J. Integr. Sci. Technol.2023, 11(4), 563

Article

Journal of Integrated
SCIENCE& TECHNOLOGY

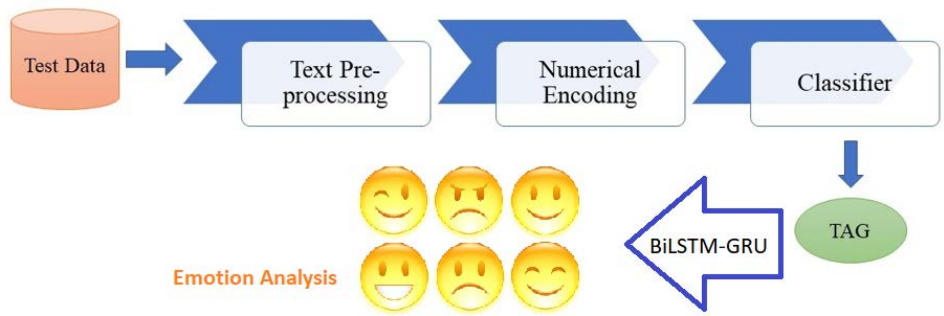
Attention-Emotion-Embedding BiLSTM-GRU network based sentiment analysis

Bhavna Kabra,* Chetan Nagar

Institute of Computer Application, Sage University, Indore, India.

Received on: 13-Dec-2022, Accepted and Published on: 11-Apr-2023

ABSTRACT



Internet-based resource, which includes social forums, review sites, blogs and networks generates enormous heaps of data in the form of views, users' feelings, thoughts, and disagreements on numerous things like brands, politics, and social events. This data can be found in the form of "user-generated content." The feelings of people who express themselves online have a significant impact not just on readers but also on those who sell products and on politicians. It is necessary to assess and well-structure the digital evidence that comes through social networks and websites, and sentiment analysis has attracted a lot of attention for this use. Sentiment analysis or text organization, is used to categorize the conveyed mentality or feelings in numerous ways, such as affirmative, favorable, unpleasant, thumbs up, hand gestures, negative, etc. Sentimental evaluation is also referred to as a "thumbs up" or "thumbs down" rating. Within the realm of natural language processing, the problem of insufficiently labeled data presents a hurdle for sentiment analysis. And as a solution to this problem, the techniques of sentiment analysis and deep learning have been combined. This was done using deep learning models because of their ability to automatically learn new information. Therefore, this paper integrated the deep learning approach for domain-independent sentiment analysis. The paper presents an Attention based Emotion-Embedding BiLSTM-GRU Network for sentiment analysis. The paper presents comparative training accuracy and loss analysis with four baseline models. The network shows an accuracy of 93% which is higher as compared to baseline models and also achieved predictive accuracy compared to cutting-edge models.

Keywords: Attention Network, Deep Learning, Emotion Embedding, Sentiment Analysis.

INTRODUCTION

The proliferation of user-generated content across social networks and websites such as Amazon, Facebook, Twitter, Instagram and Trip Advisor has led to the transformation of social networks into the preeminent forum for the expression of opinions regarding a variety of topics, including but not limited to events, products, and services.[1] The management of feelings, thoughts, and material that is subjective is what is meant by the term "sentiment analysis".[1] As a result of analyzing a variety of tweets and evaluations, sentiment analysis makes it possible to comprehend facts concerning the views held by the general public. It is a tried and tested method for forecasting a wide variety of noteworthy occurrences, including election results and the performance of movies at the box office.[2] Anyone can find public reviews of a certain entity, such as a product, location, or person, on various websites like Amazon and Yelp. These reviews are employed to assess the entity in question. The opinions might be regarded as either positive, negative, or neither. The intention behind using sentiment analysis, mechanically ascertain the expressive slant of user reviews.[3] The need for uncovering and organizing previously concealed information, which is data storage that comes from social media, has led to a growth in the demand for sentiment analysis.[4] A deep learning method subfield of machine learning that pertains to deep neural networks. It was initially conceived by G.E. Hinton in 2006 and was given its current name at that time. The human brain has a significant impact on the neural network, which consists of numerous neurons that combine to form an astonishing network. The supervised and unsupervised learning categories can both

*Corresponding Author: Ms. Bhavna Kabra, Research Scholar, Sage University, Indore
Tel: 7747883040
Email: bhavna_chandak@yahoo.com

Cite as: *J. Integr. Sci. Technol., 2023, 11(4), 563. URN:NBN:sciencein.jist.2023.v11.563*

 ISSN: 2321-4635
http://pubs.thesciencein.org/jist

benefit from the training that can be provided by deep learning networks. Deep learning incorporates a wide variety of neural networks, including RNN (Recurrent Neural Networks), Recursive Neural Networks, as well as DBN (Deep Belief Networks), CNN (Convolutional Neural Networks) and many others. Word description estimation, sentence categorization, vector depiction, phrase modeling, feature presentation, and text creation are all areas in which neural networks have proven to be extremely useful. The term "deep architecture" refers to a system that is made up of multiple tiers of non-linear operations. Due to its power of simulating, it is projected that the training algorithm would do well in semi-supervised development, such as DBN, and will achieve notable success in the field of machine learning. Deep architecture is forced to conduct well in several tasks of hard AI technologies. In-depth learning has a significant impact on both unsupervised and supervised learning; as a result, a growing number of researchers are turning to deep learning to tackle sentiment analysis. It is made up of a great number of well-known and efficient models, and all of these models are put to use to efficiently address a wide range of issues.

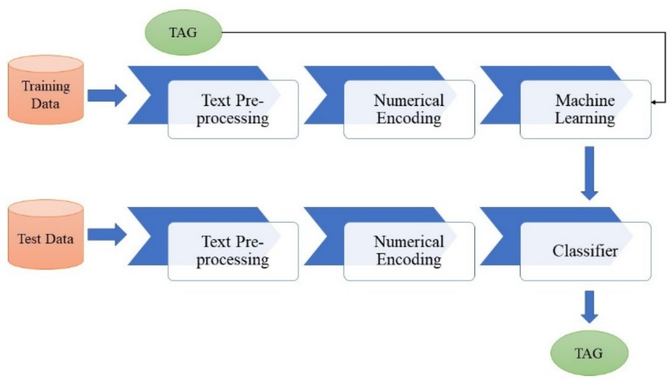


Figure 1. Sentiment Analysis Basic Workflow

LITERATURE REVIEW

The article by Habimana et. al.[1] provides an overview of deep learning previously employed algorithms to a variety of sentiment analysis applications as well as their current and future developments. After each job involving sentiment analysis, this study also presents an examination of the performance of various deep learning models applied to a specific dataset. In the final section, the review focuses on contemporary problems and suggests potential solutions that should be taken into account in subsequent research. The use of bidirectional encoder representations from transformers, cognition-based attention models, common sense knowledge, reinforcement learning, generative adversarial networks and sentiment-specific word embedding models are some of the suggestions that have been provided.

Attention-based Bidirectional Gated Recurrent Unit (BiGRU) and Convolutional Neural Network (CNN) are combined in the new model for sentiment analysis that Yang et al.[2] proposes. This model is called SLCABG and it is founded on the sentimental vocabulary. This document crawls and cleans the real book evaluation of dangdang.com for training and certification, all of which are based on Chinese, which is a well-known Chinese e-commerce website. The data have achieved a scale of 100000 orders of magnitude, which enables them to have a wide range of applications in the field of Chinese sentiment analysis. The results of the experiments indicate that the model has the potential to significantly increase the effectiveness of text sentiment analysis.

Wang et al.[3] offers their idea of sentiment as a potential solution to the issue. Then, using the multi-semantics sentiment intensity lexicon that we built for this paper to achieve correct semantics and encoding of word attitudes, as well as proper incorporation of sentiment data, we were able to obtain the information regarding the sentiments of individual words that were appropriate for the optimal sentiment concept. The reliability of the word hidden Markov method on sentiment concept that was developed in this paper is validated when it is contrasted with traditional and sentiment embeddings methods on six datasets that are typical of the whole.

Liu et al.[4] provide a synopsis of the recently presented strategies for the resolution of an aspect-based sentiment analysis problem. There are now three methods that are considered to be mainstream: lexicon-based methods and classic machine learning deep learning techniques, too. They gave a comparative examination of deep learning methods in their article paper that is considered to be state-of-the-art. Introduced here are some widely employed benchmark data sets, evaluation measures, and the performance of the many deep learning approaches that are currently in use.

Affendi et al.[5] proposed an innovative deep learning-based multilevel parallel attention neural (MPAN) model. This model makes use of a straightforward positioning binary embedding scheme (PBES) to simultaneously compute contextualized character, word, and statement embeddings. Steadily for the last but not least, the public IMDB movie review dataset is used to further validate the MPAN effectiveness of the algorithm. Its accuracy on this dataset is 96.13%, which puts it in second place overall on the IMDB scorecard.

Zhou et al.[6] provided a comprehensive overview of the deep learning-based methods that are currently considered to be state-of-the-art. These methods demonstrate the tremendous progress that has already been made in ASC. Then, we compile all of the standard datasets that are used as benchmarks by the ASC for researchers to examine, and we do extensive tests using five public datasets that are standardized according to various evaluation criteria. In conclusion, we will talk about some of the open challenges that are the most difficult to solve and point out some intriguing future research avenues in this area.

To decrease the need for manual weak tagging data is integrated into the deep learning emotions categorization model based on tagging data. Information into the training process of the model was proposed by Wang et. al.[7] According to the experiments, In the SVM classification job of hotel internet reviews, the deep learning sentiments classification algorithm based on weak labeling information performs better than the traditional deep model without increasing the required amount of human labor.

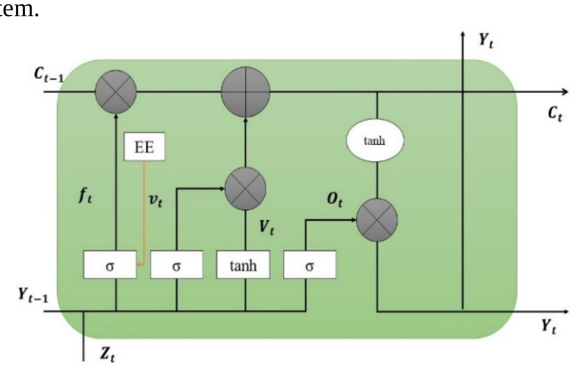
A unique approach to sentiment analysis that is based on deep learning has been proposed by Studiawn et al.[8] to determine whether or not OS logs include any unusual behaviors. Log messages are treated as sentences, and gated recurrent unit (GRU) networks are used to determine the attitudes included within the messages. OS log collections have a misclassification that is inherent in that the proportion of hurtful stereotypes is much smaller than that of positive sentiments. This means that the proportion of positive sentiments to negative sentiments is much higher. We develop a GRU layer on top of a class imbalance solution using the Tomek link approach so that we can handle the class imbalance. The experiments' findings demonstrate that the suggested technique has an overall F1 and consistency of 99.84 and 99.93%, respectively, when it comes to identifying anomalous events in operating system logs.

Shehu et al.[9] discussed to expand the training data, three data replenishment methods—Shift, Shuffle, and Hybrid are used. Then, utilizing three essential deep learning (DL) models known as recurrent neural network (RNN), convolution neural network (CNN), and hierarchical attention network (HAN) to classify the stemmed Turkish Twitter data. The TML significantly outperforms the DL models regarding both the training-time (TTM) and runtime (RTM) complexities of the methods; however, the DL models outperform the TML models in the most important quality factors as well as the average performance rankings. This conclusion is based on the algorithms, experimental, and mathematical results analysis designating identical data points. This conclusion was reached because identical datasets were used for the simulation, experimental, and statistical

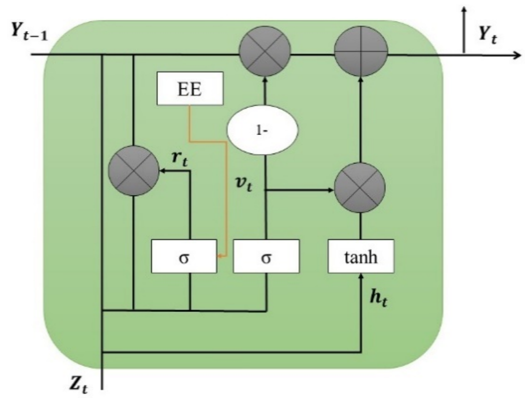
To capitalize on the indirect connection that exists between the aspect terms and the opinion terms, Fatima et al.[10] suggested a deep learning-based multilayer dual-attention model. In addition, unlike the Word2Vec model, the word embeddings are improved by assigning unique vector representations to dissimilar feelings. We overcame the issue of similar vector representations of opposing sentiments by using a word embedding model that has already been trained and viewpoint refinement. This was done to accomplish the aforementioned goal. The proposed model's performance is evaluated using three benchmark datasets from the SemEval Challenges in 2014 and 2015. The findings of our experiments suggest that our model is superior to other models currently considered to be state-of-the-art when it comes to doing aspect-based sentiment analysis. Huang et al.[11] presented a brand-new parameterized CNN for sentiment classification at the aspect level. Using SemEval 2014b datasets, experiments show that our parameterized filters and parameterized gates successfully capture the aspect-specific features and produce excellent results. Gan et al.[12] proposed a separable dilated CNN with sparse attention as its basis. Last but not least, studies on three benchmark datasets demonstrate that SA-SDCCN offers equivalent or even better performance in terms of higher parallelism and lower computational cost than state-of-the-art approaches. Wang et al.[13] suggested a novel Unified Position-aware CNN. Our model executes both ACSA and ATSA tasks successfully, according to experimental results on seven datasets. Liu et al.[14] suggested the Gated Alternate Neural Network as a new type of neural network structure. The experimental results demonstrate that GANN achieves cutting-edge results and demonstrates the language independence of our suggested model. Shuang et al.[15] introduced the Double-Gate mechanism. Two of the ABSA subtasks, the ACSA subtask, and ATSA subtask, are completed by the FDN model with state-of-the-art results. Jiang et al.[16] proposed Bidirectional Enhancement Transformation (BET) component to enhance aspect representation learning while achieving alternative aspect and context learning. To train attention in contexts and targets interactively and create different representations for targets and contexts, Ma et al.[17] suggested interactive attention networks (IAN). Experimental results show that our model works as expected findings on SemEval 2014 Datasets. Xu et al.[18] proposed a multi-attention network (MAN) whereas the local attention module takes interactions at the word level into account; this was frequently overlooked in earlier studies and focuses on the entire relation. Experiments show that the proposed model outperforms the baseline models in terms of results. Yang et al.[19] proposed a multi-task learning model, called LCF-ATEPC, for aspect-based sentiment analysis that is Chinese-oriented. Experimental results on the two SemEval-2014 task-4 datasets that are most frequently used, Eatery and Laptop, outperformed the most recent results on the ATE and APC subtask.

MODEL DESCRIPTION

In this work, we developed a RNN-based model for analyzing the tone of written content by discovering hidden patterns of coherence between words. However, Conventional RNN suffers from the issue of exploding or gradient vanishing, it is unable to effectively capture a sentiment or semantic dependence among words with large distances. A long short-term memory network (LSTM) was developed to address the problems, and it has been proven to be effective in learning sequence representations. According to the diagram in Figure 2(a), LSTMs are an artificial memory architecture designed to mimic the way the language processing functions and the human brain's memory system. However, while updating data in the LSTM block's memory cell, existing LSTM networks don't consider the impact of emotion. A different type of long short-term memory (LSTM), the Gradient Reducing Unit (GRU) is shown in Figure 2(b) to efficiently address gradient expansion and gradient disappearance in RNNs. In contrast to the LSTM, the forget gate and the input gate have been merged into a single "update gate," simplifying the GRU's construction. Due to not having to use, when dealing with large amounts of training data, the GRU can save a lot of time by using a gate and doing reduced multiplication operations. A new model called the emotion-embedded BiLSTM-GRU (EEBiLSTM-GRU) is presented to improve the simulation of the brain's memory system.



a. LSTM Unit with Emotion Embedding Unit



b. GRU Unit with Emotion Embedding Unit

Figure 2. EEBiLSTM-GRU Model Units

As part of this study, we present a stacked EEBiLSTM-GRU model for sentiment analysis prediction. The combined prediction model's network structure is depicted in Figure 2. The first step is to pick a GRU layer for training data that has a simple network structure, fewer parameters, and a more direct route to convergence. In comparison to LSTM, GRU's accuracy in making predictions is lower, but it trains faster and requires less time overall. Given that a double-layer LSTM improves upon the performance of a single-layer LSTM in terms of prediction accuracy.

The emotion state estimation EEBiLSTM-GRU model consists of the following components: a forget gate, an output gate, a hidden state, a cell state, an emotional state, a memory cell, an emotion-based embedding unit (EE), an embedding layer, and an input gate. All of the gates and the block input are looped together, and the output of EEBiLSTM-GRU is fed back into itself. It is possible to provide a more formal description of EEBiLSTM-workflow GRUs as follows:

$$\alpha_t = \tanh(W_a x_t + U_a h_{t-1} + b_a) \quad (1)$$

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \quad (2)$$

$$e_t = EE(es) \quad (3)$$

$$f_t = \mathrm{e}_t \ \mathrm{o}\ \sigma(W_f x_t + U_f h_{t-1} + b_f) \quad (4)$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \quad (5)$$

$$c_t = f_t \ o\ c_{t-1} + i_t \ o\ a_t \quad (6)$$

$$h_t = o_t \ o \tanh c_t \quad (7)$$

Where, σ = activation function such as sigmoid function

o = product operator

h_{t-1}= previously hidden layer value

x_t= current input

b = bias

i_t= input gate

f_t= forget gate

o_t= output gate

c_t= state of memory at current timestamp

e_t = emotion embedding factor from EE unit.

From this calculation, we can deduce that the cell state is crucial to EEBiLSTM-GRU. By controlling the gates, EEBiLSTM-GRU can either delete or add to the cell's state. The rule has four distinct phases: the forgetting phase, the emotion estimating phase, the updating phase, and the output phase.

When given input x t, EEBiLSTM-GRU seeks to forget some of the information about the cell state. We modify the forget gate of conventional LSTM to allow the emotion signal to strike a balance between the use of past knowledge (ht1) and the investigation of new information (x t) at each iteration of the learning process. Using the formula (4), the following is a description of the decision-making procedure: Starting with the current input x t and the output h (t-1) of the previously hidden layer, EEBiLSTM-GRU uses a sigmoid function to create a number in the range [0, 1]. Information about how likely it is that "information in the cell state is to be kept" will be output as a single number. Furthermore, the e t signal from the emotion embedding unit EE is used to further adjust the sigmoid activation function's output. In this case, the Hadamard product operator is used to combine the input and output signals (e t and e t') because of the presence of short routes, which is necessary for modulation to take place. During the emotion estimating phase, the embedding unit makes use of the proposed emotion embedding units to try to understand the distribution of its emotional state.

EEBiLSTM-GRU tries to combine the input gate and a tanh function to determine what new information should be stored in the cell state during the update phase. To update equation (2), EEBiLSTM-GRU first looks at the values of the input gate and then uses the tanh function to generate a vector of potential new values c t for the cell state. The equation (6) is then used to update the cell's state information. After processing the input sequence x t, EEBiLSTM-GRU determines the hidden state h t that best fits the data. To selectively output only the parts of the cell state that it has decided to output, EEBiLSTM-GRU first uses the output gate o t to select which parts to output, as shown in equation (5) and then applies the tanh function to the cell state and multiplies the result by the output as stated in equation (7).

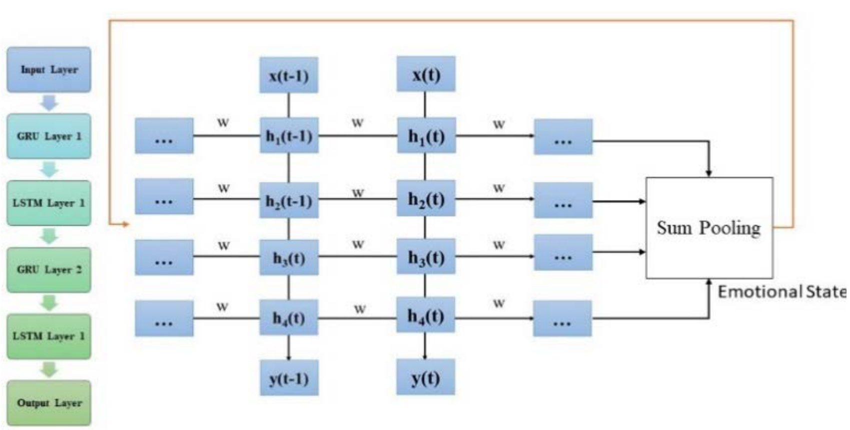


Figure 3. EEBiLSTM-GRU Architecture

Domain Attention Mechanism

To boost the effectiveness of deep learning models, attention mechanisms are now being employed to zero in on particular segments of the input sequence. Furthermore, current approaches ignore the objectively existing connection between domain and sentiment. Word embedding methods like Word2Vec and GloVe allow for a sequence of the text of length L to be represented as a feature vector. To model the domain, it is fed into an attention-based EEBiLSTM-GRU network. Domain modeling is defined as a technique for discovering latent structure in a corpus of texts within the natural language processing, field. In domain modeling, the "domain" is the resultant pattern of often occurring co-occurring terms. We provide a domain attention mechanism to modify a text's sentiment polarity membership. Calculating attention weights involves factoring in the n-gram domain distribution to the attentive representations. In particular, the domain attention weight is found by plugging in the domain vector $D_i$and the hidden vectorh_i of the i-th n-gram of the input text, respectively:

$$\boldsymbol{\eta_i} = \frac{\mathbf{exp}(\boldsymbol{h_i} \odot \boldsymbol{D_i})}{\sum_{k=1}^{\frac{L-I+1}{2}} \mathbf{exp}(\boldsymbol{h_k} \odot \boldsymbol{T_k})} \qquad (8)$$

Where,

$L - I + 1$ = Feature Map Size

In the last softmax classifier is used to perform the classification task and cross-entropy is used as the loss function.

Result And Discussion

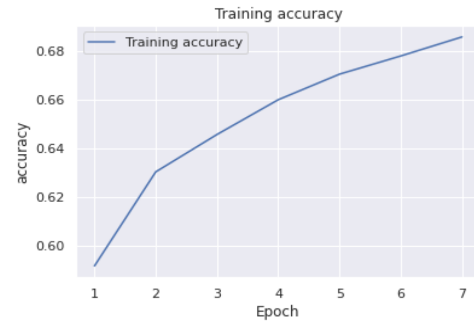
This paper has implemented and trained the models in the Keras framework with TensorFlow. The proposed model was trained using GPU on google colab. Following performance parameters are used to evaluate the model's efficiency in terms of Accuracy, precision, and recall.

$$Accuracy = \left(\frac{TP + TN}{TP + TN + FP + FN}\right) * 100 \qquad (9)$$

Where, TP= True Positive, FP= False Positive, FN= False Negative and TN = True Negative

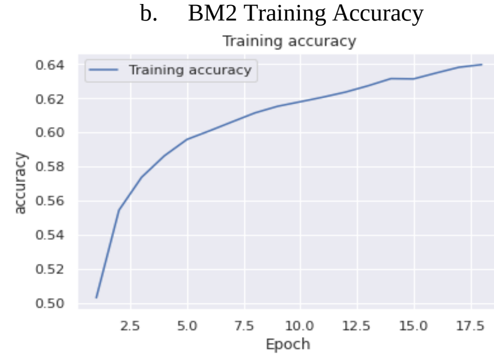
For result analysis, the paper collected reviews from multiple domains such as IMDB reviews [20], tweets [21], and rotten tomato review dataset [22] and prepared a common dataset for training and testing.

In this section, first of all, the paper presents the ablation study on baseline models. The first baseline model is presented with the implementation of the classic LSTM model (BM1). GRU network is considered the second baseline model as BM2. BiLSTM is considered BM3. Then the hybrid model of BiLSTM-GRU is considered a BM4. Finally, attention-based emotion embedding BiLSTM-GRU is considered as baseline model BM5. The training accuracy and loss graphs of all baseline models are presented below in Figure 4 and Figure 5. From Figure 4 it was observed that the BM1 model has achieved an accuracy of 68%, BM2 has achieved an accuracy of 70%, BM3 has achieved an accuracy of 64% and BM4 model has achieved an accuracy of 68% and finally, the proposed model BM5 have achieved highest accuracy of approx. 92%. Similarly, for training loss, the model has achieved the lowest loss as presented in Figure 5.

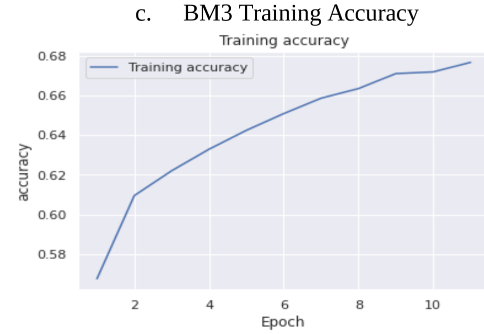


a. BM1 Training Accuracy

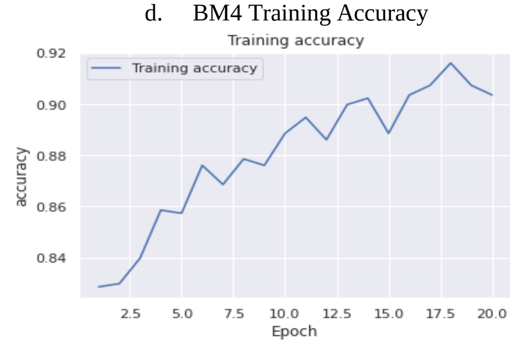
b. BM2 Training Accuracy



c. BM3 Training Accuracy



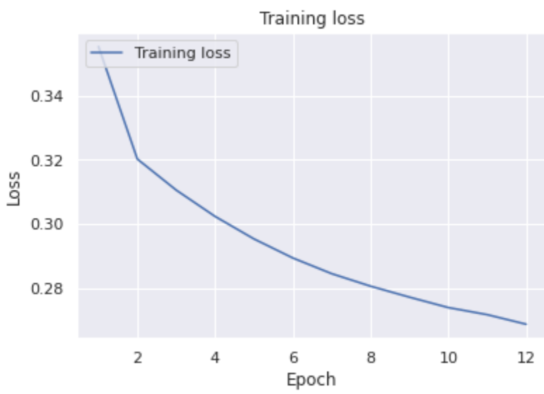
d. BM4 Training Accuracy



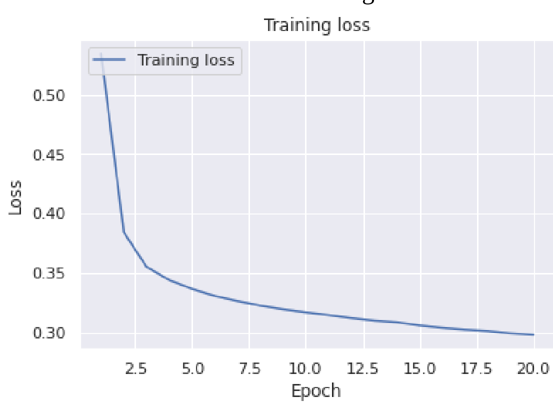
e. BM5 Training Accuracy

Figure 4. Training Accuracy of Baseline Models

a. BM1 Training Loss



b. BM2 Training Loss



c. BM3 Training Loss

d. BM4 Training Loss

e. BM5 Training Loss

Figure 5. Training Loss of Baseline Models

In Table 1, the paper presents the performance of EEBiLSTM-GRU concerning accuracy, precision, recall, and f1_score. The model achieved a testing accuracy of 93%, precision of 91%, recall of 94%, and f1_score of 92%. The comparative state-of-art of the model is presented in Table 4. In [23], the author presented the PCA algorithm with NLP for sentiment analysis and achieved approx. 76% of accuracy. In [24], DNN was used for sentiment analysis and achieved an accuracy of approx. 86%, In [25], the ensemble model is presented with an accuracy of 81%, in [26], the LSTM model was integrated with CNN for sentiment analysis and achieved an accuracy of 89% of accuracy, In [27], multi-modal sentiment analysis model was presented and achieved an accuracy of approx. 91%. In [28], BiERU model was presented that used an emotion intelligence unit but achieved an accuracy of approximately 65%. Whereas, in the developed model, the attention mechanism is added with an emotion embedding unit with a hybrid model of BiLSTM-GRU network and improvised the accuracy up to 93%. Therefore, EEBiLSTM-GRU proposed method is more accurate as compared to existing methods of sentiment analysis.[29]

Table 1. Performance Comparative of Baseline Model

Parameters	EEBiLSTM-GRU
Accuracy	93
Precision	91
Recall	94
F1_score	92

Table 4. Performance Comparative State-of-Art

Ref	Methodology	Accuracy
[23]	PCA	76.55%
[24]	DNN	85.97%
[25]	Feature ensemble model	81%
[26]	Deep Learning LSTM+ CNN	89%
[27]	Multimodal SA	91.39%
[28]	Bidirectional Emotional Recurrent Unit (BiERU)	65.93%
Ours		93%

CONCLUSION

The objective of sentiment analysis, a widely studied NLP task, is to ascertain users' opinions, feelings, and evaluations of the product, entity, or service they are reviewing. The most frequently employed techniques include word embeddings, sentiment lexicons, and even annotated data. Additionally, it takes a lot of time and effort to optimize models for each language, especially for recurrent neural network models. As a result, the deep learning methodology was incorporated into this article to do domain-independent sentiment analysis. An attention-based emotion-embedding BiLSTM-GRU Network for sentiment analysis is presented in this research. Comparative training accuracy and loss analysis using four different baseline models are presented in this research. The network has obtained better performance when compared to state-of-the-art models, and it demonstrates an accuracy of 93%, which is more than what is achieved by baseline models.

Acknowledgement

The authors are thankful to the college for providing the necessary facilities to carry out this research.

Conflict of Interest

The authors declare that they have no conflict of interest.

List of Abbreviations

BiERU = Bidirectional Emotional Recurrent Unit
BiGRU = Bidirectional Gated Recurrent Unit
CNN = Convolutional Neural Networks
DBN = Deep Belief Networks
DL = Deep Learning
DNN = Deep Neural Network
GRU = Gated Recurrent Unit
HAN = Hierarchical Attention Network
LSTM = Long Short-Term Memory
RNN = Recurrent Neural Networks
RTM = Runtime
TML = Traditional Machine Learning
TTM = Training-time

References

1. O. Habimana, Y. Li, R. Li, X. Gu, G. Yu. Sentiment analysis using deep learning approaches: an overview. *Sci. China Inf. Sci.* **2020**, 36.
2. L. Yang, Y. Li, J. Wang, R.S. Sherratt. Sentiment Analysis for E-Commerce Product Reviews in Chinese Based on Sentiment Lexicon and Deep Learning. *IEEE Access* **2020**, 8, 23522–23530.
3. Y. Wang, G. Huang, J. Li, et al. Refined Global Word Embeddings Based on Sentiment Concept for Sentiment Analysis. *IEEE Access* **2021**, 9, 37075–37085.
4. H. Liu, I. Chatterjee, M. Zhou, X.S. Lu, A. Abusorrah. Aspect-Based Sentiment Analysis: A Survey of Deep Learning Methods. *IEEE Trans. Comput. Soc. Syst.* **2020**, 7 (6), 1358–1375.
5. M.A. El-Affendi, K. Alrajhi, A. Hussain. A Novel Deep Learning-Based Multilevel Parallel Attention Neural (MPAN) Model for Multidomain Arabic Sentiment Analysis. *IEEE Access* **2021**, 9, 7508–7518.
6. J. Zhou, J.X. Huang, Q. Chen. Deep learning for aspect-level sentiment classification: Survey, vision, and challenges. *IEEE Access* **2019**, 7, 78454–78483.
7. C. Wang, X. Yang, L. DIng. Deep Learning Sentiment Classification Based on Weak Tagging Information. *IEEE Access* **2021**, 9, 66509–66518.
8. H. Studiawan, F. Sohel, C. Payne. Anomaly Detection in Operating System Logs with Deep Learning-based Sentiment Analysis. *IEEE Trans. Dependable Secur. Comput.* **2020**.
9. H.A. Shehu, M.H. Sharif, M.H.U. Sharif, et al. Deep Sentiment Analysis: A Case Study on Stemmed Turkish Twitter Data. *IEEE Access* **2021**, 9, 56836–56854.
10. S. Rida-E-Fatima, A. Javed, A. Banjar, et al. A Multi-Layer Dual Attention Deep Learning Model with Refined Word Embeddings for Aspect-Based Sentiment Analysis. *IEEE Access* **2019**, 7, 114795–114807.
11. B. Huang, K.M. Carley. Parameterized convolutional neural networks for aspect level sentiment classification. *Proc. 2018 Conf. Empir. Methods Nat. Lang. Process. EMNLP 2018* **2018**, 20182018, 1091–1096.
12. C. Gan, L. Wang, Z. Zhang, Z. Wang. Sparse attention based separable dilated convolutional neural network for targeted sentiment analysis. *Knowledge-Based Syst.* **2020**, 188, 104827.
13. X. Wang, F. Li, Z. Zhang, et al. A unified position-aware convolutional neural network for aspect based sentiment analysis. *Neurocomputing* **2021**, 450, 91–103.
14. N. Liu, B. Shen. Aspect-based sentiment analysis with gated alternate neural network. *Knowledge-Based Syst.* **2020**, 188, 105010.
15. K. Shuang, Q. Yang, J. Loo, R. Li, M. Gu. Feature distillation network for aspect-based sentiment analysis. *Inf. Fusion* **2020**, 61, 13–23.
16. B. Jiang, J. Hou, W. Zhou, et al. METNet: A Mutual Enhanced Transformation Network for Aspect-based Sentiment Analysis. *COLING 2020 - 28th International Conference on Computational Linguistics, Proceedings of the Conference*. **2020**, pp 162–172.
17. D. Ma, S. Li, X. Zhang, H. Wang. Interactive attention networks for aspect-level sentiment classification. *IJCAI International Joint Conference on Artificial Intelligence*. **2017**, pp 4068–4074.
18. Q. Xu, L. Zhu, T. Dai, C. Yan. Aspect-based sentiment classification with multi-attention network. *Neurocomputing* **2020**, 388, 135–143.
19. M. Zhao, J. Yang, L. Qu. A multi-task learning model with graph convolutional networks for aspect term extraction and polarity classification. *Appl. Intell.* **2023**, 53 (6), 6585–6603.
20. A. Oghina, M. Breuss, M. Tsagkias, M. De Rijke. Predicting IMDB movie ratings using social media. *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* **2012**, 7224 LNCS, 503–507.
21. Ankit, N. Saleena. An Ensemble Classification System for Twitter Sentiment Analysis. *Procedia Comput. Sci.* **2018**, 132, 937–946.
22. P. Tiwari, B.K. Mishra, S. Kumar, V. Kumar. Implementation of n-gram methodology for rotten tomatoes review dataset sentiment analysis. *Cognitive Analytics: Concepts, Methodologies, Tools, and Applications*. **2020**, pp 689–701.
23. N. Zainuddin, A. Selamat, R. Ibrahim. Hybrid sentiment classification on twitter aspect-based sentiment analysis. *Appl. Intell.* **2018**, 48 (5), 1218–1232.
24. Z. Jianqiang, G. Xiaolin, Z. Xuejun. Deep Convolution Neural Networks for Twitter Sentiment Analysis. *IEEE Access* **2018**, 6, 23253–23260.
25. H.T. Phan, V.C. Tran, N.T. Nguyen, D. Hwang. Improving the Performance of Sentiment Analysis of Tweets Containing Fuzzy Sentiment Using the Feature Ensemble Model. *IEEE Access* **2020**, 8, 14630–14641.
26. M.A. Ullah, S.M. Marium, S.A. Begum, N.S. Dipa. An algorithm and method for sentiment analysis using the text and emoticon. *ICT Express* **2020**, 6 (4), 357–360.
27. K. Dashtipour, M. Gogate, E. Cambria, A. Hussain. A novel context-aware multimodal framework for persian sentiment analysis. *Neurocomputing* **2021**, 457, 377–388.
28. W. Li, W. Shao, S. Ji, E. Cambria. BiERU: Bidirectional emotional recurrent unit for conversational sentiment analysis. *Neurocomputing* **2022**, 467, 73–82.
29. B. Kabra, C. Nagar. Convolutional Neural Network based sentiment analysis with TF-IDF based vectorization. *J. Integr. Sci. Technol.* **2023**, 11 (3), 503.