[1] Shilpi Gupta*
[2] Niraj Singhal
[3] Pradeep Kumar

Improved Aspect Based Sentiment Analysis with DBN-RNN model

Journal of Electrical Systems

Abstract: - Mining the sentiment target included in a sentence or text is the main aim of Aspect Based Sentiment Analysis (ABSA). This task's main challenge is to efficient extraction of a specific sentiment item's sentiment polarity. This work proposes a model namely Improved ABSA with Deep Belief Network-Recurrent Neural Network (DBN-RNN), which includes 3 working phases. Processes like stemming, stop word removal, lemmatization as well and tokenization are conducted in the initial pre-processing phase. Furthermore, in the aspect sentiment extraction phase, improved aspect term extraction (I-ATE) along with cosine similarity and word co-occurrence are used to extract the complex features from the pre-processed data. In the sentiment analysis phase, a hybrid classification model named DBN-RNN is utilized to effectively categorize the sentiments as neutral, positive, and negative polarities. The performance of proposed work is evaluated in terms of different performance measures.

Keywords: Aspect Based Sentiment Analysis, Stemming, Lemmatization, Improved aspect term extraction, Cosine similarity, Deep Belief Network (DBN), Recurrent Neural Network (RNN).

I. INTRODUCTION

In a global landscape where the pursuit of Sentiment Analysis (SA) is a field of Natural language Processing (NLP), which is also called opinion mining. This SA is an active area of research to show emotions and find the sentiments in the text automatically. Usually, SA is assumed as opinion polarity, whether anyone expressing neutral, positive, or negative sentiment about an event [6]. SA [7] has been utilized in nearly all domains like consumer products and services, social events, business organizations and influential groups, political elections and financial services, healthcare, and social events. SA[8][9] is classified at three levels: sentence level, document level, and aspect level to find which type (negative, positive, and neutral) of sentiment was expressed by a whether a sentence, document, or an aspect [10][11][12]. The ABSA assists in understanding the problem of SA reasonably since it mainly focuses on sentiments instead of the structure of language. SA is a text analysis approach that splits the text data and defines its sentiment depending on its aspects [13][14][15].

If an aspect is relevant to an entity, an aspect's basic concept is not restricted to judgment but also expands toward a way of thinking, point of view, thoughts, perspectives, or an underlying social influence or theme towards an occurrence. Therefore, a great opportunity was provided by the ABSA for analyzing the sentiments over time [16][17][18]. Designing a dependency relation among the aspect and its respective opinion expressions is the main point in the ABSA task solving. Furthermore, there might be numerous aspects along with various opinion expressions present in a sentence. For judging a certain aspect's sentiment, various RNNs were developed by previous studies, which include an attention mechanism for generating aspect-specific sentence representations [19][20][21]. Better results were achieved by them. But this attention mechanism becomes susceptible to the noise in the sentence [22][23].

More recent works developed based on Graph Convolutional Networks (GCN) and Graph Attention Networks (GAN) over dependency trees that utilize a sentence's syntactic structure. In a sentence, the relation between the words can be established by syntactic dependency. However, while applying the syntactic dependency to the ABSA tasks, two challenges occurred which were inaccuracy in the dependency parsing results and the GCNs over dependency trees will not function well with the datasets that are insensitive to the syntactic dependency [2][3][24][25]. Considering these shortcomings, a novel method named Improved ABSA with DBN-RNN model includes the following contributions.

- To propose an Improved ABSA model that utilizes the DBN-RNN hybrid classification process for effective sentiment analysis.
- To propose an Improved ATE (I-ATE) process to provide an effective sentiment analysis via extracting the primary and secondary aspects.

The proposed work is arranged as follows: Existing works are reviewed in section 2; the working process of ABSA with the DBN-RNN model is given in section 3, experiment along with comparison analysis is given in section 4, and in section this work gets concluded. Following that this work's references were provided.

II.

III. LITERATURE SURVEY

A few recent publications related to Aspect-based Sentiment analysis have been reviewed below.

[1*] Ph.D. Research Scholar, Shobhit Institute of Engineering and Technology (Deemed-to-be University), Meerut, India
[2] Director, Sir Chhotu Ram Institute of Engineering & Technology, C.C.S. University, Meerut, India
[3] JSS Academy of Technical Education, Noida, India

In 2023, Jingli Shi et al. [1] developed a new approach named Syntax-enhanced ABSA-Multilayer Attention (MA)" to manage all defined sub-tasks for ABSA. A multi-layer semantic model was created depending on the GCN, to capture the semantic relations among opinion and aspect terms. For learning the explicit dependency relations from diverse layers, a multi-layer syntax model was developed. The learned semantic features are transferred to a syntax model having superior semantic guidance to comprehensively learn the syntactic representations for facilitating the sub-tasks. Furthermore, two attention mechanisms were introduced in this work.
In 2023, Kanwal Ahmed et al. [2] proposed an Aspect-position and Entity-oriented Knowledge Convolutional Graph (APEKCG) model, which includes two modules like the Aspect position-aware module (APA) and the Entity oriented Knowledge Dependency Convolutional Graph (EKDCG). For sentiment classification, the data about aspect types into distinct parts of the context was incorporated to design an APA module to combine aspect-specific sentiment features. Utilizing a dependence graph, the dependency labels, entity-oriented knowledge, and syntactic path were incorporated by the EKDCG module. The proposed APEKCG framework's effectiveness was proved by the implementation outcomes on 5 NLP English language benchmark datasets.
In 2020, Baris Ozyurt et al. [3] proposed Sentence Segment (SS)-Latent Dirichlet Allocation (LDA), which is a topic modeling-dependent approach for aspect extraction in ABSA. There is no need for any annotated training data in this SS-LDA model since it was an unsupervised method. As a source, it needs a sentiment dictionary. Here, gathered Smartphone reviews, which were in the Turkish language are used as a dataset. Although SS-LDA was implemented in the Turkish language, this also can be applied to other languages only with little revisions. Experimental results ensured that extracting the product aspects, it is quite successful.
In 2023, Jingli Shi et al. [4] proposed the Soft Prompt-based Joint Learning (SPJL) approach for cross-domain ABSA. Especially, by integrating external linguistic features, this proposed SPJL approach learned the domain-invariant representations among source as well as target domains using several objectives that fill the gap between domains having multiple distributions. Moreover, a set of transportable soft prompts was interpolated by the proposed SPJL approach which included numerous vectors that are utilized to identify aspect terms in the target domain. The effectiveness of the proposed SPJL approach was demonstrated by the experimental results.
In 2023, Bin Liang et al. [5] proposed a GCN-based SenticNet which is named Sentic GCN, which utilizes the affective dependency in the sentence as per the specific aspect. In particular, a new solution was utilized to build the GCN via incorporating SenticNet's affective knowledge to improve the sentences' dependency graph. Depending on this, the new affective enhanced graph model considers the affective data among aspect and contextual words and the affective data among aspect and opinion words.
Some extant ABSA approaches utilized the attention mechanism for capturing the aspect-related context or GCN and dependency trees for modeling the semantic relationship among the aspect and its context. From the literature, it's proved that sentiment score determination and aspect extraction are the core challenges in ABSA. To increase the sentiment classification performance new approaches need to be developed. For that reason, an Improved ABSA with a DBN-RNN model is developed in this work to offer effective outcomes.

IV. Proposed Work

A text analysis approach that splits the text data and identifies its sentiment depending on its aspect is named ABSA. Text data can be analyzed in the ABSA, using DL approaches, and classified as different sentiment polarities like negative, positive, or neutral. As an input, raw text data i_x is given in our work. With this raw data pre-processing is conducted, which include processes like stemming, stop word removal, lemmatization as well and tokenization. Afterward, aspect sentiment extraction takes place, which uses improved aspect term extraction (I-ATE) along with semantic similarity and word co-occurrence matrix to extract complex features from that pre-processed data. Finally, the sentiment analysis process is conducted using the hybrid classification model named DBN-RNN, which classifies the sentiment using these extracted features. The mean of this classifier output was considered as the outcome. The detailed ABSA process is provided below and its architecture is illustrated in Fig. 1.

3.1 Pre-processing

Generally, a technique that can be performed on the raw data by including converting and cleaning processes to make it suitable for further processing. In this work, raw text data in_R is preprocessed via the processes like stemming, stop word removal, lemmatization, and tokenization. The detailed pre-processing phase has been described below.

1) Stemming

To convert the raw text data in_R into its root (stem) or base forms, the stemming process is conducted. A lookup table is included in this process, which assists in identifying the word and its corresponding base form. Due to its count computation drop in required levels, this is considered an advantageous process. For instance, the words walk, walked, walking and walks have the same base word walk.

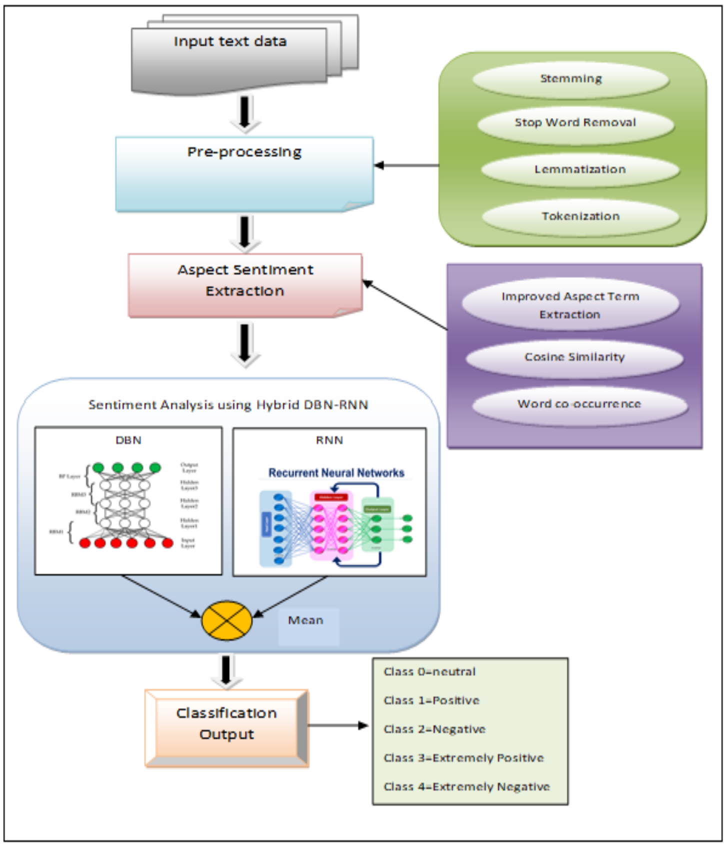


Fig. 1. Architecture of Improved ABSA with DBN-RNN model

2) Stop word removal

From raw text data in_R, commonly utilized words are removed in this process. Also, for SA, words aren't very informative. Furthermore, these commonly utilized words have less or no meaning. For example, *the, a, an, are, is, was, were, etc.*

3) Lemmatization

All words in in_R, are converted into their base form which is called as lemma. This is an advancement of stemming. This approach functions by considering language as well as contextual information that is more beneficial for our improved ABSA with the DBN-RNN model.

4) Tokenization

Transforming a sentence in the in_R, into tokens (collection of words). These tokens are the fundamental building blocks. Tokenization commonly interprets the way the word is expressed via analyzing the word sequence. Finally, the pre-processed text data is attained which is symbolized as in_R^P. As a result of this pre-processing phase, we got normalized, dimension-reduced, and noise-removed data in_R^P. With this in_R^P, an aspect sentiment extraction process is conducted.

3.2 Aspect Sentiment Extraction

From the pre-processed data in_R^P, crucial aspects and sentiments were extracted in this phase. This phase is important because it provides the important features to classify the sentiments. In this aspect sentiment extraction phase, features like improved aspect term extraction (I-ATE), cosine similarity, and word-occurrence-based features were extracted.

1) Improved Aspect term Extraction (I-ATE)

Extracting the aspect term is the initial step in the aspect sentiment extraction phase. If a sentence is given, then the ATE discovers all the aspect terms from that sentence. To provide an effective aspect term extraction, we have proposed I-ATE. Our (I-ATE) has two stages which are aspect term extraction and aspect groping.

a. Aspect Term Extraction

The task of discovering and retrieving terms relevant to SA is named aspect term extraction. Aspect term (AT) can be extracted by the following Eq. (1) [26].

$$AT = \frac{\exp(S_{sc})}{\sum_{k=1}^{n} \exp(S_{sc})} \tag{1}$$

In Eq. (1), Here S_{sc} is the score of the sentence and k=1,2,…n.
Our improved aspect term uses the following Eq. (2), instead of Eq. (1) to extract the aspect term.

$$AT = \frac{\exp(S_{sc}) \times Rank(S_{sc}) \times \max(S_{sci})}{\sum_{k=1}^{n} \exp(S_{sci})} \tag{2}$$

In (I-ATE), the score S_{sc} is calculated using the hybrid similarity function, which is the hybridization of the Sorensen-Dice Coefficient [27] and Levenshtein distance (edit distance) [28].
A statistic utilized to measure two sample's similarity is referred to as Sorensen-Dice Coefficient. Sorensen-Dice Coefficient has been expressed numerically in Eq. (3).

$$Sorensen - Dice = \frac{2 \times |J \cap K|}{|J| + |K|} \tag{3}$$

Here, J and K are two sets of input features, $|J \cap K|$ denoting the intersection of sets J and K Also, $|J|$, $|K|$ were the size of sets J and K.
The minimum count of single-character edits needed to convert one word into another is considered the Levenshtein distance between two words. To effectively measure the difference between two sequences, it is more useful. Levenshtein Distance has been expressed numerically in Eq. (4).

$$d[p,q] = \begin{cases} \max(p,q) & if \ \min(p,q) = 0 \\ d(p-1,q-1) & if \ Ap = Bq \\ \min \begin{cases} d(p-1,q-1)+1 \\ d(p,q-1)+1 \\ d(p-1,q)+1 \end{cases} & otherwise \end{cases} \tag{4}$$

After evaluating the Sorensen-Dice Coefficient and Levenshtein Distance values, they were averaged to get the hybrid similarity score.

$$Hybrid\ Similarity\ score(S_{sc}) = \frac{Sorensen - Dice\ Coefficient + Levenshtein\ Distance}{2} \tag{5}$$

By substituting this S_{sc} value in Eq. (2), we can get an effective aspect term extraction.
After the score calculation, the score rank $Rank(S_{sc})$ is calculated. To calculate the score rank, we have used the dense rank function in our (I-ATE). Dense rank is a window function, that provides the given row's number starting at 1 and afterward follows the window function's order by sequence with the same values to get the same outcomes. Score rank is calculated using Eq. (6).

$$Rank(S_{sc}) = dense\ rank(z_{i-1}) + 1 \tag{6}$$

Here, z_i denotes the current value and z_{i-1} denotes the previous distinct value. The steps used in evaluating the dense rank are listed below.
Step 1: Arrange the score values in ascending order
Step 2: Assign the rank of first to the smallest value (score)
Step 3: For tied value assign the same rank
Step 4: Increase the rank for the next distinct value.

By substituting Eq. (6) in Eq. (2), we can get the improved aspect terms. Afterward, these aspect terms are grouped. The groping process is described below.

b. Aspect grouping

After extracting the aspects, similar aspects will be grouped. To offer an effective aspect-based opinion summary, similar words should be grouped. In our work aspect grouping consists of two groups which were primary and secondary aspects. These primary and secondary aspects are differentiated based on the weights. These weights are estimated based on term frequency and mutual information function.

Term frequency is the measurement of how frequently a term occurs within a document. The easiest calculation is simply counting the number of times a word appears.

Mutual information is the amount of information that one random variable contains about another. In other words, it's a measure of their mutual dependence.

The weight is evaluated using Eq. (7).

$$Weight = TF\left(\delta_i, c_i\right) \times MI\left(\delta_i, a_i\right) \tag{7}$$

Eq. (7), c_i denotes the context of the aspect phrase a_i. Here, a_i denotes the improved aspect term extraction. The term $MI\left(\delta_i, a_i\right)$ denotes the mutual information between δ_i, a_i and $TF\left(\delta_i, c_i\right)$ denotes the count of δ_i in c_i.

Higher-weighted aspect phrases are considered primary aspects and lower-weighted aspect aspects are considered secondary aspects.

Afterward, with every input aspect word, the next aspect word is compared via evaluating the semantic similarity by the utilization of the co-reference PMI_score which is given in the following Eq. (8).

$$score\left(\omega_1, \omega_2\right) = \frac{S^{\beta}\left(\omega_1\right)}{\beta_1} + \frac{S^{\beta}\left(\omega_2\right)}{\beta_2} \tag{8}$$

Where, $S^{\beta}\left(\omega_1\right)$, $S^{\beta}\left(\omega_2\right)$ are the summation of all positive scores of the whole collection of semantically similar terms. Also, β_1 β_2 are summation of all positive scores of the whole collection of semantically similar terms, implies the existence of a word ω in the text. The score will be between (0, 1). If the score ranges between 0 to 0.5, that is selected as the secondary aspect and the score ranging between 0.5 to 1 is chosen as the primary aspect. Extracted I-ATE features were represented as $F_{(I-ATE)}$. Along with I-ATE features, cosine similarity, and word co-occurrence-based features were also extracted.

2) Semantic similarity or cosine similarity

The similarity between an inner product's two vectors is measured as cosine similarity. In our work, we have utilized this to measure the work similarity in the text.

$$F_{SS} = \frac{(D.E)}{\|D\|\|E\|} \tag{9}$$

In Eq. (9), $D.E$ =Vectors of D and E and $\|D\|\|E\|$ =Euclidean norms of vectors D and E. The extracted cosine similarity-based feature is symbolized as F_{SS}.

3) Word co-occurrence matrix

A basic concept of NLP which can be utilized to represent the relation among the elements in a text is known as word co-occurrence matrix. These elements might be words or phrases. As a result, the extracted word co-occurrence matrix-based feature is symbolized as F_{WC}.

The feature set which includes the extracted features is $FS = \left\{F_{(I-ATE)}, F_{SS}, F_{WC}\right\}$ and it is sent to the sentiment analysis phase for sentiment classification.

3.3 Sentiment Analysis using the DBN-RNN model

The process of analyzing the text to discover whether the emotional tone of the message is neutral, positive, or negative. We have utilized DBN-RNN, which is the hybridization of both DBN and RNN models, to get an accurate sentiment classification. For problems that involve sequential data, RNN is more useful. At the same time, for handling high-dimensional data and learning the data's hierarchical representations. RNN remembers past inputs due to an internal memory which is useful for sentiment analysis tasks. For that reason, DBN and RNN are utilized. Considering these advantages, we have hybridized these DBN and RNN techniques in our work. The feature set FS is given as the input to the DBN-RNN model.

1) DBN Architecture

DL models are created by the inspiration of the biological nervous system. This DL model has an input and an output layer along with multiple hidden layers. Through neurons or nodes, each layer has been linked. The previous layer's output is considered as each hidden layer's input. In DL, one of the classical algorithms is DBN. From the raw data, low to high-level and concrete-to-abstract features were automatically extracted by this DBN. Multiple Restricted Boltzmann Machine (RBM)s were included in this DBN. To initially fix the feed-forward neural network's parameters for enhancing the model's generalization ability, RBMs were generally utilized. Also, n neurons and m hidden layer neurons were included in an RBM network. The links among the nodes are present only between layers. From the classical thermal theory, RBMs were generated. If the energy function is small, then the system will be more stable. To attain the network's optimal parameters, the least energy of the network has been trained. Following Eq. (10), express the energy function.

$$E\left(L,M\right)=-\sum_{i=1}^{n}G_iL_i-\sum_{j=1}^{m}b_jM_j-\sum_{i,j}L_iM_jw_{ij} \tag{10}$$

Where L_i i and M_j represents the random state of the visible layer's *i*-th unit and the hidden layer's *j*-th unit. Respective biases were symbolized as G_i and b_j. The weight among two units was represented as w_{ij}. By training the network, we got the optimal parameters (w_{ij}, G_i, b_j).

RBM employs a greedy learning algorithm for optimizing the deep neural network's connection weight. DBN also uses the training process as RBM. During DBN training, it considers the network's every two layers, regarding it as a single RBM. A single RBM's weight and biases were trained afterwards hidden variables were created. These hidden variables are taken as the next RBM network's visible variable. When training of all RBM stacks is completed, we can assume the DBN's whole training process is completed. Finally, we can get the classified sentiments in the output layer, which are neutral, positive, negative, extremely positive, and extremely negative sentiments.

2) RNN Architecture

RNN is a kind of Neural Network (NN), where the previous step's output is given as the input to the current step. In the case of traditional NN, input and output were independent of each other. But in the case of predicting a word's next sentence, the previous word is needed. For that reason, it's necessary to remember the prior words. This issue was solved by RNN using a hidden layer. Then RNN's main and crucial feature is its hidden state. This hidden state remembers some sequence information. Since it remembers the network's previous input, it can be named as a memory state. For each input, it utilizes the same parameters and executes the same tasks on all hidden or input layers to offer the output, which lowers the difficulty of parameters.

Like other NNs, RNNs also have input and output architecture. The differences might occur depending on the data travel from input to output. For every dense network in RNN, there are different weight matrices. Also, the weight across the network is the same.

Consider SM as state matrix, which has an element SM_{Ti} as the network's state at Ti timestep. The hidden state H_{Ti} for all input X_{Ti} is evaluated using Eq. (11).

$$HI=\eta\left(UX+WtHI_{-1}+BI\right) \tag{11}$$

Here, the network's parameters were Wt, U, BI, and that were shared across Ti.

Using Eq. (12), we can calculate the current state:

$$HI_t=fn\left(HI_{t-1},X_t\right) \tag{12}$$

Where, HI_t denotes the current state, HI_{t-1} denotes the previous state and the input state is symbolized as X_t

For applying the Activation function σ, the following Eq. (13) can be utilized.

$$HI_t=\sigma\left(Wt_{HIHI}HI_{t-1}+Wt_{XHI}X_t\right) \tag{13}$$

In Eq. (13), Wt_{HIHI} denotes the weight at the recurrent neuron, and the weight at the input neuron is symbolized as Wt_{XHI}.

To evaluate the output, Eq. (14) can be utilized.

$$Y_t = Wt_{HIY} HI_t \tag{14}$$

Where, Y_t is output and the weight at output layer is symbolized as Wt_{HIY} .

The hybrid Formation: The mean is calculated for those output from RNN and DBN to provide accurate classification outputs, that is considered as the final output. The obtained output is neutral, positive, negative, extremely positive, and extremely negative sentiments.

V. Result And Discussion

4.1 Experimental Setup

A new Sentimental Analysis model was implemented using an Intel i5 CPU with 16 GB of RAM in Python 3.7.9. The DBN + RNN model uses coronavirus tweets NLP - Text Classification for training and testing. A few conventional techniques are contrasted with the DBN + RNN model, including Gated Recurrent Unit (GRU), Convolutional Neural Network (CNN), Bidirectional-Long Short Term Memory (Bi-LSTM), PyramidNet, and LSTM. The ultimate outcomes have been successfully validated.

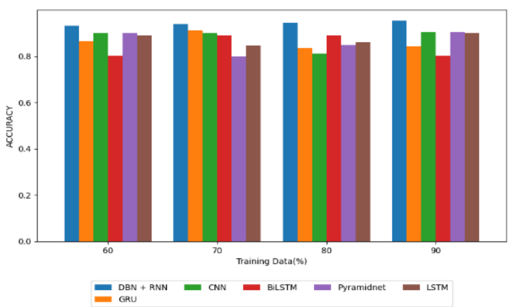
4.2 Dataset Description

NLP - Text Classification is a Corona Virus Tagged Dataset: Coronavirus tweets. The data is utilized to do text classification. After the tweets were taken down from Twitter, manual tagging was completed. To allay any privacy worries, codes have been assigned to the identities and names. Location, Tweet At, Original Tweet, and Label Make Up It. The dataset is 11.5 MB in total.

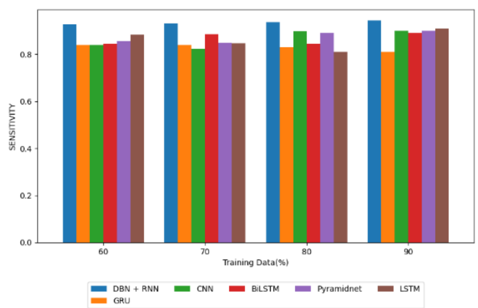
4.3 Comparative analysis on DBN + RNN and conventional methods

The effectiveness of the suggested strategy is assessed by contrasting it with conventional approaches and noting the outcomes based on several performance metrics. The DBN + RNN method is contrasted with pre-existing techniques for sentiment analysis, namely GRU, CNN, Bi-LSTM, PyramidNet, and LSTM. The DBN + RNN model is evaluated using a variety of positive and negative measures, including accuracy, sensitivity, specificity, precision, F-measure, Negative Predictive (NPV), and Mathew Correlation Co-efficient (MCC). The negative metrics are False Positive Rate (FPR) and False Negative Rate (FNR).

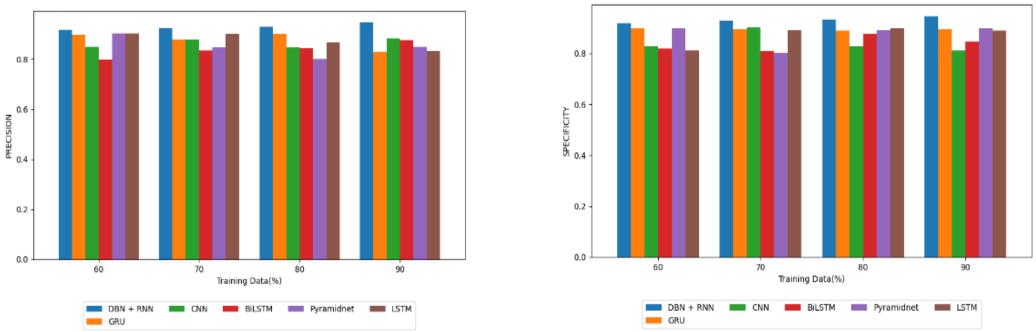
A few favorable metrics are used to compare the DBN + RNN model to the traditional techniques. The pre-existing models GRU, CNN, Bi-LSTM, PyramidNet, and LSTM are compared with the DBN + RNN method at varying learning percentages of 60, 70, 80, and 90. Fig. 2 demonstrates that the DBN + RNN model is clearly more accurate at evaluating feelings. While most conventional processes have accuracy rates below 91%, the DBN + RNN methodology achieves the greatest accuracy rate of 93% at learning percentage 70. Our ability to attain a higher accuracy rate than previous models can be attributed to the hybrid model that integrates RNN and DBN. In comparison to previous models, the DBN + RNN model obtains a 5% greater accuracy rate of sentimental analysis system at 90th learning percentage. In comparison, the specificity of the DBN + RNN approach is higher at 94% learning percentage 90. The models outperform conventional models in terms of accuracy, sensitivity, and precision. The DBN + RNN model achieves a sensitivity rate of 93% at learning percentage 80, while the other methods provide a sensitivity rate of less than 84%. When the DBN + RNN solution is compared to the traditional models, a higher percentage is found.



(a)



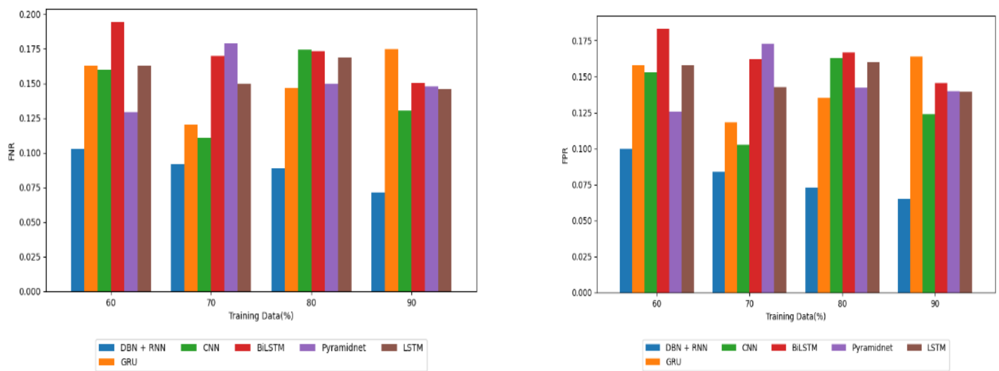
(b)



(c) (d)

Fig. 2. Assessment of the DBN + RNN approach's performance in terms of (a) accuracy, (b) sensitivity, (c) precision, and (d) specificity compared to existing schemes

FNR and FPR are useful in this study in terms of negative metrics to ascertain the efficacy of the DBN + RNN sentimental analysis model. Using both RNN and DBN, the model uses both to classify and identify the sentiment for error-free analysis. When compared to other conventional models, Fig. 3 clearly shows that the DBN + RNN approach produces low FNR and FPR values. Fig. 3(a) shows the FNR values for the different learning percentages of 60, 70, 80, and 90. At learning % 90, all earlier techniques perform better than the DBN + RNN model, suggesting that the DBN + RNN model has a smaller total error. At learning percentage 80, the FNR and FPR values are approximately 0.07. At the 90-learning percentage, the DBN + RNN model's FNR and FPR values are 0.07 and 0.06, respectively, while the FNR and FPR of earlier methods are 6% higher. Consequently, the DBN + RNN approach is more efficient and has an extremely low chance of error when interpreting human emotions.



(a) (b)

Fig. 3. Assessment of the DBN + RNN approach's performance in terms of (a) FNR (b) FPR

4.4 Performance Analysis

The DBN + RNN model for Sentimental analysis is developed with hybrid model DBN and RNN. The performance of the DBN + RNN model is assessed based on conventional technique. DBN + RNN model is contrasted to model with Conventional Aspect Term. Ablation assessment of the DBN + RNN approach is shown in Table 1.

The suggested model outperforms the model created with conventional methods. The suggested method, which preprocesses the input text by removing stop words, tokenizing, lemmatizing, and stemming, analyzes the sentiments in the text. DBN and RNN are utilized in the classification process after a few stages. The outcomes of the DBN + RNN ablation analysis are shown in Table 1. The accuracy of the DBN + RNN model is 93%, whereas the typical aspect term has an accuracy of about 85%. In addition, the sensitivity of the DBN + RNN model is 93% higher than that of the conventional model. The specificity of the DBN + RNN model is 92%, which is 10% higher than that of the conventional techniques. With a FNR error value of 0.09 and an FPR of 0.08 in comparison to the standard methods, the DBN + RNN model outperforms them in negative measurements. The DBN + RNN model outperforms the conventional techniques in sentiment analysis.

Table I: Ablation assessment of the DBN + RNN approach

Parameter	**Conventional Aspect Term**	**DBN + RNN**
Sensitivity	0.89219	0.93188
Specificity	0.81989	0.92876
Accuracy	0.85667	0.93827
Precision	0.90022	0.92436
F-Measure	0.84688	0.92838

MCC	0.82091	0.92737
NPV	0.87788	0.93287
FPR	0.13998	0.08387
FNR	0.14201	0.09187

VI. CONCLUSION

A higher amount of opinionated text was utilized by internet users, which makes the SA popular in both business and research. This ABSA is a complex task because it identifies both aspects and sentiments. The progress in this field urges the researchers to develop new approaches. Although DL techniques developed for ABSA, it's still in the early stage. As a result, an Improved ABSA with a DBN-RNN model was developed in this work, which includes 3 phases pre-processing, aspect sentiment extraction phase, and SA phase. Stemming, stop word removal, lemmatization as well and tokenization processes were conducted in the initial pre-processing phase to get clean and reduced dimension data. An improved aspect term extraction (I-ATE) was proposed in the aspect sentiment extraction phase along with cosine similarity and word co-occurrence matrix to extract complex features from that pre-processed data. In the sentiment analysis phase, DBN-RNN was utilized to effectively categorize the sentiments as neutral, positive, and negative polarities. To ensure the classification accuracy, the mean of this classifier output was taken as the final output.

REFERENCES

[1] Shi, J., Li, W., Bai, Q., Yang, Y. and Jiang, J., "Syntax-enhanced aspect-based sentiment analysis with multi-layer attention". Neurocomputing, 557, p.126730, (2023).

[2] Ahmed, K., Nadeem, M.I., Zheng, Z., Li, D., Ullah, I., Assam, M., Ghadi, Y.Y. and Mohamed, H.G., "Breaking down linguistic complexities: A structured approach to aspect-based sentiment analysis". Journal of King Saud University-Computer and Information Sciences, 35(8), p.101651, (2023).

[3] Ozyurt, B. and Akcayol, M.A., "A new topic modeling based approach for aspect extraction in aspect based sentiment analysis: SS-LDA". Expert Systems with Applications, 168, p.114231, (2021).

[4] Shi, J., Li, W., Bai, Q., Yang, Y. and Jiang, J., "Soft prompt enhanced joint learning for cross-domain aspect-based sentiment analysis". Intelligent Systems with Applications, 20, p.200292, (2023).

[5] Liang, B., Su, H., Gui, L., Cambria, E. and Xu, R., "Aspect-based sentiment analysis via affective knowledge enhanced graph convolutional networks". Knowledge-Based Systems, 235, p.107643, (2022).

[6] Mowlaei, M.E., Abadeh, M.S. and Keshavarz, H., "Aspect-based sentiment analysis using adaptive aspect-based lexicons". Expert Systems with Applications, 148, p.113234, (2020).

[7] Li, R., Chen, H., Feng, F., Ma, Z., Wang, X. and Hovy, E., "Dual graph convolutional networks for aspect-based sentiment analysis. In Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers) (pp. 6319-6329), (2021).

[8] Zhao, A. and Yu, Y., "Knowledge-enabled BERT for aspect-based sentiment analysis". Knowledge-Based Systems, 227, p.107220, (2021).

[9] Karimi, A., Rossi, L. and Prati, A., "Adversarial training for aspect-based sentiment analysis with Bert". In 2020 25th International conference on pattern recognition (ICPR) (pp. 8797-8803). IEEE, (2021).

[10] Song, W., Wen, Z., Xiao, Z. and Park, S.C., "Semantics perception and refinement network for aspect-based sentiment analysis. Knowledge-Based Systems, 214, p.106755, (2021).

[11] Liu, N. and Shen, B., "Aspect-based sentiment analysis with gated alternate neural network". Knowledge-Based Systems, 188, p.105010, (2020).

[12] Al-Dabet, S., Tedmori, S. and Mohammad, A.S., "Enhancing Arabic aspect-based sentiment analysis using deep learning models". Computer Speech & Language, 69, p.101224,(2021).

[13] Abdelgwad, M.M., Soliman, T.H.A., Taloba, A.I. and Farghaly, M.F., "Arabic aspect based sentiment analysis using bidirectional GRU based models". Journal of King Saud University-Computer and Information Sciences, 34(9), pp.6652-6662,(2022).

[14] Alamoudi, E.S. and Alghamdi, N.S., "Sentiment classification and aspect-based sentiment analysis on yelp reviews using deep learning and word embeddings". Journal of Decision Systems, 30(2-3), pp.259-281,(2021).

[15] Geetha, M.P. and Renuka, D.K., "Improving the performance of aspect based sentiment analysis using fine-tuned Bert Base Uncased model". International Journal of Intelligent Networks, 2, pp.64-69, (2021).

[16] Li, X., Fu, X., Xu, G., Yang, Y., Wang, J., Jin, L., Liu, Q. and Xiang, T., "Enhancing BERT representation with context-aware embedding for aspect-based sentiment analysis". IEEE Access, 8, pp.46868-46876, (2020).

[17] Kastrati, Z., Imran, A.S. and Kurti, A., "Weakly supervised framework for aspect-based sentiment analysis on students' reviews of MOOCs". IEEE Access, 8, pp.106799-106810, (2020).

[18] Ishaq, A., Asghar, S. and Gillani, S.A., "Aspect-based sentiment analysis using a hybridized approach based on CNN and GA". IEEE Access, 8, pp.135499-135512, (2020).

[19] Mohammadi, A. and Shaverizade, A., "Ensemble deep learning for aspect-based sentiment analysis". International Journal of Nonlinear Analysis and Applications, 12(Special Issue), pp.29-38, (2021).

[20] Aydin, C.R. and Güngör, T., "Combination of recursive and recurrent neural networks for aspect-based sentiment analysis using inter-aspect relations". IEEE Access, 8, pp.77820-77832, (2020).

[21] García-Díaz, J.A., Cánovas-García, M. and Valencia-García, R., "Ontology-driven aspect-based sentiment analysis classification: An infodemiological case study regarding infectious diseases in Latin America". Future Generation Computer Systems, 112, pp.641-657, (2020).

[22] Zhang, B., Li, X., Xu, X., Leung, K.C., Chen, Z. and Ye, Y., "Knowledge guided capsule attention network for aspect-based sentiment analysis". IEEE/ACM Transactions on Audio, Speech, and Language Processing, 28, pp.2538-2551, (2020).

[23] Shams, M., Khoshavi, N. and Baraani-Dastjerdi, A., "LISA: language-independent method for aspect-based sentiment analysis". IEEE Access, 8, pp.31034-31044, (2020).

[24] Li, N., Chow, C.Y. and Zhang, J.D., "SEML: A semi-supervised multi-task learning framework for aspect-based sentiment analysis". IEEE Access, 8, pp.189287-189297, (2020).

[25] Meškelė, D. and Frasincar, F., "ALDONAr: A hybrid solution for sentence-level aspect-based sentiment analysis using a lexicalized domain ontology and a regularized neural attention model". Information Processing & Management, 57(3), p.102211, (2020).

[26] Yang, H., Zeng, B., Yang, J., Song, Y. and Xu, R.,"A multi-task learning model for chinese-oriented aspect polarity classification and aspect term extraction". Neurocomputing, 419, pp.344-356, (2021).

[27] Binanto, I., Warnars, H.L.H.S., Abbas, B.S., Heryadi, Y., Sianipar, N.F. and Sanchez, H.E.P., "Comparison of similarity coefficients on morphological rodent tuber". In 2018 Indonesian Association for Pattern Recognition International Conference (INAPR) (pp. 104-107). IEEE, (2018),

[28] Khalid, M., Yousaf, M.M. and Sadiq, M.U., "Toward Efficient Similarity Search under Edit Distance on Hybrid Architectures". Information, 13(10), p.452, (2022).