the pigeon has the ability for navigating the homeland derived from small magnetic particles situated at its peak. This particle transmits indicators to its brain via

trigeminal nerve [22]. The homing knowledge of pigeons is mathematically expressed, which depending upon 2 major operators namely landmark, map, and compass. It utilized their capability of magnetic function for sensing the magnetic field (MF) of the earth, also the pigeon can maintain the sun rise as compass to adapt the direction. When the pigeon gets nearer to its destination, it becomes low that is based on the compass and map operator.

To facilitate this operator, it is expressed mathematical by adapting the velocity V_i and location X_i of pigeon i in all iterations. The values of V_i and X_i are upgraded to the succeeding iteration $(t + 1)^{th}$ based on present iteration value t^{th}

$$V_j(t + 1) = V_i(t).e^{-Rt} + rand.(X_g - X_i(t))\tag{1}$$

$$X_i(t + 1) = X_i(t) + V_i(t + 1)\tag{2}$$

whereas R represents compass and map feature, whereas $rand$ denotes uniform arbitrary number in the extent of zero and one, X_g indicates global optimum solution, $X_i(t)$ indicates present pigeon location at instance t, and $V_i(t)$ indicates present pigeon at iteration t. In landmark operators, every pigeon is rated based on its fitness value. In all generations, the pigeon count is upgraded by Eq. (3), while half pigeon counts are assumed for calculating the desirable location of the center pigeon, whereas each pigeon adapts its target by succeeding the desired destination location. The location of the desired target is estimated in Eq. (4), whereas every pigeon upgrades their location by Eq. (5). The desired destination location is denoted as black pigeon, whereas the pigeon from the circle represents half pigeon counts estimated in Eq. (3).

$$N_p(t + 1) = \frac{N_p(t)}{2}\tag{3}$$

where N_p denotes pigeon counts in the present iteration t.

$$X_c(t + 1) = \frac{\sum X_i(t + 1).Fitness(X_i(t + 1))}{N_p \sum Fitness(X_i(r + 1))}\tag{4}$$

Here X_c denotes location of the center pigeon (i.e., desired destination), whereas X_i indicates present location of every pigeon.

$$X_i(t + 1) = X_i(t) + rand.(X_c(t + 1) - X_i(t))\tag{5}$$

The initial while loop is assumed as compass and map operator, and the secondary loop begins later the initial destination for calculating their direction and create modifications.

The initial form of the presented PIO based FS determines the pigeon/solution by a vector with length equivalents to the feature counts. Since the PIO based process handling the pigeon position in a continued method, the presented PIO for FS determines the solution depiction as a vector of certain length (i.e., feature counts), while the value of the velocity and location vectors primarily denote arbitrary number among zero and one. A conventional method is utilized for calculating the velocity of all pigeons in Eq. (1), later sigmoid operator is utilized for transferring the velocity to the binary form by Eq. (6). To binaries, in a SI method, every pigeon's location is upgrade based on sigmoidal operator value and possibility of an arbitrary uniform number among zero and one in Eq. (7). The remaining method would be the process as the conventional PIO excepting the upgrade location in the landmark operators [23]. Similarly, the sigmoidal operator is utilized in MPIO algorithm for transferring the velocity, later the locations are upgraded as follows.

$$S(V_i(t)) = \frac{1}{1 + e^{\frac{-v_i}{2}}}\tag{6}$$

$$X (t)_{(i,p)} [i] = \begin{cases} 1, if \; (S \; (V_i (t)) > r) \\ 0, \; \text{otherwise} \end{cases} \tag{7}$$

in which $V_i (t)$ represents pigeon velocity in iteration t as well as r denotes uniformly arbitrary number.

Cost function/Fitness function (FF)/objective represents terminology for a process to calculate the fitness solution. The FF calculates the solution that is a subset of chosen features regarding, False Positive Rate (FPR), feature counts, and True Positive Rate (TPR). The feature counts are involved in the FF thus when there is any feature occurs but doesn't impact the FPR or TPR, it chooses to remove. Eq. (6) existing the equation utilized for evaluating the solution/pigeon fitness.

$$FF = w_1 * \frac{SF}{NF} + w_2 * FPR + w_3 * \frac{1}{TPR} \tag{8}$$

where SF denotes chosen feature count, NF represents overall feature counts. The values of the weight are fixed, $w_1 = 0.1$, and $w_2 = w_3 = 0.45$ as FPR and TPR are essential. For $KDDCUP^{99}$ datasets the overall feature count is 41. Fig. 2 demonstrated the flow diagram of PIO algorithm [24].
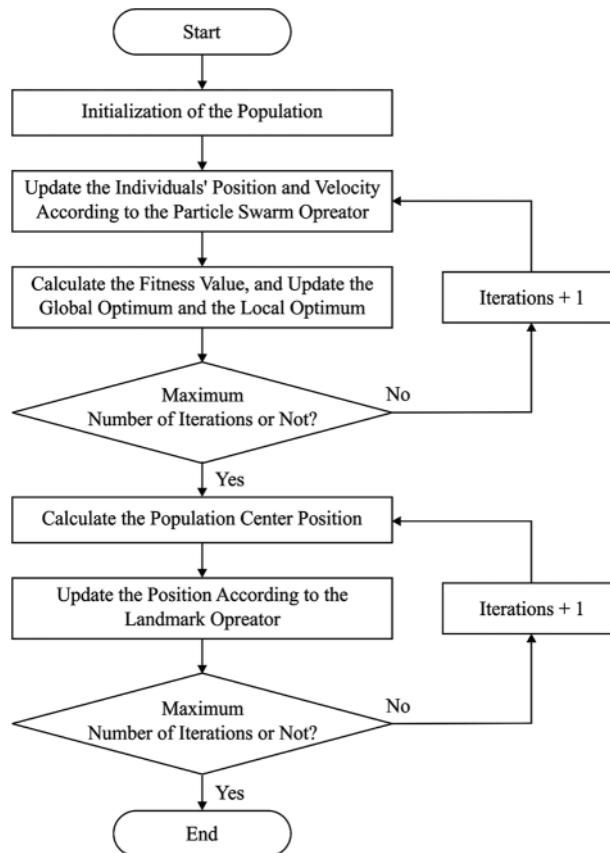

Figure 2: Flowchart of PIO algorithm

3.4 Data Classification

The chosen feature subset of the user reviews is feed as to the BDL based classifier model to determine the class labels of the applied reviews. Where, supervised classification datasets

$D = \{d_1, \ldots, d_N\}$ is signified by the group of tuples which comprises feature and label $d = (x, y)$. The activation function for softmax layer is vector of possibilities $[\phi_1(\theta), \ldots, \phi_k(\theta)]$ and individuals of elements expressed by:

$$\phi_i(\theta) = \frac{\exp\left(x^T w_i + b_i\right)}{\sum_{j=1}^{K} \exp\left(x^T w_j + b_i\right)} \tag{9}$$

where $\theta = \{(w_1, b_1), \ldots, (w_K, b_K)\}$, where $w_i \in R^D$, b_i refers to the scalar, D implies the dimensional of feature space x and $K = 38$ the count of classes. Lastly, an objective function relates to negative cross-entropy model (*log-loss*) determined by:

$$\log p(D|\theta) = -\frac{1}{|D|} \sum_{d \in D} \sum_{k=1}^{K} 1_{y==k} \log \phi_k(\theta) \tag{10}$$

3.4.1 Stochastic Gradient Descent (SGD)

Generally, ML techniques need an enormous number of instances for generalizing, and numerical models to optimize the empirical loss need to calculate the gradient in Eq. (10). In batch optimization techniques is utilize full dataset D for computing the gradient. On the other hand, SGD considers as gradient is the possibility that is evaluated in small batch of data B [25]:

$$\theta_i = \theta_{i-1} + \epsilon \left(\nabla_\theta logp(\theta) + \tilde{\nabla}_\theta logp(D|\theta)\right) \tag{11}$$

where ϵ implies the constant learning rate and $\tilde{\nabla}_\theta \log p(D|\theta) = -\frac{N}{|B_i|} \sum_{j \in B_i} \nabla_\theta logp(d_j|\theta)$ indicates the stochastic gradient estimated utilizing mini-batch B_i.

3.4.2 Stochastic Gradient MCMC (SG-MCMC)

The standard Monte Carlo (MC) techniques for Bayesian data analysis utilize complete trained data D. But this function is very difficult for large-scale datasets $(0 \ll N)$. Conversely, the SG-MCMC techniques calculate to get loser posterior utilizing smaller data batches B which suitable for the restricted calculation cost. Particularly, the stochastic gradient Langevin dynamics (SGLD) method utilizes inspired by the dynamical model for generating schemes. The condition of this model (current location of parameter θ) moves beside the trajectory based on the Langevin dynamics, whereas a time-varying learning rate is utilized for ensuring which the schemes is forever recognized as $\epsilon_i \to 0$.

$$\theta_i = \theta_{i-1} + \frac{\epsilon_i}{2} \left(\nabla_\theta logp(\theta) + \tilde{\nabla}_\theta logp(D|\theta)\right) + \xi_i \tag{12}$$

where $\xi_i \sim \mathcal{N}(0, \epsilon_i)$ refers to the isotropic Gaussian random variable and ϵ_i represents the time-varying learning rate. To create the forecast y^* at new test input x^*, it is required to compute the predictive distribution utilizing a limited group of instances $\{\theta_t\}_{t=1}^{T}$ like:

$$p(y^*|x^*, D) \approx \int p(y^*|x^*, \theta) P(\theta|D) d\theta \tag{13}$$

$$\approx \frac{1}{T} \sum_{t=1}^{T} p(y^*|x^*, \theta_t) P(\theta_t|D) \tag{14}$$

3.4.3 MC Dropout

In recent times, it is demonstrated that DL methods trained by using dropout regularization are corresponding to variational near Bayesian inference [26]. The empirical predictive distribution is

evaluated by MC sampling and due to their simplicity, it can be an optimal candidate to evaluate posterior distributions.

$$z \sim B(p) \tag{15}$$

$$\theta_d = \Theta_{MAP} \odot z \tag{16}$$

$$p\left(y^*|x^*, \mathrm{D}, p\right) \approx \frac{1}{T} \sum_{t=1}^{T} p\left(y^*|x^*, \theta_d\right) P\left(\theta_d|D, p\right) \tag{17}$$

where θ_{MAP} refers the MAP estimation and θ_d implies the parameter vector θ when it executing the Bernoulli mask z by dropout rate p.

4 Experiments

This section validates the analysis of the MPIO-BDL method with other present technologies with respect to different measures. Table 1 explores the results analysis of the MPIO-BDL model with other existing models under distinct types of classes.

Table 1: Result analysis of existing with proposed MPIO-BDL Method using different types of features

		TF-IDF feature								
Techniques	Acc.	Negative class			Neutral class			Positive class		
		Prec.	Rec.	F1	Prec.	Rec.	F1	Prec.	Rec.	F1
MPIO-BDL	94.00	93.00	86.00	89.00	88.00	92.00	88.00	95.00	98.00	97.00
GBSVM	92.00	92.00	84.00	88.00	85.00	90.00	87.00	94.00	96.00	95.00
GBM	91.00	92.00	82.00	87.00	80.00	94.00	86.00	95.00	95.00	95.00
SVM	91.00	89.00	83.00	86.00	84.00	86.00	85.00	93.00	95.00	94.00
LR	88.00	91.00	75.00	82.00	85.00	71.00	77.00	88.00	97.00	92.00
RF	87.00	87.00	71.00	78.00	79.00	82.00	81.00	89.00	94.00	92.00
		TF-IDF (Bi-gram) feature								
Methods	Acc.	Negative class			Neutral class			Positive class		
		Prec.	Rec.	F1	Prec.	Rec.	F1	Prec.	Rec.	F1
MPIO-BDL	77.00	76.00	58.00	65.00	63.00	29.00	38.00	78.00	94.00	85.00
GBSVM	75.00	75.00	56.00	64.00	60.00	26.00	36.00	76.00	92.00	83.00
GBM	70.00	76.00	36.00	49.00	46.00	18.00	26.00	71.00	94.00	81.00
SVM	75.00	83.00	49.00	61.00	90.00	20.00	32.00	74.00	97.00	84.00
LR	68.00	88.00	21.00	34.00	92.00	3.00	6.00	67.00	99.00	80.00
RF	73.00	78.00	49.00	60.00	55.00	26.00	35.00	75.00	93.00	83.00

(Continued)

Table 1: Continued

Methods	Acc.	Negative class			Neutral class			Positive class		
		Prec.	Rec.	F1	Prec.	Rec.	F1	Prec.	Rec.	F1
MPIO-BDL	74.00	86.00	40.00	54.00	73.00	21.00	35.00	72.00	98.00	85.00
GBSVM	72.00	83.00	37.00	51.00	69.00	19.00	30.00	71.00	96.00	82.00
GBM	68.00	84.00	18.00	30.00	77.00	11.00	19.00	67.00	99.00	80.00
SVM	72.00	96.00	32.00	48.00	94.00	17.00	29.00	70.00	99.00	82.00
LR	66.00	97.00	11.00	20.00	100.0	1.00	3.00	65.00	100	79.00
RF	72.00	41.00	32.00	47.00	84.00	18.00	30.00	70.00	99.00	82.00

TF-IDF (Tri-gram) feature

On the other hand, it is also noticed that the MPIO-BDL model has accomplished superior classification results with a higher acc_y of 94% whereas the other methods such as Gradient Boosted Support Vector Machine (GBSVM), gradient boosting algorithm (GBM), SVM, logistic regression (LR), and random forest (RF) methods have demonstrated low acc_y of 92%, 91%, 91%, 88%, and 87% respectively.

Fig. 3 explores the results analysis of MPIO-BDL model under the existence of TF-IDF features. From the figure, it is obvious that the presented MPIO-BDL model has effectually classified the data under several measures. For instance, the presented MPIO-DBL model has shown higher categorized by classifying the negative class instances with the maximum $prec_n$ of 93%, $recal_l$ of 86%, and F_{score} of 89%. Besides, the presented MPIO-DBL model successfully classified the neutral class instances with the $prec_n$ of 88%, $recal_l$ of 92%, and F_{score} of 88%. Likewise, the positive class instances are classified with the $prec_n$ of 95%, $recal_l$ of 98%, and F_{score} of 97%.
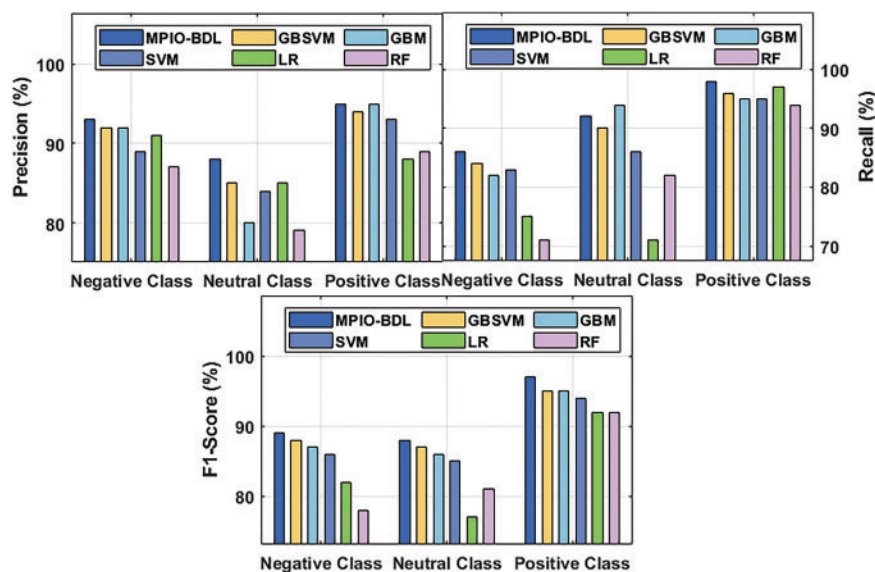
Figure 3: Classification results analysis of MPIO-BDL model under TF-IDF features

Fig. 4 investigative the outcomes analysis of the MPIO-BDL method under the existence of TF-IDF (Bi-gram) features. From the figure, it is stated that the proposed MPIO-BDL technique has effectually classified the data under several measures. For sample, the proposed MPIO-DBL technique has demonstrated superior outcomes by classifying the negative class instances with a higher $prec_n$ of 76%, $recal_l$ of 58%, and F_{score} of 65%. Besides, the presented MPIO-DBL manner successfully classified the neutral class instances with the $prec_n$ of 63%, $recal_l$ of 29%, and F_{score} of 38%. Simultaneously, the positive class instances are classified with the $prec_n$ of 78%, $recal_l$ of 94%, and F_{score} of 85%. On the other hand, it is also noticed that the MPIO-BDL technique has accomplished higher classification results with the maximum acc_y of 77% whereas the other methods such as GBSVM, GBM, SVM, LR, and RF models have outperformed minimal acc_y of 75%, 70%, 75%, 68%, and 73% correspondingly.
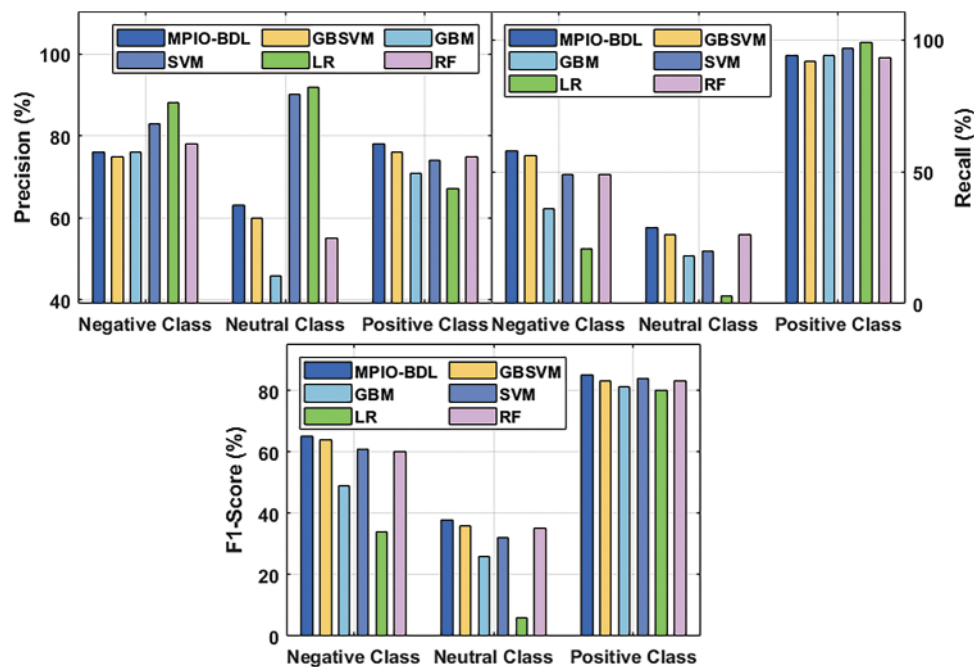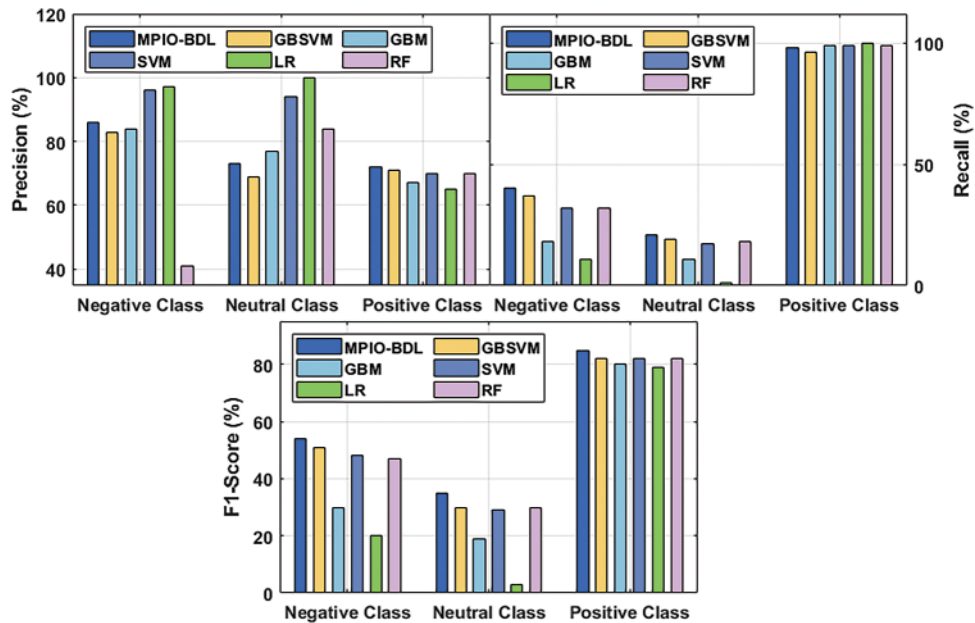

Figure 4: Classification results analysis of MPIO-BDL model under TF-IDF (Bi-gram) features

Fig. 5 examines the analysis of the MPIO-BDL approach under the existence of TF-IDF (Tri-gram) features. From the figure, it is clear that the presented MPIO-BDL methodology has efficiently classified the data under several measures. For instance, the projected MPIO-DBL methodology has showcased superior outcomes by classifying the negative class instances with the maximal $prec_n$ of 86%, $recal_l$ of 40%, and F_{score} of 54%. Followed by, the presented MPIO-DBL approach successfully classified the neutral class instances with the $prec_n$ of 73%, $recal_l$ of 21%, and F_{score} of 35%. Concurrently, the positive class instances are categorized with the $prec_n$ of 72%, $recal_l$ of 98%, and F_{score} of 85%. On continuing with, it is also noticed that the MPIO-BDL algorithm has accomplished superior classification results with the maximal acc_y of 74% whereas the other techniques namely GBSVM, GBM, SVM, LR, and RF processes have showcased less acc_y of 72%, 68%, 72%, 66%, and 72% correspondingly.

Figure 5: Classification analysis of MPIO-BDL method under TF-IDF (Tri-gram) features

A comprehensive analysis of the MPIO-BDL approach takes place with the recently presented models in Table 2 and Fig. 6. From the result, it is evident that the Rocchio+Naïve Byes+KNN model has failed to achieve effective outcomes by offering a lower acc_y of 73%. At the same time, the NABoE model has accomplished certainly raised acc_y of 86% whereas the SGDC+LR and GCN models have demonstrated moderately equal acc_y of 88%. Next to that, the SGCN model has obtained even increased acc_y of 89%. Though the GBSVM model has resulted in a competitive acc_y of 93%, the presented MPIO-BDL model has showcased better results with a higher acc_y of 94%.

Table 2: Comparison of MPIO-BDL with recently proposed methods

Methods	Accuracy (%)
Proposed MPIO-BDL	94.00
GBSVM	93.00
SGDC+LR	88.00
Rocchio+Naïve Byes+KNN	73.00
GCN	88.00
SGCN	89.00
NABoE	86.00
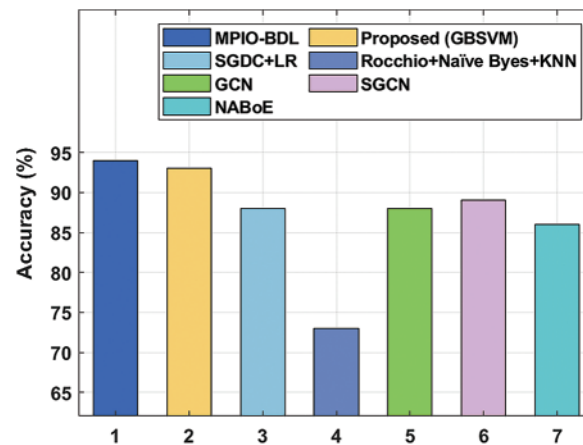
Figure 6: Accuracy analysis of MPIO-BDL model with existing methods

5 Conclusion

This paper presents a MPIO-BDL approach for SA on WI applications. The presented MPIO-BDL model initially involved preprocessing and feature extraction to derive a useful set of information from the user reviews. The presented MPIO-BDL model involves preprocessing, feature extraction, FS, and classification. Then, the optimal feature subset selection process is carried out using the MPIO-FS model. The application of MPIO algorithm for FS helps to enhance classification accuracy and reduce computation complexity. Finally, the classification of user reviews takes place using BDL model and proper class labels are allocated to the applied user reviews. Extensive simulation analysis is conducted for validating the enhanced classification performance of the MPIO-BDL model. The attained simulation values emphasized the enhanced classification accuracy of the suggested approach over the compared methods interms of various measures. In future, it can be applied in real-time e-commerce websites to rate online products.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] R. Pareek, "Web intelligence-an emerging vertical of artificial intelligence," *International Journal of Engineering and Computer Science*, vol. 3, pp. 9430–9436, 2012.

[2] A. Weichselbraun, J. Steixner, A. Braşoveanu, A. Scharl, M. Göbel *et al.*, "Automatic expansion of domain-specific affective models for web intelligence applications," *Cognitive Computation*, vol. 14, no. 1, pp. 228–245, 2022.

[3] M. Khalid, I. Ashraf, A. Mehmood, S. Ullah, M. Ahmad *et al.,* "GBSVM: Sentiment classification from unstructured reviews using ensemble classifier," *Applied Sciences*, vol. 10, no. 8, pp. 2788, 2020.

[4] L. Chen, W. Wang, M. Nagarajan, S. Wang and A. P. Sheth, "Extracting diverse sentiment expressions with target-dependent polarity from twitter," in *Proc. of the Sixth Int. AAAI Conf. on Weblogs and Social Media*, Dublin, Ireland, vol. 6, 2012.

[5]   B. Liu, "Sentiment analysis and subjectivity," in *Handbook of Natural Language Processing*, 2nd ed., London: Chapman and Hall/CRC, 2009.

[6]   J. Bollen, H. Mao and X. Zeng, "Twitter mood predicts the stock market," *Journal of Computational Science*, vol. 2, no. 1, pp. 1–8, 2011.

[7]   J. Uthayakumar, N. Metawa, K. Shankar and S. K. Lakshmanaprabu, "Intelligent hybrid model for financial crisis prediction using machine learning techniques," *Information Systems and e-Business Management*, vol. 18, no. 4, pp. 617–645, 2020.

[8]   F. Rustam, I. Ashraf, A. Mehmood, S. Ullah and G. Choi, "Tweets classification on the base of sentiments for US airline companies," *Entropy*, vol. 21, no. 11, pp. 1078, 2019.

[9]   L. M. Qaisi and I. Aljarah, "A twitter sentiment analysis for cloud providers: A case study of Azure vs. AWS," in *2016 7th Int. Conf. on Computer Science and Information Technology (CSIT)*, Amman, Jordan, pp. 1–6, 2016.

[10] S. E. Shukri, R. I. Yaghi, I. Aljarah and H. Alsawalqah, "Twitter sentiment analysis: A case study in the automotive industry," in *2015 IEEE Jordan Conf. on Applied Electrical Engineering and Computing Technologies (AEECT)*, Amman, Jordan, pp. 1–5, 2015.

[11] H. Hakh, I. Aljarah and B. A. Shboul, "Online social media-based sentiment analysis for us airline companies," in *Proc. of the New Trends in Information Technology*, Amman, Jordan, pp. 176–181, 2017.

[12] Y. Liu, J. W. Bi and Z. P. Fan, "Multi-class sentiment classification: The experimental comparisons of feature selection and machine learning algorithms," *Expert Systems with Applications*, vol. 80, no. 3, pp. 323–339, 2017.

[13] S. K. Lakshmanaprabu, S. N. Mohanty, S. S. Rani, S. Krishnamoorthy, J. Uthayakumar *et al.,* "Online clinical decision support system using optimal deep neural networks," *Applied Soft Computing*, vol. 81, pp. 105487, 2019.

[14] A. Hasan, S. Moin, A. Karim and S. Shamshirband, "Machine learning-based sentiment analysis for twitter accounts," *Mathematical and Computational Applications*, vol. 23, no. 1, pp. 11, 2018.

[15] A. C. Pandey, D. S. Rajpoot and M. Saraswat, "Twitter sentiment analysis using hybrid cuckoo search method," *Information Processing & Management*, vol. 53, no. 4, pp. 764–779, 2017.

[16] C. Catal and M. Nangir, "A sentiment classification model based on multiple classifiers," *Applied Soft Computing*, vol. 50, no. 9, pp. 135–141, 2017.

[17] D. Eler, D. Grosa, I. Pola, R. Garcia, R. Correia *et al.,* "Analysis of document pre-processing effects in text and opinion mining," *Information*, vol. 9, no. 4, pp. 100, 2018.

[18] J. Li, G. Huang, C. Fan, Z. Sun and H. Zhu, "Key word extraction for short text via word2vec, doc2vec, and textrank," *Turkish Journal of Electrical Engineering Computer Sciences*, vol. 27, no. 3, pp. 1794–1805, 2019.

[19] R. Dzisevič and D. Šešok, "Text classification using different feature extraction approaches," in *Open Conf. of Electrical, Electronic and Information Sciences (eStream)*, Vilnius, Lithuania, pp. 1–4, 2019.

[20] A. C. Pandey, D. S. Rajpoot and M. Saraswat, "Twitter sentiment analysis using hybrid cuckoo search method," *Information Processing & Management*, vol. 53, no. 4, pp. 764–779, 2017.

[21] H. Duan and P. Qiao, "Pigeon-inspired optimization: A new swarm intelligence optimizer for air robot path planning," *International Journal of Intelligent Computing and Cybernetics*, vol. 7, no. 1, pp. 24–37, 2014.

[22] A. Varun and M. S. Kumar, "A comprehensive review of the pigeon inspired optimization algorithm," *International Journal of Engineering & Technology*, vol. 7, no. 3.29, pp. 758–761, 2018.

[23] H. Alazzam, A. Sharieh and K. E. Sabri, "A feature selection algorithm for intrusion detection system based on Pigeon Inspired Optimizer," *Expert Systems with Applications*, vol. 148, no. 13, pp. 113249, 2020.

[24] H. Liu, X. Yan and Q. Wu, "An improved pigeon-inspired optimisation algorithm and its application in parameter inversion," *Symmetry*, vol. 11, no. 10, pp. 1291, 2019.

[25] S. Mandt, M. D. Hoffman and D. M. Blei, "Stochastic gradient descent as approximate Bayesian inference," *The Journal of Machine Learning Research*, vol. 18, no. 1, pp. 4873–4907, 2017.

[26] S. Hernández and J. L. López, "Uncertainty quantification for plant disease detection using Bayesian deep learning," *Applied Soft Computing*, vol. 96, pp. 106597, 2020.