

RESEARCH

Open Access



Sentiment analysis from textual data using multiple channels deep learning models

Adepu Rajesh^{1,2*} and Tryambak Hiwarkar¹

*Correspondence:
adeapurajeshadepu@gmail.com

¹ Department of Computer Science and Engineering, Sardar Patel University, Balaghat, MP, India

² Department of Computer Science and Engineering, Guru Nanak Institute of Technology, Ibrahimpatnam, Hyderabad, Telangana, India

Abstract

Text sentiment analysis has been of great importance over the last few years. It is being widely used to determine a person's feelings, opinions and emotions on any topic or for someone. In recent years, convolutional neural networks (CNNs) and long short-term memory (LSTM) have been widely adopted to develop such models. CNN has shown that it can effectively extract local information between consecutive words, but it lacks in extracting contextual semantic information between words. However, LSTM is able to extract some contextual information, where it lacks in extracting local information. To counter such problems, we applied the attention mechanism in our multi-channel CNN with bidirectional LSTM model to give attention to those parts of sentence which have major influence in determining the sentiment of that sentence. Experimental results show that our multi-channel CNN model with bidirectional LSTM and attention mechanism achieved an accuracy of 94.13% which outperforms the traditional CNN, LSTM+CNN and other machine learning algorithms.

Keywords: LSTM, CNN, Sentiment analysis, Attention, Deep models

Introduction

Sentiment analysis, also called opinion mining, is the field of study that examines people's opinions, sentiments, attitudes and emotions toward entities such as products, services, organizations, individuals, issues, events, topics and their attributes [1]. Due to wide availability of Internet and mobile networks, there is more and more textual data being published over the Internet. People are expressing their views or comments over the Internet, either on social media or on sites in the form of reviews. Every organization wants to analyze the reviews it is getting online, mostly, the product's success depends on the reviews, generally, positive reviews. For instance, the IMDB [16, 41] dataset contains reviews of people regarding different movies. Everyone wants to see a film, only if it has positive reviews. So, nowadays more research is being focused to analyze these reviews and predict sentiment with minimum errors.

Many sentiment analysis models use machine learning (ML) methods such as naïve Bayes and support vector machines [2]. But recently, due to increased amount of data or textual information, researchers have been greatly utilizing deep neural networks (DNNs) for this task. The CNN is basically applied for the image handling task such as mental disorder [37] and video surveillance [38]. Moreover, LSTM utilization apart

from text processing can be identified in fault detection [39] in induction motor, electro-oculogram signal classification [40], etc. In deep neural networks, CNNs have emerged as very effective tool in analyzing text, especially the local information of sentence. The multi-channel CNN has provided even greater results in this specific field. But, the CNNs lack behind in extracting the semantic contextual information between words. In addition, LSTMs lack behind CNN in terms of getting the local information. To tackle these problems, we have developed a multi-channel CNN + BiLSTM model with attention mechanism; after experimentation, it is observed as the revolutionary change in text analysis, i.e., sentiment analysis task.

In this paper, we have first used the GloVe embeddings [3] to vectorize the representation of words in a meaningful way; then, this embedding is fed to multi-channel CNN with each CNN layer having separate BiLSTM and attention layer. Later, these results after all processing are passed through a multilayer perceptron network to classify the text as positive or negative. The main objective of this paper is to offer a better classifier model to predict sentiment from IMDB dataset.

The paper organization is as follows—"Related work" section highlights the literature review related to sentiment analysis. Model development is presented in section "Methodology." The section "Experimental setup and dataset" demonstrates experimental setup, dataset and data preprocessing. Results and discussions are shown in section "Results and discussion." Finally, conclusion and future scope are presented in section "Conclusion and future scope."

Related work

Sentiment analysis is an important and interesting field of research in natural language processing (NLP). In recent years, many deep learning techniques have been developed for effective and efficient prediction of sentiments. In this section, we discuss the recent work accomplished in the field of sentiment analysis.

One of the most important breakthroughs in the NLP is the word embeddings [4]. In 2013, Google released a tool called "Word2Vec" to calculate word vectors. It helped to relate the similar words with each other and grab the semantic information from a sentence effectively. Another similar word embedding, namely, "GloVe" vector embedding, was researched and developed by the researchers of Stanford University [3]. These embeddings helped us to overcome major flaws in one-hot encoding, i.e., sparse distribution of words and increased dimensionality of the dataset.

The deep learning techniques have emerged in recent years, and showed excellent results in this field. Most commonly used models in NLP are CNNs [5], RNN [6] and long short-term memory [7]. The usage of word embeddings has made it possible to use CNN for texts and enables the CNN to extract local information in the sentence [8]. A two-layer CNN model was proposed by Attardi and Santos to classify the text sentences on the basis of extracted features [9], whereas Yin and Schütze used a multi-channel CNN network by forming combinations of words [10]. The detailed review on the application of CNN is presented in [32], wherein challenges in sentiment analysis are highlighted for the textual, visual and multimodal data.

LSTM and CNN-LSTM are other types of models which have been widely applied for the task of sentiment analysis. A tree-structured LSTM model was introduced by Tai

et al. [11] which achieved good results in sentiment classification, whereas Wang et al. introduced a tree-structured CNN–LSTM model for dimensional sentiment analysis [12] which captured both local and long distance dependencies between words in sentences. Other than this, Wang et al. [13] also proposed an attention-based LSTM network that focuses on various parts of the sentences. Further, attention model is improved for sentiment analysis in [30]. The cognition grounded data is utilized for training the model, wherein contextual information from sentence and document level is extracted at the time of model training. The accuracy of 66.80% is achieved on Yelp14 dataset by the proposed hybrid model.

A CNN–LSTM-based model was also proposed by Zhou et al. [14] which first utilized the CNN layer to extract the sentence local features and later applied the LSTM layer in place of pooling layer to obtain the desired classification results. Sun et al. [15] applied CNN–LSTM model and hybrid deep learning algorithms to classify the Tibetan blogs which achieved the good results.

In recent years, there has been serious adaptation of attention mechanism in text analysis-based models. Yang et al. [18] combined the bidirectional RNN (BRNN) and the attention layer for the text-level classification task, whereas Long et al. [31] proposed an improved CNN and multilayered attention mechanism for the classification task of sentiment analysis. In [33], a systematic survey for the textual, visual and multimodal data is discussed for the RNN and variants of RNN such as LSTM and gated recurrent unit (GRU). The challenges and issues pertaining to sequence modeling in different modalities are highlighted.

On the basis of aforementioned literature review, we identified the potential of CNN, LSTM and the attention mechanisms for the task of sentiment analysis as per the past research. In this paper, we propose a multi-channel CNN-BiLSTM model with an attention mechanism to extract more relevant and important features to predict the sentiments from the given input. Table 1 presents the summary of the related work and highlights the benefits and drawbacks of the existing systems.

Methodology

Before passing the dataset through defined model, the data needs to be preprocessed, cleaned and represented in vector form for the model to train or test. Other than that, the data also needs to be split and tokenized. Figure 1 shows an overall dataflow of the proposed system. Firstly, dataset preprocessing and splitting is performed. Word embedding and tokenization is applied on the dataset. Training is performed using proposed classification model, and finally, accuracy and losses are visualized. Further, Fig. 2 presents the preprocessing of raw input text for the task of sentiment analysis.

Word embedding

To better encode the semantic and co-occurrence information, word embeddings are utilized. A word embedding is an n -dimensional dense vector of floating values and these values are generated on the basis of cosine similarity between these words. Here, n represents the vector dimensions of each embedded word. Hence, overcome the problem of sparse matrix and increased dimensions in case of one-hot encoding.

Table 1 Summarization of related works

Ref	Model used	Benefits	Issues	Dataset used
[10]	CNN + attention	Integrate related impact between sentences Counterpart of each sentence is also considered	Limited dataset	WikiQA SICK
[11]	Tree LSTM	Superior representation of sentence meaning Tree LSTM worked good for shorter sentences	Challenge in depicting the part of structure sentences	SemEval 2014 Stanford Sentiment Treebank
[12]	Tree CNN-LSTM	Text divided in several regions and affective facts are extracted Task-specific clauses and phrases for constructing structured information	More training time due to attention module	Stanford Sentiment Treebank EmoBank
[14]	C-LSTM	Best strengths from CNN and LSTM are utilized for model design	Lacks in contextual information extraction	Stanford Sentiment Treebank TREC
[26]	LSTM	Efficient data preprocessing and partitioning for post-classification LSTM used for feature extraction	Less accuracy	IMDB
[35]	B-MLCNN	Prepare single document for complete textual review and categorizes into offered sentiments BERT used for representation of feature vector and capturing global features MLCNN performs feature extraction	Struggling to determine the contextual sense of sentences Large set of parameters restrict to find the optimal combination	IMDB Amazon Review

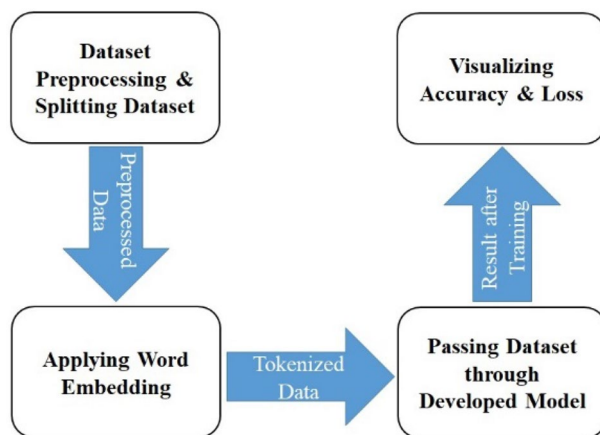


Fig. 1 Overall system model

There are mainly two pretrained word embeddings—(1) Word2Vec and (2) GloVe. Here, we have used the Glove embeddings of 300 dimensions having 42 billion tokens with vocab size of 1.9 million. Glove vector model obtains meaning from word-to-word co-occurrence statistical information.

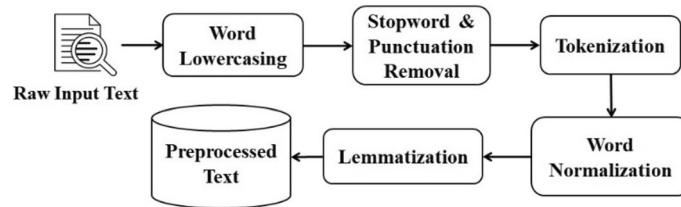


Fig. 2 Preprocessing of raw input text

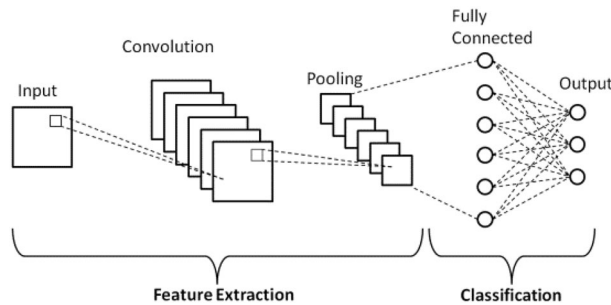


Fig. 3 Convolutional neural network

Convolutional neural network

It was the Yoon Kim [17] who first used the CNN for sentence classification. It has the ability to capture semantic local information from a sentence. It can do that by using the different numbers and sizes of filters or kernels. These filters are applied to the input to extract the feature maps. Filters with different sizes are able to extract variety of information from a sentence. We usually refer to it as multi-channel CNN, i.e., we define a model with different input channels for processing n -grams of text, where n is the number of words a kernel reads at a time. Here, we have defined the model for five channels to process 3-g, 4-g, 5-g, 6-g and 8-g with equal padding. Usually, after this a pooling layer is utilized to consolidate the results from convolutional layer. Later, these pooled layers are concatenated, flattened and passed through a dense layer to classify the text. A general CNN architecture is shown in Fig. 3 [19].

Bidirectional LSTM

This paper implements the multi-channel CNN method except that in place of using pooling layer, all of the convoluted outputs are passed into two bidirectional LSTM layer to obtain the contextual information from each of the extracted local data. A long short-term memory model has the great ability to remember important information about the input as compared to its predecessor RNN, which will for the longer period of time. It overcomes the problem of vanishing gradient in RNN by using the gating mechanism. LSTM uses the input gate, output gate, forget gate and memory cell to remember the contextual information for longer period of time. These gates are composed of sigmoid layer which decides how much each component of information should go through a gate. The value of 1 means, let everything go and the value of 0 means, to block everything. A LSTM architecture is presented in Fig. 4 [20].

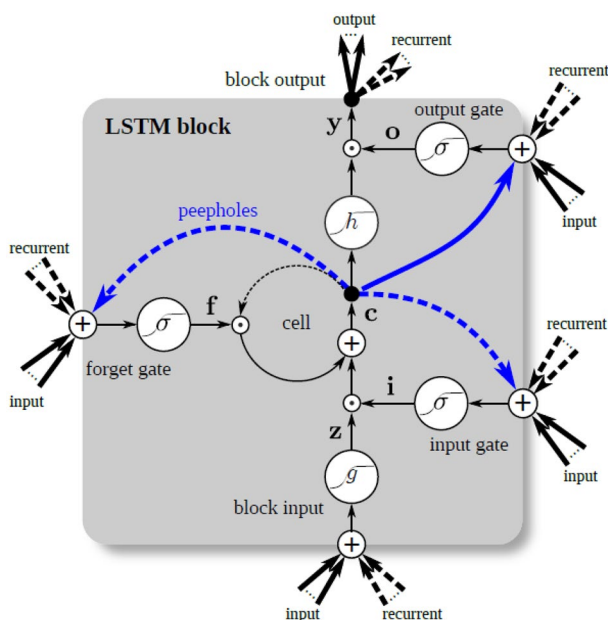


Fig. 4 LSTM Cell architecture [20]

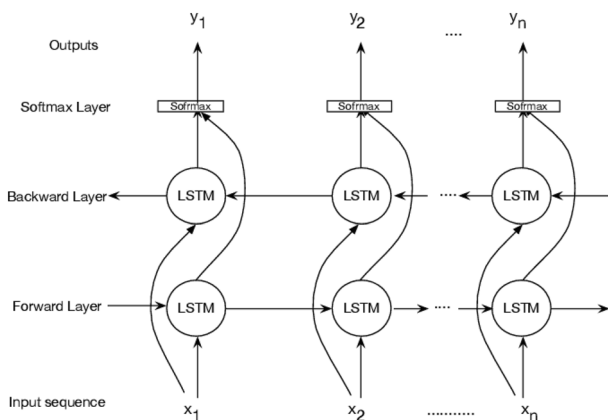


Fig. 5 BiLSTM layer

Later, it was observed that LSTM layers also tend to lose information as the sentence gets long and they only allow the information to propagate forward, i.e., the current state T output depended only on the states before T . Therefore, to extract more contextual information from a sentence, a BiLSTM model was proposed. It was inspired from the bidirectional RNN model which was proposed in 1997 by Schuster and Paliwal [21]. It not only considers the previous states but also the future states that will come after state T , which undoubtedly gives the excellent results. Figure 5 represents the BiLSTM layer model [22].

Here, we implemented the BiLSTM layer, after the convoluted output generated from CNN and process this data to extract contextual information from the input values. Further, BiLSTM output is passed through an attention layer to generate the word scores.

Attention mechanism

It is certain that not every word of the sentence contributes equally to the meaning of a sentence. Some words have higher importance and others do not. The attention mechanism checks for such words and assigns a score to each word on the basis of their importance.

Here, we have implemented Bahdanau [23] attention. All the outputs from the BiLSTM layer are passed through this layer to assign score to each output. Bahdanau attention takes all the hidden states of last layer at once and multiply them with attention layer's hidden state weights to assign an attention score to each input hidden state by taking a softmax of output values. After multiplying each input hidden state with attention score, a weighted sum is taken for all input hidden states. This weighted sum is called context vector and it is later passed onto the next layer. After receiving the weighted sum or adjusted hidden states, then we passed them through a single dense layer.

After performing the aforementioned operations for each channel of CNN (here in this case 5 channels), all results of each dense layer are concatenated and passed through a dense layer network to classify text either, positive or negative. The input and output dimensions for each layer for the proposed model are shown in Table 2; moreover, the proposed deep learning model architecture for the task of sentiment analysis is described in Fig. 6. Here, five channels of CNNs with different kernel size and filter of 128, followed by the BiLSTM layers to extract the temporal features from text, are identified. Next, attention mechanism is applied to generate the score for the important words and all these vectors are supplied to dense layer, and further classification is performed on the final data.

Experimental setup and dataset

To build and train the proposed Google Colab is utilized with access to free TPU's. In addition, Tensorflow 2.0+ and keras software libraries are used to prepare the model. To fine-tune the proposed model, various hyperparameters values are adopted to achieve the better performance. The model is trained for 20 epochs. To avoid overfitting, dropout of 0.2 is applied at second last dense layer and early-stopping is implemented with

Table 2 Input/output dimensions of each layer of proposed model

Layers	Input shape	Output shape
Input shape (<i>with sequence length = 300</i>)	(None, 300)	(None, 300)
Embedding	(None, 300)	(None, 300, 300)
Convolutional layer (<i>for each layer out of 5</i>)	(None, 300, 300)	(None, 300, 128)
BiLSTM layer I	(None, 300, 128)	(None, 300, 64)
BiLSTM layer II	(None, 300, 64)	(None, 300, 64)
Batch normalization	(None, 300, 64)	(None, 300, 64)
Attention layer	(None, 300, 64)	(None, 64)
Dense layer	(None, 64)	(None, 128)
Concatenation of five dense layers for each channel	(None, 128) * 5	(None, 640)
Dense layer	(None, 640)	(None, 16)
Dropout layer	(None, 16)	(None, 16)
Output layer	(None, 16)	(None, 1)

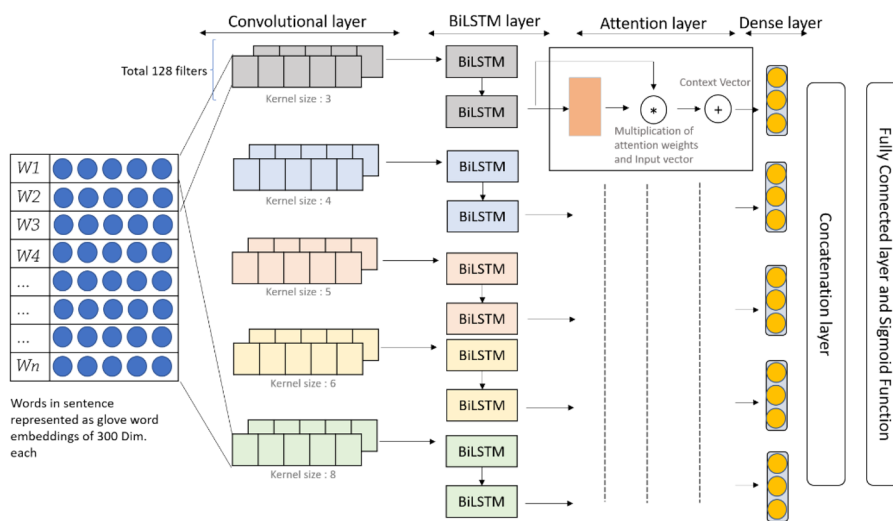


Fig. 6 Proposed deep learning model for sentiment analysis

Table 3 Hyperparameters values

Sr. No	Hyper parameters	Values
1	Batch size	64
2	Learning rate	0.00005
3	Optimizer	Adam
4	No. of epochs	20
5	Clipnorm	1
6	Dropout	0.2
7	No. of CNN filters	128
8	CNN kernel size	(3,4,5,6,8)

patience of three epochs. Adam optimizer with learning rate of 0.00005 and clipnorm with value “1” are used to handle the problem of exploding gradient. For loss calculation, binary cross-entropy function is applied. In case of CNN, to obtain rich collection of information, we used kernels of five different sizes—3, 4, 5, 6 and 8, wherein number of each kernel are 128. The summary of hyperparameters and their values are presented in Table 3.

Dataset description

For this work, we have selected IMDB movie review dataset [16]. It consists of around 50 K movie reviews in which 25 k are positive labeled and remaining 25 k are negative labeled. The dataset is divided into a train/valid split ratio of 80–20. For validation, 20% of validation data is taken.

Data preprocessing

Data preprocessing is an important step before the development of any deep learning model. Before passing text into neural networks, it is necessary to clean the dataset and remove any noise in the dataset, thus to improve the performance of developed model.

Table 4 Performance analysis on IMDB dataset

Sr. No	Model	Accuracy
1	CNN	0.91
2	LSTM	0.88
3	CNN + LSTM	0.89
4	Multi-channel CNN + LSTM + attention (proposed model)	0.94

Table 5 Comparison with existing models on IMB Dataset

Year	Model name	Accuracy on IMDB dataset (%)
2021	LSTM + CGAG [24]	89.4
2019	Hybrid CNN + LSTM model [25]	91
2020	LSTM [26]	89.9
2020	BiLSTM [27]	91.96
2021	TF-IDF + SVM-RFE [28]	86.85
2021	CSA + CNN + LSTM [29]	89.88
2023	BERT network [33]	90.67
2022	DNN [34]	89
2023	BERT + CNN [35]	93
2022	BiLSTM [36]	88.12
Our (2023)	Proposed model	94.13

The IMDB movie review dataset consists sentences with URL's, hashtags, slang words, abbreviations, emojis, stopwords, etc. The dataset persists with lots of noise. Therefore, to remove such noises such as punctuations, URL's, hashtags and stopwords are removed from the dataset. In addition, all uppercase letters are converted into lowercase; all emojis are converted into text. Other than that, lemmatization is performed on the dataset to normalize the words. Lemmatization uses the context in which the word is being used and replaces the word with the word of similar context. Later, the data is tokenized using *keras_tokenizer* and padded with the sequence length of 300.

Results and discussion

We compared the proposed multi-channel CNN + BiLSTM + attention mechanism model with multi-channel CNN, LSTM, multi-channel CNN and machine learning algorithms, and found that the proposed model gives better results than the state-of-the-arts. Our model achieved a highest accuracy of 94.13% with early-stopping with patience of three epochs to prevent overfitting. Table 4 shows the performance metric in the form of accuracy based on CNN, LSTM, CNN + LSTM and proposed model on the IMDB dataset. From Table 4, we depict that the proposed model outperforms the other models. It is also seen from Table 5 that proposed sentiment classification model's accuracy is higher than the existing models in [24–29] and [33–36]. The increased accuracy is reported by the proposed model, i.e., more accuracy of 3.46% from [33], 5.13% from [34], 1.13% from [35] and 6.01% from [36], respectively.

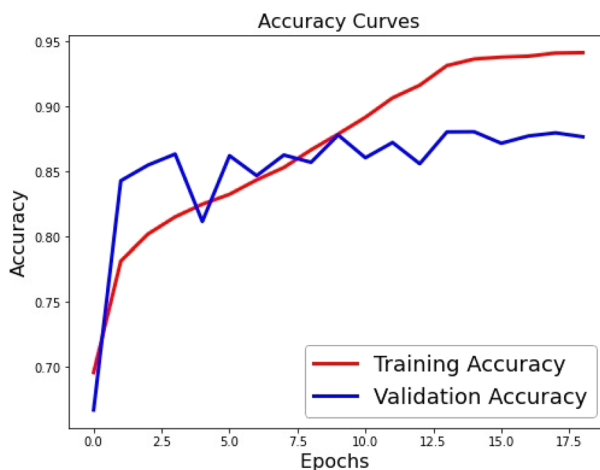


Fig. 7 Training and validation accuracy of proposed model

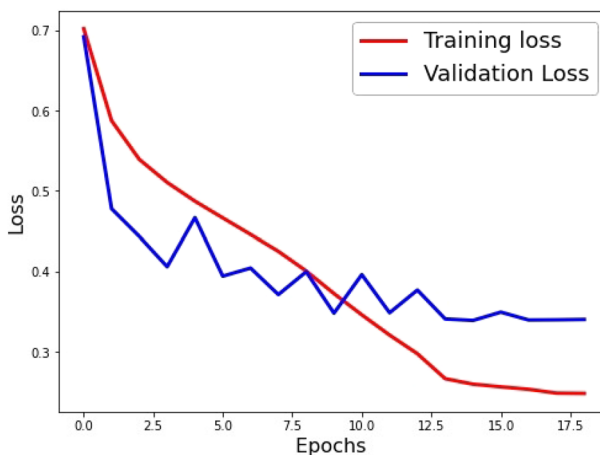


Fig. 8 Training and validation loss of proposed model

Accuracy and loss of the proposed models are shown in Figs. 7 and 8, respectively. The validation accuracy is improved after 10th epochs and reaches to the maximum of 94.31%; moreover, the validation loss is reduced to approximately 0.1.

Conclusion and future scope

In this paper, we present a multi-channel CNN with BiLSTM and attention layer. The proposed model captures both local and contextual information from the sentence. The effective and efficient features are extracted from the proposed model wherein the attention layer significantly improves the accuracy of classification model as compared with the state-of-the-arts. The proposed model achieved an accuracy of 94.13% on the IMDB dataset which shows the superiority among the other models developed in the past. The limitation of proposed model would be challenged in training using attention module which needs more data and computing resources to learn attention weights. In future, we will focus on more advanced models in CNNs and RNNs to develop the generalized

model to predict the more accurate sentiments from the different datasets available in public domain.

Author contributions

Adepu Rajesh performed the design and implementation of proposed work and prepared the initial draft of manuscript. Tryambak Hiwarkar conceptualized the work, examined and suggested the changes in the manuscript. All authors reviewed and finalized the manuscript.

Funding

No funding was received for conducting this study.

Data availability

All data generated or analyzed during this study are included in this published article.

Declarations

Competing interests

The authors have no competing interests to declare that are relevant to the content of this article.

Received: 15 August 2023 Accepted: 10 November 2023

Published online: 23 November 2023

References

- Zhao J, Liu K, Xu L (2016) Sentiment analysis: mining opinions, sentiments, and emotions
- Pang B, Lee L, Vaithyanathan S (2002) Thumbs up? Sentiment classification using machine learning techniques. arXiv preprint cs/0205070
- Pennington J, Socher R, Manning CD (2014) Glove: global vectors for word representation. In: Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP), pp 1532–1543
- Mikolov T, Chen K, Corrado G, Dean J (2013) Efficient estimation of word representations in vector space. arXiv preprint [arXiv:1301.3781](https://arxiv.org/abs/1301.3781)
- Fukushima K (1980) neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biol Cybern* 36(4):193–202
- Sherstinsky A (2018) Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. *CoRR* abs/1808.03314 (2018). arXiv preprint [arXiv:1808.03314](https://arxiv.org/abs/1808.03314)
- Tang D, Qin B, Feng X, Liu T (2015) Effective LSTMs for target-dependent sentiment classification. arXiv preprint [arXiv:1512.01100](https://arxiv.org/abs/1512.01100)
- Pang B, Lee L (2005) Seeing stars: exploiting class relationships for sentiment categorization with respect to rating scales. arXiv preprint cs/0506075
- Attardi G, Sartiano D (2016) UniPI at SemEval-2016 Task 4: convolutional neural networks for sentiment classification. In: Proceedings of the 10th international workshop on semantic evaluation (SemEval-2016), pp 220–224
- Yin W, Schütze H, Xiang B, Zhou B (2016) Abcnn: Attention-based convolutional neural network for modeling sentence pairs. *Trans Assoc Comput Linguist* 4:259–272
- Tai KS, Socher R, Manning CD (2015) Improved semantic representations from tree-structured long short-term memory networks. arXiv preprint [arXiv:1503.00075](https://arxiv.org/abs/1503.00075)
- Wang J, Yu LC, Lai KR, Zhang X (2019) Tree-structured regional CNN-LSTM model for dimensional sentiment analysis. *IEEE/ACM Trans Audio Speech Lang Process* 28:581–591
- Wang Y, Huang M, Zhu X, Zhao L (2016) Attention-based LSTM for aspect-level sentiment classification. In: Proceedings of the 2016 conference on empirical methods in natural language processing, pp 606–615
- Zhou C, Sun C, Liu Z, Lau F (2015) A C-LSTM neural network for text classification. arXiv preprint [arXiv:1511.08630](https://arxiv.org/abs/1511.08630)
- Sun B, Tian F, Liang L (2018) Tibetan micro-blog sentiment analysis based on mixed deep learning. In: 2018 international conference on audio, language and image processing (ICALIP). IEEE, pp 109–112
- Maas A, Daly RE, Pham PT, Huang D, Ng AY, Potts C (2011) Learning word vectors for sentiment analysis. In: Proceedings of the 49th annual meeting of the association for computational linguistics: human language technologies, pp 142–150
- Kim Y (2014) Convolutional neural networks for sentence classification proceedings of the 2014 conference on empirical methods in natural language processing, EMNLP 2014, october 25–29, 2014, doha, qatar, a meeting of sigdat, a special interest group of the acl. Association for Computational Linguistics, Doha, Qatar
- Yang Z, Yang D, Dyer C, He X, Smola A, Hovy E (2016) Hierarchical attention networks for document classification. In: Proceedings of the 2016 conference of the North American chapter of the association for computational linguistics: human language technologies, pp 1480–1489
- Gurucharan M (2020) Basic CNN architecture: explaining 5 layers of convolutional neural network. <https://www.upgrad.com/blog/basic-cnn-architecture>
- Greff K, Srivastava RK, Koutnik J, Steunebrink BR, Schmidhuber J (2016) LSTM: a search space odyssey. *IEEE Trans Neural Netw Learn Syst* 28(10):2222–2232
- Schuster M, Paliwal KK (1997) Bidirectional recurrent neural networks. *IEEE Trans Signal Process* 45(11):2673–2681

22. Tavakoli N (2019) Modeling genome data using bidirectional LSTM. In: 2019 IEEE 43rd annual computer software and applications conference (COMPSAC), vol 2, pp 183–188. IEEE
23. Bahdanau D, Cho K, Bengio Y (2014) Neural machine translation by jointly learning to align and translate. arXiv preprint [arXiv:1409.0473](https://arxiv.org/abs/1409.0473)
24. Bui V, Le NT, Nguyen VH, Kim J, Jang YM (2021) Multi-behavior with bottleneck features LSTM for load forecasting in building energy management system. *Electronics* 10(9):1026
25. Rehman AU, Malik AK, Raza B, Ali W (2019) A hybrid CNN-LSTM model for improving accuracy of movie reviews sentiment analysis. *Multimed Tools Appl* 78:26597–26613
26. Qaisar SM (2020) Sentiment analysis of IMDb movie reviews using long short-term memory. In: 2020 2nd international conference on computer and information sciences (ICCIS). IEEE, pp 1–4
27. Dong Y, Fu Y, Wang L, Chen Y, Dong Y, Li J (2020) A sentiment analysis method of capsule network based on BiLSTM. *IEEE Access* 8:37014–37020
28. Nafis NSM, Awang S (2021) An enhanced hybrid feature selection technique using term frequency-inverse document frequency and support vector machine-recursive feature elimination for sentiment classification. *IEEE Access* 9:52177–52192
29. Al Bataineh A, Kaur D (2021) Immunocomputing-based approach for optimizing the topologies of LSTM networks. *IEEE Access* 9:78993–79004
30. Long Y, Xiang R, Lu Q, Huang CR, Li M (2019) Improving attention model based on cognition grounded data for sentiment analysis. *IEEE Trans Affect Comput* 12(4):900–912
31. Diwan T, Tembhurne JV (2022) Sentiment analysis: a convolutional neural networks perspective. *Multimed Tools Appl*, pp 1–25
32. Tembhurne JV, Diwan T (2021) Sentiment analysis in textual, visual and multimodal inputs using recurrent neural networks. *Multimed Tools Appl* 80:6871–6910
33. Domadula PSSV, Sayyaparaju SS (2023) Sentiment analysis of IMDB movie reviews: a comparative study of Lexicon based approach and BERT Neural Network model. BS Thesis, May 2023
34. Sabba S, Chekired N, Katab H, Chekkai N, Chalbi M (2022). Sentiment analysis for IMDb reviews using deep learning classifier. In 2022 7th international conference on image and signal processing and their applications (ISPA). IEEE, pp 1–6
35. Atandoh P, Zhang F, Adu-Gyamfi D, Atandoh PH, Nuhoho RE (2023) Integrated deep learning paradigm for document-based sentiment analysis. *J King Saud Univ Comput Inf Sci* 35(7):101578
36. Basarslan MS, Kayaalp F (2022) Sentiment analysis with various deep learning models on movie reviews. In 2022 international conference on artificial intelligence of things (ICAIoT). IEEE, pp 1–5
37. Hussein SA, Bayoumi AERS, Soliman AM (2023) Automated detection of human mental disorder. *J Electric Syst Inf Technol* 10(1):1–10
38. Khairy M, Al-Makhlasawy RM (2022) A reliable image compression algorithm based on block luminance adopting deep learning for video surveillance application. *J Electric Syst Inf Technol* 9(1):21
39. Vanga J, Ranimekhala DP, Jonnala S, Jamalapuram J, Gutta B, Gampa SR, Alluri A (2023) Fault classification of three phase induction motors using Bi-LSTM networks. *J Electric Syst Inf Technol* 10(1):1–15
40. Hassanein AM, Mohamed AG, Abdullah MA (2023) Classifying blinking and winking EOG signals using statistical analysis and LSTM algorithm. *J Electric Syst Inf Technol* 10(1):44
41. IMDB Dataset. <https://developer.imdb.com/non-commercial-datasets/>. Accessed 13 May 2023

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
