Journal of Intelligent Systems and Internet of Things
Vol. 16, No. 01, PP. 233-244, 2025
DOI: https://doi.org/10.54216/JISIoT.160120

Natural Language Processing Driven Applied Linguistics for Sarcasm Detection Using Artificial Hummingbird Algorithm with Deep Learning

Maryam Alsolami[1,*]
[1]Department of Computer Science and Artificial Intelligence, College of Computing, Umm Al-Qura University, Makkah, Saudi Arabia
Emails: mksolami@uqu.edu.sa

Abstract

Natural Language Processing (NLP)-driven applied linguistics for sarcasm detection includes computational models to understand and identify sarcastic expressions within text. This interdisciplinary method integrates linguistics principles with advanced NLP techniques to identify subtle and nuanced cues indicative of sarcasm correctly. It includes computational approaches like linguistic feature extraction, machine learning models, and sentiment analysis. Furthermore, deep learning (DL) algorithms, including transformers and recurrent neural networks (RNNs), hold significant potential in capturing complex linguistic nuances inherent in sarcastic expression. These approaches can learn the hierarchical representation of text, which enables capturing context dependency, which is crucial for accurately detecting sarcasm. The applications of NLP-driven applied linguistics for sarcasm detection show great potential in various domains namely social media analysis, online content moderation, and customer feedback interpretation. By automating sarcasm detection, this system can enhance communication understanding, improve sentiment analysis accuracy, and contribute to better decision-making processes in various contexts. This study develops automated Sarcasm Detection using the Artificial Hummingbird Algorithm with Deep Learning (ASD-AHADL) technique. The ASD-AHADL technique applies the optimal DL model for detecting sarcastic content. To achieve this, the ASD-AHADL technique undergoes data preprocessing and the BERT-based word embedding process at the initial stage. Followed by the ASD-AHADL technique uses attention-gated recurrent unit long short-term memory (AGRU-LSTM) for the sarcasm detection process. At last, the AHA-based parameter tuning process is involved to fine-tune the parameters based on the DL algorithm. The experimental study of the ASD-AHADL technique has been tested under a social media dataset. The outcomes indicated that the solution of the ASD-AHADL technique was significant compared to others.

Received: November 12, 2024 Revised: January 17, 2025 Accepted: February 18, 2025

Keywords: Natural Language Processing; Deep Learning; Sarcasm Detection; Artificial Hummingbird Algorithm; Sentiment Analysis

1. Introduction

Sarcasm is extensively utilized every day through social media networks such as Facebook and X (Twitter) [1]. Sarcasm is a turn of expression for sharing bad feelings, disregard, and comic by the linguistic creation. It is a category of fake respect employed to surge anger unintentionally. Sarcasm could have appeared to thinly hidden maliciousness [2]. The sarcastic tags and commentaries have been majorly toward celebrities and political groups as they are considered influencers. Sarcasm has a connection to psychological characteristics, namely depression and anxiety [3]. The study shows that teasers feel their reports are less difficult than their victims are but more destructive in real time. Sarcasm can simply be identified in one-to-one communication by noting the speaker's gestures, tone, and facial expressions [4]. Nevertheless, recognizing sarcasm is challenging since no indicators will easily exist in written communication. Recognizing the sarcastic commentaries for videos, text, or images transferred through social media networks is still highly complex as context belongs to the image key comments texts or headlines [5]. Sarcasm recognition in virtual interactions from e-commerce, discussion media, and social media platforms becomes vital for opinion mining, sentiment analysis (SA), fake news detection, and identifying cyberbullies and online trolls. Nowadays, identifying sarcasm is a major research field [6].

Due to a large quantity of data, organizations have a tremendous opportunity to learn more about individuals' feelings, viewpoints, and other components, but there are also numerous complexities [7]. For example, sarcasm normally employs positive language, and then the contexts will be transferred to the negative approaches. These small problems are affected by improper assessment of product reviews or inappropriate categorization in false news recognition [8]. Researchers and businesses are now gaining precise information from text data, comprising statements created sarcastically, arising from these problems. Various Natural language processing (NLP) techniques are provided, and sarcasm detection is considerably trained for such methods. The understanding of words supports learning the contextual module. Various techniques have numerous temporal proximities and are strongly related to training contexts [9]. Implementing the machine learning (ML) method can generate efficient outcomes for identifying sarcasm. Making an efficient classification technique dependent upon numerous factors. The main aspect is the attributes employed and independent features in the learning model that must be easily combined with the class model [10]. With deep learning (DL) development, a neural network (NN) is employed to learn lexical and context features and it removes the essentials for handcrafted features. Once DL-based techniques achieve notable results, they have a shortage of interpretability.

This study bridges the gap between theory and practice using an Artificial Hummingbird Algorithm with Deep Learning (ASD-AHADL) to develop an automated Sarcasm Detection technique. The ASD-AHADL technique applies the optimal DL model for detecting sarcastic content, increasing the accuracy and efficiency of detecting sarcasm in written and spoken language. To accomplish this, the ASD-AHADL technique undergoes data preprocessing and the BERT-based word embedding process at the initial phase. Followed by the ASD-AHADL technique uses attention-gated recurrent unit long short-term memory (AGRU-LSTM) for the sarcasm detection. Finally, the AHA-based parameter fine-tuning process is required to modify the parameters based on the DL algorithm. The experimental study of the ASD-AHADL technique has been tested under a social media dataset.

2. Related Works

In [11], a multimodal graph contrastive learning approach was developed to combine and differentiate the sarcastic signs for visual and textual methods. Particularly, the system initially employs object detection to raise the vital visual regions accompanied by their image descriptions. Moreover, the technique utilizes optical character identification for extraction. Subsequently, a multi-modal graph was made for every sample to model the complex sarcastic relationships between modes. Rosid et al. [12] presented the Hybrid Pretrained Word Embedding method. The developed method integrates two predominant pre-trained word-embedding methods: Fasttext and Glove. Fasttext could be used for extracting semantic context for Indonesian words, although Glove has been employed to remove the semantic context vectors in English words. The classification method leverages the DL method. In [13], a DL method was developed. The technique developed a Convolution and Attention with BiGRU (CAT-BiGRU) method, encompassing input, two considerations embedding, BiGRU, and convolutional layers. The convolution layer removed the SDS-associated syntactic and semantic features in the embedding layer, the Bi-GRU layer got context data, and attention layers are deployed for recovering the SDS-based wide-ranging context in the input texts. In conclusion, the sigmoid function was implemented for classification purposes.

Fu et al. [14] developed a hierarchical architecture in this study. This technique employs sentiment dictionaries followed by integrating them with every modality. Additionally, for extracting the combined semantic data inferred in the modality and increasing dimension of emotional unpredictability, the sensitive data acquired by combining every mode’s data will be integrated to the sentimental vector. Then, the sarcasm was identified by combining lower-level data at the cross-modal fusion layer. Madani et al. [15] presented a two-stage model employing NLP and ML methods. In the primary stage, 2 novel architectural features with alternative key features have been removed from news instances. During secondary stage, a hybrid technique dependent upon program approach, comprising statistical data, and a k-nearest neighbor (K-NN) method has been presented for increasing the effectiveness of DL methods.

In [16], a hierarchical fusion system that incorporates sentiment data was presented. Mainly, the technique utilized feature-object matching in the image modalities. Sentiment information has been removed from every mode and integrated to accomplish a further wide-ranging model with modalities. Likewise, the method describes the connections of inter-modal inconsistencies employing a cross-modal Transformer. The sentiment-aware image-text contrastive loss method was also exploited for synchronizing the text and images semantics. In [17] a stance-level sarcasm detection (SLSD) technique was developed. The technique designed a fundamental architecture that encompasses BERT and innovative stance-centered graph attention network (SCGAT) methods. Particularly, the BERT has been utilized for capturing the sentence illustration, and SCGAT was developed to take the data in particular objective.

3. The Proposed Method

In this work, we have developed an ASD-AHADL technique. The ASD-AHADL technique applies optimal DL model for the detection of sarcastic content. To accomplish this, the ASD-AHADL technique comprises different kinds of sub-processes involved as data preprocessing, BERT-based word embedding, AGRU-LSTM-based sarcasm detection, and AHA-based hyperparameter tuning. Fig. 1 illustrates the working flow of ASD-AHADL model.

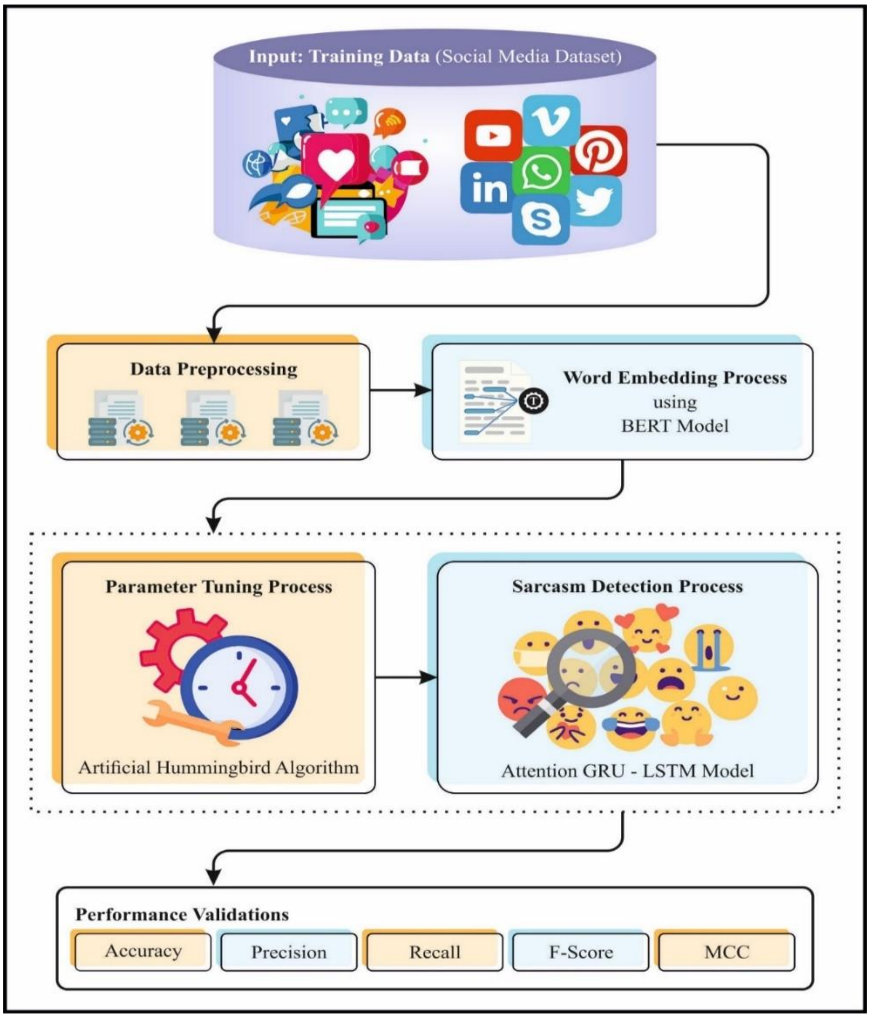

Figure 1. Workflow of ASD-AHADL technique

A. Data preprocessing

Initially, the ASD-AHADL method undergoes data preprocessing and BERT-based word embedding process. The pre-processed phase comprises cleaning and removing noises in the databases beforehand the transformation and representation techniques can be performed. The text processing aims to transform and express the social media post from the approach, which is analyzed and classified using DL algorithms. However, the database requires that they be cleaned using the succeeding stages:

- Remove the punctuations (e.g., ?!:; ‘) in all the social media posts to appear cleaner. “The,” “a,” “an,” and “in” are the stop words removed from the databases.
- Transform all the words from the text to lowercase text.
- Eliminate redundant words, emojis, characters, white space, and digits in the social media posts from the database.
- Tokenization is used to break sentence down into its integral parts such as words, phrases, etc.
- DL-NN approach is used to identify the social media post as cyberbullying or not, each text order from the database has real-value vector0. The post-padding series is used for completing these tasks.

BERT is an NLP preprocessing model based on neural network [18]. BERT can be used as a substitute for $word2vec$ and is more advantageous and flexible than the classical language model. $word2vec$ denotes no words changes with context and in the fixed form. Specifically, the representation has no connection with the context where the word appear, however owing to the complicated semantic features of natural language, the similar word meaning in diverse contexts might vary, hence using $word2vec$ for the representation of word vector may decrease the accuracy of downstream task. It has new records in eleven NLP tasks, including QQP, MultiNLI, etc. In addition, the word vector attained through the BERT has high-level features. The word vector is inputed with high-level features may accomplish improved outcomes in downstream tasks.

B. Sarcasm Detection using AGRU-LSTM Model

Next, the ASD-AHADL technique uses AGRU-LSTM for the sarcasm detection process. The information is transmitted to the GRU layer as a normalized window [19]. During this phase, the amount of GRU later blocks can be equivalent to the feature counts. An input of GRU layer can vectors attained by sliding windows and the output was computed by the subsequent formulas.

$$z_t = o\left(w_s x_t + U_s h(t-1)\right) + b_Z$$

$$r_t = o(w_\gamma x_L + U_\gamma h(t-1) + b_r$$

$$h^{|} = o(w_h x_L + U_n\left(r_t O h(t-1)\right) + b_h$$

$$h_t^t = (1-z)h(t-1) + z_t h_t^J \tag{1}$$

Whereas z_t implies the update gate, r_t signifies the rest gate, $h^j t$ denotes the candidate gate, and h_t indicates the resultant activation. W_Z, W_R, W_h, U_Z, U_R, UN defines the learnable matrixes, b_n, b_s, b_r represents the learnable biases, 0 stands for the sigmoid activation function, and $\odot$ represents the element-by-element multiplication.

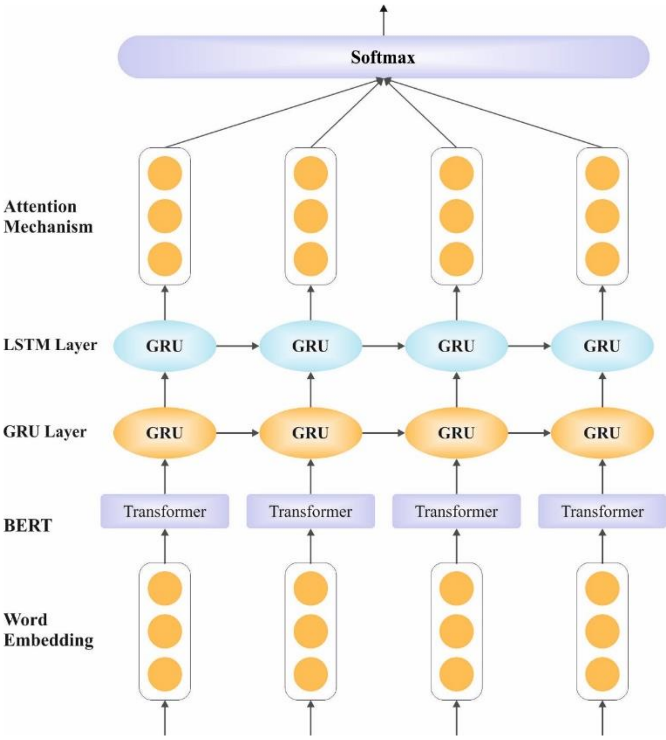

Figure 2. Structure of AGRU-LSTM

The following step in the presented system is to send the resultant of GRU layer as input to LSTM layer. The LSTM input in step T is the vector $X \in R^E$.

$$f^{<k>} = o\left(W_f r^{<k>} + U_f h^{<k-1>} + b_f\right)$$

$$q^{<\kappa>} = \sigma(W_i r^{<k>} + U_i h^{<k-1>} + b_i)$$

$$g^{<k>} = tanh\left(W_g r^{<k>} + U_g h^{<k-1>} + b_g\right) \tag{2}$$

$$o^{<k>} = \sigma(W_o r^{<k>} + U_o h^{<k-1>} + b_0)$$

Whereas i implies the input gate, 0 represents the output gate, f denotes the forget gate, and g stands for the update gate, $[W_i, W_f, W_q t, W_o, b_i, b_f, b_{qt}, b_o]$ defines the group of parameters that are learned. $q^{<k>}$ is upgraded by the next connection:

$$q^{<k>} = f^{<k>} \odot q^{<k-1>} + i^{<k>} \odot g^{<k>} \quad (3)$$

Whereas, $\odot$ defines the element-by-element product between 2 vectors.

At last, the activation of cell is attained by the succeeding connection:

$$h^{<k>} = o^{<k>} \odot \ tanh\ (q^{<k>}) \quad (4)$$

To define the class of every point, it can be appropriate to execute a sigmoid to encoder stage outcome as follows:

$$z = sigmoid\ (W_s E + b_s) x \quad (5)$$

While processing an enormous data quantity, the global attention model has been included to adjust the weight of important feature data and extract essential features, to boost the prediction outcomes and reinforce the impact of key features on the degradation trend [20]. The computation formula of target state information, attention weight vector and context vector are given below:

$$x_t = h_s^T h_t \quad (6)$$

$$a_t = \frac{exp(x_t)}{\sum_{j=0}^{t} e\ xp(x_j)} \quad (7)$$

$$y_t = softmax\left(W_s \widetilde{h_t}\right) \quad (8)$$

$$c_t = \sum_{t=0}^{T} a_t\, h_t \quad (9)$$

Where the weight matrix of global attention training is represented as W_s. Fig. 2 demonstrates the structure of AGRU-LSTM.

C. Hyperparameter Tuning using AHA

At last, the AHA-based parameter fine-tuning process is involved to change the parameters related to the DL method. In the year 2022, the authors developed the AHA as a nature-inspired swarm-based metaheuristic optimization method [21]. AHA is based on the combat and foraging strategies of hummingbirds (HBs). The explanations and mathematical model of HB foraging and fighting abilities are in this section. Next, the technique contributed to resolving the ODNR and CBS issues in AND can be discussed.

1. Modelling of foraging approaches of hummingbird

The foraging approaches of HBs have been demonstrated in three various ways territorial, guided, and migration foraging. As same to every swarm-based method, AHA starts with the randomly generated population in candidate outcome (x_{ii}) by applying their upper and lower bounds. For n-HBs allocated to d food sources, the population can be initialized arbitrarily by,

$$x_{ij} = l_j + r \times \left(u_j - l_j\right) i = 1, \dots, n\ and\ j = 1, \dots, d \quad (10)$$

Here d describes number of dimensions and n denotes the total population or candidate outcomes, correspondingly; r means a uniform distribution random number, u_j defines the upper bounds, andl_j represents lower bounds of the jth candidate, correspondingly.

During the initial population, firstly, find the fitness values, fittest population, and fitness function (FF).

Then, it will be considered that every HB is proficient in recognizing its food source with the help of memory. A visit table will be made in AHA, as followed by,

$$vt_{ij} = \begin{cases} 0 & if\ i \neq j \\ null & if\ i = j \end{cases} i = 1, \dots, n\ and\ j = 1, \dots, d \quad (11)$$

whereas $vt_{ij} = 0$ for $i \neq j$ specifies that the ith HB to the jth food source with the iteration; $vr_{ii} = null$ for $i = j$ pointed out that an HB is to be eaten from a given food source.

2. Modeling of guided foraging

In order to gain the major liquid from the food source, it should be significant nectar-replenishing supremacy and neglected by the HB for a long time. It can identify what kind of food sources obtain the majority of visiting for directed foraging strategy, and choose any one that takes the liquid refills the aimed food sources. Next, to identify the food sources, the HB is flying. For purposes of this, the combative abilities of HBs could be sampled into 3 diverse approaches in the AHA as axial, diagonal, and omnidirectional fights by presenting the switch vector. The omnidirectional, diagonal, and axial combat abilities of HBs have been represented in Eq. (14).

$$S_{(i)} = \begin{cases} 1 \; if \; i = randi(1, d) \\ 0 \; else \end{cases} i = 1, \dots, d \tag{12}$$

$$S_{(i)} = \begin{cases} 1 \; if \; i = p(j), j \in [1, k], p = randperm(k), k \in [2, r_1.(d-2)+1] \\ 0 \; else \end{cases} \tag{13}$$

$$S_{(i)} = 1 i = 1, \dots, d \tag{14}$$

By implementing these combat capabilities, HBs will detect the food sources in a multi-dimensional search vector. The calculated formula to simulate the fixed foraging behavior and possible food sources will be represented:

$$v_i(t+1) = x_i^*(t) + g_f.S.[x_i(t) - x_i^*(t)] = 1, \qquad g_f \sim N(0,1) \tag{15}$$

Now $x_i^*(t)$ represents the preferred target food sources, $x_i(t)$ denotes existing food source locations by ith HB in time t, g_f refers to a directed parameter with normal distribution $mean = 0$ and $srd = 1$. By employing various fighting abilities, the position of ith food source can be upgraded by,

$$x_i(t+1) = \begin{cases} x_i(t) & for \; f(x_i(t)) \leq f(v_i(t+1)) \\ v_i(t+1) & for \; f(x_i(t)) > f(v_i(t+1)) \end{cases} \tag{16}$$

Whereas $f(\cdot)$ specifies fitness function value. Eq. (16) states that once the nectar-refilling rate of candidate food source is greater than the existing one, the HB ends the existing food source and provides candidate food sources denoted in Eq. (15).

3. Modelling of territorial foraging

Later an HB feeds honey from the flower; it is highly possible to appear for novel food sources compared with previous food sources. When the existing food sources are not enough for the HB, it will be moved to alternative place in its region and search for an enhanced one. The territorial foraging approach of HBs and potential food sources have been demonstrated in the mathematical form:

$$v_i(t+1) = x_i(t) + t_f.S.x_i(t), \qquad t_f \sim N(0,1) \tag{17}$$

Now t_f is a territorial parameter created same as g_f. Eq. (17) will support any HB to rapidly detect a novel food source in its community due to its distinctive combative abilities. It must be altered later the territorial foraging approach will be applied.

4. Modelling of migration foraging

When an HB's normal serving position goes nearby food, it moves to a faraway feeding place. The AHA has described the migration coefficient. Once the count of iterations surpasses the coefficient of migration, HBs move to an arbitrarily produced food source within the search range. The HB can unrestraint the previous source and provide the new source that would be updated in the visiting table. The population movement of an HB forage from single food source to another will be shown in given below

$$x_w(t+1) = l + r \times (u - l) \tag{18}$$

Here x_w denotes the food sources with poorer nectar-refill levels. The population number and migration coefficient will be correlated by $m = 2n$.

The AHA derives an FF to attain enhanced classifier outcomes. It defines the positive integer to characterize the better performance of the solution candidate. Now, the minization of the classifier rate of error has been implicit as the FF.

$$fitness(x_i) = ClassifierErrorRate(x_i) = \frac{No.of \; misclassified \; samples}{Total \; No.of \; samples} * 100 \tag{19}$$

4. Experimental Validation

The stimulation outcomes of the ASD-AHADL technique has been tested under social media dataset containing 2 datasets such as Twitter and dialogues datasets [23]. The Twitter dataset comprises 1956 samples and dialogues dataset involves 4692 samples under 2 classes as illustrated in Table 1.

Table 1: Details of Twitter and Dialogues datasets

Class	Twitter Dataset	Dialogues Dataset
Sarcastic	308	2346
Non Sarcastic	1648	2346
Total Samples	1956	4692

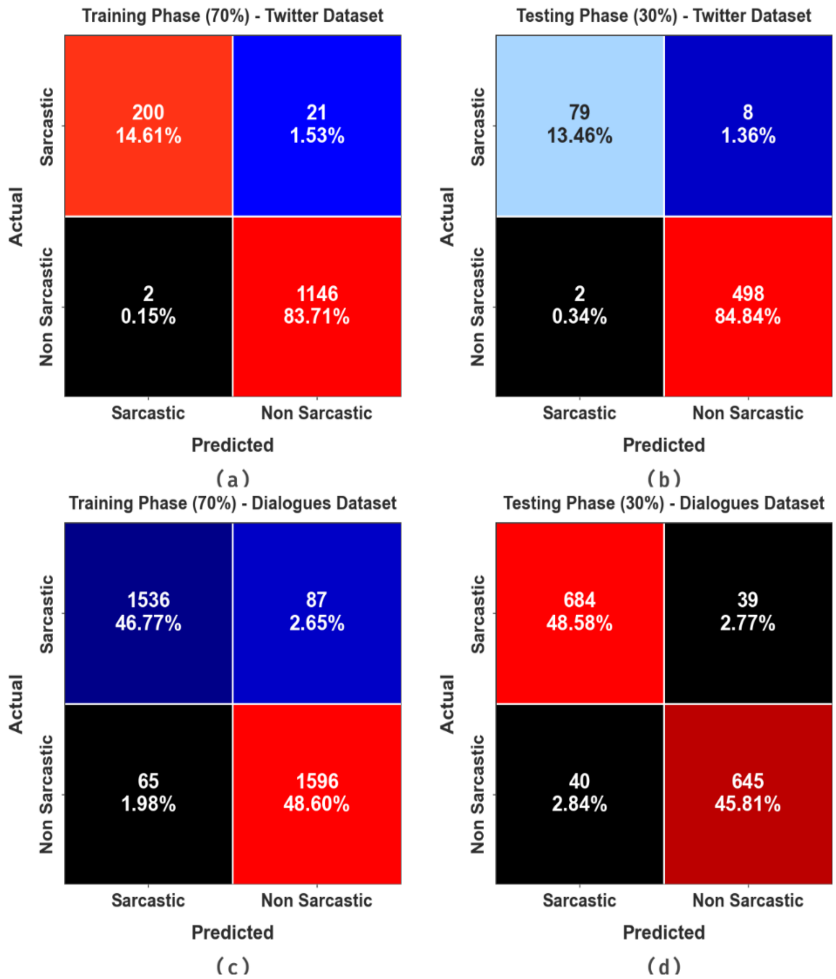

Figure 3. Confusion matrices under 70:30TRAS/TESS (a-b) Twitter dataset and (c-d) Dialogues dataset

Fig. 3 represents the confusion matrices acquired by the ASD-AHADL method at Twitter and Dialogues datasets. These results pointed out that the ASD-AHADL approach has effectual recognition with sarcastic and non-sarcastic classes correctly.

Table 2 and Fig. 4 examine the sarcasm detection outcomes of ASD-AHADL method in Twitter dataset. These experimental values have shown that the ASD-AHADL algorithm appropriately recognized the sarcastic and non-sarcastic classes. Based on 70%TRAS, the ASD-AHADL method gives the $accu_y$, $prec_n$, $reca_l$, F_{score}, and AUC_{score} of 98.32%, 98.61%, 95.16%, 96.78%, and 93.70%. Meanwhile, with 30%TESS, the ASD-AHADL system accomplishes the $accu_y$, $prec_n$, $reca_l$, F_{score}, and AUC_{score} of 98.30%, 97.97%, 95.20%, 96.53%, and 93.14%, respectively.

Table 2: Sarcasm detection outcome of ASD-AHADL model under Twitter dataset

Classes	$Accu_y$	$Prec_n$	$Reca_l$	F_{score}	MCC
TRAS (70%)					
Sarcastic	98.32	99.01	90.50	94.56	93.70
Non Sarcastic	98.32	98.20	99.83	99.01	93.70
Average	98.32	98.61	95.16	96.78	93.70
TESS (30%)					
Sarcastic	98.30	97.53	90.80	94.05	93.14
Non Sarcastic	98.30	98.42	99.60	99.01	93.14
Average	98.30	97.97	95.20	96.53	93.14

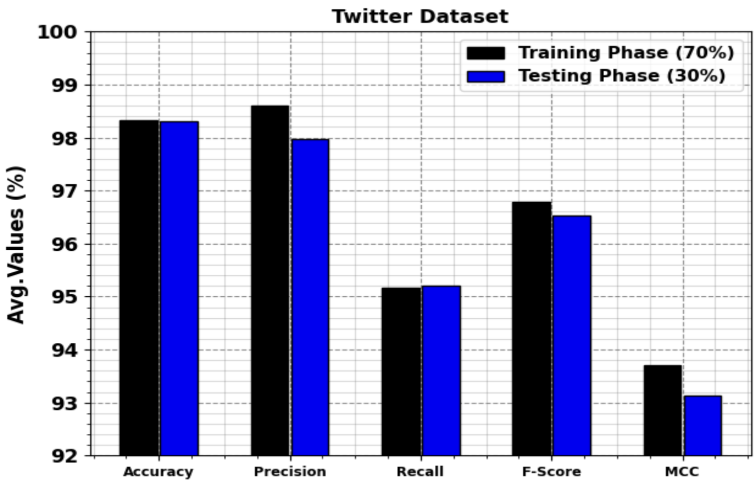

Figure 4. Average of ASD-AHADL system at Twitter dataset

Investigative the PR analysis, as represented in Fig. 5, the outcomes confirmed that the ASD-AHADL method at Twitter dataset gradually achieves increased values of PR at each 2 classes. It proves the enhanced capabilities of the ASD-AHADL technique in the detection of diverse classes, presenting ability in the identification class labels.

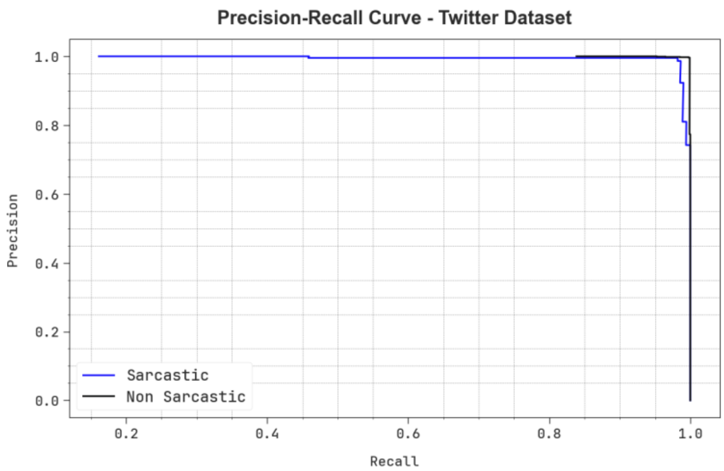

Figure 5. PR curve of ASD-AHADL methodology on Twitter dataset

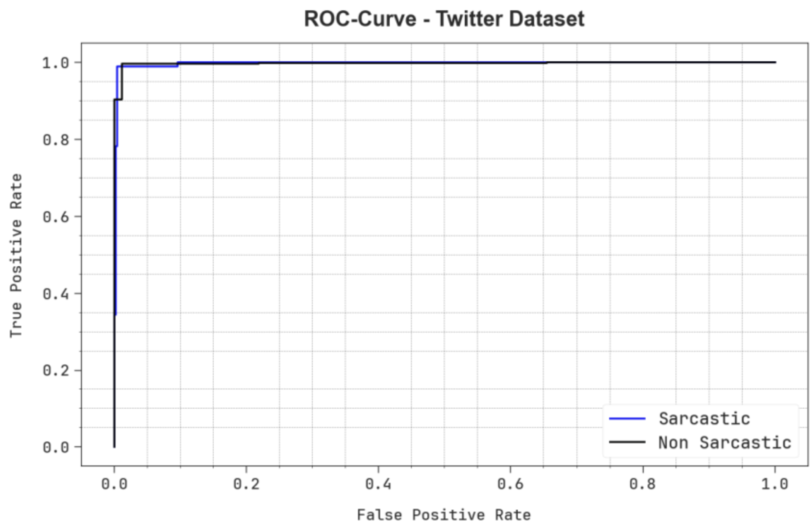

Figure 6. ROC curve of ASD-AHADL model under Twitter dataset

Likewise, in Fig. 6, ROC inspection completed through the ASD-AHADL model with Twitter dataset exceeded the classification of various labels. It offers complete knowledge of the tradeoff between FRP and TPR at distinct recognition epoch counts and threshold values. This figure underscored the boosted classification outcomes of the ASD-AHADL system with each class labels, showing the efficiency in overcoming several classification complexities.

Table 3 demonstrate the comparative assessment of the QSSA-MLHAR method under Twitter dataset [22]. These achieved experimentation outcomes emphasized that the NBOW and ELMo-BiLSTM techniques obtained minimum values respectively. Meanwhile, the Fracking Sarcasm, A2Text-Net, IMHSAA, and DLNLP-SA systems get reasonably closer values. However, the QSSA-MLHAR algorithm exhibits higher performance with$accu_y$, $prec_n$, $reca_l$, and F_{score} of 98.30%, 97.97%, 95.20%, and 96.53%.

Table 3: Comparative outcomes of QSSA-MLHAR method with other systems under Twitter dataset

Twitter Dataset				
Methods	$\boldsymbol{Prec_n}$	$\boldsymbol{Reca_l}$	$\boldsymbol{Accu_y}$	$\boldsymbol{F_{score}}$
NBOW	70.87	62.03	64.26	64.18
ELMo-BiLSTM	78.04	73.72	77.68	75.05
Fracking Sarcasm	88.04	87.63	88.57	88.29
A2Text-Net	91.59	91.06	91.49	89.65
IMHSAA	95.19	94.47	95.40	94.76
DLNLP-SA	97.06	92.43	97.61	94.99
ASD-AHADL	97.97	95.20	98.30	96.53

Table 4 and Fig. 7 display the sarcasm detection outcomes of ASD-AHADL system in Dialogues dataset. These experimental values pointed out that the ASD-AHADL technique correctly recognized the sarcastic and non-sarcastic classes. According to 70%TRAS, the ASD-AHADL algorithm gains the $accu_y$, $prec_n$, $reca_l$, F_{score}, and AUC_{score} of 95.36%, 95.39%, 95.36%, 95.37%, and 90.75%. Finally, based on 30%TESS, the ASD-AHADL approach gives the $accu_y$, $prec_n$, $reca_l$, F_{score}, and AUC_{score} of 94.38%, 94.39%, 94.38%, 94.38%, and 88.77%.

Table 4: Sarcasm detection outcomes of ASD-AHADL model under Dialogues database

Class	$Accu_y$	$Prec_n$	$Reca_l$	F_{score}	MCC
TRAS (70%)					
Sarcastic	94.64	95.94	94.64	95.29	90.75
Non Sarcastic	96.09	94.83	96.09	95.45	90.75
Average	95.36	95.39	95.36	95.37	90.75
TESS (30%)					
Sarcastic	94.61	94.48	94.61	94.54	88.77
Non Sarcastic	94.16	94.30	94.16	94.23	88.77
Average	94.38	94.39	94.38	94.38	88.77

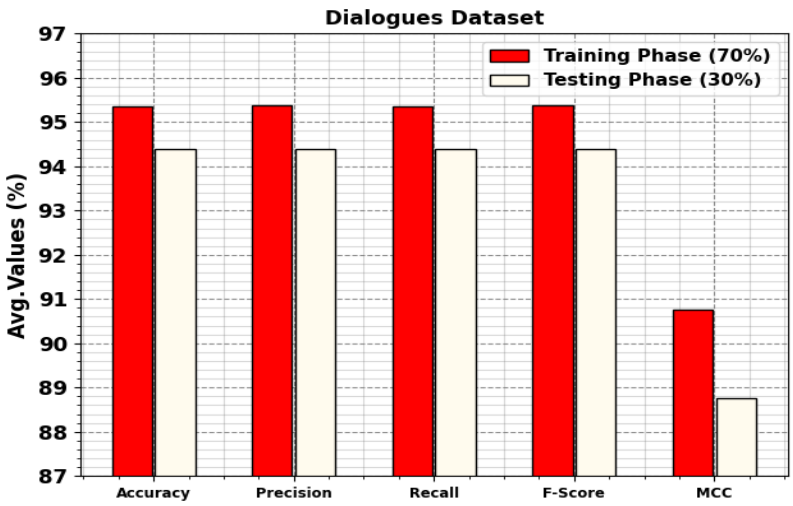

Figure 7. Average of ASD-AHADL model with Dialogues dataset

Table 5 exemplify the comparison assessment of the QSSA-MLHAR system on Dialogues dataset. These accomplished values denoted that the Attention-LSTM, SIARN, and MIARN algorithms get diminished values. Moreover, the ELMo-BiLSTM, IMHSAA, and DLNLP-SA methods acquire somewhat considerable values. But the QSSA-MLHAR approach shows maximized performance with $accu_y$, $prec_n$, $reca_l$, and F_{score} of 95.36%, 95.39%, 95.36%, and 95.37%, respectively.

Table 5: Comparative outcomes of QSSA-MLHAR model with recent methods under Dialogues dataset

Dialogues Dataset				
Methods	$Prec_n$	$Reca_l$	$Accu_y$	F_{score}
Attention-LSTM	69.93	69.31	70.18	69.92
SIARN	72.62	72.18	71.74	72.19
MIARN	73.10	73.12	72.78	72.62

ELMo-BiLSTM	75.68	75.62	76.02	75.82
IMHSAA	77.92	77.57	77.22	77.66
DLNLP-SA	94.53	94.43	94.32	94.32
ASD-AHADL	95.39	95.36	95.36	95.37

5. Conclusion

In this paper, we have established an ASD-AHADL technique. The ASD-AHADL methodology applies optimal DL model for the detection of sarcastic content. To accomplish this, the ASD-AHADL technique comprises distinct kinds of sub-processes involved in data preprocessing, BERT-based word embedding, AGRU-LSTM-based sarcasm detection, and AHA-based parameter fine-tuning. At the initial stage, the ASD-AHADL technique undergoes data preprocessing and BERT-based word embedding process. Followed by, the ASD-AHADL technique uses AGRU-LSTM for the sarcasm detection process. Finally, the AHA-based parameter tuning process is involved to fine-tune the parameters based on the DL approach. The empirical investigation of the ASD-AHADL technique is tested on social media database. The experimental results highlighted the significant performance of the ASD-AHADL technique compared to other ones.

Acknowledgment: "The author extends her appreciation to Umm Al-Qura University, Saudi Arabia, for funding this research work through Grant Number 25UQU4310245GSSR02R"

Funding: "This research was funded by Umm Al-Qura University, Saudi Arabia, under Grant Number 25UQU4310245GSSR02R"

Conflicts of Interest: "The authors declare no conflict of interest."

References

[1] D. Vinoth and P. Prabhavathy, "An intelligent machine learning-based sarcasm detection and classification model on social networks," *J. Supercomput.*, vol. 19, no. 1–2, p. 288, 2022.

[2] L. K. Ahire, S. D. Babar, and G. R. Shinde, *Sarcasm detection in online social network: Myths, realities, and issues*, in *Security Issues and Privacy Threats in Smart Ubiquitous Computing*, Studies in Systems, Decision and Control, vol. 341. Singapore: Springer, pp. 227–238, 2021.

[3] L. H. Son, A. Kumar, S. R. Sangwan, A. Arora, A. Nayyar et al., "Sarcasm detection using soft attention-based bidirectional long short-term memory model with convolution network," *IEEE Access*, vol. 7, pp. 23319–23328, 2019.

[4] K. Rothermich, A. Ogunlana, and N. Jaworska, "Change in humor and sarcasm use based on anxiety and depression symptom severity during the COVID-19 pandemic," *J. Psychiatr. Res.*, vol. 140, pp. 95–100, 2021.

[5] C. I. Eke, A. A. Norman, L. Shuib, and H. F. Nweke, "Sarcasm identification in textual data: Systematic review, research challenges and open directions," *Artif. Intell. Rev.*, vol. 53, pp. 4215–4258, 2020.

[6] S. M. Sarsam, H. A. Samarraie, A. I. Alzahrani, and B. Wright, "Sarcasm detection using machine learning algorithms in Twitter: A systematic review," *Int. J. Mark. Res.*, vol. 62, pp. 578–598, 2020.

[7] A. Kumar, V. T. Narapareddy, V. A. Srikanth, A. Malapati, and L. B. M. Neti, "Sarcasm detection using multi-head attention-based bidirectional LSTM," *IEEE Access*, vol. 8, pp. 6388–6397, 2020.

[8] R. A. Potamias, G. Siolas, and A. G. Stafylopatis, "A transformer-based approach to irony and sarcasm detection," *Neural Comput. Appl.*, vol. 32, no. 23, pp. 17309–17320, 2020.

[9] F. B. Kader, N. H. Nujat, T. B. Sogir, M. Kabir, H. Mahmud, and K. Hasan, "Computational sarcasm analysis on social media: A systematic review," *arXiv preprint arXiv:2209.06170*, 2022.

[10] Y. Du, T. Li, M. S. Pathan, H. K. Teklehaimanot, and Z. Yang, "An effective sarcasm detection approach based on sentimental context and individual expression habits," *Cogn. Comput.*, vol. 14, no. 1, pp. 78–90, 2022.

[11] B. Liang, L. Gui, Y. He, E. Cambria, and R. Xu, "Fusion and discrimination: A multimodal graph contrastive learning framework for multimodal sarcasm detection," *IEEE Trans. Affect. Comput.*, vol. 1, pp. 1–15, 2024.

[12] M. A. Rosid, D. Siahaan, and A. Saikhu, "Improving sarcasm detection in mash-up language through hybrid pretrained word embedding," in *Proc. IEEE 8th Int. Conf. Softw. Eng. Comput. Syst. (ICSECS)*, 2023, pp. 58–63.

[13] A. Kamal and M. Abulaish, "Cat-BiGRU: Convolution and attention with bi-directional gated recurrent unit for self-deprecating sarcasm detection," *Cogn. Comput.*, vol. 14, no. 1, pp. 91–109, 2022.

[14] H. Fu, H. Liu, H. Wang, L. Xu, J. Lin, and D. Jiang, "Multi-modal sarcasm detection with sentiment word embedding," *Electronics*, vol. 13, no. 5, p. 855, 2024.

[15] M. Madani, H. Motameni, and R. Roshani, "Fake news detection using feature extraction, natural language processing, curriculum learning, and deep learning," *Int. J. Inf. Technol. Decis. Mak.*, pp. 1–36, 2023.

[16] H. Liu, R. Wei, G. Tu, J. Lin, C. Liu, and D. Jiang, "Sarcasm driven by sentiment: A sentiment-aware hierarchical fusion network for multimodal sarcasm detection," *Inf. Fusion*, p. 102353, 2024.

[17] Y. Zhang, D. Ma, P. Tiwari, C. Zhang, M. Masud, M. Shorfuzzaman, and D. Song, "Stance-level sarcasm detection with BERT and stance-centered graph attention networks," *ACM Trans. Internet Technol.*, vol. 23, no. 2, pp. 1–21, 2023.

[18] H. Li, Y. Ma, Z. Ma, and H. Zhu, "Weibo text sentiment analysis based on BERT and deep learning," *Appl. Sci.*, vol. 11, no. 22, p. 10774, 2021.

[19] R. Mousa, M. Khezli, and S. Hesaraki, "Classifying objects in 3D point clouds using recurrent neural network: A GRU LSTM hybrid approach," *arXiv preprint arXiv:2403.05950*, 2024.

[20] W. Xu, Y. Shen, L. Jing, and X. Sun, "A life prediction method based on MDFF and DITCN-ABiGRU mixed network model," *Heliyon*, vol. 10, no. 2, 2024.

[21] L. Kondisetti and S. Katragadda, "A multi-objective artificial hummingbird algorithm for dynamic optimal volt-var controls for high electric vehicle load penetration in a photovoltaic distribution network," *e-Prime Adv. Electr. Eng. Electron. Energy*, p. 100474, 2024.

[22] A. R. W. Sait and M. K. Ishak, "Deep learning with natural language processing-enabled sentimental analysis on sarcasm classification," *Comput. Syst. Sci. Eng.*, vol. 44, no. 3, pp. 2553–2567, 2023.

[23] IBM, "Twitter conversations dataset for conversational document prediction dataset," *GitHub Repository*, 2023. [Online]. Available: https://github.com/IBM/twitter-customer-care-document-prediction