Journal of Smart Internet of Things (JSIoT)
VOL 2023, No.02 | 114-132 | 2023
DOI: 10.2478/jsiot-2022-0008

Article

Modeling of Optimal Fully Connected Deep Neural Network based Sentiment Analysis on Social Networking Data

Zaid Alsalami[1]

[1] Information technology Department, Imam Reza International University, Mashhad, Iran

*Corresponding author: (email: zaid.alsalami@imamreza.ac.ir)

Received 29 July 2022; Accepted 21 October 2022; Issue Published 15 December 2023

Abstract: In recent times, the user reviews and opinions postedon social-media like Twitter, Facebook, and Google+, and so on have received considerable attention in the domain of sentiment analysis (SA), offering valued feedback to public as well as private organizations. The examination of reviews roles an essential role for enhancing the product and service qualities. Although numerous SA methods are available in the literature, it is still needed to boost the outcome of the SA for understanding the customer feedbacks and thereby enhances the product quality. The presented article proposesan optimum fully connected deep neural network (OFCDNN) based SA, called OFCDNN-SA technique on social networking data. The OFCDNN-SA technique encompasses distinct procedureslike feature extraction, classification, pre-processing, and parameter optimization. Additionally, Glove technique is applied for the transformation of input data into feature vectors. Moreover, salp swarm optimization (SSO) based hyperparameter optimization technique is derived for maximally selecting the hyperparameters utilized in the DL method. Finally, the FCDNN methodis employed for classification purposes. To investigate the supreme achievement of the OFCDNN-SA method, a wide-ranging simulation assessment is performed and the experimental outputs highlighted the betterment over current methods by means of diverse measures.

Keywords: Social Media; Deep Neural Network; Optimization; Natural Language Processing; Sentiment Analysis;

Copyright © 2023 Journal of Smart Internet of Things (JSIoT) published by Future Science for Digital Publishing and Sciendo . This is an open access article license CC BY (https://creativecommons.org/licenses/by/4.0/)

1. Introduction

SA allowing us to manually weigh the opinion of an individual towards services, entities, and other products [1]. These facts could assist in making better conclusions looking to enhance key performances indicator. Additionally, the huge adaption of social-media platforms like Twitter and Facebook, platforms for services and e-commerce like Amazon, and even review dedicated

websites like rotten tomatoes, resulted in a massive number of things to be assessed [2]. These dataare generally multi-cultural and multi-lingual, and thus, analysing based on an individual language might bring the challenge of not considering the complete insights [3]. Furthermore, significant problems can avoid fully leveraging this information. Except for some cases, viz., English, most of the language lacks well-preserved resources extensively employed for SA like lexicons and annotated corpus. Next, it could be not direct to adopt the similar SA method for distinct languages, e.g., because of the variation in words usage/order, or the noise presented as a machine translation [4]. As well, they possess code switching content, while user exposes their opinion with a mixture of language in the similar sentence. The sentiments can be defined as a view of or an attitude towards an event/situation. It generally has different kinds of self-indulgent feelings: tenderness, happiness, and nostalgia/sadness. One might determine the sentimental label as a valence/polarity (for example neutral, negative, and positive) or different types of emotional feelings (for example, proud, angry, happy, sad). The definition of sentimental labels would affect the result of the sentimental analyses, hence, it is necessary to carefully define the sentimental labels [5].

The research determined three or more sentimental labels (for example emotional feelings, opinion rating scores), and few researchers adapted 2D labels (for example, negative and positive) [6]. Even though there have been greater performances in the area of sentimental analyses, the binary classification of sentiments remains a significant challenge; e.g., the performances (for example accuracy) of current research vary from 70%–90% based on the characteristic of data. The study aims at the ternary classification (negative, positive, and neutral) and the binary classification (negative/positive). There have been several researches on categorizing sentiments via machine learning (ML)techniques like Support Vector Machine (SVM), Stochastic Gradient Descent (SGD), ensemble, Naive Bayes (NB), and Maximum Entropy (ME). There exists a broad range of methods for SA, e.g., [7] that is based on supervised and unsupervised approaches which exploit syntactic patterns, sentiment lexicons, and grammatical analysis [8]. In recent years, deep learning (DL) methods have become more popular which leads to advanced results, with researchers examining Adversarial Networks, RNNs, and CNN amongst other methods [9].

The presented article proposes an optimum fully connected deep neural network (OFCDNN) based SA, called OFCDNN-SA technique on social networking data. The OFCDNN-SA technique encompasses distinct procedures like feature extraction, classification, pre-processing, and parameter optimization. Additionally, Glove technique is applied for the transformation of input data into feature vectors. Moreover, salp swarm optimization (SSO) based hyperparameter optimization technique is derived for maximally selecting the hyperparameters utilized in the DL method. Finally, the FCDNN method is employed for classification purposes. To investigate the supreme achievement of the OFCDNN-SA method, a wide-ranging simulation assessment is performed and the experimental outputs highlighted the betterment over current methods by means of diverse measures.

2. Literature Review

Abd El-Jawad et al. [10] compare the performances of various DL and ML models, as well as present a novel hybrid method which employs neural networks and text mining for sentiments classification. The datasets employed in this study comprise over 1 million Twitter gathered in 5 fields. The training of the method is performed by 75% of the datasets and examined by the rest 25%. The result shows the highest precision rate of 83.7% that displays the efficacy of the hybrid learning model employed by the scheme on the typical supervised models.Nemes and Kiss [11] analyses the manifestations and sentiments (tweets, comments, hashtags, posts) ofTwitter users, depending on the major trend (using keyword, i.e., mainly ‘covid’ and coronavirus theme) with NLP and Sentimental Classification with RNN model. Wherever we compile, analyses, summarize,

and visualize statistics for additional processes. The trained method works more precisely, with a small errormargin, in defining sentimental polarization in the presentfrequent ‘modern’ obstruse twitters.

Yoo et al. [12] proposed Polaris, a method for predicting and analyzing user's sentiment trajectory for events explored in realtime out of the huge social networking content, and shows the result of preliminary validation work. They display SA and trajectory evaluation such that users could attain the insights in a glimpse. Besides, they increase the precision in SA and prediction with the help of modern DL techniques.Vashishtha and Susan [13] implement relative analyses of this technique on 9 public tweets datasets, 3 sentimental lexicons, 4 advanced techniques for unsupervised SA, and one advanced technique for supervised ML. Conventionally, SA of twitter data is implemented by a single lexicon. The result could provide insights to the researcher for selecting which lexicons are better for social-media platforms. The combination of FL with lexicon for sentimental classification offers a novel model in SA.Asif et al. [14] focus on the SA of social-media multi-lingual textual data for discovering the intensity of the sentiment of radicalism. This research classifies the integrated textual insights to any of 4 classes, containing low extreme, high extreme, neutral, and moderate, according to their level of extremism. In the beginning, multi-lingual lexicons with the intensity weight are generated. This lexicon is authenticated by domain specialists and it achieves 88% precision for authentication. Then, Linear Support Vector Classifier and Multinomial NB models are applied for classification purposes.

Rogers et al. [15] present RuSentiment, a novel datasets for SA of social-media post in Russian, and a further complete annotation guidelineset, which is extendable to another language. Currently, RuSentiment is the larger in its class for Russian consisting 31,185 tweets annotated with Fleiss’ kappa of 0.58 (3 annotations per post). To differentiate the dataset, 6,950 tweets have been pre-elected with an active learning model. Jeong et al. [16] identify product opportunities on the basis of sentiment analysis and topic modeling of social network data. For multi-functional products, this method could recognize latent product subjects deliberated by product consumers in social-media with topic modelling, thus quantify the significance of every topic of the product. Then, the satisfactory level of every product subject is estimated by SA. Lastly, the improvement direction and opportunity value of every product subject from a consumer centered opinionis recognized as a prospect method according to the product topic's satisfaction and importance.

3. The Proposed Model

In the present study, a novel OFCDNN-SA methodology is presented to detect and classify sentiments using social networking data. Fig. 1 shows the working procedure of the OFCDNN-SA methodology. The OFCDNN-SA method comprises diverse procedures called pre-processing, feature extraction, Glove’s embedding, FCDNN based classification, and SSA based parameter optimization. These processes are detailed in the subsequent following:

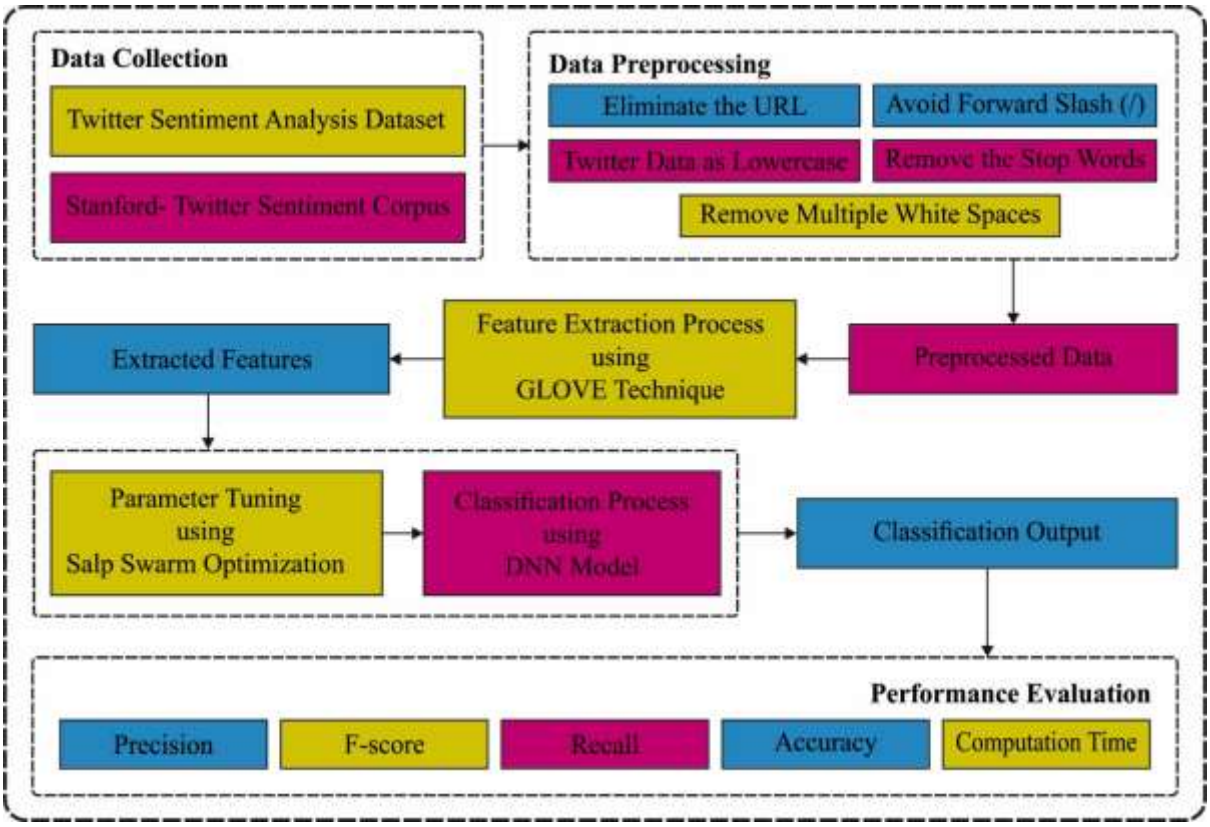


Figure 1: Working principle of proposed model

3.1. Pre-processing

The actual tweets collected from Twitter are embedded with artifacts such as stopwords, fuzzy words, URLs and etc. that have to be limited beforehand processing the feature extraction. Therefore, the recently proposed method enforces the pre-processing method in 2 stages before computing the feature extraction. In the initial phase, the noise is removed from the tweets by applying the certain processes:

Eliminate the URL via routine expression mapping. A standard expression is determined as text patterning that illustrates the exploration of strings and text patterns.

Restore "@Username" with "usr" through regular expression matching.

As "hash-tag(#)" provides useful information, it eliminates # and maintains the word as same. i.e., " #Lee" is restored using "Lee".

Avoid forward slash (/), parentheses, backward slash (\), and - from Twitter contexts.

Remove multiple white spaces via replacing single white spaces.

Now, 2 dictionaries have been applied; stop word and acronym for enhancing the performance of concluding Twitter datasets of Phase 1. The procedure of Phase 2 isportrayed in the following:

Change the Twitter data as lowercase.

Remove the stop words such as a, is, the, etc. by linking to stop word phrasebook.

Restore the sequence of duplicate words (more than 3) by single character i.e., "hellooooo" is converted to "Hello".

Retaining only the word starts with alphabet.

Restore the short form with concerned abbreviation by using acronym dictionary as abovementioned.

3.2.Feature Extraction

When preprocessing is employed, tweets are transformed into feature vectors by evaluating the provided feature of Tweet datasets.

- Overall Features: It illustrates the total number of words existing in twitters.
- Positive Emoji: Emojis as :), ;), : D, etc., are the signs employed for representing happy situations. This feature applies a positive emoticon dictionary.
- Negative Emoji: Specific symbols employed to exhibits sad and negative feelings, such as : (, : (, > : (, etc., are called negative emoji. Negative emoticon dictionary is employed to acquire negative emojis.
- Neutral Emoji: This metric (straight-faced emoji) doesn't provide a particular emotion.
- Positive Exclamation: Exclamatory words, such as hurrah! wow! are employed for changing a stronger opinion/feeling or about the topic. Likewise, a positive exclamation dictionary was used to count the positive exclamation.
- Negative Exclamation: This exclamation is estimated by relating the tweets with a negative exclamation dictionary.
- Negation: to represent the negative suggestion, negation words like no, not, and so on., were employed.
- Positive Words: It computes positive words such as, achieve, confidence, and so on., by employing a positive word dictionary. When there are 2 negative words (dual negation) then, this word is termed as a single positive word.
- Negative Words: It shows the entire negative words such as bad, lost, and several other words in a tweet.
- Neutral Words: The words such as okay, rarely don't denote a certain emotion.
- Intense Words: The phrases such as very, much, and so on. was employed in the sentence for making a productive meaning. The overall intense words are calculated under the application of intense word dictionary.

3.3. Feature Vector Transformation

The glove embedded manner changes the pre-processing data as to a helpful group of feature vectors. It could be effectual unsupervised trained count-based manner. It makes the embedded vector dependent upon the existence of words. During the official, the space vectors are defined as utilize of weighted least squares manner as determined under.

$$J = \sum_{i=1}^{V} \sum_{j=1}^{V} f\left(X_{ij}\right)\left(w_i^T w_j + b_i + b_j - \log\left(X_{ij}\right)\right)^2 \quad (1)$$

where V and f(Xij) refer the vocabulary size and weighting function correspondingly [17]. The GloVe manner is endured trained on dataset with 1 billion words with dictionary of 400000 words with changing embedded vector sizes of 50, 100, 200, and 30. During this analysis, the embedded vector size of 100.

3.4. Data Classification using OFCDNN Technique

To optimally choose the set of hyperparameters involved in the OFCDNN technique, a comprehensive parameter tuning process is performed using SSA technique. Finally, the FCDNN technique accepts the feature vector as input and accomplishes the process of classification.The FCDNN is a class of techniques which utilize several layers for extracting data in the input data. The fundamental layers are FC as well as activation layers. The FC layer contains several neurons. All neurons in FC layers are related to every neuron under the next layer. The outcome of FC layer has bene computed as Eq. (2).

$$y = Wx + b \tag{2}$$

The FC layer is only controlling with linear issues. For adding the non-sequential feature to the technique, the method of activation layers is established [18]. Around extremely utilized activation function contains sigmoid, hyperbolic tangent (Tanh), and ReLU functions. The ReLU has been utilized in Eq. (3).

$$y = \max(0, x) \tag{3}$$

The loss function has been utilized for measuring the efficiency of methods. During this work, the binary cross-entropy loss function was utilized that is computed in Eq. (4):

$$L = -\frac{1}{N}\sum_{i=1}^{N}(y_i \log(p(y_i)) + (1 - y_i)log(1 - p\ (yi)) \tag{4}$$

where N refers the entire amount of instances, i represents the index of instance, y_j signifies the label of ith instance and $p(y_j)$ demonstrates the forecasted probabilities of instance fitting to class. Fig. 2 illustrates the DNN framework [19].One of the general trained techniques has SGD that computes the gradient and next upgrades the weight as well as bias correspondingly. The last aim of trained is for minimizing the loss functions. The FCDNNs are dealing with massive datasets and apply feature engineering with no explicit programming.The primary data for the SSA simulate the behavior of salp family. This swarm ensembleexhibits movement alike to jellyfish. Like other metaheuristic models, the SSA starts with making a set of N solutions X, this solution (x_1 , x_2 , x_N) is kept in a 2D array Dim. The solution is separated into 2 groups, the 1st group is the salp in from, i.e., known as leader salp. Where the 2nd group is the remaining salps i.e., known as followers, and every group is upgraded based on distinct behaviors.

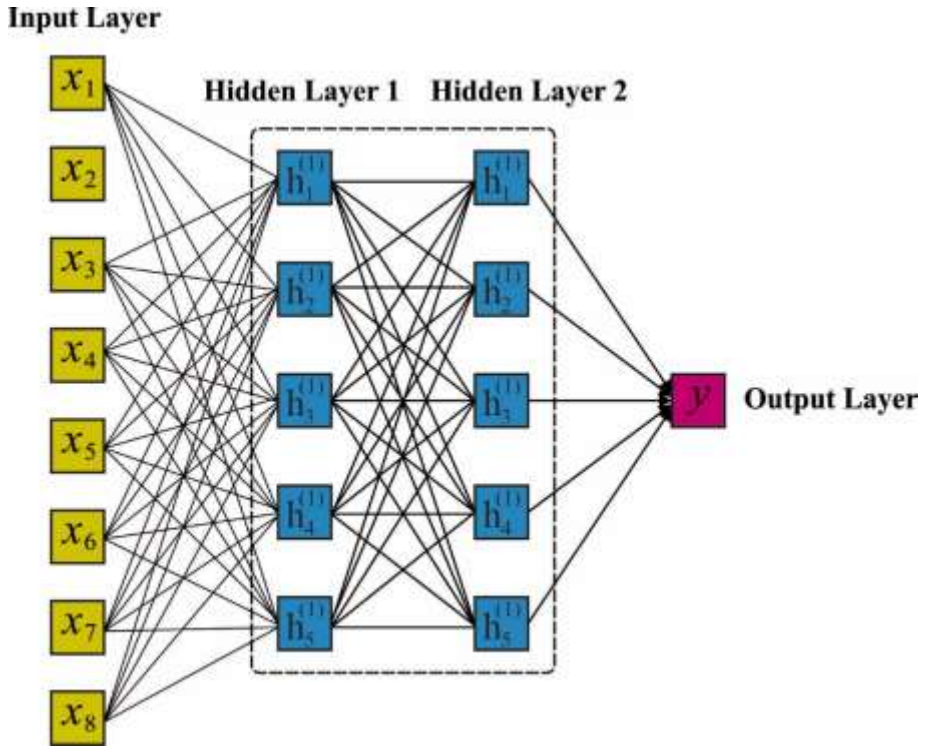


Figure 2: DNN structure

The salp in the leader group is upgraded by,

$$x_j^1(\mathrm{t}+1) = \begin{cases} x_j^b(\mathrm{t}) + r_1\left(\left(ub_j - lb_j\right) \times r_2 + lb_j\right) r_3 \le 0, \\ x_j^b(\mathrm{t}) - r_1\left(\left(ub_j - lb_j\right) \times r_2 + lb_j\right) r_3 > 0, \end{cases} \tag{5}$$

In the equation $x^b(\mathrm{t})$ represent the optimal solution (known as food source), whereas, $x^1(\mathrm{t})$ represent the leader's location. The present iteration is represented as t. The lower and upper limits are represented as lb_j and ub_j at dimension j, correspondingly. r_2&r_3 indicates the arbitrary number in the range of zero and one which preserves the search area.The parameter r_1 is utilized for balancing among the exploration and exploitation by upgrading its value as [20-22]:

$$\alpha_1 = 2e^{-\left(\frac{4t}{\mathrm{t}_{max}}\right)^2}, \tag{6}$$

In which $t_{\max}$ represent the maximal amount of iterations.

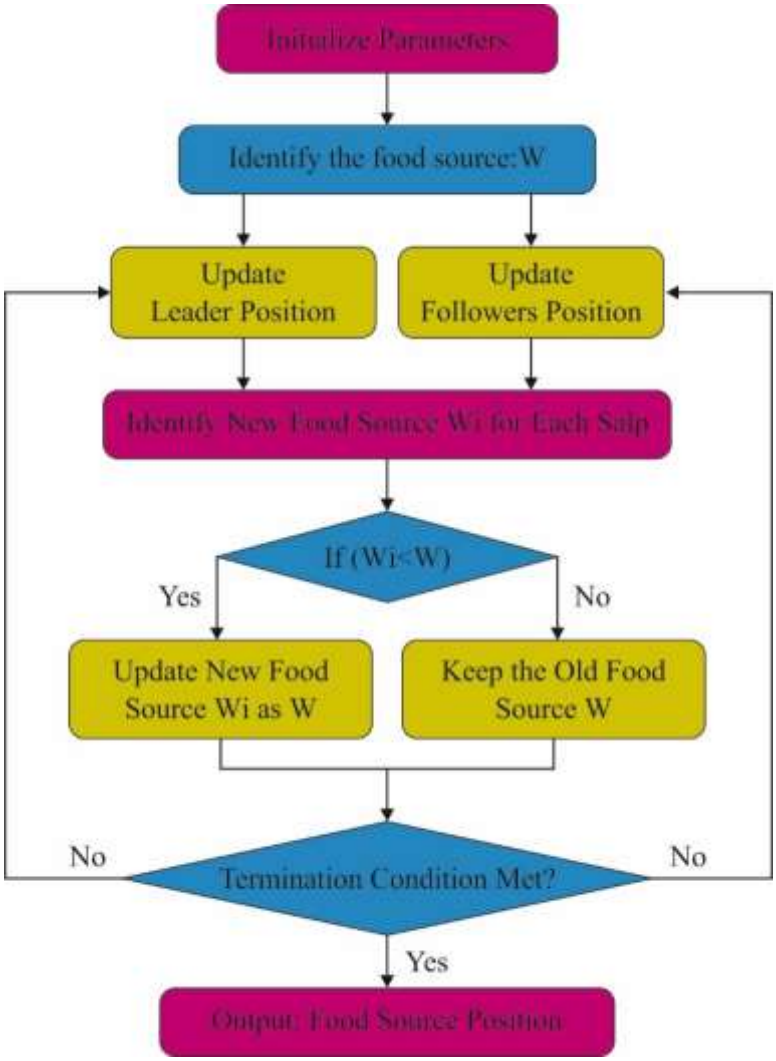


Figure 3: Flowchart of SSA

Additionally, the follower upgrades their position by Eq. (7):

$$x_j^i(\mathrm{t}+1) = \frac{1}{2}\left(x_j^i(\mathrm{t}) + x_j^{i-1}(\mathrm{t})\right), i = 2, \ldots, N \tag{7}$$

The step of the SSA is shown in Algorithm 1. Fig. 3 showcases the flowchart of SSA.

Algorithm 1: Salp swarm algorithm.
Create a group of N solutions x with dimension Dim.
while $(\mathrm{t} < \mathrm{t}_{\max})$ do
Compute the fitness function for every solution.
Define the leader salp (x^{b}).
Upgrade the r_1 by Eq. (3)
for $i = 1, \ldots, N$ do
if $(i == 1)$ then
Upgrade the leader's location by Eq. (2).

else
Upgrade the follower's location by Eq. (4).
end if
end for
end while
Return the leaders location (x^b).

4. Performance Validation

This process of the OFCDNN-SA approach take place on the test Twitter Sentiment Analysis Dataset (TSAD) and Stanford- Twitter Sentiment Corpus (STSC) dataset. These two datasets contain samples under two classes namely positive and negative.

4.1. Results Analysis on STSC Dataset

Table 1 offers a brief results evaluation of the OFCDNN-SA technique on the enforced STSC dataset. The outputs depicted that the OFCDNN-SA technique has achievedgreater classification under different folds. As a sample, with fold-1, the OFCDNN-SA approach has categorized the positive class with the $Prec_n$, rec_l, and $F1_{score}$ of 0.816, 0.922, and 0.817. Besides, the negative class instances are categorized with the $Prec_n$, rec_l, and $F1_{score}$ of 0.855, 0.839, and 0.813 appropriately. Subsequently, with fold-3, the OFCDNN-SA approachhas classified the positive class with the $Prec_n$, rec_l, and $F1_{score}$ of 0.856, 0.902, and 0.817. In line with, the negative class instances are categorized with the $Prec_n$, rec_l, and $F1_{score}$ of 0.855, 0.819, and 0.813 appropriately. Besides, with fold-4, the OFCDNN-SA approach has categorizedthe positive class with the $Prec_n$, rec_l, and $F1_{score}$ of 0.856, 0.912, and 0.847. Furthermore, the negative class instances are classified with the $Prec_n$, rec_l, and $F1_{score}$ of 0.895, 0.839, and 0.843 appropriately.

Table 1: Results assessment of OFCDNN-SA on STSC dataset

Fold Numbers	$Prec_n$	$Reca_l$	$F1_{score}$
Fold-1			
Positive Class	0.816	0.922	0.817
Negative Class	0.855	0.839	0.813
Fold-2			
Positive Class	0.826	0.912	0.867
Negative Class	0.875	0.839	0.843
Fold-3			
Positive Class	0.856	0.902	0.817
Negative Class	0.855	0.819	0.813
Fold-4			
Positive Class	0.816	0.882	0.857
Negative Class	0.895	0.829	0.803
Fold-5			
Positive Class	0.856	0.912	0.847
Negative Class	0.895	0.839	0.843

A brief relative examination of the OFCDNN-SA approach with current methods occurs in Table 2. Figs. 4-5 offers an elaborated output evaluation of the OFCDNN-SA approachwith conventional methods on the enforced STSC dataset. The figure has shown that the LR, MV, and ensemble approaches attained worse outcomesthan the other techniques. Simultaneously, the SVM, ensemble, and NB approaches gave an outcome in a moderate performance. However, the

OFCDNN-SA technique has accomplished maximum outcome with the $Prec_n$, rec_l, and $F1_{score}$ of 0.834, 0.906, and 0.841 under positive class and $Prec_n$, rec_l, and $F1_{score}$ of 0.875, 0.833, and 0.823 subsequently.

Table 2: Comparative results assessmentof OFCDNN-SA approach with recent methods on STSC dataset

	Positive Class			Negative Class		
Methods	$Prec_n$	$Reca_l$	$F1_{score}$	$Prec_n$	$Reca_l$	$F1_{score}$
Proposed OFCDNN-SA	0.834	0.906	0.841	0.875	0.833	0.823
Naive Bayes	0.756	0.745	0.751	0.748	0.759	0.753
Random Forest	0.678	0.832	0.747	0.781	0.603	0.681
SVM	0.740	0.792	0.765	0.775	0.720	0.747
Logistic Regression	0.724	0.783	0.752	0.763	0.700	0.730
Majority Voting	0.714	0.828	0.767	0.795	0.667	0.726
Ensemble Model	0.748	0.780	0.764	0.769	0.736	0.752

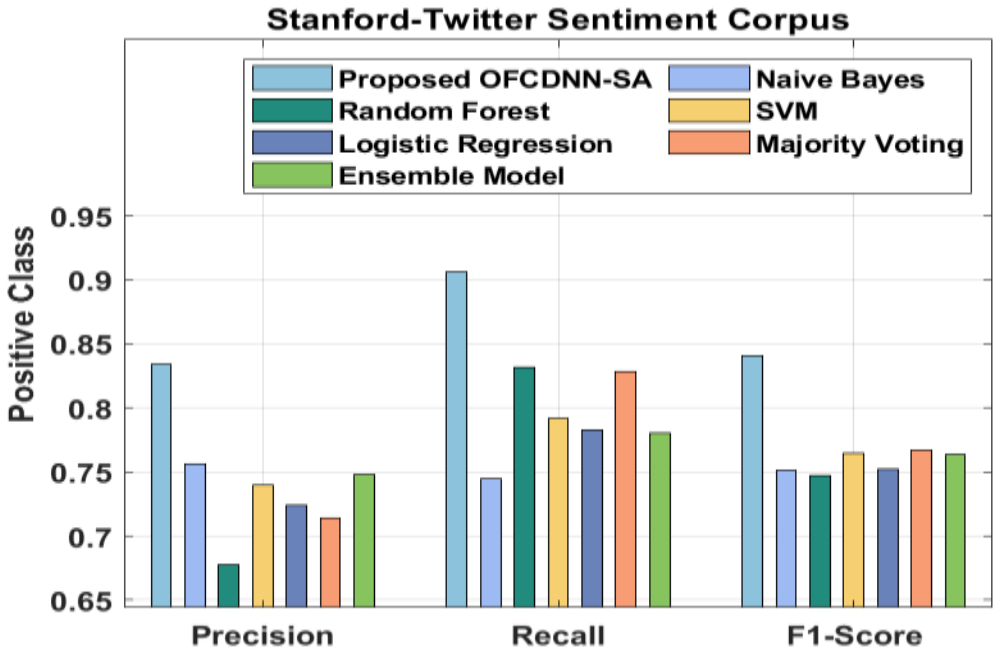


Figure 4: Positive class analysis of OFCDNN-SA model on STSC dataset

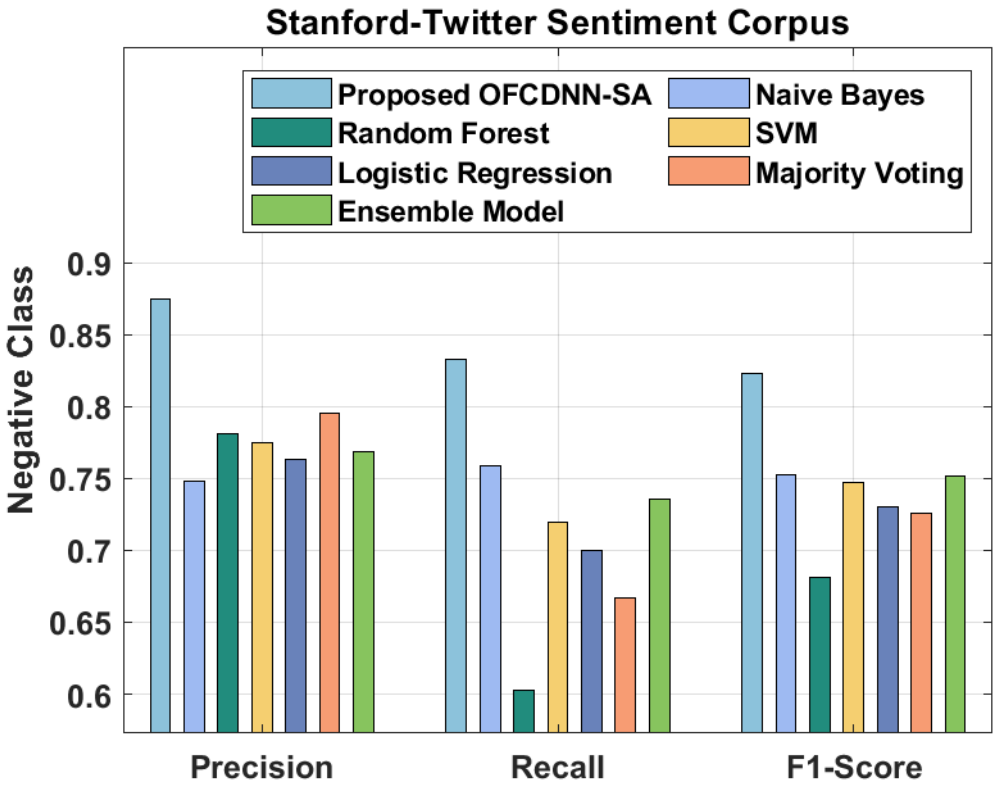


Figure 5: Negative class analysis of OFCDNN-SA method on STSC dataset

Fig. 6 shows the accuracy graph assessment of the OFCDNN-SA model on the enforced STSC dataset. The figure has shown that the accuracy values tend to increase with a rise in epoch count. Additionally, the validation accuracy is observed to be better than the training accuracy.

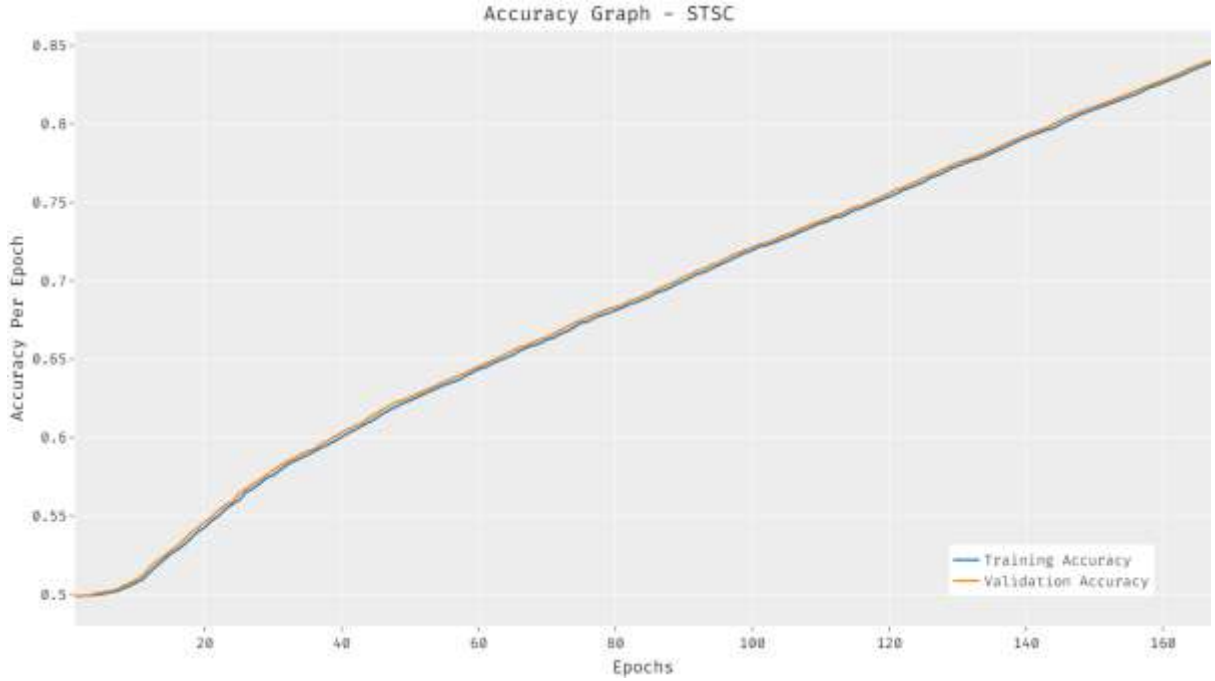


Figure 6: Accuracy Graph assessment of OFCDNN-SA methodon STSC dataset

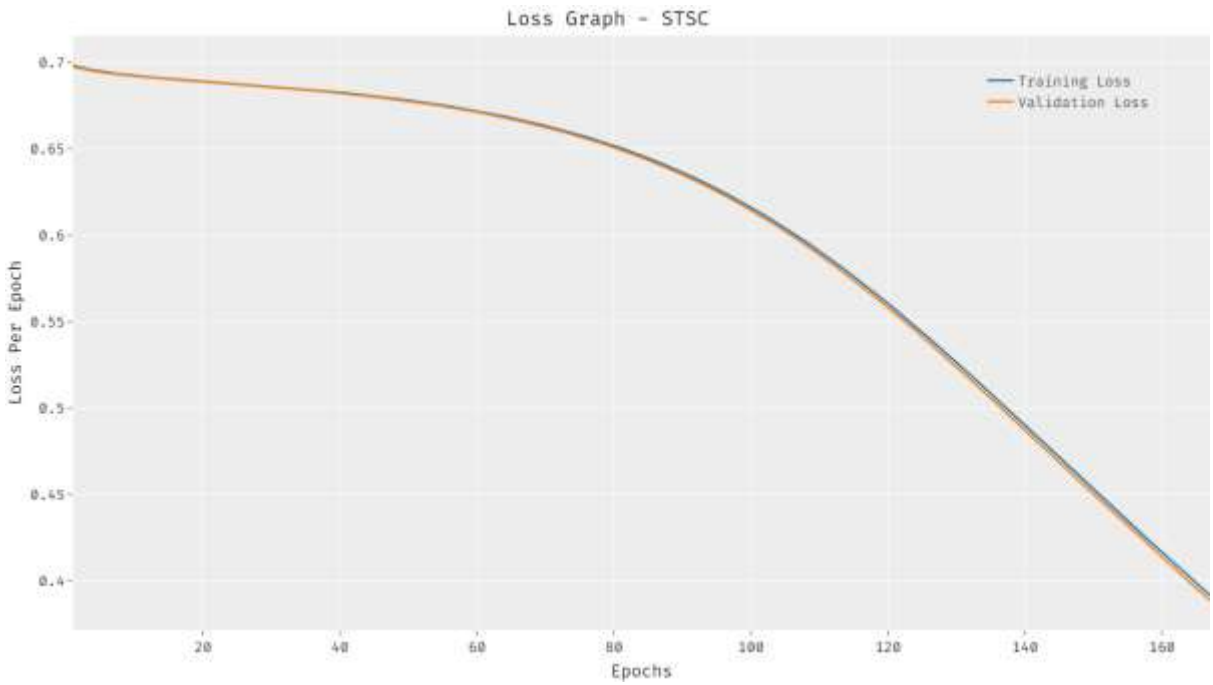


Figure 7: Loss Graph assessmentof OFCDNN-SA method on STSC dataset

Fig. 7 displays the loss graph evaluation of the OFCDNN-SA model on the applied STSC dataset. The figure revealed that the loss values are found to be reduced with a rise in epoch count. It is also noticed that the validation loss is considerably lesser than the training loss.

4.2. Results Analysis on TSAD Dataset

Similarly, Table 3 shows the proficient results assessment of the OFCDNN-SA approach on the applied TSAD dataset.The outcomes showcased that the OFCDNN-SA approachhas accomplished maximal classification under distinct folds. Also, with fold-1, the OFCDNN-SA manner has classified the positive class with the $Prec_n$, rec_l, and $F1_{score}$ of 0.850, 0.951, and 0.869. Also, the negative class instances are categorized with the $Prec_n$, rec_l, and $F1_{score}$ of 0.827, 0.736, and 0.741correspondingly. In line with, with fold-3, the OFCDNN-SA technique has classified the positive class with the $Prec_n$, rec_l, and $F1_{score}$ of 0.840, 0.961, and 0.859. Also, the negative class instances are categorized with the $Prec_n$, rec_l, and $F1_{score}$ of 0.837, 0.776, and 0.731correspondingly. In addition, with fold-4, the OFCDNN-SA technique has classified the positive class with the $Prec_n$, rec_l, and $F1_{score}$ of 0.860, 0.961, and 0.889. Finally, the negative class instances are classified with the $Prec_n$, rec_l, and $F1_{score}$ of 0.867, 0.776, and 0.731correspondingly.

Table 3: Output assessment of OFCDNN-SA on TSAD dataset

Class	$Prec_n$	$Reca_l$	$F1_{score}$
Fold-1			
Positive Class	0.850	0.951	0.869
Negative Class	0.827	0.736	0.741
Fold-2			
Positive	0.840	0.961	0.859

Class			
Negative Class	0.857	0.776	0.731
Fold-3			
Positive Class	0.840	0.941	0.899
Negative Class	0.837	0.776	0.731
Fold-4			
Positive Class	0.816	0.882	0.857
Negative Class	0.895	0.829	0.803
Fold-5			
Positive Class	0.860	0.961	0.889
Negative Class	0.867	0.776	0.731

A brief relative research of the OFCDNN-SA methodology with recent methods occurs in Table 4. Figs. 8-9 gives a short output evaluation of the OFCDNN-SA methodologywith conventional methods on the enforced TSAD dataset.

Table 4: Comparative results analysis of OFCDNN-SA technique on TSAD dataset

	Positive Class			Negative Class		
Methods	$Prec_n$	$Reca_l$	$F1_{score}$	$Prec_n$	$Reca_l$	$F1_{score}$
Proposed OFCDNN-SA	0.844	0.957	0.883	0.847	0.768	0.741
Naive Bayes	0.780	0.772	0.776	0.676	0.686	0.681
Random Forest	0.687	0.921	0.787	0.777	0.396	0.525
SVM	0.761	0.826	0.792	0.713	0.626	0.667
Logistic Regression	0.739	0.851	0.791	0.725	0.567	0.636
Majority Voting	0.730	0.884	0.799	0.759	0.529	0.623
Ensemble Model	0.766	0.822	0.793	0.713	0.639	0.674

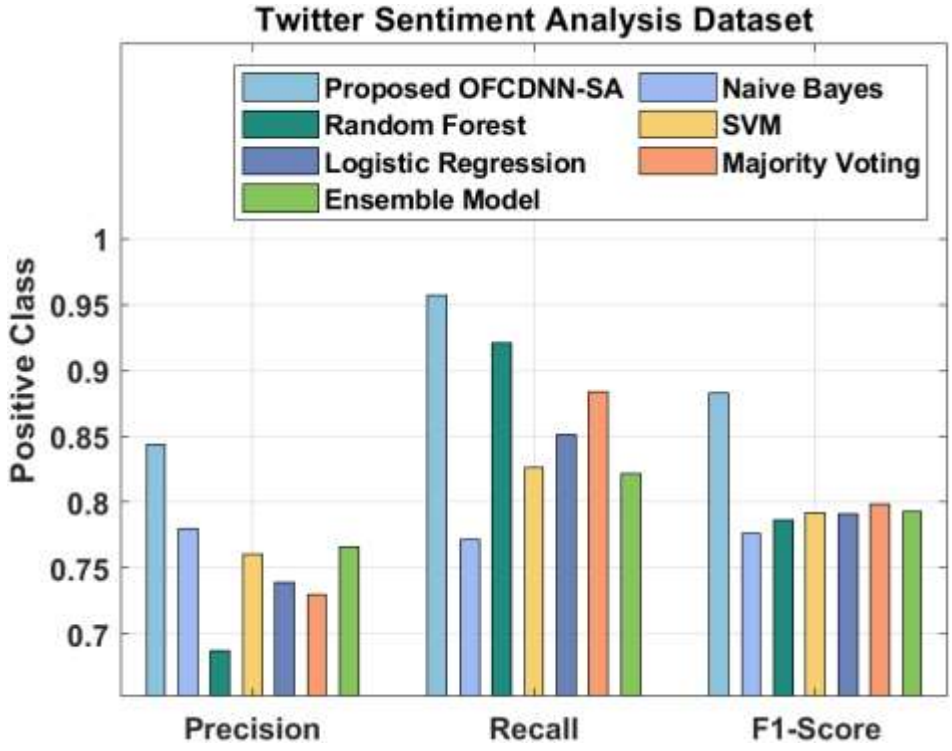


Figure 8: Positive class assessment of OFCDNN-SA method on TSAD dataset

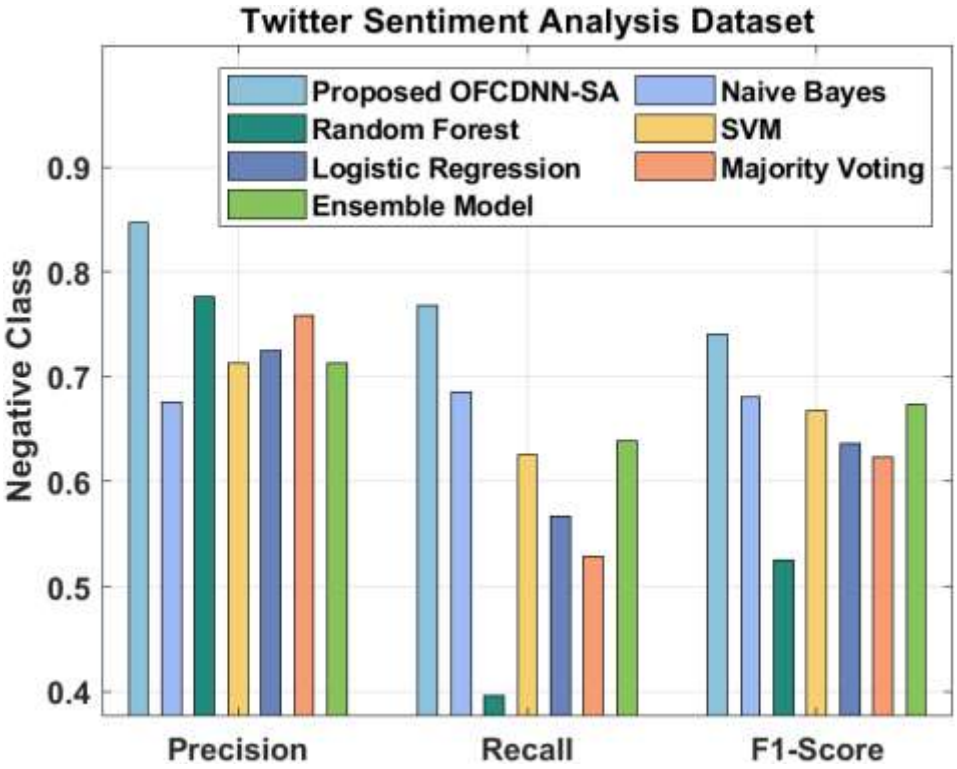


Figure 9: Negative class assessmentof OFCDNN-SA methodon TSAD dataset

The figure portrayed that the LR, MV, and ensemble methods attained worst outcome over the other manners. At the same time, the SVM, ensemble, and NB methodologies have resulted in a moderate performance. Followed by, the OFCDNN-SA method has accomplished maximal result

with the $Prec_n$, rec_l, and $F1_{score}$ of 0.844, 0.957, and 0.883 under positive class and $Prec_n$, rec_l, and $F1_{score}$ of 0.847, 0.768, and 0.471 correspondingly.

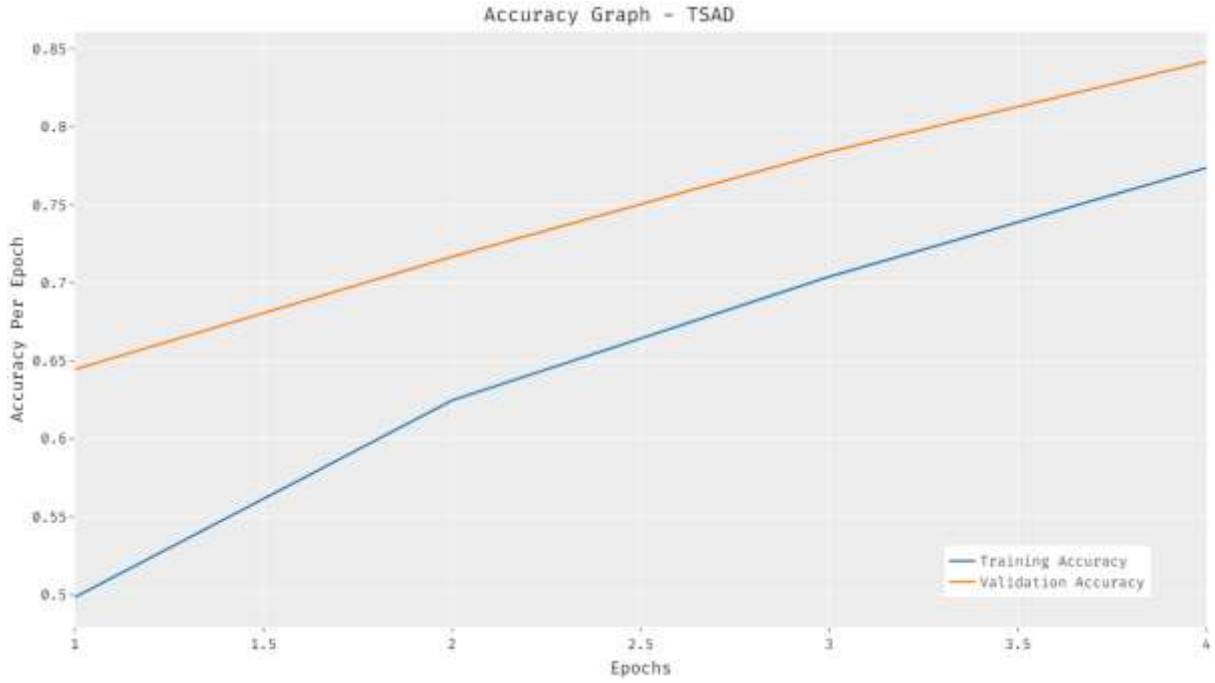


Figure 10: Accuracy Graph assessment of OFCDNN-SA methodon TSAD dataset

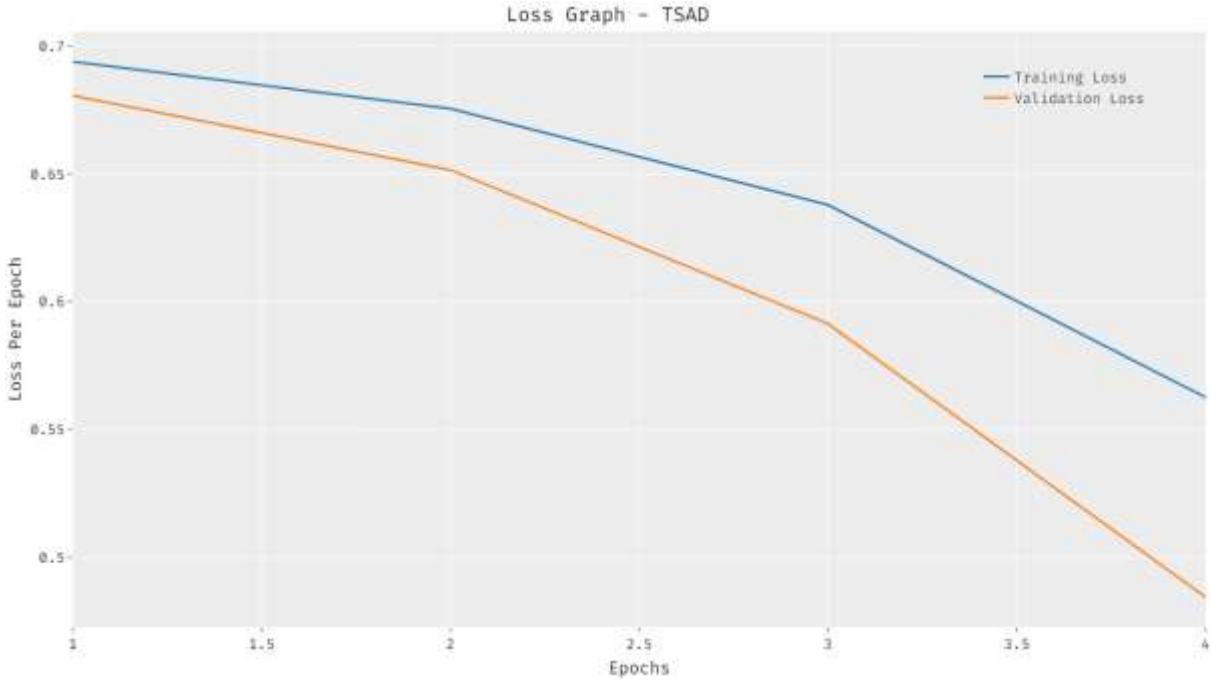


Figure 11: Loss Graph assessmentof OFCDNN-SA methodon TSAD dataset

Fig. 10 illustrates the accuracy graph assessment of the OFCDNN-SA method on the enforced TSAD dataset. The figure exhibited that the accuracy values tend to improve with growth in epoch count. Besides, the validation accuracy is detected to begreater to the training accuracy.

Fig. 11 showcases the loss graph assessmentof the OFCDNN-SA method on the enforced

TSAD dataset. The figure exhibited that the loss values tend to decreasewith growth in epoch count. Besides, the validation loss is detected to belesser to the training loss.

4.3. Discussion

Table 5 demonstrates the computation time analysis of proposed OF CDNN-SA model with TSAD and STSC dataset with existing approach. A detailed CT analysis of the OFCDNN-SA with current method soccurs in Fig. 12. The figure portrayed that the ensemble and majority voting methods resulted in a higher CT of 186s and 169s respectively. Also, the LR, SVM, and RF approaches have tried to obtain moderate CT of 164s, 145s, and 137s respectively. Moreover, the NB model has accomplished near optimal CT of 121s. Moreover, the presented OFCDNN-SA approach gained effective outputs with lower CT of 112s.

Table 5: Computation time analysis of OFCDNN-SA model

Methods	TSAD-Computational Time (s)	STSC-Computational Time (s)
Proposed OFCDNN-SA	112	113
Naive Bayes	121	120
Random Forest	137	138
SVM	145	148
Logistic Regression	164	161
Majority Voting	169	170
Ensemble Model	186	178

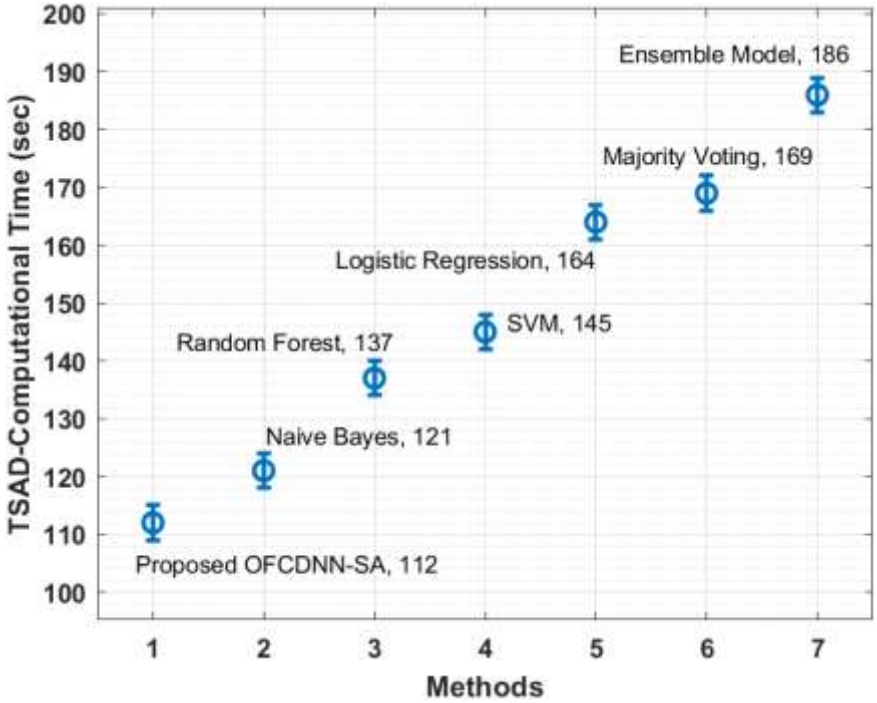


Figure 12: Computational time analysis of OFCDNN-SA technique on TSAD dataset

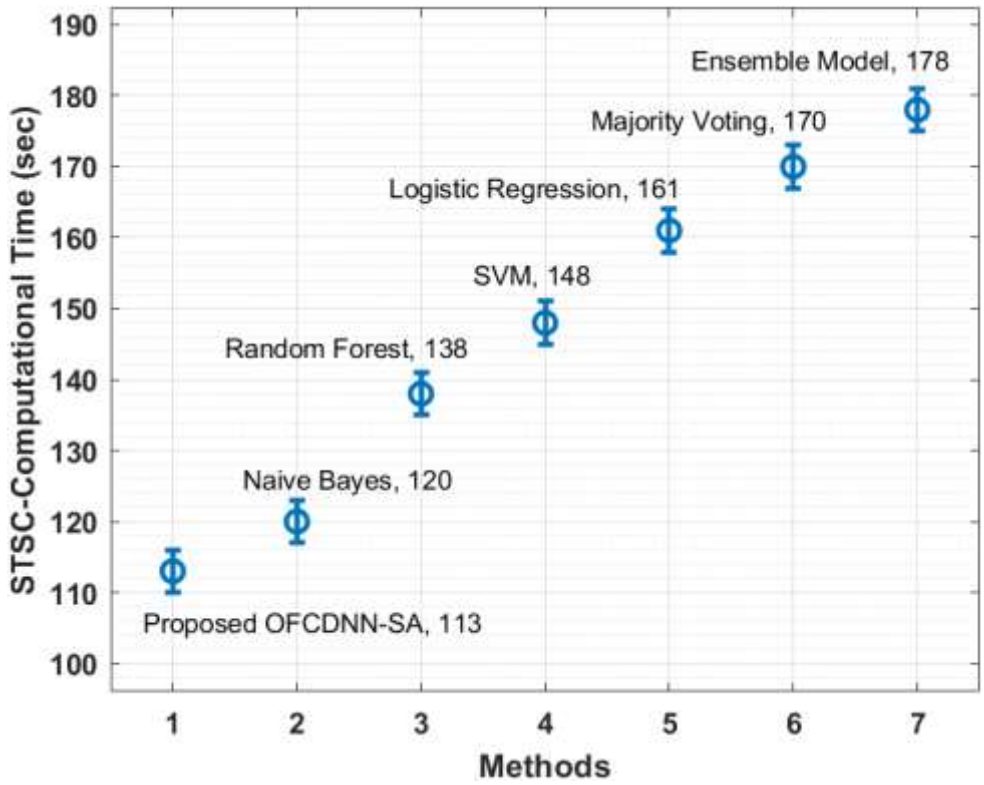


Figure 13: Computational time analysis of OFCDNN-SA technique on STSC dataset

A comprehensive CT analysis of the OFCDNN-SA with existing approaches occured in Fig. 13. The figure demonstrated that the ensemble and majority voting manners have resulted in a maximum CT of 178s and 170s correspondingly. Besides, the LR, SVM, and RF algorithms have tried to attain moderate CT of 161s, 148s, and 138s correspondingly. Likewise, the NB model has accomplished near optimal CT of 120s. At last, the projected OFCDNN-SA methodology has reached effectual outcome with the lower CT of 113s.

Table 6: Relative assessment of OFCDNN-SA method with recent technique

Methods	TSAD-Accuracy	STSC-Accuracy
Proposed OFCDNN-SA	0.841	0.839
Naive Bayes	0.737	0.752
Random Forest	0.706	0.718
SVM	0.744	0.756
Logistic Regression	0.734	0.742
Majority Voting	0.738	0.748
Ensemble Model	0.747	0.758

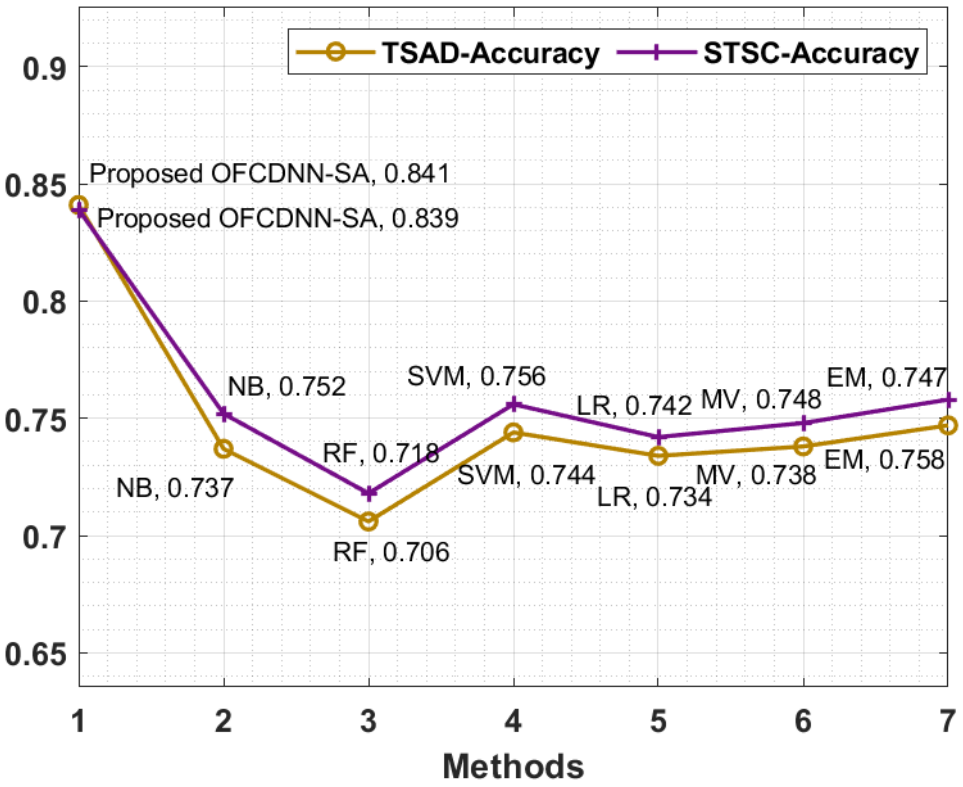


Figure 14: Relative assessmentof OFCDNN-SA methodwith recent techniques

Finally, an elaborate relative research of the OFCDNN-SA with existing approaches on the enforced dataset is in Table 6 and Fig. 14. The outputs portrayed that the existing methods have accomplished lower performance on the applied TSAD and STSC datasets. On compared to existing techniques, the proposed OFCDNN-SA approach has achieved maximum accomplishment with the higher accuracy of 0.841 and 0.839 on the applied TSAD and STSC datasets. Therefore, the OFCDNN-SA approachcan be employed as an effectual tool to determine sentiments on social networking data.

5. Conclusion

In the present study, a novel OFCDNN-SA methodology is presented to detect and classify sentiments using social networking data. The OFCDNN-SA method comprises diverse procedures called pre-processing, feature extraction, Glove's embedding, FCDNN based classification, and SSA based parameter optimization.The design of SSA for the optimal adjustment of the hyperparameters involved in the FCDNN technique outputs in considerableclassifying accomplishment. For inspecting the ultimate accomplishment of the OFCDNN-SA technique, a wide-ranging simulation assessment is carried out. The investigational outputs accentuated the betterment of the OFCDNN-SA technique over current approaches by means of diverse measures. In future, hybrid metaheuristics-based feature selection and feature reduction manners are incorporated to boost the overall classification outcomes.

References

[1] Agüero-Torales, M.M., Salas, J.I.A. and López-Herrera, A.G., 2021. Deep learning and multilingual sentiment analysis on social media data: An overview. Applied Soft Computing, p.107373.

[2] S.L. Lo, E. Cambria, R. Chiong, D. Cornforth, Multilingual sentiment analysis: from formal to informal and scarce resource languages, Artif. Intell. Rev. 48 (4) (2017) 499–527.

[3] D. Tang, B. Qin, T. Liu, Deep learning for sentiment analysis: successful approaches and future challenges, Wiley Interdiscip. Rev.: Data Min. Knowl. Discov. 5 (6) (2015) 292–303.
[4] P. Singhal, P. Bhattacharyya, Sentiment Analysis and Deep Learning: A Survey, Center for Indian Language Technology, Indian Institute of Technology, Bombay, 2016.
[5] D. Vilares, Compositional Language Processing for Multilingual Sentiment Analysis (Ph.D. thesis), Universidade da Coruña, 2017.
[6] B. Liu, Sentiment analysis and opinion mining, Synth. Lect. Human Lang. Technol. 5 (1) (2012) 1–167.
[7] S. Wang, C.D. Manning, Baselines and bigrams: Simple, good sentiment and topic classification, in: Proc. of the 50th Annual Meeting of the ACL: Short Papers, Vol. 2, ACL, 2012, pp. 90–94
[8] Y. Kim, Convolutional neural networks for sentence classification, in: Proc. of the 2014 Conf. on Empirical Methods in NLP, EMNLP, ACL, Doha, Qatar, 2014, pp. 1746–1751.
[9] L.M. Rojas-Barahona, Deep learning for sentiment analysis, Lang. Linguist. Compass 10 (12) (2016) 701–719.
[10] Abd El-Jawad, M.H., Hodhod, R. and Omar, Y.M., 2018, December. Sentiment analysis of social media networks using machine learning. In 2018 14th international computer engineering conference (ICENCO) (pp. 174-176). IEEE.
[11] Nemes, L. and Kiss, A., 2021. Social media sentiment analysis based on COVID-19. Journal of Information and Telecommunication, 5(1), pp.1-15.
[12] Yoo, S., Song, J. and Jeong, O., 2018. Social media contents based sentiment analysis and prediction system. Expert Systems with Applications, 105, pp.102-111.
[13] Vashishtha, S. and Susan, S., 2019. Fuzzy rule based unsupervised sentiment analysis from social media posts. Expert Systems with Applications, 138, p.112834.
[14] Asif, M., Ishtiaq, A., Ahmad, H., Aljuaid, H. and Shah, J., 2020. Sentiment analysis of extremism in social media from textual information. Telematics and Informatics, 48, p.101345.
[15] Rogers, A., Romanov, A., Rumshisky, A., Volkova, S., Gronas, M. and Gribov, A., 2018, August. RuSentiment: An enriched sentiment analysis dataset for social media in Russian. In Proceedings of the 27th international conference on computational linguistics (pp. 755-763).
[16] Jeong, B., Yoon, J. and Lee, J.M., 2019. Social media mining for product planning: A product opportunity mining approach based on topic modeling and sentiment analysis. International Journal of Information Management, 48, pp.280-290.
[17] Alessandro, C., Daniela, M., Michele, M., Andrea, T., Gianmarco, G., Massimo, S., Orazio, Z., Fabio, G. and Giuseppe, T., 2012. Glove port technique for transanal endoscopic microsurgery. International journal of surgical oncology, 2012.
[18] Bi, L., Hu, G., Raza, M.M., Kandel, Y., Leandro, L. and Mueller, D., 2020. A Gated Recurrent Units (GRU)-Based Model for Early Detection of Soybean Sudden Death Syndrome through Time-Series Satellite Imagery. Remote Sensing, 12(21), p.3621.
[19] Ramesh, S. and Vydeki, D., 2020. Recognition and classification of paddy leaf diseases using Optimized Deep Neural network with Jaya algorithm. Information processing in agriculture, 7(2), pp.249-260.
[20] Ewees, A.A., Al-qaness, M.A. and Abd Elaziz, M., 2021. Enhanced salp swarm algorithm based on firefly algorithm for unrelated parallel machine scheduling with setup times. Applied Mathematical Modelling, 94, pp.285-305.
[21] Abualigah, L., Shehab, M., Alshinwan, M. and Alabool, H., 2020. Salp swarm algorithm: a comprehensive survey. Neural Computing and Applications, 32(15), pp.11195-11215.
[22] Zhang J, Wang Z, Luo X (2018) Parameter estimation for soil water retention curve using the salp swarm algorithm. Water 10:815