Retraction Notice

The Editor-in-Chief and the publisher have retracted this article, which was submitted as part of a guest-edited special section. An investigation uncovered evidence of systematic manipulation of the publication process, including compromised peer review. The Editor and publisher no longer have confidence in the results and conclusions of the article.

PP and VD either did not respond directly or could not be reached.

Anas platyrhynchos optimizer with deep belief network-based sarcasm detection and classification model

Vinoth Dakshnamoorthy and Panneer Prabhavathy*
Vellore Institute of Technology, School of Information Technology and Engineering, Vellore, Tamil Nadu, India

Abstract Sarcasm is a state of speech where the speaker says something that is externally unfriendly to wound or deride the listener. The design of an automated sarcasm detection model in the domain of natural language processing (NLP) is mainly based on the context of the statement/statement and sometimes even human beings cannot identify the intrinsic sarcasm in the statement. The recently developed deep learning (DL) models can accomplish better performance than the other traditional models. With this motivation, this study introduces an efficient Anas Platyrhynchos optimizer (APO) with deep belief network (DBN) based sarcasm detection and classification (APODBN-SDC) technique. The proposed APODBN-SDC model intends to properly identify the existence of sarcasm or not. Primarily, the data preprocessing stage encompasses different subprocesses and the preprocessed data can be transformed into the feature vectors using TF-IDFs technique. Finally, the normalized feature vectors are given as input to the DBN model for the detection and classification of sarcasm. Finally, the hyperparameter tuning of the DBN is optimally tuned by the use of APO technique and shows the novelty of the work. The proposed method is tested against the benchmark dataset and the results reported the enhancements of the proposed model with maximum precision of 98.88% and recall of 98.90%. © *2022 SPIE and IS&T* [DOI: 10.1117/1.JEI.32.5.052302]

Keywords: sarcasm detection; deep learning; parameter tuning; natural language processing; sentiment analysis.

Paper 220705SS received Jul. 20, 2022; accepted for publication Oct. 11, 2022; published online Dec. 1, 2022.

1 Introduction

The fast development of microblogging websites like Twitter[1] has efficiently created a greater basis to comprehend the likes and dislikes of the community. Additionally, the public reply on specific incidents could also be estimated by this microblogging site. For instance, people share incidents regarding natural disasters, and become acquainted with the effect of recently released films[2] and the reviews based on recently manufactured products. Sarcasm is determined by a positively expressed declaration that possesses fundamental negative sentiments.[3,4] The author carried out research on documented constraint accuracy and human judges in recognizing sarcastic text. The major aim behind an ironic comment is to demean or mock separate people regarding an altered opinion for certain topics among more than two separate entities. Research by Gibbs reveals that ironic declaration is remembered more than the genuine and literal use of the similar expression expressed in a nonsarcastic content. Also, the research reveals that the memory is fully based on the intensity of beliefs and addresses opinion. The mining of text information to distinguish sarcasm has been a common technique for building a smart technology.[5–7] The method of sarcasm recognition is dependent on supervised learning algorithm. Sarcasm on a topic that possesses critical emotion for some set of individuals is regarded as an offensive behavior and that individual who is being sarcastic is noticeable for improper behaviors.[8] With the emergence of deep learning, current studies[9] leverage NNs for learning lexical and contextual characteristics, and eliminate the requirement for handcrafted features.[10] Since the DL-based approach achieves remarkable outcomes, they lack interpretability.

*Address all correspondence to Panneer Prabhavathy, pprabhavathy@vit.ac.in

The authors in Ref. 11 constructed an automated Twitter analyzer with the transfer learning approach to manage the unsupervised method. Then employed the linear support vector classifiers technique in this ML method, as well as the term frequency inverse document frequency (TF-IDF) process to handle the text information of Twitter. In Ref. 12, word embedding utilized for extracting features is incorporated with context-aware language model to offer automated feature engineering abilities and classification accuracy more than the fundamental model. The authors in Ref. 13 enhance the BERT model by fine-tuning intermediate task beforehand fine-tuning the target task. Especially, based on the correlation between emotions and sarcasm and (implied negative) sentiment, the study explored a TL architecture that employs sentimental classification and emotion recognition as single intermediate task to pervade knowledge into the targeted task of detecting sarcasm. Ding et al.[14] presented a multilevel late-fusion learning architecture with residual connection, a moderate dataset splitter, and two method variants according to distinct experiment sceneries.

Bhardwaj and Prusty[15] proposed a method for detecting sarcasm by using deep learning and machine learning. This method employs bidirectional encoder representation from the transformer (BERT) for preprocessing the sentence and feed toward a hybrid DL method for classification and training. This hybrid method employs LSTM and CNN. Xie et al.[16] developed a green artificial intelligence (AI) solution for a GreenSAP for gaining an effective consideration of altering public opinion on social networks. Especially, we proposed an effective ability of the pretrained Transformer encoder, and utilized various open-source SA data sets from domain and scenario to design a presented method.

This study introduces an efficient Anas Platyrhynchos Optimizer (APO) with deep belief network (DBN) based sarcasm detection and classification (APODBN-SDC) technique. Primarily, data preprocessing stage encompasses different subprocesses and the preprocessed data can be transformed into the feature vectors using TF-IDFs technique. Finally, the normalized feature vectors are given as input to the DBN model for the detection and classification of sarcasm. Finally, the hyperparameter tuning of the DBN is optimally tuned by the use of APO algorithm. The proposed method is tested against the benchmark dataset and the results are inspected in terms of different measures.

The rest of the paper is organized as follows. Section 2 introduces the proposed model and Sec. 3 validates the performance of the proposed model. Finally, Sec. 4 concludes the work.

2 Proposed Model

In this study, a novel APODBN-SDC model has been developed for sarcasm detection and classification. The presented APODBN-SDC model undergoes data preprocessing stage and TF-IDF-based feature vector. Finally, the normalized feature vectors are given as input to the DBN model for the detection and classification of sarcasm. Finally, the hyperparameter tuning of the DBN is optimally tuned by the use of APO algorithm. Figure 1 illustrates the block diagram of APODBN-SDC technique.

2.1 *Data Preprocessing*

Once the dataset was presented, the initial stage is to preprocess the textual data. Text preprocessing is the technique of cleaning the original textual data. A strong text preprocessing approach is crucial for application on natural language processing (NLP) process. Since every textual component obtained after preprocessing serves as the important component of input, i.e., fed into textual data application. Preprocessing involves various methods for translating the original text into a well-defined method: removal of stop-word, lemmatization, and lexical analyses (removal of punctuations, special symbols or characters, word tokenization, and ignores case sensitivity).

2.2 *Feature Extraction*

When the input dataset is preprocessed, the following phase is the feature extraction with the TF-IDF techniques. TF-IDF is a statistical value that defines the primary part of word to the

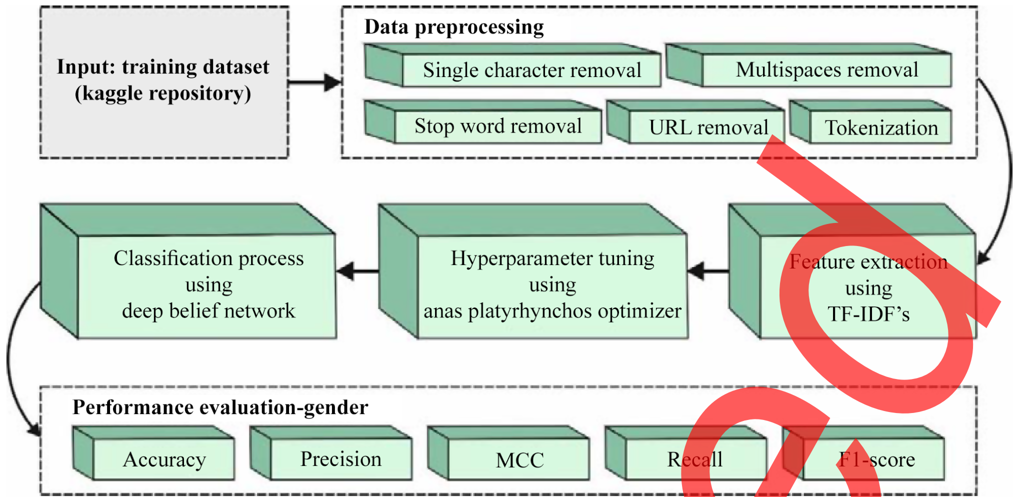

Fig. 1 Block diagram of APODBN-SDC technique.

document based on the word that displays for maximal time to provide more data based on the document; instantaneously they offer data or no details regarding the word. TF-IDF maximize the significance of the word with different iteration in a document and minimizes the importance of word that is looked at in different documents[17]

$$tf - idf = tf(t,f) \times idf(d,D). \tag{1}$$

An easier selection to select the term frequency $tf(t,f)$is to employ frequency $f(t,d)$, the amount of incidences t in a document d. When the length of document differs, then it is an improved recommendation to normalize the term frequency as long as documents consist of maximal occurrence of word. It can be upgraded as

$$tf(t,d) = 0.05 + \frac{0.5 \times f(t,d)}{\max\{f(w,d: w \in d\}}, \tag{2}$$

where Idf denotes the significance of a word. The $tf(t,f)$process each word as significant feature. Some words are general like "and" and "to," etc. which contain high frequency but provide smaller number of data. Therefore, $idf(t,d)$, Eq. (4.3), illustrates the significance of word that occurs in the document by categorizing the technique of document N by the amount of documents d in D that comprise the word t

$$idf(t,D) = \log \frac{N}{|\{d \in D: t \in d\}|}. \tag{3}$$

2.3 *DBN-based Classification*

During the classification process, the normalized feature vectors are given as input to the DBN model for the detection and classification of sarcasm. An artificial neural network (ANN) system is modeled by different layers of outputs and hidden unit that is called DL. It comprises two levels, such as

- Pretraining
- Fine-tuning phases

2.3.1 *Pretraining stage*

A DBN was applied that is crucial feed-forward network and deep structure in which the sample is derived from input to output layers via maximal amount of hidden nodes that have further nodes. According to the application of the DBN[18] method, the system generates activation function based on hidden unit, which distinguishes DBN. Furthermore, restricted Boltzmann machine (RBM) can be determined to resolve the problem of feasible activation function. An RBM is a kind of Markov subjective field that comprises individual layers of stochastic hidden units including the individual layer of stochastic visible units.

Step1: Initialize the clear unit ν for training the RBM vector

$$F(v,h) = -\sum_{k=1}^{K}\sum_{l=1}^{L} W_{kl} v_k h_l - \sum_{k=1}^{K} \alpha_k v_k - \sum_{l=1}^{L} \beta_l h_l, \tag{4}$$

where W_{kl} represents the symmetric transmission amongst visible layer $v_k$and the hidden layerh_l, α, βindicates the bias term, K, L indicates the amount of visible and hidden layers. The subordinate of the log possibility of a preparation vector concerns a weight that is irregularly simpler. From the hidden unit of RBM, it does not have unexpected replies that tend to attain impartial instance from $(V_k, h_l)_{\text{data}}$

$$\rho(h_l = 1|v) = \delta(\sum_{k=1}^{K} W_{kl} v_k + \alpha_l), \tag{5}$$

where $\delta(x)$denotes the logistic sigmoid function, $\frac{1}{(1+\exp(x))}$, v_k, h_l represent the unbiased sample.

2.3.2 *Updating process*

The hidden layer is upgraded and visible unit is considered as simultaneous from hidden and visible layer. It results in a complicated method as

$$\Delta W_{kl} \theta (v_k h_l)_{\text{data}} - (v_k h_l)_{\text{reconstruction}}, \tag{6}$$

where RBM endures training, a differing RBM might be stacked through a frame with multilayer approach. Generally, there is dissimilar RBM that is stacked and input visible unit was organized as quality and vector for unit that is efficiently placed RBM layer that shared by the application of shared model in existing weight and bias. Therefore, the finishing layer that is officially trained is protected to be in RBM. Thus, the attained DNN weight is occupied by the fine-tuning stage.

2.4 *Hyperparameter Optimization*

At the last stage, the hyperparameter tuning of the DBN is optimally tuned by the use of APO algorithm. Previously the warning performance, based on the anas platyrhynchos performance method, the population in APO was established as

$$\text{Pop}_i = \text{rand} \times (\text{up} - \text{low}) + \text{low}, \tag{7}$$

whereas Pop_i refers to the i'th population, N implies the population size, low and up are the lower and upper bounds of searching space, and rand refers to the arbitrary value chosen in the range of zero and one. During the warning performance, the fly in danger function recognized by chosen probability Pc was presented.[19] An essential procedure of warning performance was established as follows.

Step1: Compute the probability of distress Pc as

$$\text{Pc}_i = \frac{\text{rank}(\text{fit}(\text{Pop}_i))}{N}, \tag{8}$$

whereas $\text{fit}(\text{Pop}_i)$ refers to the fitness value of Pop_i and $\text{rank}(\text{fit}(\text{Pop}_i))$ has assumed as the rank of individuals Pop_i amongst the other individuals from the population.

Step2: When the probability Pc was fulfilled, create a novel individual as

$$\mathrm{Pop}_i(t+1) = \mathrm{Pop}_i(t) + \mathrm{sign}\left(\mathrm{rand} - \frac{1}{2}\right) \times \alpha_0 \times |\mathrm{Pop}_i(t) - \mathrm{Pop}_{\mathrm{best}}(t)| \times \mathrm{Levy}(s). \quad (9)$$

In which t denotes the existing iterations and $\mathrm{Pop}_{\mathrm{est}}$ stands for the leading duck, $\alpha_0 > 0$ refers to the step length scale factors, and sign denotes the sign function. Levy flight (LF) offers the arbitrary walks for Eq. (6), and their distribution formula is written as

$$\mathrm{Levy} \sim \mu = t^{-\lambda}, 1 < \lambda \leq 3. \quad (10)$$

LF is a different kind of arbitrary walk, and the probability distribution of their step length obeys a heavy-tailed distribution that is demonstrated as

$$s = \frac{\mu}{|\nu|^{\frac{1}{\beta}}}, \quad (11)$$

whereas s signifies the LF step lengths. In addition, in Eq. (4), $\lambda = 1 + \beta$, it can be utilized $\alpha_0 = 0.01$ and $\beta = 3/2$ as CS set. μ and ν are chosen in the normal distribution $\mu = N(0, \delta_\mu^2)$ and $\nu = N(0, \delta_v^2)$, whereas

$$\delta_\mu = \left[\frac{\Gamma(1+\beta)\sin\left(\frac{\pi\beta}{2}\right)}{\beta \times \Gamma\left(\frac{1+\beta}{2}\right) \times \frac{2(\beta-1)}{2}}\right]_{\delta_v=1}^{\frac{1}{\beta}}. \quad (12)$$

There is an intelligent technique to simulate biological migration. Related to the SOMA, APO mostly concentrated on anas platyrhynchos, simulates their migration movement and methods. It can be distinct in the place upgrade approach and technique flow in SOMA. During this performance of APO, an essential procedure was established as follows.

Step 1: Afterward, the optimum particle was determined, and another search particle is aimed at moving to optimum particles. The mathematical technique of this performance is provided as

$$\mathrm{Pop}_i(t+1) = \mathrm{Pop}_i(t) - A|C \times \mathrm{Pop}_{\mathrm{best}}(t) - \mathrm{Pop}(t)|, \quad (13)$$

whereas A and C signify the coefficient vectors that are attained as

$$A = 2a \times \mathrm{rand} - a, \quad (14)$$

$$C = 2 \times \mathrm{rand}. \quad (15)$$

In which a signifies the coefficient vector, which is linearly reduced with iterations. The value of a is attained as

$$a = 2 - t\frac{2}{T}. \quad (16)$$

In which T refers to the maximal amount of iterations. Figure 2 illustrates the flowchart of APO technique.

Step2: If the novel individual is worse than the old solution, choose another particle $\mathrm{Pop}_{\mathrm{and}}$ at arbitrary.

Step3: If $\mathrm{Pop}_{\mathrm{and}}$ is superior to Pop, the i'th individual transfers to arbitrary particle $\mathrm{Pop}_{\mathrm{and}}$ as

$$\mathrm{Pop}_i(t+1) = (\mathrm{Pop}_{\mathrm{rand}}(t) - \mathrm{Pop}_i(t)) \times e^{-l^2} + \mathrm{Pop}_i(t), \quad (17)$$

whereas l signifies the distance of arbitrary particle and i'th individual.

Step4: If $\mathrm{Pop}_{\mathrm{and}}$ is equivalent to Pop_i, it keeps unchanged.

Step5: If $\mathrm{Pop}_{\mathrm{rad}}$ is lesser than Pop_i, the arbitrary particle move to i'th individual by Eq. (17).

$$\mathrm{Pop}_{\mathrm{rand}}(t+1) = (\mathrm{Pop}_i(t) - \mathrm{Pop}_{\mathrm{rand}}(t)) \times e^{-l^2} + \mathrm{Pop}_{\mathrm{rand}}(t). \quad (18)$$

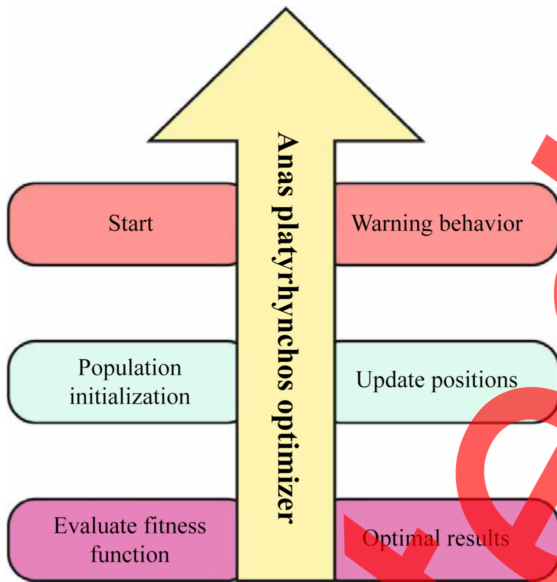

Fig. 2 Flowchart of APO technique.

3 Performance Validation

In this section, a brief sarcasm classification outcome of the APODBN-SDC model using benchmark Kaggle dataset.[20] The dataset comprises 14,949 samples under nonsarcasm (NSAM) class and 13552 samples under sarcasm (SAM) class. The proposed model is simulated using Python tool.

Figure 3 reports a pair of confusion matrices offered by the APODBN-SDC model on 70% of training set (TRS) and 30% of testing set (TSS). On 70% of TRS, the APODBN-SDC model has detected 10,355 samples under NSAM class and 8950 samples under SAM class. Besides, on 30% of TSS, the APODBN-SDC model has detected 4416 samples under NSAM class and 3876 samples under SAM class.

Table 1 and Fig. 4 highlight the overall sarcasm classification outcomes of the APODBN-SDC model on 70% of TRS and 30% of TSS. On 70% of TRS, the APODBN-SDC model has offered average $accu_y$, $prec_n$, $reca_l$, F_{score}, and MCC of 96.77%, 96.92%, 96.66%, 96.76%, and

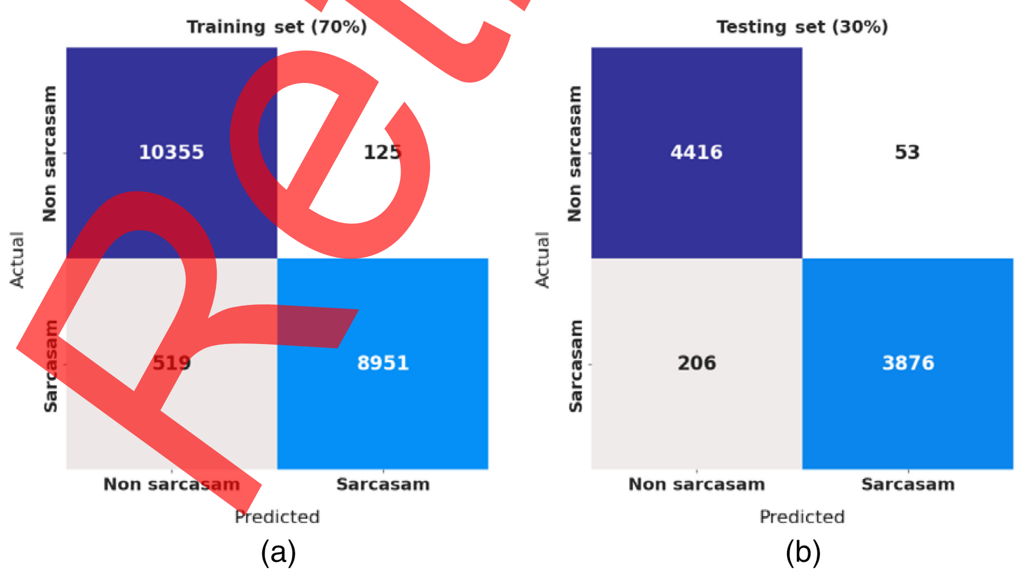

Fig. 3 Confusion of APODBN-SDC technique on TR/TS of 70:30 dataset. (a) Training Dataset (b) Testing Dataset.

Table 1 Result analysis of APODBN-SDC technique with distinct measures on TR/TS of 70:30 dataset.

Training/Testing (70:30)					
Class labels	Accuracy	Precision	Recall	F-score	MCC
Training phase					
Nonsarcasm	96.77	95.23	98.81	96.98	93.59
Sarcasm	96.77	98.62	94.52	96.53	93.59
Average	96.77	96.92	96.66	96.76	93.59
Testing phase					
Nonsarcasm	96.97	95.54	98.81	97.15	93.98
Sarcasm	96.97	98.65	94.95	96.77	93.98
Average	96.97	97.10	96.88	96.96	93.98

Fig. 4 Result analysis of APODBN-SDC technique on TR/TS of 70:30 dataset.

93.59%, respectively. Besides, on 30% of TSS, the APODBN-SDC model has offered average $accu_y$, $prec_n$, $reca_l$, F_{score}, and MCC of 96.97%, 97.10%, 96.88%, 96.96%, and 93.98%, respectively.

The training accuracy (TA) and validation accuracy (VA) attained by the APODBN-SDC model on TR/TS of 70:30 dataset is demonstrated in Fig. 5. The experimental outcomes implied that the APODBN-SDC model has gained maximum values of TA and VA. In specific, the VA is seemed to be higher than TA.

The training loss (TL) and validation loss (VL) achieved by the APODBN-SDC model on TR/TS of 70:30 dataset is established in Fig. 6. The experimental outcomes inferred that the APODBN-SDC model has accomplished least values of TL and VL. In specific, the VL is seemed to be lower than TL.

Figure 7 defines a pair of confusion matrices offered by the APODBN-SDC approach on 80% of TRS and 20% of TSS. On 80% of TRS, the APODBN-SDC model has detected 11,822

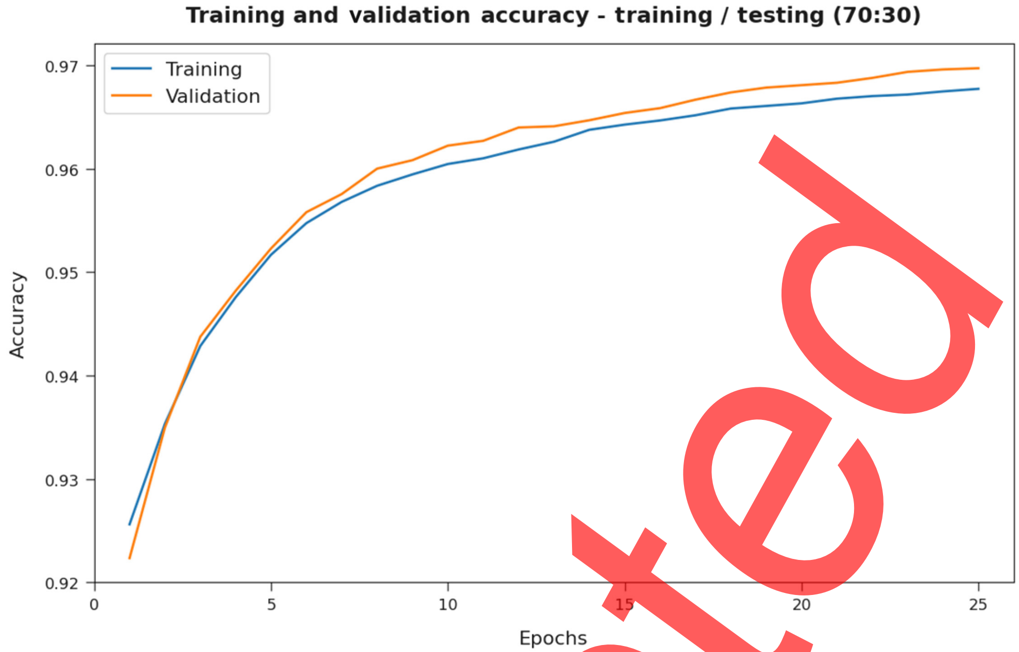

Fig. 5 TA and VA analysis of APODBN-SDC technique on TR/TS of 70:30 dataset.

Fig. 6 TL and VL analysis of APODBN-SDC technique on TR/TS of 70:30 dataset.

samples under NSAM class and 10709 samples under the SAM class. In addition, on 30% of TSS, the APODBN-SDC model has detected 2931 samples under NSAM class and 2707 samples under the SAM class.

Table 2 and Fig. 8 examine the overall sarcasm classification outcomes of the APODBN-SDC model on 80% of TRS and 20% of TSS. On 80% of TRS, the APODBN-SDC technique has offered average $accu_y$, $prec_n$, $reca_l$, F_{score}, and MCC of 98.82%, 98.81%, 98.83%, 98.82%, and 97.64% respectively. Moreover, on 20% of TSS, the APODBN-SDC model has offered average $accu_y$, $prec_n$, $reca_l$, F_{score}, and MCC of 98.89%, 98.88%, 98.90%, 98.89%, and 97.79% correspondingly.

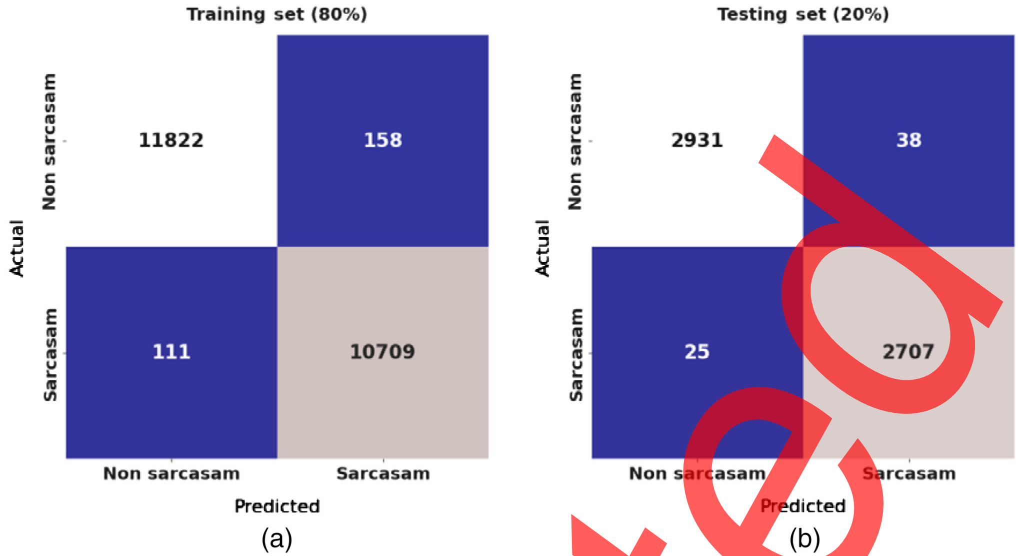

Fig. 7 Confusion of APODBN-SDC technique on TR/TS of 80:20 dataset.

Table 2 Result analysis of APODBN-SDC technique with distinct measures on TR/TS of 80:20 dataset.

Training/testing (80:20)					
Class labels	Accuracy	Precision	Recall	F-score	MCC
Training phase					
Nonsarcasm	98.82	99.07	98.68	98.88	97.64
Sarcasm	98.82	98.55	98.97	98.76	97.64
Average	**98.82**	**98.81**	**98.83**	**98.82**	**97.64**
Testing phase					
Nonsarcasm	98.89	99.15	98.72	98.94	97.79
Sarcasm	98.89	98.62	99.08	98.85	97.79
Average	**98.89**	**98.88**	**98.90**	**98.89**	**97.79**

The TA and VA attained by the APODBN-SDC model on TR/TS of 80:20 dataset are demonstrated in Fig. 9. The experimental outcomes implied that the APODBN-SDC technique has gained maximum values of TA and VA. In specific, the VA is seemed to be higher than TA.

The TL and VL achieved by the APODBN-SDC model on TR/TS of 80:20 dataset are established in Fig. 10. The experimental outcomes inferred that the APODBN-SDC model has accomplished least values of TL and VL. In specific, the VL is seemed to be lower than TL.

For ensuring the enhanced outcomes of the APODBN-SDC model, a comparative examination with recent models is shown in Table 3.[21,22] On measuring the results in terms of $prec_n$, the results revealed that the CNN-LSTM model has obtained lower $prec_n$ value of 67.30%. At the same time, the SIARN and MIARN models have attained slightly enhanced performance with closer $prec_n$ values of 71.70% and 72%, respectively. Along with that, the ELM-BiLSTM model has shown moderately improved $prec_n$ of 76.30%. Though the IMH-SA and IMLB-SDC models have resulted in reasonable $prec_n$ of 84% and 95.40%, the presented APODBN-SDC model has showcased superior $prec_n$ of 98.88%. The presented APODBN-SDC model has showcased superior $reca_l$ of 98.90%.

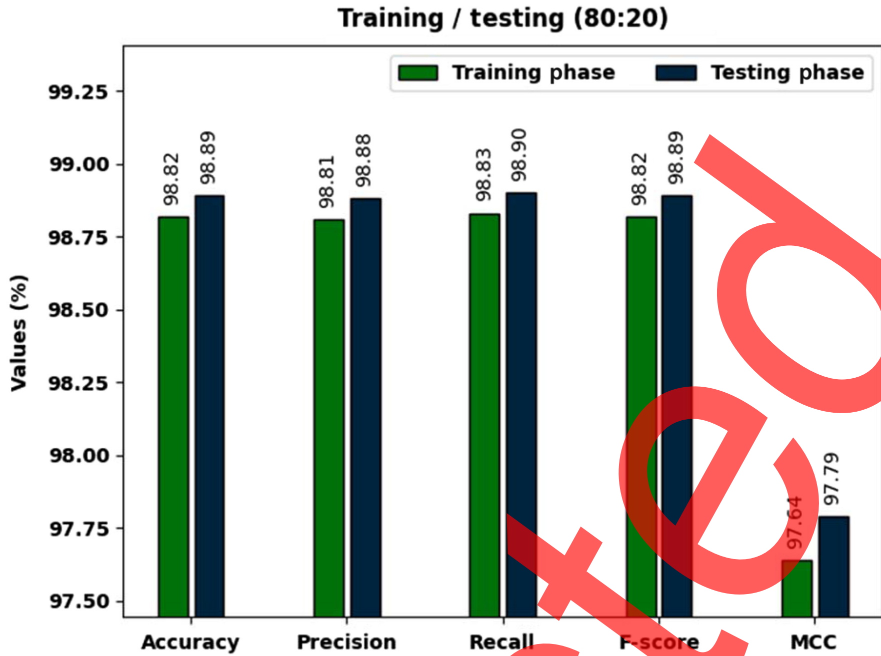

Fig. 8 Result analysis of APODBN-SDC technique on TR/TS of 80:20 dataset.

Fig. 9 TA and VA analysis of APODBN-SDC technique on TR/TS of 80:20 dataset.

Finally, the presented APODBN-SDC model has showcased superior F_{score} of 98.89%. The above-mentioned tables and results reported the enhanced sarcasm classification outcomes of the APODBN-SDC model.

4 Conclusion

An APODBN-SDC model has been developed for sarcasm detection and classification. The presented APODBN-SDC model undergoes data preprocessing stage which encompasses different subprocesses and the preprocessed data can be transformed into the feature vectors

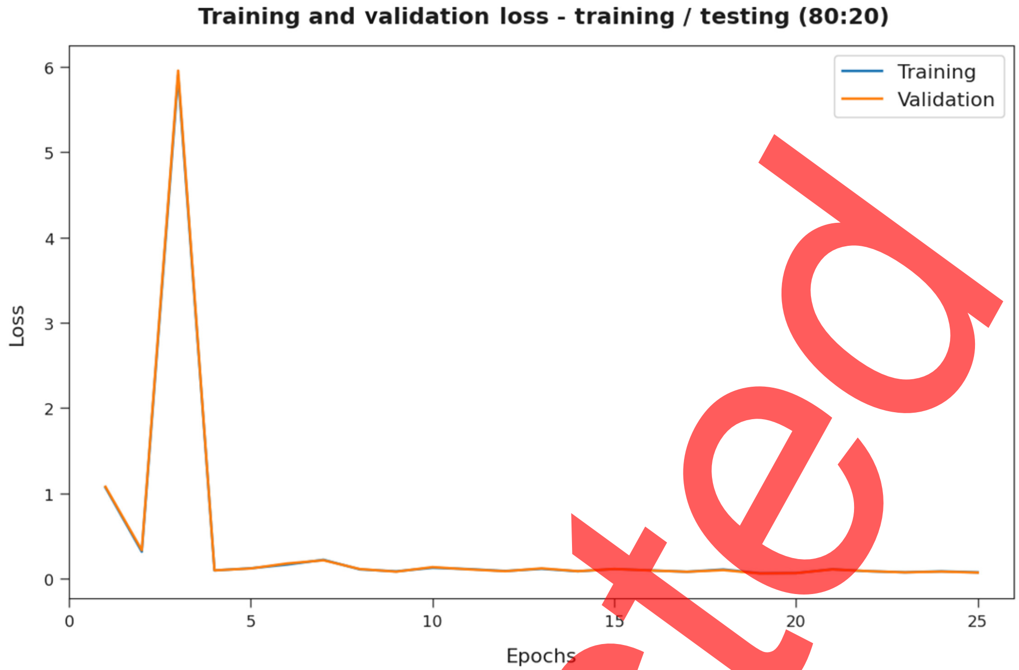

Fig. 10 TA and VA analysis of APODBN-SDC technique on TR/TS of 80:20 dataset.

Table 3 Comparative analysis of APODBN-SDC technique with recent approaches.

Models	Precision	Recall	F-Score
CNN-LSTM	67.30	70.60	70.10
SIARN	71.70	79.60	78.90
MIARN	72.00	73.40	74.60
ELM-BiLSTM	76.30	73.00	79.40
IMH-SA	84.00	80.70	84.80
IMLB-SDC	95.40	95.00	94.30
APODBN-SDC	98.88	98.90	98.89

using TF-IDFs technique. Finally, the normalized feature vectors are given as input to the DBN model for the detection and classification of sarcasm. Finally, the hyperparameter tuning of the DBN is optimally tuned by the use of APO algorithm. The proposed method is tested against the benchmark dataset and the results attained the maximum precision of 98.88% and recall of 98.90%. Thus, the APODBN-SDC model can be employed for the identification of sarcasm and nonsarcasm. In the future, hybrid DL models can be utilized for enhanced sarcasm detection results.

References

1. Z. Yin and F. You, "Multimodal sarcasm semantic detection based on inter-modality incongruity," *Proc. SPIE* **12168**, 501–505 (2022).
2. A. G. Prasad et al., "Sentiment analysis for sarcasm detection on streaming short text data," in *2nd Int. Conf. Knowl. Eng. and Appl. (ICKEA)*, October, IEEE, pp. 1–5 (2017).
3. D. K. Nayak and B. K. Bolla, "Efficient deep learning methods for sarcasm detection of news headlines," in *Machine Learning and Autonomous Systems*, J. I. Z. Chen et al., eds., pp. 371–382, Springer, Singapore (2022).

4. A. Kamal and M. Abulaish, "Cat-BIGRU: convolution and attention with bi-directional gated recurrent unit for self-deprecating sarcasm detection," *Cogn. Comput.* **14**(1), 91–109 (2022).
5. B. Moores and V. Mago, "A survey on automated sarcasm detection on Twitter," arXiv: 2202.02516 (2022).
6. A. Rustagi et al., "Toward sarcasm detection in reviews—a dual parametric approach with emojis and ratings," in *Soft Computing in Interdisciplinary Sciences*, S. Chakraverty, ed., pp. 245–257, Springer, Singapore (2022).
7. A. Joshi, P. Bhattacharyya, and M. J. Carman, "Automatic sarcasm detection: a survey," *ACM Comput. Surv.* **50**(5), 1–22 (2017).
8. M. S. M. Prasanna et al., "Sarcastic sentiment detection and polarity classification of tweets using supervised learning," in *Proc. Int. Conf. Recent Trends in Comput.*, Springer, Singapore, pp. 161–172 (2022).
9. M. I. Al-mashhadani, K. M. Hussein, and E. T. Khudir, "Sentiment analysis using optimized feature sets in different Facebook/Twitter dataset domains using big data," *Iraqi J. Comput. Sci. Math.* **3**(1), 64–70 (2022).
10. R. Srivastava, P. K. Bharti, and P. Verma, "A review on multipolarity in sentiment analysis," in *Inf. and Commun. Technol. for Competitive Strategies (ICTCS)*, Springer, Singapore, pp. 163–172 (2022).
11. A. Khare et al., "Sentiment analysis and sarcasm detection of indian general election tweets," arXiv:2201.02127 (2022).
12. P. Kumar and G. Sarin, "WELMSD–word embedding and language model based sarcasm detection," *Online Inf. Rev.* **46**(7), 1242–1256 (2022).
13. E. Savini and C. Caragea, "Intermediate-task transfer learning with BERT for sarcasm detection," *Mathematics* **10**(5), 844 (2022).
14. N. Ding, S. W. Tian, and L. Yu, "A multimodal fusion method for sarcasm detection based on late fusion," *Multimedia Tools Appl.* **81**(6), 8597–8616 (2022).
15. S. Bhardwaj and M. R. Prusty, "BERT pre-processed deep learning model for sarcasm detection," *Natl. Acad. Sci. Lett.* **45**, 203–208 (2022).
16. M. Xie et al., "A green pipeline for out-of-domain public sentiment analysis," in *Int. Conf. Adv. Data Mining and Appl.*, February, Springer, Cham, pp. 190–202 (2022).
17. W. Zhang, T. Yoshida, and X. Tang, "A comparative study of TF* IDF, LSI and multi-words for text classification," *Expert Syst. Appl.* **38**(3), 2758–2765 (2011).
18. N. Tazeen and K. S. Rani, "A novel ant colony based DBN framework to analyze the drug reviews," *Int. J. Intell. Syst. Appl.* **13**(6), 25–39 (2021).
19. Y. Zhang et al., "Novel swarm intelligence algorithm for global optimization and multi-UAVs cooperative path planning: Anas platyrhynchos optimizer," *Appl. Sci.* **10**(14), 4821 (2020).
20. Dataset, https://www.kaggle.com/rmisra/news-headlines-dataset-for-sarcasm-detection
21. R. Akula and I. Garibay, "Interpretable multi-head self-attention architecture for sarcasm detection in social media," *Entropy* **23**(4), 394 (2021).
22. D. Vinoth and P. Prabhavathy, "An intelligent machine learning-based sarcasm detection and classification model on social networks," *J. Supercomput.* **78**(8), 10575–10594 (2022).

Vinoth Dakshnamoorthy completed his BE degree in computer science in Sri Balaji Chockalingam Engineering College, Arani, and his ME degree in Sathyabama University, Chennai. Currently pursuing his PhD in Vellore Institute of Technology, Vellore, Tamil Nadu, India. His areas of research include soft computing, social media analysis, and security etc.

Panneer Prabhavathy is working as an associate professor in the School of Information Technology and Engineering, VIT University, Vellore. Her research area includes computational intelligence, data mining, data science, machine learning, and deep learning. She has published around 20 journal papers in her research field. She is a life member of CSI and IEEE. She is also part of various school activity committees. She has published number of papers in international conferences. She had completed funded project from ISRO, SAC, Ahmedabad, India.