IJREAT International Journal of Research in Engineering & Advanced Technology, Volume 8, Issue 5, October - November, 2020
ISSN: 2320 – 8791 (Impact Factor: 2.317)
www.ijreat.org

Score Level Fusion of two CNN using Co-Occurrence CBoW and Skip Gram Word Embedding models for Sentiment Analysis

[1]Swamy L N, Research Scholar, Department of CSE, Sahyadri College of Engineering & Management, VTU-Belagavi, Mangalore, Karnataka, India

[2]Dr. J V Gorabal, Department of Computer Science & Engineering, Sahyadri College of Engineering & Management, Mangalore, Karnataka, India

***Abstract*— Sentiment classification or opinion mining is an important domain research in Machine Learning and Text Mining. The research in this domain is having direct impact on the societal business, which is very much essential in today's technological life. In this article, we present the score level fusion of two convolutional neural network models based on different word embedding methods. Each method is having its own advantages and disadvantages. But when these methods are fused or merged, the understanding is that we will be the better result compared to conventional method. It is a unique kind of experimentation presented here. The proposed model is evaluated with publically available data sets. Large number of experimental trails reveals the effectiveness of the model. The details are presented in the respective section.**

Keywords: Sentiment Classification, Deep Learning.

I. Introduction

Opinion or sentiment is a kind of human opinion about an entity or about an object. It will represent the state of mindset related to an object or towards entity. I will show the customers view point with respect to a product. Opinion mining and sentiment analysis pays an important role and this will be represented in the form a review or face expression or emoji's [1,2]. To address the problem of sentiment analysis and opinion mining, artificial based solutions like machine learning and deep learning are the two solutions. These techniques includes prediction algorithms which help us to focus on the product's opinions from the customers. In deep learning, convolutional neural network is the popular layered network which is widely accepted by the research community. In CNN, major components include convolution layer, activation functions, optimizers and word embedding. The concept of embedding is defines the method or procedure for loading the sentiments in the form of text data to the system. Basically, effectiveness of the model depends on all the above parameters.[3]

Sentiment or opinion classification is the major application of natural language processing. It aims in extracting the sentiments from the text samples. State of the text mining techniques are used in this application. Here the mode of repressing the feelings is the form of textual data. State of the art preprocessing algorithms were used to prepare the data ready to the system. This domain is all about

presenting the polarities about a movie or about a product. This study will also help us to understand the publics view or opinion about the product. How publics has received it. In which direction they are looking forward for change in the product. One of the challenging task here is, understanding different language from the review is a challenging task. Every individual express his or her feelings in a different way. Understanding them and building the model to tackle them is a challenging task. In some few rare cases, a single sentence will have information related to both polarities. For example. This mobile is very good, but the camera quality is not upto the mark. In this example, first past is related to positive polarity and the second half is related to second half.

This research paper address the problem of sentiment analysis and opinion mining. Here single dimensional CNN model is designed. Two different approaches are presented based on the two different word embedding methods. Later outputs of these two models is fused to obtain global result of the model presented in this article. The details of the model is presented in the respective section.conference page limits.

II. LITERATURE SURVEY

In the history of sentiment analysis, many articles are found in this domain. But we mainly select the works related to the sentiment analysis or opinion mining using deep learning. In deep learning algorithms, it is obvious, term embedding is the major step in the processing of sentiments or emotions using deep learning models. Hence, this section collectively present some the major contributions related to it. We will start from the sentiment encoded term embeddings. For addressing a sentiment classification problem, if conventional representation models are considered like continues bag of word, skip gram, then the context of the terms will not be preserved. It is because, the terms fails to capture the proper contexts and it will be encoded to nearby vectors of the embedding space. In [5], word is related to the capturing of semantic and emotional information from the sentence is proposed. In [6], presented that conventional n-gram representation model in combination of latent representation captured more information and it is suitable for sentiment analysis problem. In [7] proposed embedding in the double space with logistic regression for sentiment of the each sentence for regularization. Let us focus on aspect level sentiment classifications. In this survey section, our main objective is to go for sentence level and document level approaches. In case of aspect level approach, emotional and target object both entities will be considered during classification. [8]presented Ada – RNN for opinion classification of twitter sentiments. In this technique, terms will be learnt based on the object or target by considering the context into account. All the features will be represented at the root node and softmax learning algorithm is considered for observation of samples between the classes. In [9], a LSTM based algorithm is proposed for sentiment classification using target information. The target information is also fed into the system as features. Hence the target information is considered which will give us the better performance in terms of classification rate. This approach is based on sentence level LSTM, which is having the capacity to handle interclass and intraclass opinions at the sentence level. The model is compared with state of the art techniques for its performance. The model implemented and tested in many real time situations. [10]proposed an algorithm which uses many features for the process. These features are extracted unconventionally using clustering algorithms. Multiple term embeddings with many pooling functions with lexicons are used for obtaining better performance in terms of accuracy. Among many deep learning architecture, LSTM is having the capability to capture both semantic and emotional features from text. [11]proposed LSTM with attention networks. It is proved for its own effectiveness. It forces the network to understand the targets dependent sentences present at the class level. Not only that, the attention features present in the network, enforces the network focus only on the important terms of the sentence. Hence this model address both sentence level and term level for classification of sentiments in a unique manner. The model is compared with state of the art techniques for its performance. The model implemented and tested in many real time situations.

III. PROPOSED MODEL

This section of the article present the details of the proposed score level fusion based model for sentiment analysis. Here two different CNN's will be designed. Later the results are fused to get the global result of the model. The stages of the fusion based approached is illustrated in Fig 1.

Convolutional neural network is one of the popular and widely accepted model, which works on layered architecture. The proposed model built using keras sequential layer development stage. Two different word embedding us considered here for the effective classification of the sentiments.

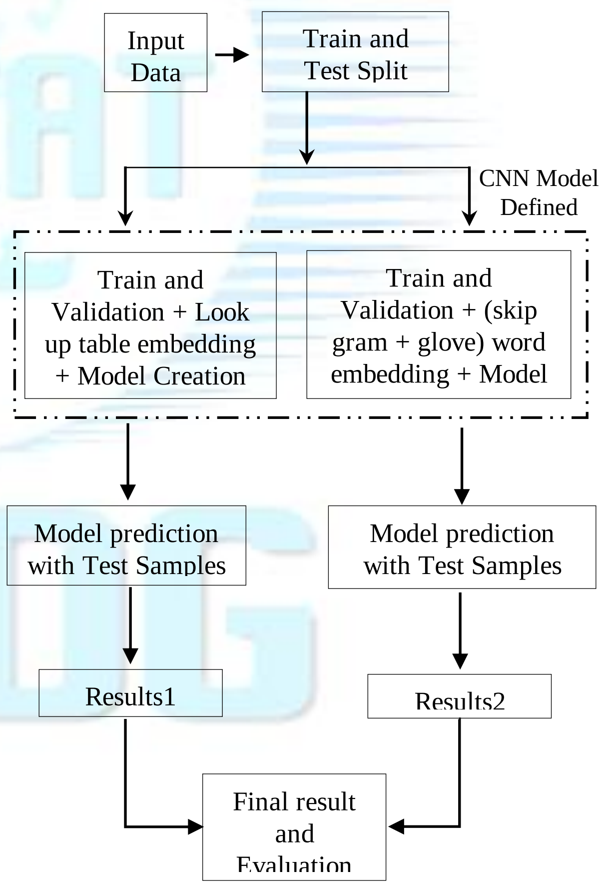

Fig 1. Stages present in the Proposed Score level Fusion Model.

The design of the layer is mainly got inspired from [1]. Based on it, two embedding layers namely co-occurrence and skip gram + glove models based cnn is designed. The proposed model is initialized with sentence matrix, where the training samples will be converted into vector and embedding into the system. This will make network to understand the importance of each token present in the network.

Co-occurrence word embedding: In co-occurrence word embedding model, the sliding window of size 3 will be created. This sliding window is made to go through all the sentences of the training samples. This will help to extract the relationship of ith word with i- and i+ word. For instance: This movie is good. Here 'good' term depends on the term movie. Such relationships can be easily maintained using co-occurrence word embedding. In this embedding, each review will be represented as a distributional term vector. This vector will be loaded to the network for training purpose. Later during testing stage, the unknown samples are brought to the same representation for processing purpose.

Skip Gram + Glove Embedding: In the second cnn, skip gram + glove word embedding is used to represent the review in the form a vector. The main purpose of the skip gram is, for a given term it will preserve the semantics of the term with other vectors or reviews. This way of representation will help to main the relationships with source term to the target terms. In this very important to know, what is the impact of termi with rest of the vectors of the data samples. This is kind of representation of semantics among the input reviews. Here, skip gram is used uniquely with glove to represent the terms with similar meaning. Advantage of this, terms with relationships will be preserved, glove help us to correlate the terms with similar meaning. For an instance: 'good' and 'satisfactory' are the two different terms which represent positive polarities. But both words not same syntactically. Such situation can be easily handled using skip + glove embedding.

Based on the above word embedding models, two different cnn's are defined. Other hyper parameters of the network include the activation function. The two cnn models are used relu activation. It is because, relu has given very good results compared to other activation functions. The outputs of the convolution layer will go through the pooling layer. This is kind of dimensionality reduction process. Since we are handling review in the form of text data and lot of information is required for it, max pooling is considered in the system.

At the last stage of the processing, optimizers pays a vital role in handling nonlinear data for the liner outputs. Generally, cnn will stay as concave optimization zone. Which means, vertical movements of the gradient need to be reduced and the oscillation to be moved to the global result or global minimum. This can be accomplished used RMS prop optimization algorithm. The gradient at the optimization stage need to be pushed towards the solutions. To achieve good optimization, learning rate and batch size of the network matters a lot. The complete model is algorithmically represented as follows.

Based on the designed two cnn models, the results are fused to obtain the global results. Both the models will provide the probability scores for each samples. The probability used for the global result estimation as follows.

$$Final_{output} = \frac{\text{probability}(\text{co} - \text{occurrence model}) + \text{probability}(\text{skip gram} + \text{glove model})}{n}$$

Where n: *in our case it is 2 (Number of models involved in generation of the results)*

IV. EXPERIMENTATION AND RESULTS

This is the experiment section of the article. Here details about the experiments conducted and results of the experiments are presented here. To carefully examine the model, huge and variety data sample is required. To create a variety of data sample, three different sets are generated based on the primary data sample. This model is evaluated using Twitter dataset. Based on it three other samples were created. Three trails of experiments conducted for the model. The details are as follows.

TABLE 1 : DETAILS OF THE DATASETS USED FOR EXPERIMENTS

Dataset name	**Trail 1**	**Trail 2**	**Trail 3**
Train and Test Ration	**50Train : 50Test**	**50Train : 50Test**	**50Train : 50Test**

The details of the experiments are presented in the following figures. Fig2. Present the model with co-occurrence word embedding. Fig3. Present the model with skip gram+glove word embedding and Fig4. Present the global result from both the models.
Present the global result from both the models.

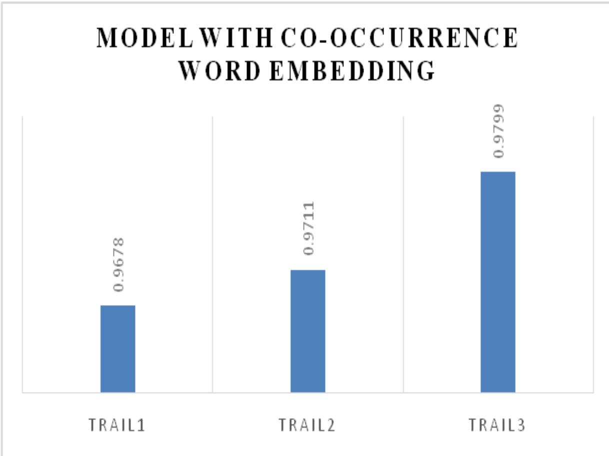

Fig 2. Model Performance with model with co-occurrence word embedding.

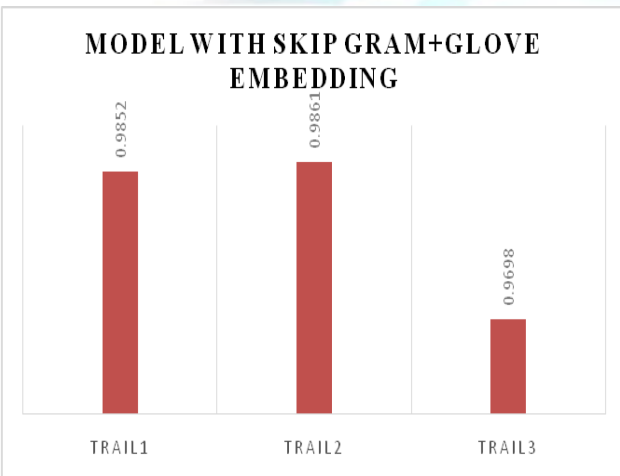

Fig 3. Model Performance with model with co-occurrence word embedding.

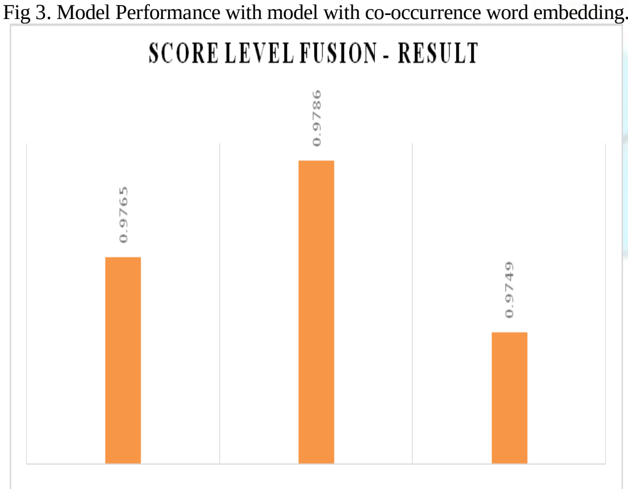

Fig 4. Model Performance of score level fusion.

V. CONCLUSION

This research article present a different manner of using multiple CNN for sentiment classification task. Here, authors have made a different novel way of using multiple cnn's for opinion mining operation. Two different cnn's were designed with various word embedding methods for effective way of capturing the semantic word relationships in the user reviews. . Each method is having its own. But when these methods are fused or merged, the understanding is that we will be the better result compared to conventional method. It is a unique kind of experimentation presented here. The proposed model is evaluated with publically available data sets. The results of the models have shown the good performance of the model. Due to multiple cnn's the model is well trained and it will behave well for the unseen samples. In future more focus can be given to the developing word representation models. It is one of the major score.

REFERENCES

[1] Kalchbrenner, N., Grefenstette, E. and Blunsom, P., 2014. A convolutional neural network for modelling sentences. arXiv preprint arXiv:1404.2188.

[2] Liu, P., Qiu, X., Chen, X., Wu, S. and Huang, X.J., 2015, September. Multi-timescale long short-term memory neural network for modelling sentences and documents. In Proceedings of the 2015 conference on empirical methods in natural language processing (pp. 2326-2335).

[3] Prasad, D.K., Liu, S., Chen, S.H.A. and Quek, C., 2018. Sentiment analysis using EEG activities for suicidology. Expert Systems with Applications, 103, pp.206-217.

[4] Liu, B., 2012. Sentiment analysis and opinion mining. Synthesis lectures on human language technologies, 5(1), pp.1-167.

[5] Maas, A.L., Daly, R.E., Pham, P.T., Huang, D., Ng, A.Y. and Potts, C., 2011, June. Learning word vectors for sentiment analysis. In Proceedings of the 49th annual meeting of the association for computational linguistics: Human language technologies-volume 1 (pp. 142-150). Association for Computational Linguistics.

[6] Bespalov, D., Bai, B., Qi, Y. and Shokoufandeh, A., 2011, October. Sentiment classification based on supervised latent n-gram analysis. In Proceedings of the 20th ACM international conference on Information and knowledge management (pp. 375-382).

[7] Labutov, I. and Lipson, H., 2013, August. Re-embedding words. In Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers) (pp. 489-493).

[8] Yin, Y., Wei, F., Dong, L., Xu, K., Zhang, M. and Zhou, M., 2016. Unsupervised word and dependency path embeddings for aspect term extraction. arXiv preprint arXiv:1605.07843.

[9] Ma, D., Li, S., Zhang, X. and Wang, H., 2017. Interactive attention networks for aspect-level sentiment classification. arXiv preprint arXiv:1709.00893.

[10] Liu, J. and Zhang, Y., 2017, April. Attention modeling for targeted sentiment. In Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers (pp. 572-577).

[11] Tan, C., Wei, F., Wang, W., Lv, W. and Zhou, M., 2018, July. Multiway Attention Networks for Modeling Sentence Pairs. In IJCAI (pp. 4411-4417).

[12] Swamy L N, Dr. J V Gorabal, "Concept of TF-IDF, Common Bag of Word and word embedding for effective sentiment classification." International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-9 Issue-6, April 2020.

[13] Swamy L N, Dr. J V Gorabal, "Convolutional neural network with Co-occurrence CBoW for Sentiment Analysis " Journal of Xi'an University of Architecture & Technology ISSN No : 1006-7930, Volume XII, Issue V, May 2020.

[14] Swamy L N, Dr. J V Gorabal, "Logistic Regression based Classification for reviews analysis on E-commerce based Applications" 7th International Conference on Frontiers of Intelligent Computing: Theory And Application (FICTA 2018), Springer AISC series on November 29th - 30th 2018.