A COMBINED CNN-LSTM MODEL FOR IMPROVING ACCURACY OF MOVIE REVIEWS SENTIMENT ANALYSIS

Dr. V Kavidha, Professor, Department of Computer Science and Engineering, Holy Cross Engineering College, Vagaikulam -628851, Thoothukudi, TamilNadu, India. Email: vkavidha1@gmail.com

ABSTRACT
Long Short Term Memory (LSTM) model and Convolutional Neural Network (CNN) model have applied to different Natural Language Processing (NLP) tasks with effective results. The CNN model proficiently removes higher level features utilizing convolutional layers and max-pooling layers. The LSTM model is effective to catch long-term dependencies between word sequences. In this paper, we propose a model using LSTM and very deep CNN model named as Combined CNN-LSTM Model to overcome the sentiment analysis problem. First, we use Word to Vector (Word2Vc) detain to train initial word embeddings. The Word2Vc translates the text strings into a vector of numeric values, determines distance linking words and constructs groups of similar words established on their meanings. Afterword embedding is represented in which the proposed model collaborated set of features that are removed by convolution and global max-pooling layers with long term dependencies. The proposed model also uses dropout technology, normalization and a rectified linear unit for accuracy enhancement. Our results show that the proposed Combined CNN-LSTM model outperforms traditional deep learning and machine learning techniques in terms of precision, recall, f-measure, and accuracy. The proposed approach attained competitive results utilizing state-of-the-art techniques on the IMDB movie review dataset and Amazon movie reviews dataset.

Keywords:
Sentiment Analysis, Long Short Term Memory, Natural Language Processing, Convolutional Neural Network, Deep Learning, Recurrent Neural Network.

I. Introduction

There is a need to get useful information from huge amount of data using different machine learning techniques [1]. A new architecture is presented for NLP in [2] to employ at character level directly and utilized small convolution and pooling operations to learn a high-level representation of sentences. We observed that our proposed Combined CNN-LSTM model improved accuracy with respect to baseline algorithms [3]. The authors proposed in [4-6] represents the rating prediction recommended system utilizing deep learning. Weights are observed binary which decrease memory consumption [7]. Authors in [8] proposed LSTM based approach for product based sentiment analysis. They executed conditional random field classifier with bidirectional LSTM and aspect based LSTM for identification of polarity on Hotel's review dataset. The authors in [9] compared Naïve Bayes (NB) and Support Vector Machine (SVM) techniques for Arabic tweets and text classification utilizing WEKA tool. TF-IDF and cosine measure approaches were used for weighting scheme and similarity calculation among documents respectively. The study [10] focused on the basic problem in Shuffle Attention (SA) which is sentiment polarity categorization. It uses products reviews dataset from amazon. Random Forest (RF), SVM and NB techniques are utilized and generated better results.
This paper contributes as follows :
1.Word embedding is created using Word2Vec model, which is an unsupervised model and is trained on a large collection of words. This model is able to capture semantics of words.
2.To represent sentiment polarity from texts, the LSTM model is used to detect deeper semantics of words and this model efficiently learning long-term dependencies between word sequences in long texts.
3.Further refinements are done for embeddings, a very deep CNN model is used on a supervised

dataset and generate a number of features, used many weight matrices also with windows of different length.
4.The advantages are taken for the CNN model in extracting local features and long distance dependencies. They are captured by the LSTM model and mix these features into one single proposed Combined CNN-LSTM model. The proposed model achieved efficient experimental results.
The paper is organized as follows. Section 2 describes Literature review, Section 3 describes the architecture of CNN and LSTM model and Word to Vector Model. The methodology of our research work is described in Section 4. Section 5 describes experimental setup. Section 6 shows the results and discussion. Section 7 concludes the study.

II. Literature Review

Wide amount of data is created through web on a daily basis which requires to be processed to prevail meanings from data and according to Govindarajan [11] Deep learning and Machine learning techniques have been used for SA. A combination of CNN and LSTM (ConvLstm) technique proposed by Hao [12]. Hassan and Mahmood [13] said the proposed model utilized 10 convolutional layers to removes local information in an well organized way as compared to the networks. Himelboim etal [14] described the hierarchy of the network is a symbol of its individual flow of information . Islam and Zhang [15] provided an overview of forecasting the sentiments of visual content in visual SA, CNN framework approach. Kaur and Gupta [16] suggested NLP field, researchers have been evoluting many different techniques to solve SA issues and these techniques utilize a bag of words representation. Li et al [17] offered a description of Sentiment Analysis a pretty challenging task and many issues appear during the processing of social media content. Lia et al [18] demonstrated business, sentiment analysis is a procedure of cataloging and identifying a slice of text according to the business . Manek et al [19] presented a Two-Parse algorithm for product review analysis with approximate 7000 keywords training dataset. Pang and Lee [20] conducted a combination of CNN and Word2Vc framework. Ruangkanokmas et al [21] developed Dropout technique. Sing et al [22] created the benchmark Internet Movie Data Base (IMDB) dataset of movie reviews for sentiment analysis was first time published .

3. Deep Learning Model

3.1 Word to vector model

In 2013 Google proposed Deep learning model word2Vc. This model creates vector numeric values using sentence of words. Word meaning uses Word2Vc to compute the distance between words. The huge amount of data is given, usage and content, therefore Word2Vc can generate exceedingly accurate estimates about a word's meaning. Word2Vc runs faster even for a huge dataset and utilizes google news dataset for training. The google news dataset contains pre-trained vectors and dropout technique which is used in our proposed model to block from overfitting and to drop the irrelevant information from network to strengthen the performance. Dropout techniques is used and it selects top most related words from google news dataset are associated with "good", "bad" and "terrible". Positive words like "good" and "fantastic" appear in the second group while negative words like "terrible", "bad" and "horrible" disappear on one side of the graph. The Deconvolutional Neural Network (DN) demonstrate that Word2Vc can absolutely notice the similar words in vector space. We use the same size of input data in next and input data of CNN cannot change in next layer, it means the input sentence contains same number of words.

3.2 Convolutional neural network model

The field of image processing is employed that the CNN is a special type of neural network and in text classification, CNN model has been effectively used. A subset of input to its preceding layers is connected in CNN model using a convolutional layer that is why CNN layers are called feature map. To reduce the complexity the CNN model uses polling layer. In CNN the polling techniques reduces the output size of one stack layers to next in such a way it preserved the important information. Many polling techniques are available, but max-polling is mostly used in which pooling window contains max value element. The flattened layer is utilized to sustain he output of polling layer and maps it to next layers and the final layer in CNN typically is fully connected. Figure 1 shows the basic architecture of CN.

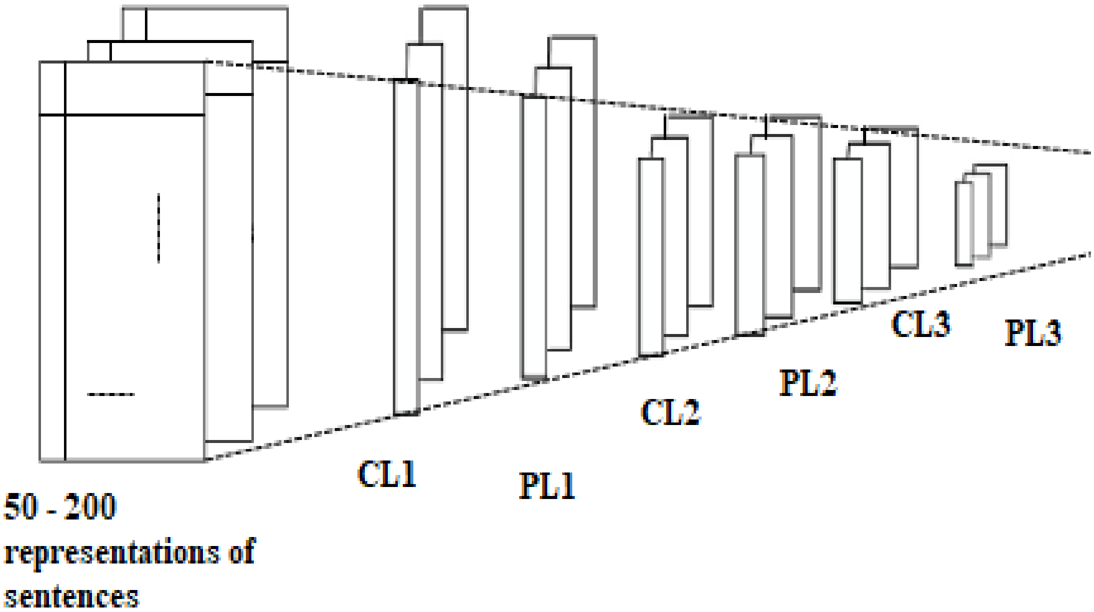

Figure 1: Basic Architecture of CNN

3.3 Recurrent network model

The proposed model uses LSTM that is a special type of RNN. The neurons are connected in RNN with each other in the form of directed cycle and processes the information in a sequential manner because to process a sequence of words or input it uses internal memory. For each element RNN performs the same task because output is based on all previous nodes inputs and remember information.

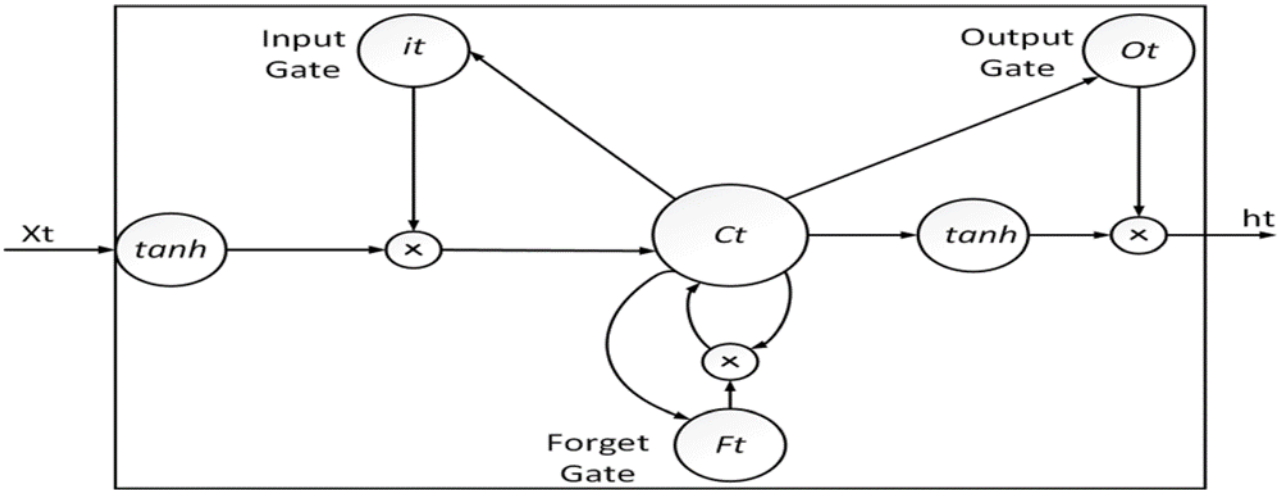

Figure 2: The Architecture of LSTM Model

To handle the vanishing gradient problem a special type of RNN model which is called LSTM is used. Using three different gates in an effective way the LSTM model saves long-term dependencies. Figure 2 shows the architecture of LSTM model. Similar to RNN the structure of LSTM is chain like, however LSTM utilizes three gates to regulate and conserve information into every node state. The LSTM gates are input gate, forget gate and output gate is provided with memory cell. For further processing represent general RNN model, where xt is the input vector at time t, ht is the new state at time t, tanh is the activation function.

4. Methodology

In this section we describe the models that we used in our experiments. The proposed Combine CNN-LSTM model is described and compared the results of three models with each other and also with the traditional machine learning techniques.

4.1 The proposed Combined CNN-LSTM model

The Figure 3 show the main architecture of the proposed Combined CNN-LSTM model. Taking a corpus as input and in preprocessing phase it performs sentence segmentation, tokenization, stemming tasks and stop word removal. It also applies word embedding layer using Word2Vec and convolutional layer extracts the high level features. LSTM layer detects long term dependencies between words. We apply classification layer using sigmoid function in the end.

4.2 The proposed model using CNN

We describe the proposed CNN model in this section that uses Word2Vc technique for word embedding. First Word2Vc translates the text into vector numeric values and then applied CNN model to train vector numeric values. We use 3 pooling layers, 3 convolutional layers and one fully connected layer. The systematic diagram of CNN is shown with 7 configuration layers and for experiments, use tensor flow open source python library for numerical computation. Pooling layers, convolutional layers, dropout out layers is used in CNN and RLU for accuracy improvement. In deep learning dropout is an important trick becauseit prevents machine learning algorithms from over fitting. Dropout algorithms in back propogation do not contribute to skip the neurons. During training to prevent neurons from co-adaptation the dropout technique drops the neurons and each hidden neuron gives output with 0.5 probability. In following subsections, the proposed CNN model is described with different layers which are used in the model.

Pre-Processing

Pre-processing is a process helps to assemble the dataset by performing basic operations on dataset before passing it to a model such as moving of spaces and meaningless word. It converts different forms of a word into theirs roots words and removal of duplicate words. Raw dataset converts dataset into a useful and organized dataset for further use.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 8, No.2, August : 2024

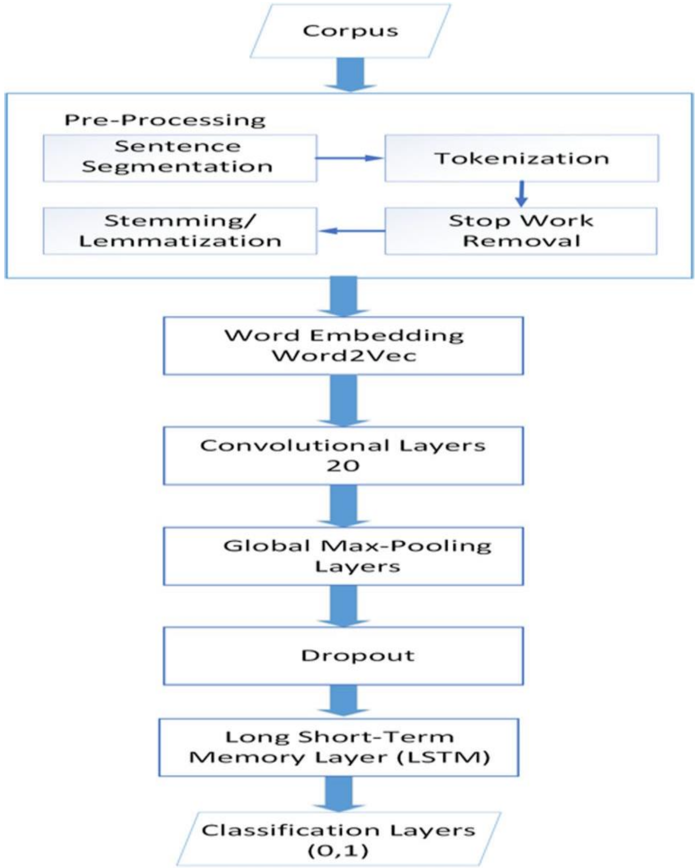

Figure 3: Methodology of Proposed Combined CNN-LSTM Model

Embedding layer

Pre-processed dataset provides a meaningful sequence of words and unique, every word has unique ID. Embedding layer initializes the words to assign random weights. It also learns the embedding to embed all words in the training dataset. Embedding layer is used in different ways and is mostly utilized to learn embedding of words that can be saved to use in another model. The pre-trained Word2Vc model is used in this study for words embedding.

Convolution layer

The proposed CNN model consist of seven layers where three convolutional layers, three pooling layers and fully connected layer is the last one. It passes the word in the form of sentences to convolutional layers. Convolution layer convolve the input using pooling layers where pooling layer helps decrease the representation of input sentences, input parameters, computation in the network which check the overfitting in the network.

Global max-pooling

The global max-pooling is applied at the end of network layers and it provides the global best results after applying different convolution layers from the whole network.

Activation Function

The RELU activation function is used in our model and RELU gives zero atnegative values and it increases with positive values.

Dense layer

Dense layer which is also called fully connected layered to perform classification onthe extracted features of the convolutional layers. The dense layer is used every current input (neuron) in the layer of the network is connected to every input (neuron) in the proceeding layer of the network.

SoftMax

SoftMax is a function which is mostly used in the final layer of the neural network and it takes the average of the random results into 1 and 0 form.

5. Experimental Setup

The two standard datasets are used for evaluation of our proposed Combined CNN-LSTM model. First one is the IMDB movie reviews dataset available on ttp//rottentomatoes.com and second one is amazon movie reviews dataset available on https://www.kaggle.com/bittlingmayer/amazonreviews. More experiments are performed using the proposed model on both IMDB and Amazon datasets. The proposed model obtained better results with high precision, recall, f-measure, and accuracy as compared to traditional machine learning algorithms such as Naïve Base, Support Vector Machine.

5.1 MDB Movie Reviews Dataset

In [22] the benchmark IMDB dataset of movie reviews for sentiment analysis was first time is published. IMDB dataset contains 50,000 binary labeled reviews and we divide the datasetinto 80:20 training and testing cases. Each subset of data is balanced by the label distribution. And for the validation set, we used 10% labeled from training documents.

5.2 Amazon Movie Review Dataset

We remove irrelevant Hyper Text Markup Language (HTML) tags in the start from the dataset and normalize the dataset. Then we perform pre-processing on the dataset which include tokenization, space removal, punctuation removal and irrelevant words as stop word. There are 5000 examples of movie reviews, half of which is negative and the other half positive in the original dataset. To train the model 1900 examples are used and it is tested by 500 examples. In the dataset, 1 represents positive comment and 0 represents negative comment about the movie. Deep learning model takes input in thevector forms as mentioned above and using word2vec and changes the text into vector. The parameter settings of proposed CNN-LSTM model are shown in Table 1.

Table 1: Model Parameters

Tuning Parameters	CNN Model	LSTM Model
Learning Rate	0.01	0.01
Dropout	0.2	0.2
Embed size	300	300
Step size	20	20
No filters	256	256
Batch size	64	64

Table 2 shows the IMDB dataset comprising two columns: review and the corresponding sentiment value. This data set encompasses a total of 50,000 records.

Table 2: IMDB Dataset encompasses a total of 50,000 records

50,000 rows	Review	Sentiment
0	One of the other reviewers has mentioned that after watching just 1 Oz episode you'll be hooked. The...	positive
1	I thought this was a wonderful way to spend time on a too hot summer weekend, sitting in the air con...	positive
2	Basically there's a family where a little boy (Jake) thinks there's a zombie in his closet & his par...	negative
3	Petter Mattei's "Love in the Time of Money" is a visually stunning film to watch. Mr. Mattei offers ...	positive
4	Encouraged by the positive comments about this film on here I was looking forward to watching this f...	negative
….		
49995	The cast played Shakespeare. Shakespeare lost. I appreciate that this is tryin...	negative

49996	I had the terrible misfortune of having to view this "b-movie" in it's entirety. All I ha...	negative
49997	This film tried to be too many things all at once: stinging political satire, Hollywood blockbuster,...	negative
49998	"Ardh Satya" is one of the finest film ever made in Indian Cinema. Directed by the great director Go...	positive
49999	I just watched The Dresser this evening, having only seen it once before, about a dozen years ago.<b...	positive

6. Results and Discussion

Two common deep learning models (CNN, LSTM) and proposed Combined CNN-LSTM Model are implemented using two datasets IMDB and Amazon. Many experiments performed on IMDB sentiment analysis dataset to attempt a candid comparison with competitive techniques. In IMDB dataset, many sentences are contained in one movie review. The Combined CNN-LSTM model on IMDB dataset is applied and using word2vc technique to initialize the words as vector space and word2vec use skip gram and to convert the words in vector representation use bag to word techniques. The f-measure, recall and precision of our Combined CNN-LSTM model and different other techniques on two datasets is shown in Figure 4. Our results are initialized highlights on IMDB dataset and that the Combined model improves the f-measure score up to 4-8% when compared with CNN and LSTM individually. The Combined model used 10 convolutional layers to extracts local information in an efficient way as compared to the networks proposed in [13, 20]. Figure 5 show the accuracy of the proposed Combined CNN-LSTM model and traditional approaches (NB, SVM, GA). A machine learning hybrid approach NB-SVM performed better in term of accuracy, however it was applied on a small IMDB dataset with more parameters.

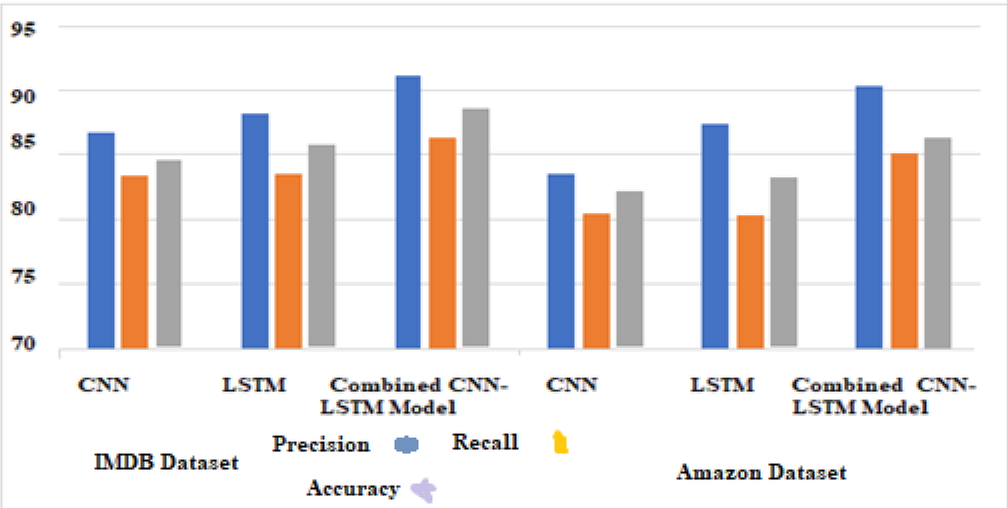

Figure 4: Comparison of Combined CNN-LSTM model with CNN and LSTM

Many experiments are performed on amazon movie reviews dataset and compared the results with traditional models. Figure 4 shows the f-measure, precision and recall ofmany different models using Amazon movie review dataset. The experimental results show that performance of our proposed deep learning models is better than traditional machine learning techniques. The proposed model used dropout technique which improved the execution time and also observed that our proposed Combined CNN-LSTM model improved accuracy with respect to baseline algorithms [3, 20, 22]. Figure 5 shows the performance of the proposed model that outperformed traditional machine learning techniques in term of accuracy on the IMDB movie reviews dataset.

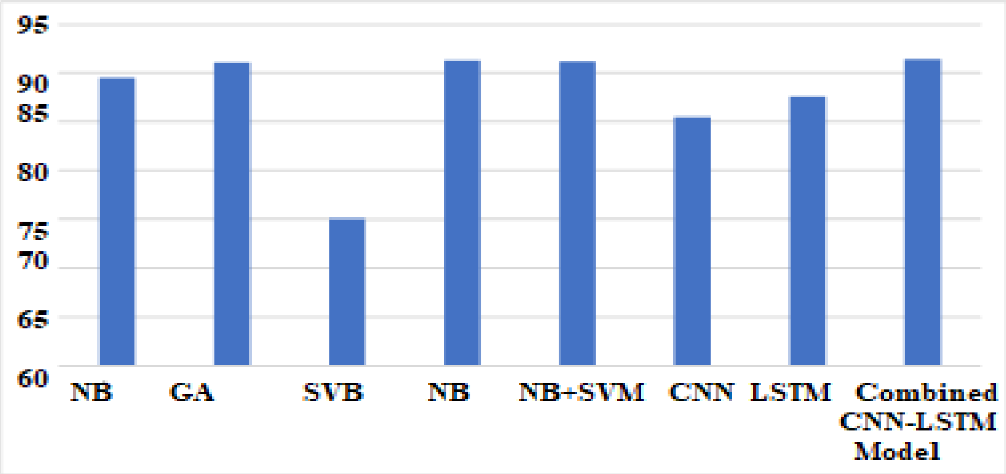

Figure 5: Accuracy Comparison of Combined CNN-LSTM model with traditional approaches on IMDB dataset

Conclusion

CNN assists to learn how to extract features from the data and it also needs many convolution layers to captures the long-term dependencies. Capturing dependencies becomes worse with the increase of input sequence of length in a neural network. Basically it accelerate towards a very deep layer of convolution neural networks and the LSTM model is effective to capture long-term dependencies between word sequences. The Combined CNN-LSTM model is proposed for sentiment analysis and the proposed model performed very well on two benchmark movie reviews datasets as compared to single CNN and LSTM models in terms of accuracy. As compared to traditional machine learning and deep learning models the proposed model achieved 91% accuracy.

References

[1] Ain QT, Ali M, Riaz A, Noureen A, Kamran M, Hayat B, Rehman A, “Sentiment Analysis using deep learning techniques: a review”, International Journal Advanced Computer Science Applications, 2017, 8(6),pp.425-433.
[2] Al-Smadi M, Talafha B, Al-Ayyoub M, Jararweh Y, “Using long short-term memory deep neural networks for aspect-based sentiment analysis of Arabic reviews”, International Journal of Machine Learning and Cybernetics, 2019, 1-13, https://link.springer.com/article/10.1007/s13042-018-0799-4.
[3] Amolik A, Jivane N, Bhandari M, Venkatesan M, “Twitter sentiment analysis of movie reviews using machine learning techniques”, International Journal of Engineering Technology,2016, 7(6),pp.1–7.
[4] Cheng Z, Ying D, Lei Z, Mohan K, “Aspect-aware latent factor model: Rating prediction with ratings and reviews”, In: Proceedings of the World Wide Web Conference on World Wide Web, 2018,pp. 639-648.
[5] Cheng Z, Ying D, Xiangnan H, Lei Z, Xuemeng S, Mohan K , “ A 3NCF: An Adaptive Aspect Attention Model for Rating Prediction”, International Journal IJCAI, 2018, pp. 3748-3754.
[6] Cheng Z, Xiaojun C, Lei Z, Rose C, Mohan K, “MMALFM: Explainable recommendation by leveraging reviews and images”, ACM Transactions Information System, 2019,37(2)pp.16.
[7] Collobert R, Weston J, “A united architecture for natural language processing: Deep neural networks with multitask learning”, in Proceedings 25th International Conference Machine Learning, 2018, pp. 160-167.
[8] Conneau A, Schwenk H, Barrault L, Lecun Y, “Very deep convolutional networks for text

classification", In Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics,2017, pp. 1107-1116.
[9] Elghazaly T, Mahmoud A, Hefny HA, "Political sentiment analysis using twitter data", In Proceedings of the ACM International Conference on Internet of things and Cloud Computing, 2016,pp.1-5.
[10] Fang X, Zhan J, "Sentiment analysis using product review data", Journal Big Data, 2015,2(1), http://dx.doi.org/10.1186/s40537-015-0015-2
[11] Govindarajan M, "Sentiment analysis of movie reviews using hybrid method of naive bayes and genetic algorithm", International Journal Advanced Computer Research, 3(4), pp.139-145.
[12] Hao H, " Recursive deep learning for sentiment analysis over social data", In Proceedings of the 2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT), 2014,pp. 180-185.
[13] Hassan A, Mahmood A, "Convolutional Recurrent Deep Learning Model for Sentence Classification", IEEE Access,2018, 6:13949–13957.
[14] Himelboim I, Smith MA, Rainie L, Shneiderman B, Espina C, "Classifying twitter topic-networks using social network analysis", Social Media+ Society,2017,pp. 1-13.
[15] Islam, J, Zhang Y, "Visual sentiment analysis for social images using transfer learning approach", In IEEE International Conferences on Big Data and Cloud Computing (BDCloud),2016, pp. 124-130.
[16] Kaur A, Gupta V, "A survey on SA and opinion mining techniques", Journal of Emerging Technology Web Intelligence,2017,5(4),pp.367–371.
[17] Li G, Hoi SC, Chang K, Jain R, "Micro-blogging sentiment detection by collaborative online learning", In IEEE 10th International Conference on Data Mining (ICDM), 2010,pp. 893-898.
[18] Liao S, Wang J, Yu R, Sato K, & Cheng Z, "CNN for situations understanding based on SA of twitter data", Procedia Computer Science, 2017,pp.376-381.
[19] Manek AS, Shenoy PD, Mohan MC, Venugopal KR, "Aspect term extraction for sentiment analysis in large movie reviews using Gini Index feature selection method and SVM", 2017, World wide web, 20(2),pp.135–154.
[20] Pang B, Lee L , "Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales". In Proceedings of the 43rd annual meeting on association for computational linguistics, 2005,pp. 115-124
[21] Ruangkanokmas P, Achalakul T, Akkarajitsakul K, "Deep belief networks with feature selection for sentiment classification", In IEEE 7th International Conference on Intelligent Systems, Modelling and Simulation (ISMS),2016, pp. 9-14.
[22] Singh J, Singh G, Singh R , "Optimization of sentiment analysis using machine learning classifiers", Human centric Computer Information Science,2017,7(1),pp.32.