Tech Science Press

Check for updates

Dragonfly Optimization with Deep Learning Enabled Sentiment Analysis for Arabic Tweets

Aisha M. Mashraqi and Hanan T. Halawani[*]

College of Computer Science and Information Systems, Najran University, Najran, 61441, Saudi Arabia
*Corresponding Author: Hanan T. Halawani. Email: hthalawani@nu.edu.sa
Received: 13 April 2022; Accepted: 16 June 2022

Abstract: Sentiment Analysis (SA) is one of the Machine Learning (ML) techniques that has been investigated by several researchers in recent years, especially due to the evolution of novel data collection methods focused on social media. In literature, it has been reported that SA data is created for English language in excess of any other language. It is challenging to perform SA for Arabic Twitter data owing to informal nature and rich morphology of Arabic language. An earlier study conducted upon SA for Arabic Twitter focused mostly on automatic extraction of the features from the text. Neural word embedding has been employed in literature, since it is less labor-intensive than automatic feature engineering. By ignoring the context of sentiment, most of the word-embedding models follow syntactic data of words. The current study presents a new Dragonfly Optimization with Deep Learning Enabled Sentiment Analysis for Arabic Tweets (DFODL-SAAT) model. The aim of the presented DFODL-SAAT model is to distinguish the sentiments from opinions that are tweeted in Arabic language. At first, data cleaning and pre-processing steps are performed to convert the input tweets into a useful format. In addition, TF-IDF model is exploited as a feature extractor to generate the feature vectors. Besides, Attention-based Bidirectional Long Short Term Memory (ABLSTM) technique is applied for identification and classification of sentiments. At last, the hyperparameters of ABLSTM model are optimized using DFO algorithm. The performance of the proposed DFODL-SAAT model was validated using the benchmark dataset and the outcomes were investigated under different aspects. The experimental outcomes highlight the superiority of DFODL-SAAT model over recent approaches.

Keywords: Natural language processing; sentiment analysis; arabic tweets; deep learning; metaheuristics; lexicon approach

1 Introduction

Sentiment Analysis (SA) is a domain of research that is associated with several other research domains such as text mining, computational linguistics, and Natural Language Processing (NLP). It is regarded as the extraction of particular information from textual data fed by the individuals. Indeed, numerous terminologies exist for SA with different purposes while the commonly used names include subjective analysis, emotion

extraction (based on affective computing) [1], and opinion mining. In Opinion Mining (OM), the difference between positive and negative opinions is analyzed through the text fed by the individuals [2], whereas emotion extraction is regarded as the process that tracks different kinds of emotions such as sadness, angry, happy, etc., The current research aims at performing SA in line with opinion mining approach. Typically, SA targets at analyzing the individuals' attitudes, emotions, opinions, appraisals, and sentiments with respect to entities namely, services, issues, products, events, topics, organizations, and individuals.

Social Networking Sites (SNS) like Facebook, Twitter and LinkedIn are useful platforms to fetch data and conduct SA, since individuals use these SNS to share their opinions on a broad array of matters. Opinion Mining or SA focuses on understanding the attitude of a writer, speaker, or other subjects about a specific event or topic [3]. SA consists of various trending applications under different domains. In business operations, it automatically permits the enterprises or companies to gather the views of their consumers about their services or products. From a political perspective, it is useful in inferring the orientation of public and predicting their reaction for political activities. Such insights are highly helpful in making political decisions [4]. SA can be conducted on different levels such as topic, sentence, document, and so on. In spite of its significance, there is only a handful of studies available with regards to SA on Arabic language. This might be due to multiple reasons such as challenging script and morphological and ambiguity of Arabic language. These characteristics, along with inadequate resources and the changing dialects, require more inputs to conduct SA in Arabic language [5].

The complexity arises in SA of Arabic tweets due to rich structure and informal characteristics of Arabic language on Twitter [6]. SA methods involve supervised learning methods which exploit Machine Learning (ML) techniques with feature engineering and unsupervised learning methods. These techniques leverage sentiment lexicons and rule-related approaches [7]. The superiority of these approaches that employ ML techniques, lie in the manual extraction of features for categorization. Manual extraction of features is labor-intensive and time-consuming in nature while features that are manually extracted are cited as 'surface features' [8]. At present, Deep Learning (DL) techniques depict incredible enhancements in SA for English language.

Alsayat [9] proposed a personalized DL technique using state-of-the-art word embedding technique and developed a Long Short Term Memory (LSTM) model. Then, the study established a strong architecture based on LSTM word and embedding models which decode the contextual relationship amongst words and understand rare or unseen words in comparatively emergent circumstances. For example, the outbreak of COVID-19 pandemic introduced prefixes and suffixes and these words were used in the training dataset. Wazrah et al. [10] introduced a Bi-direction Gated Recurrent Unit (SBi-GRU) and Stacked GRU (SGRU) model for opining mining from embedded words in Arabic language. The study presented a novel method to discard the stop words using Automated Sentimental Refinement (ASR) instead of using low quality Arabic stop words or automatic collection of stop word lists.

Alharbi et al. [11] developed a DL-based sentimental analysis technique to predict the divergence of sentiments and opinions. Two kinds of recurrent neural networks were leveraged in this study to learn about high-level representation. In order to enhance the robustness of the method and mitigate data dependency problems, three separate classification approaches were applied and the results were achieved. In literature [12], the authors proposed two DL techniques to accomplish crucial aspect-based sentimental analysis processes such as aspect-sentiment classifier and aspect-category recognition. Initially, a recognition method was presented as per stacked independent LSTM and Convolutional Neural Network (CNN). Next, a classifier method was presented on the basis of stacked bi-directional independent LSTM. Gandhi et al. [13] recognized all the words in tweets and assigned implication for the words. This featured work was integrated with stop word, tweeter word and word2vec and was

incorporated with DL technologies of LSTM and CNN. This way, the technique gained the potential to recognize the pattern of stop word count, using a specific approach.

The current research work presents a new Dragonfly Optimization with Deep Learning Enabled Sentiment Analysis for Arabic Tweets (DFODL-SAAT) model. The aim of the presented DFODL-SAAT model is to distinguish the sentiments from opinions found in Arabic language tweets. At first, data cleaning and pre-processing steps are performed to convert the input tweets into a useful format. In addition, TF-IDF model is exploited as a feature extractor to generate feature vectors. Besides, Attention based Bidirectional Long Short Term Memory (ABLSTM) technique is applied for identification and classification of sentiments. At last, the hyperparameters of ABLSTM model are optimized using DFO approach. The proposed DFODL-SAAT model was experimentally validated for its performance using benchmark dataset and the outcomes were investigated under several aspects.

Rest of the paper is organized as follows. Section 2 discusses the proposed model, Section 3 provides the experimental validation, and Section 4 concludes the study.

2 The Proposed Model

In this study, a new DFODL-SAAT model has been developed to distinguish the sentiments from opinions that exist in Arabic language tweets. The proposed DFODL-SAAT model encompasses pre-processing, TF-IDF feature extraction, ABLSTM classification, and DFO hyperparameter optimizer. Fig. 1 depicts the overall processes involved in DFODL-SAAT technique.

2.1 Data Pre-Processing

In this initial stage, data cleaning and pre-processing steps are performed to convert the input tweets into a useful format. The tweets, gathered in the form of text input from humans, comprise of orthographic mistakes, unstructured language, slang words, and abbreviations. In this context, the infrastructure design should be altered by executing some pre-processed approaches so that ML technique can be enabled to analyze the text and produce reliable outcomes with high accuracy. The analysis of tweets in Arabic language presents further challenges than the analysis of tweets from other languages. Arabic natural language lacks a robust tool and sufficient resources to support the extraction of sentiment from Arabic text [14]. The subsequent steps describe the stages involved in data pre-processing, which are executed upon two datasets.

- Eliminating unrelated tweets i.e., tweets that contain ads and tweets that do not discuss about distance learning in Saudi Arabia are removed manually. Subsequently, the number of tweets in primary data set got decreased to 5,096 tweets whereas in second dataset, the number of tweets got decreased to 9,160 tweets.
- Eliminating non-Arabic letters.
- Eliminating symbols that represent emotions and contain hashtag sign, emoticons, symbols, and numbers.
- Eliminating URL and user mentions.
- Eliminating Tashkeel that utilizes the symbol "-"to increase the length of some characters.
- Eliminating punctuation.
- Eliminating repeated characters.
- Eliminating stop words i.e., the stop words extracted contains a typical group of Arabic stop words, listed by Python's NLTK library.
- Executing Arabic normalization.

• Executing word stemming with the help of stemmer to reduce the Arabic word to its word stem.
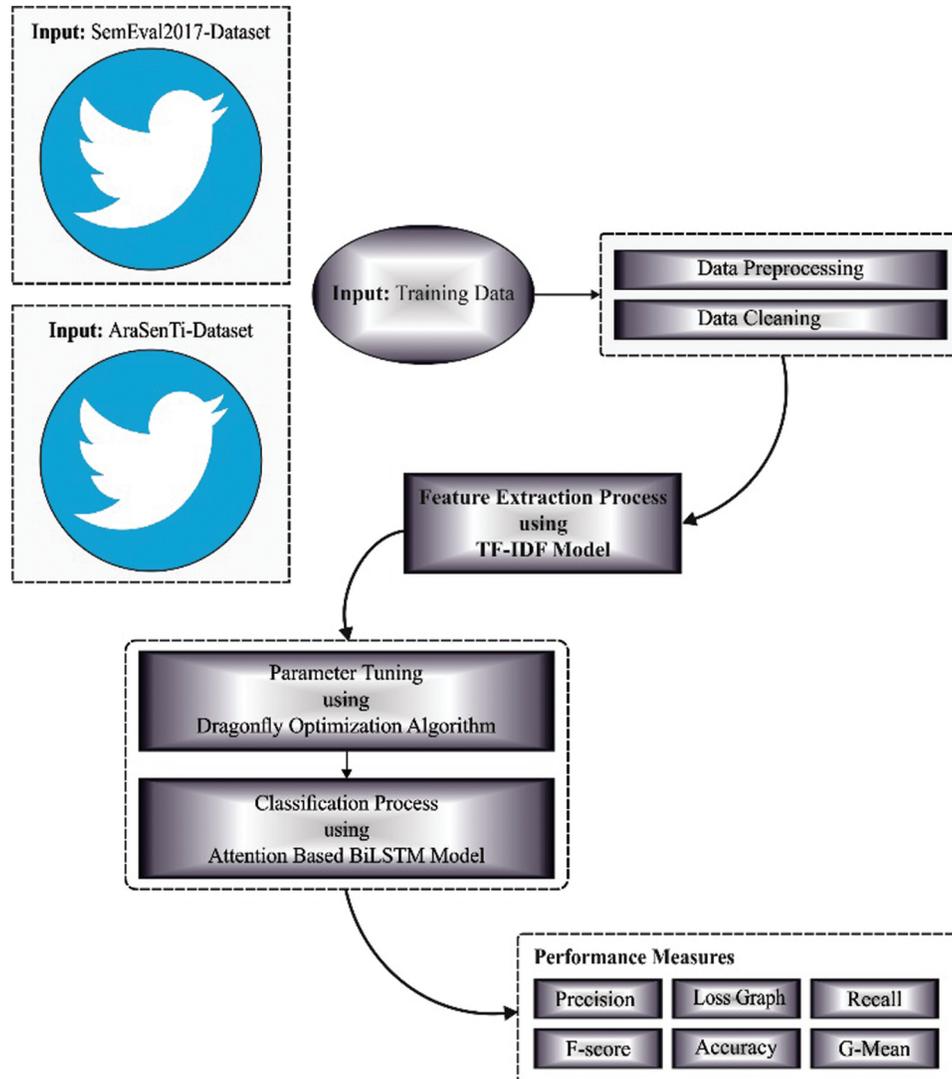• Executing tokenization process that separates the text to lesser piece or token.

Figure 1: Overall processes involved in DFODL-SAAT technique

2.2 TF-IDF Based Feature Extraction

After pre-processing, TF-IDF model, a combination of Term Frequency (TF) and Inverse Document Frequency (IDF), is exploited as a feature extractor to extract the feature vectors. As the original values of TF are utilized from the document directly, TF depiction remains the easiest TWS. The capacity of TF, to distinguish every relevant document from irrelevant document, is very low since it ignores the gathered frequency. In order to resolve this issue, IDF is presented by the gathered frequency that improves the discriminatory capacity of 'term to text' classifier. IDF, extended in Document Frequency (DF), represents the amount of documents in which the term takes place. Based on the assumption that when a term occurs in fewer number of documents, it is considered to be increasingly relevant than its

occurrence in multiple documents [15]. The IDF values of a particular term are attained using the equation given below.

$$TF - IDF(t, \ d, \ D) = \log \frac{|D|}{DF(t, \ D)} \tag{1}$$

whereas DF(t, D) indicates the DF value of term t in corpus D. The symbol in Eq. (2) denotes the overall amount of documents from corpus D. To prevent infinity in a certain number of exceptional cases, the equation is enhanced as given below.

$$TF - IDF(t, \ d, \ D) = \log \frac{|D| + 1}{DF(t, \ D) + 1} \tag{2}$$

Like IDF, TF-IDF indicates a global statistical measure. The traditional infrastructure of TF-IDF is given herewith.

$$TF - IDF(t, \ d, \ D) = TF(t, \ d) * IDF(t, \ d, \ D). \tag{3}$$

Now TF−IDF(t, d, D) signifies the weight of term t of document d in corpus D, whereas TF(t, d) denotes the TF value of term t in document d.

2.3 ABiLSTM Based Classification

At this stage, ABLSTM model is applied for both identification and classification of sentiments. LSTM is a Recurrent Neural Network (RNN) which resolves the gradient problems like disappearance and explosion by learning long- and short-term dependencies [16]. Every time, in step t, a hidden forward state with hidden unit purpose $\vec{h}$ is calculated based on the preceding Hidden Layer (HL) $\overrightarrow{h_{t-1}}$ and input at the existing step x_t whereas hidden backward state with hidden unit function $\overleftarrow{h}$ are calculated based on future HL $\overleftarrow{h_{t+1}}$ and input at the present step x_t. Both forward as well as backward context representations, created by $\overrightarrow{h_t}$ and $\overleftarrow{h_t}$ correspondingly, are concatenated as long vectors. The integrated output is the forecast of teacher-provided target signal.

Attention Machine (AM) is used to enhance the performance of Bi-LSTM by paying attention to specific input features, using one of the discriminative data. In order to capture the significance of all the input segments, AM is determined as given in Eqs. (4)–(6):

$$u_t = \ \tanh \ (W_w e_t + b_w) \tag{4}$$

$$h_t = \frac{exp(u_t^T u_w)}{\Sigma_t exp(u_t^T u_w)} \tag{5}$$

$$v_t = \sum_t h_t \cdot e_t \tag{6}$$

Here, v_t implies the resultant attention layer, but W_w, u_w, and b_w signify the two trainable weights and bias correspondingly. By multiplying e_t and h_t, temporal and spatial data are chosen and extracted in e_t which provides the highly-important decoded tasks.

2.4 Hyperparameter Optimization

Finally, the hyperparameters involved in ABLSTM model are optimized using DFO algorithm. DFO is a newly presented swarm-based method. It stimulates the migration and hunting behaviors of dragonflies [17]. Migration behavior is named after dynamic swarming (migratory) characteristics of dragonflies. Here, the

dragonflies fly in one direction as a swarm i.e., large groups, during migration. The hunting behavior of dragonflies is named as static swarm (feeding) in which the dragonflies fly as a small group over a small region to seek food sources. Similar to other nature-inspired approaches, DFO algorithm also comprises of two stages such as exploitation and exploration. In initial phase, it is stimulated by dynamic swarm mechanism whereas in second phase, it is stimulated by static swarm mechanism.

Here, X denotes the location of present searching agent, X_j indicates the j-th neighbor of X searching agent, and N indicates neighborhood size. In order to model the swarm-based mechanism of dragonflies, five individual behaviors are employed.

- Separation is a process in which a searching agent avoids its neighboring searching agents. It can be arithmetically expressed as follows.

$$S_i = -\sum_{j=1}^{N} X - X_i \tag{7}$$

- Alignment represents an individual that matches the velocity of neighboring individuals. It can be arithmetically modelled using the following equation.

$$A_i = \frac{\sum_{j=1}^{N} V_j}{N} \tag{8}$$

- Here, V_j indicates the velocity of jth neighbor.
- Cohesion denotes the tendency of an individual to fly near the neighboring center of mass. The mathematical expression of the dragonfly is given below.

$$C_i = \frac{\sum_{j=1}^{N} x_j}{N} - X \tag{9}$$

- Attraction represents the tendency of individuals to fly towards the food sources. The attraction between i^{th} solution and food source is shown below.

$$F_i = F_{loc} - X \tag{10}$$

- Now F_{loc} signifies the location of food sources.
- Distraction represents the tendency of an individual to fly away from its enemies. The distraction between i^{th} solution and the enemy is arithmetically expressed in the following equation.

$$E_i = E_{loc} + X \tag{11}$$

Here, E_{loc} indicates the location of the energy.

The food source fitness and position are assumed to be upgraded by an optimal candidate (searching agent). Furthermore, location and fitness of the enemy are upgraded by the worst candidate. This causes divergence outside the non-promising region and convergence towards the promising area of searching space. Fig. 2 demonstrates the steps involved in DFO technique.

In PSO model, DFO algorithm employs two vectors to upgrade the location of dragonfly: Step vector (X) is the same as position vector and velocity vector in Particle Swarm Optimization (PSO) algorithm. Step vector signifies the direction of motion of dragonfly as shown in the following equation.

$$\triangle X_{t+1} = (sS_i + aA_i + cC_i + fF_i + eE_i) + wX_t \tag{12}$$

Now, s, w, a, c, f, and e denote the weight of separation S_i, alignment A_i, cohesion C_i, attraction towards the food source F_i, and distraction in the enemy E_i of ith individual correspondingly. Here, weight enables the DFO algorithm to accomplish exploration and intensification behaviors.
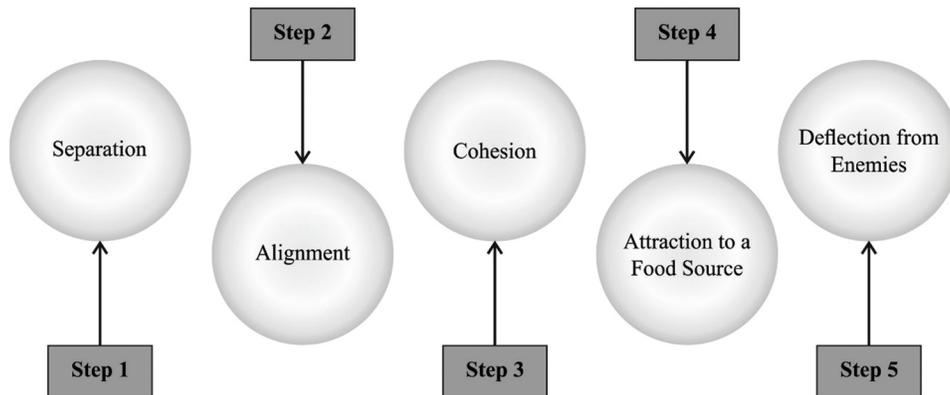
Figure 2: Steps in DFO technique

The location of an individual can be upgraded using the following equation:

$$X_{t+1} = X_t + \triangle X_{t+1} \tag{13}$$

Here, t denotes the existing iteration.

The process is initiated by generating a random population. Both location and step vectors of the dragonfly are arbitrarily determined. During all the iterations, the process repeatedly performs the subsequent step until the end condition is met. Initially, every individual from the population can be estimated using Fitness Function (FF). Next, the main coefficient is upgraded. Afterwards, enemy (E), separation (S), food source (F), cohesion (C), and alignment (A) are upgraded using the Eqs. (7) and (11). At last, the step and location vectors are upgraded using the Eqs. (12) and (13) correspondingly.

3 Results and Discussion

In this section, the proposed DFODL-SAAT model was experimentally validated using SemEval2017-dataset [18] and AraSenTi-dataset [19]. SemEval2017-dataset includes 2235 samples in which 965 are positive tweets and 1270 are negative tweets. AraSenTi-dataset comprises of 9750 tweets in which 4235 tweets are positive and 5515 tweets are negative. Table 1 describes the details of the dataset.

Table 1: Dataset details

Dataset	SemEval2017-dataset	AraSenTi-dataset
Positive tweet	965	4235
Negative tweet	1270	5515
Total no. of instances	2235	9750

Fig. 3 shows the confusion matrix generated by the proposed DFODL-SAAT model on SemEval2017 dataset. With entire dataset, the proposed DFODL-SAAT model classified 855 samples under positive class and 1,196 samples under negative class. Moreover, with 70% of training set (TRS) dataset, DFODL-SAAT approach categorized 594 samples under positive class and 844 samples under negative class. Furthermore, with 30% of testing set (TSS) dataset, the proposed DFODL-SAAT system classified 261 samples under positive class and 352 samples under negative class.
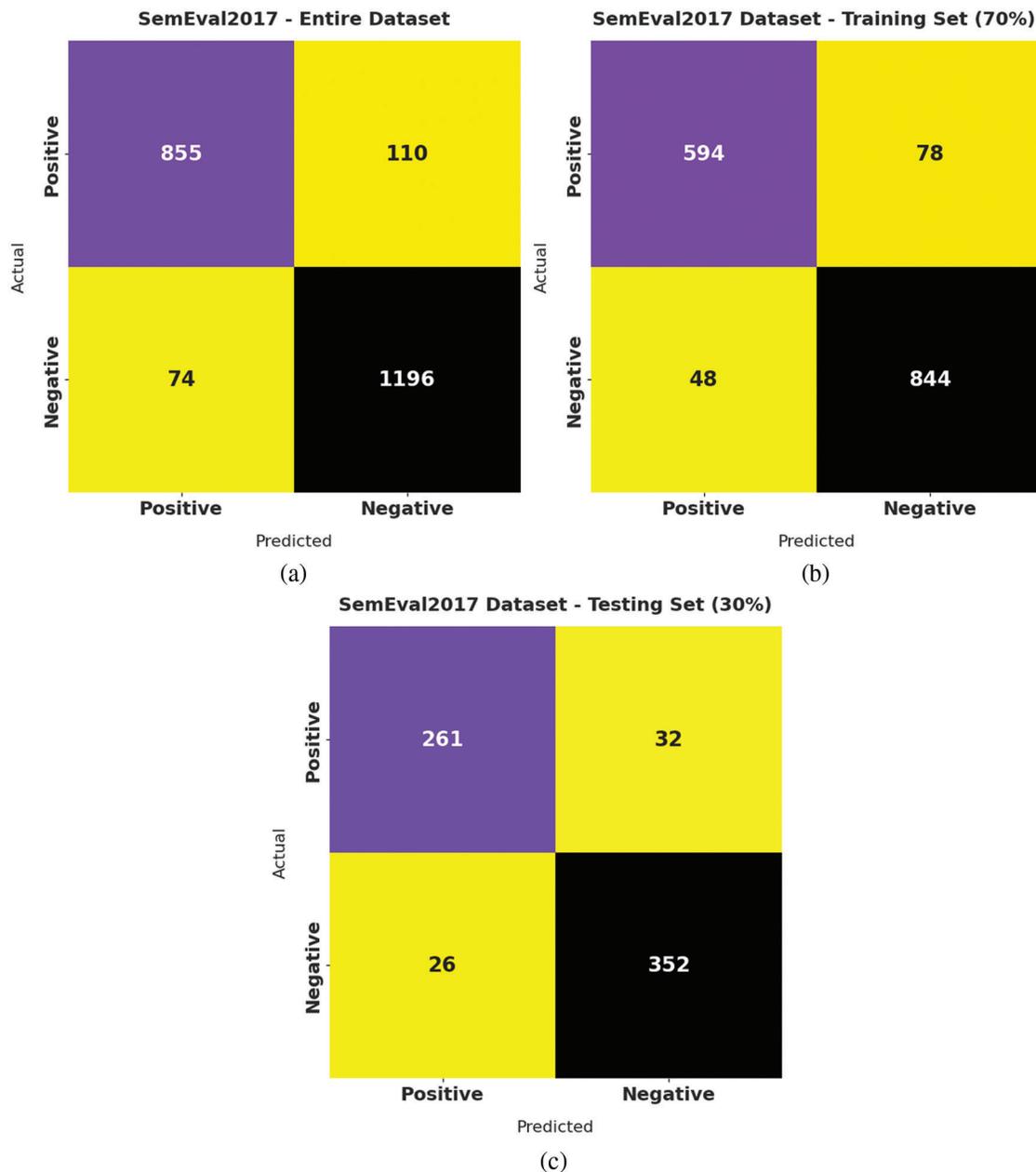
Figure 3: Confusion matrix of DFODL-SAAT technique on SemEval2017 dataset

Table 2 and Fig. 4 provide an overview on the classification outcomes achieved by DFODL-SAAT model on test SemEval2017 dataset. The experimental outcomes imply that the proposed DFODL-SAAT model reached the maximum outcome under all classes. For instance, with entire dataset, DFODL-SAAT model reached $accu_y$, $prec_n$, $reca_l$, F_{score}, and Geometric Mean (GM) values such as 91.77%, 91.81%, 91.39%, 91.57%, and 91.34% respectively. Also, with 70% of TRS dataset, the proposed DFODL-SAAT method reached $accu_y$, $prec_n$, $reca_l$, F_{score}, and GM values such as 91.94%, 92.03%, 91.51%, 91.73%, and 91.45% correspondingly. Besides, with 30% of TSS dataset, the proposed DFODL-SAAT model reached $accu_y$, $prec_n$, $reca_l$, F_{score}, and GM values such as 91.36%, 91.30%, 91.10%, 91.19%, and 91.08% correspondingly.

Table 2: Results of the analysis of DFODL-SAAT model under different measures on SemEval2017 dataset

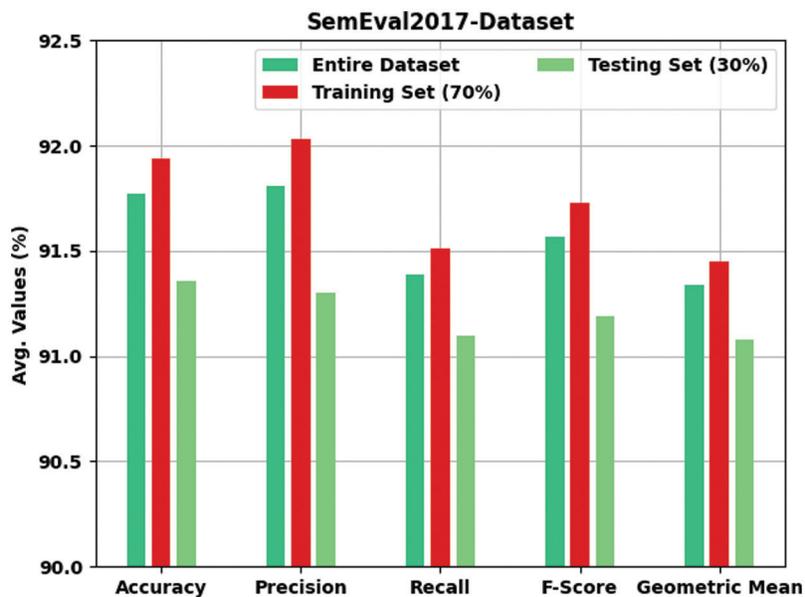| Class labels	Accuracy	Precision	Recall	F-score	Geometric mean
SemEval2017-dataset					
Entire dataset					
Positive	91.77	92.03	88.60	90.29	91.34
Negative	91.77	91.58	94.17	92.86	91.34
Average	91.77	91.81	91.39	91.57	91.34
Training set (70%)					
Positive	91.94	92.52	88.39	90.41	91.45
Negative	91.94	91.54	94.62	93.05	91.45
Average	91.94	92.03	91.51	91.73	91.45
Testing set (30%)					
Positive	91.36	90.94	89.08	90.00	91.08
Negative	91.36	91.67	93.12	92.39	91.08
Average	91.36	91.30	91.10	91.19	91.08

Figure 4: Results of the analysis of DFODL-SAAT technique on SemEval2017 dataset

Both Training Accuracy (TA) and Validation Accuracy (VA) values, attained by DFODL-SAAT model on SemEval2017 dataset, are portrayed in Fig. 5. The experimental outcome imply that the presented DFODL-SAAT model gained the maximum TA and VA values. To be specific, VA seemed to be higher than TA.
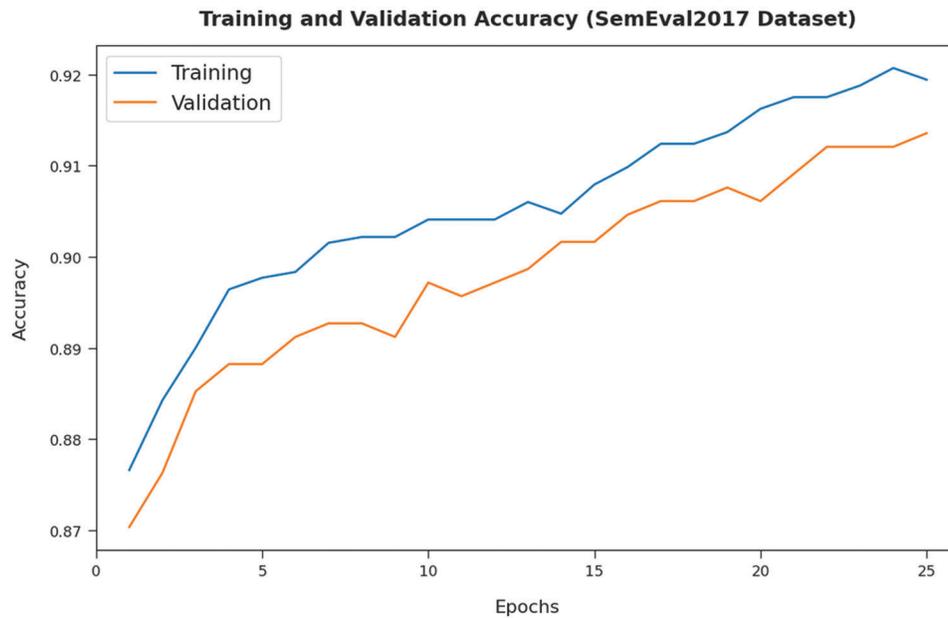
Figure 5: TA and VA analyses results of DFODL-SAAT technique on SemEval2017 dataset

Both Training Loss (TL) and Validation Loss (VL) values, achieved by the proposed DFODL-SAAT model on SemEval2017 dataset, are shown in Fig. 6. The experimental result infer that DFODL-SAAT method achieved the least TL and VL values. Specifically, VL seemed to be lower than TL.

Figure 6: TL and VL analyses results of DFODL-SAAT technique on SemEval2017 dataset

Table 3 and Fig. 7 portrays the comparative examination results accomplished by DFODL-SAAT model and other recent models on SemEval2017-dataset. The results imply that surface features model offered the least classification outcome. In line with this, both ASEH and Bidirectional Encoder Representations from

Transformers (BERT) models achieved slightly enhanced classifier results. Next, generic embeddings and hybrid models exhibited reasonable classification performance. However, the proposed DFODL-SAAT model accomplished an effectual outcome with $prec_n$, $reca_l$, and F_{score} values such as 91.30%, 91.10%, and 91.19% respectively.

Table 3: Comparative analysis results of DFODL-SAAT technique and other existing algorithms on SemEval2017 dataset

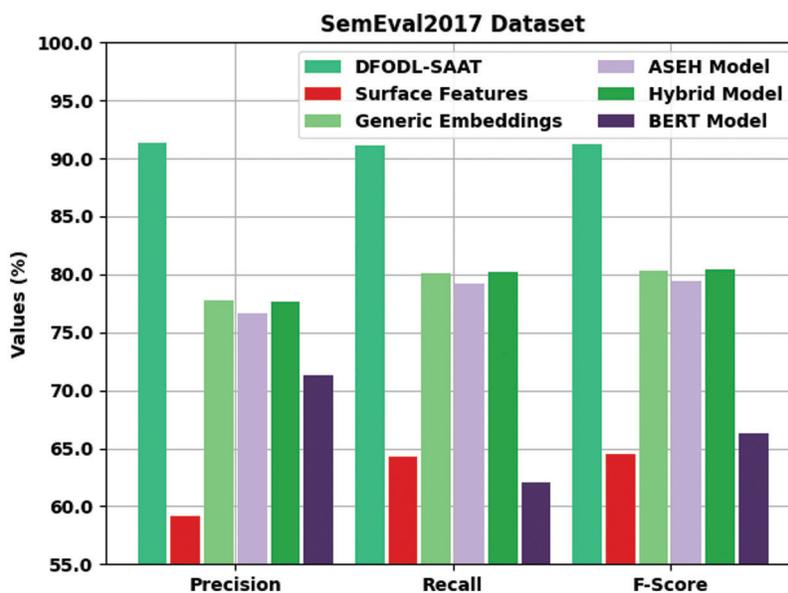| SemEval2017 dataset			
Method	Precision	Recall	F-score
DFODL-SAAT	91.30	91.10	91.19
Surface features	59.18	64.33	64.49
Generic embeddings	77.71	80.10	80.28
ASEH model	76.59	79.26	79.42
Hybrid model	77.68	80.22	80.38
BERT model	71.30	62.03	66.27

Figure 7: Comparative analysis of DFODL-SAAT technique on SemEval2017 dataset

The confusion matrices generated by the proposed DFODL-SAAT method on AraSenTi dataset are shown in Fig. 8. With entire dataset, the proposed DFODL-SAAT algorithm classified 4016 samples under positive class and 5347 samples under negative class. Additionally, with 70% of TRS dataset, DFODL-SAAT model categorized 2819 samples under positive class and 3728 samples under negative class. Besides, with 30% of TSS dataset, the proposed DFODL-SAAT model recognized 1197 samples under positive class and 1619 samples under negative class.
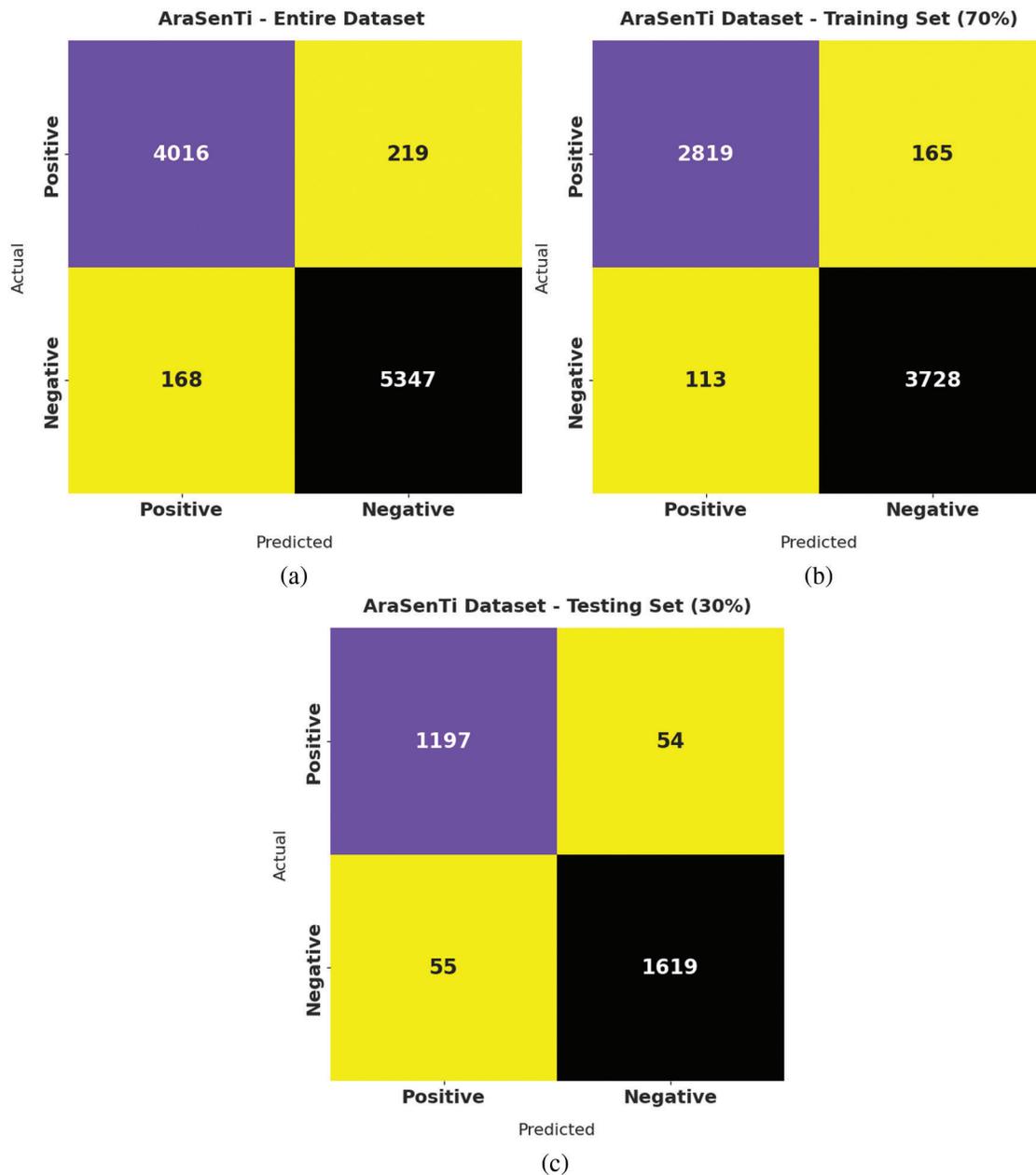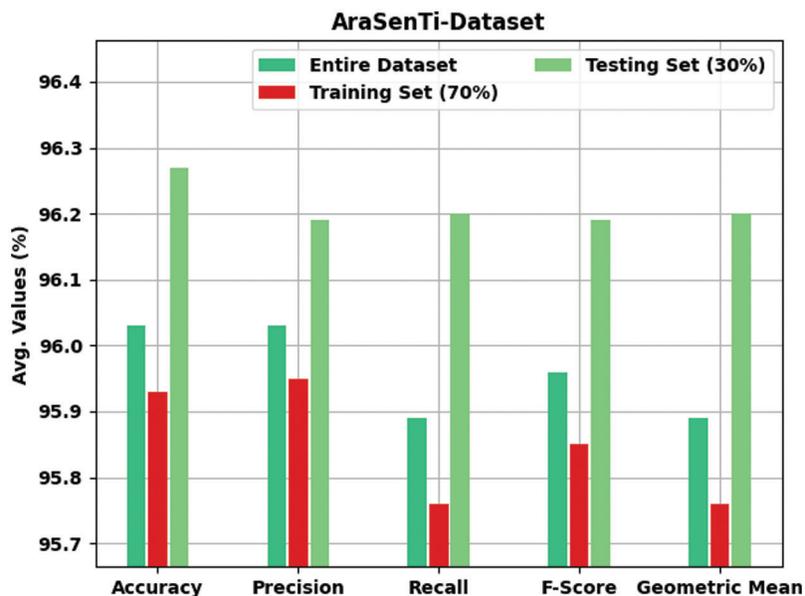
Figure 8: Confusion matrix of DFODL-SAAT approach on AraSenTi dataset

Table 4 and Fig. 9 provide the detailed classification outcomes accomplished by DFODL-SAAT algorithm on test AraSenTi dataset. The experimental outcomes infer that DFODL-SAAT model attained the maximum outcomes under all classes. For instance, with entire dataset, the proposed DFODL-SAAT technique reached $accu_y$, $prec_n$, $reca_l$, F_{score}, and GM values such as 96.03%, 96.03%, 95.89%, 95.96%, and 95.89% respectively. In addition, with 70% of TRS dataset, DFODL-SAAT model reached $accu_y$, $prec_n$, $reca_l$, F_{score}, and GM values such as 95.93%, 95.95%, 95.76%, 95.85%, and 95.76% correspondingly. Eventually, with 30% of TSS dataset, the proposed DFODL-SAAT technique reached $accu_y$, $prec_n$, $reca_l$, F_{score}, and GM values such as 96.27%, 96.19%, 96.20%, 96.19%, and 96.20% correspondingly.

Table 4: Results of the analysis of DFODL-SAAT model under different measures on AraSenTi dataset

Class labels	Accuracy	Precision	Recall	F-score	Geometric mean
AraSenTi-dataset					
Entire dataset					
Positive	96.03	95.98	94.83	95.40	95.89
Negative	96.03	96.07	96.95	96.51	95.89
Average	96.03	96.03	95.89	95.96	95.89
Training set (70%)					
Positive	95.93	96.15	94.47	95.30	95.76
Negative	95.93	95.76	97.06	96.41	95.76
Average	95.93	95.95	95.76	95.85	95.76
Testing set (30%)					
Positive	96.27	95.61	95.68	95.65	96.20
Negative	96.27	96.77	96.71	96.74	96.20
Average	96.27	96.19	96.20	96.19	96.20

Figure 9: Results of the analysis of DFODL-SAAT model on AraSenTi dataset

Both TA and VA values, attained by the proposed DFODL-SAAT model on AraSenTi dataset, are demonstrated in Fig. 10. The experimental outcomes imply that DFODL-SAAT model gained the maximal TA and VA values. To be specific, VA seemed to be higher than TA.

TL and VL values, attained by the proposed DFODL-SAAT model on AraSenTi dataset, are portrayed in Fig. 11. The experimental outcome infer that the proposed DFODL-SAAT model accomplished the least TL and VL values. To be specific, VL seemed to be lower than TL.
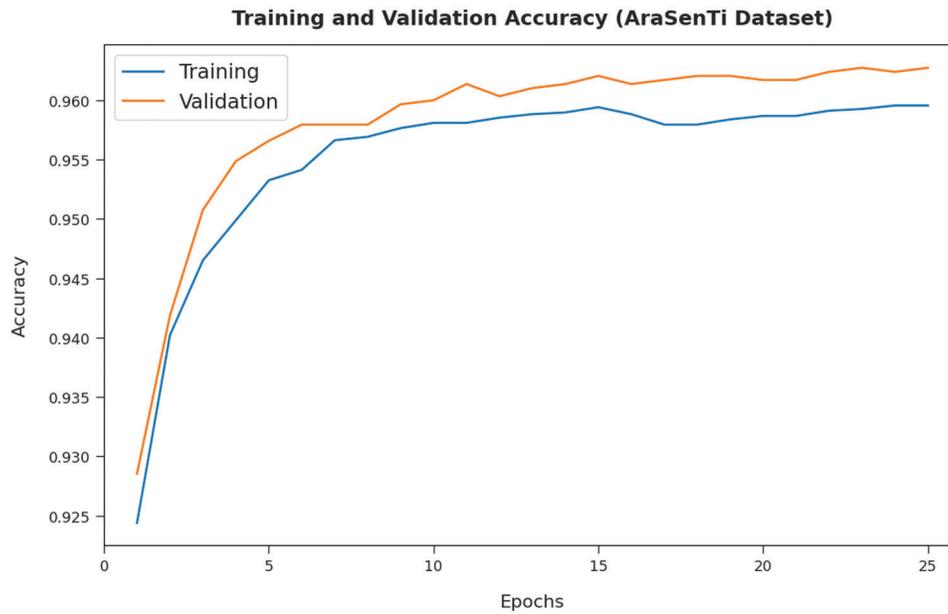
Figure 10: TA and VA analyses results of DFODL-SAAT technique on AraSenTi dataset
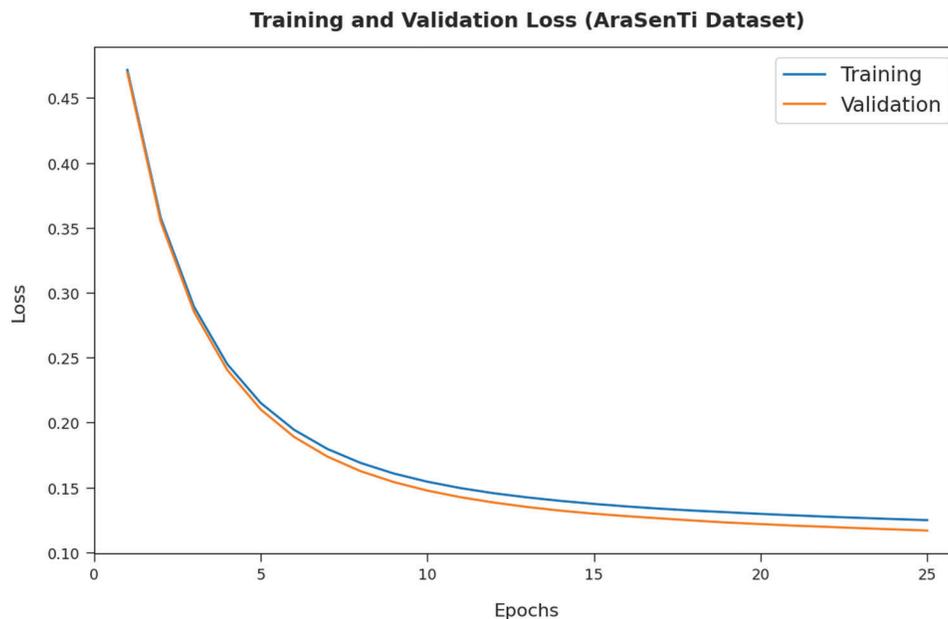

Figure 11: TL and VL analyses results of DFODL-SAAT technique on AraSenTi dataset

Table 5 demonstrate the comparative examination results accomplished by DFODL-SAAT model and other recent models on AraSenTi-dataset. The results imply that surface features approach offered the least classification outcomes. Besides, ASEH and BERT approaches accomplished somewhat improved classifier results. Afterward, generic embeddings and hybrid models exhibited reasonable classification performance. At last, the proposed DFODL-SAAT approach accomplished an effectual outcome with $prec_n$, $reca_l$, and F_{score} values such as 96.19%, 96.2%, and 96.19% correspondingly. From the

aforementioned tables and results, it is clear that the proposed DFODL-SAAT model is a proficient tool to conduct sentiment analysis compared to other existing models.

Table 5: Comparative analysis results of DFODL-SAAT and other existing techniques on AraSenTi dataset

AraSenTi dataset			
Method	Precision	Recall	F-score
DFODL-SAAT	96.19	96.2	96.19
Surface features	73.61	70.16	70.00
Generic embeddings	78.04	71.51	70.95
ASEH model	70.74	69.02	69.03
Hybrid model	75.85	74.81	74.84
BERT model	64.34	78.39	70.88

4 Conclusion

In this study, a novel DFODL-SAAT model has been developed to distinguish the sentiments from opinions that exist in Arabic language tweets. At first, data cleaning and pre-processing steps are performed to convert the input tweets into a useful format. In addition, TF-IDF model is exploited as a feature extractor to generate feature vectors. Besides, ABLSTM model is applied for both identification and classification of sentiments. At last, the hyperparameters involved in ABLSTM model are optimized using DFO algorithm. The proposed DFODL-SAAT model was experimentally validated for its performance using benchmark dataset and the outcomes were investigated under different aspects. The experimental outcomes established the superiority of the proposed DFODL-SAAT model over recent approaches. In future, hybrid DL models can be applied to enhance the classification outcomes.

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

References

[1] N. A. Twairesh and H. Al-Negheimish, "Surface and deep features ensemble for sentiment analysis of arabic tweets," *IEEE Access*, vol. 7, pp. 84122–84131, 2019.

[2] M. Heikal, M. Torki and N. El-Makky, "Sentiment analysis of arabic tweets using deep learning," *Procedia Computer Science*, vol. 142, pp. 114–122, 2018.

[3] T. K. Tran and T. T. Phan, "Capturing contextual factors in sentiment classification: An ensemble approach," *IEEE Access*, vol. 8, pp. 116856–116865, 2020, https://doi.org/10.1109/ACCESS.2020.3004180.

[4] T. K. Tran, H. Dinh, H. Nguyen, D. N. Le, D. K. Nguyen *et al.,* "The impact of the COVID-19 pandemic on college students: An online survey," *Sustainability*, vol. 13, no. 19, pp. 10762, 2021, https://doi.org/10.3390/su131910762.

[5] D. Zhang, J. Hu, F. Li, X. Ding, A. K. Sangaiah *et al.,* "Small object detection via precise region-based fully convolutional networks," *Computers, Materials and Continua*, vol. 69, no. 2, pp. 1503–1517, 2021.

[6] J. Wang, Y. Wu, S. He, P. K. Sharma, X. Yu *et al.,* "Lightweight single image super-resolution convolution neural network in portable device," *KSII Transactions on Internet and Information Systems (TIIS)*, vol. 15, no. 11, pp. 4065–4083, 2021.

[7] J. Wang, Y. Zou, P. Lei, R. S. Sherratt and L. Wang, "Research on recurrent neural network based crack opening prediction of concrete dam," *Journal of Internet Technology*, vol. 21, no. 4, pp. 1161–1169, 2020.

[8] J. Zhang, J. Sun, J. Wang and X. G. Yue, "Visual object tracking based on residual network and cascaded correlation filters," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 8, pp. 8427–8440, 2021.

[9] A. Alsayat, "Improving sentiment analysis for social media applications using an ensemble deep learning language model," *Arabian Journal for Science and Engineering*, vol. 47, no. 2, pp. 2499–2511, 2022.

[10] A. A. Wazrah and S. Alhumoud, "Sentiment analysis using stacked gated recurrent unit for arabic tweets," *IEEE Access*, vol. 9, pp. 137176–137187, 2021.

[11] A. Alharbi, M. Kalkatawi and M. Taileb, "Arabic sentiment analysis using deep learning and ensemble methods," *Arabian Journal for Science and Engineering*, vol. 46, no. 9, pp. 8913–8923, 2021.

[12] S. A. Dabet, S. Tedmori and M. A. Smadi, "Enhancing arabic aspect-based sentiment analysis using deep learning models," *Computer Speech & Language*, vol. 69, pp. 101224, 2021.

[13] U. D. Gandhi, P. M. Kumar, G. C. Babu and G. Karthick, "Sentiment analysis on twitter data by using convolutional neural network (cnn) and long short term memory (LSTM)," *Wireless Personal Communications*, vol. 2021, pp. 1–10, 2021, https://doi.org/10.1007/s11277-021-08580-3.

[14] T. H. Alenazi, N. F. B. Dhim, M. H. Alenazi, H. Tamim, R. S. Almagrabi *et al.* "Prevalence and predictors of anxiety among healthcare workers in Saudi Arabia during the COVID-19 pandemic," *Journal of Infection and Public Health*, vol. 13, no. 11, pp. 1645–1651, 2020.

[15] Z. Jiang, B. Gao, Y. He, Y. Han, P. Doyle *et al.* "Text classification using novel term weighting scheme-based improved tf-idf for internet media reports," *Mathematical Problems in Engineering*, vol. 2021, pp. 1–30, 2021.

[16] Y. Wang, Y. Dai, Z. Liu, J. Guo, G. Cao *et al.* "Computer-aided intracranial eeg signal identification method based on a multi-branch deep learning fusion model and clinical validation," *Brain Sciences*, vol. 11, no. 5, pp. 615, 2021.

[17] S. Mirjalili, "Dragonfly algorithm: A new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems," *Neural Computing and Applications*, vol. 27, no. 4, pp. 1053–1073, 2016.

[18] S. Rosenthal, N. Farra and P. Nakov, "SemEval-2017 task 4: Sentiment analysis in twitter," in *Proc. of the 11th Int. Workshop on Semantic Evaluation (SemEval-2017)*, Vancouver, Canada, pp. 502–518, 2017.

[19] N. Al-Twairesh, H. A. Khalifa, A. A. Salman and Y. A. Ohali, "Arasenti-tweet: A corpus for arabic sentiment analysis of Saudi tweets," *Procedia Computer Science*, vol. 117, pp. 63–72, 2017.