ROLE OF SENTIMENT ANALYSIS USING DEEP LEARNING

[1]V. Roseline ,[2]Dr. HerenChellam G
[1]Research Scholar, Rani Anna Govt. College, Manonmaniam Sundaranar University,Tirunelveli - 8.
[2]Assistant Professor of Computer Science, Rani Anna Govt. College for Women,Tirunelveli–8
[1]rose_vasee@yahoo.com,[2]hchellam@gmail.com

ABSTRACT

Sentiment Analysis is a recent topic in the area of Natural Language Processing. Sentiment Analysis analyses the problem of forums, discussions, likes, comments, reviews uploaded on micro blogging platforms regarding about the views that they have an idea about a person, product, or event. This paper identifies the role of sentiment analysis with deep learning to classify the polarity of given information or the expressed view is positive, negative or neutral. Deep Learning is the up-to-date term in the area of machine learning. Machine Learning is a process to construct intelligent systems. Deep Learning is a method to utilize machine learning. The main goal of this paper is to find out the recent updates that relate to text classification of sentiment analysis. Sentiment Analysis is implemented in different approaches of deep level representation and also to find out the approach that generate output with high accurate results.

Keywords: Deep Learning, Machine Learning, Natural Language Processing, Sentiment Analysis.

I. INTRODUCTION

Sentiment Analysis is an increasingly emerging field whose objective is to classify sentiments and opinions expressed in text, image, video and audio generated by human. The amount of data that has been used on regular basis is the text data seems to be commonly used with other datum like audio, video and image. By utilizing deep learning techniques we can understand the sentiment better than traditional methods. Deep learning algorithms allow us to understand statement structure and semantics. Deep learning networks provide training in both supervised and unsupervised manners. It provides automated feature extraction and avoid involving human involvement therefore it can minimize time.

Sentiment analysis is the study of individuals' sentiments, mentalities and feelings. The most essential purpose in sentiment analysis is classifying the polarity of a text positive, negative or neutral. This purpose can be described as classifying the text into two classes i.e objective or subjective. This problem can be more challenging when classifying the polarity. When the text has an objective content, the text depicts regular statements

without expressing any emotions. Subjective text has generally expressed by a human having regular moods or feelings. The subjective depend on content and an objective may contain subjective sentences.

Text mining is the technique by examining large collections of written resources to generate new information and to transform the unstructured text into structured text for use in analysis. The accuracy of sentiment analysis is in how well it has the same opinion with human judgement. This is typically measured with the aid of various measures based on precision and recall of positive and negative texts.

Machine learning (ML) is a study of artificial intelligence uses mathematical techniques that allow computers to "learn" from data, without any detailed program. Machine learning explores the construction of algorithms that can learn and make estimations on data through constructing a prototype from sample inputs. Machine learning is used in a variety of tasks where programming algorithms with high performance was difficult. Machine learning is closely connected to computational statistics, which focus on predictions through computers.

Deep Learning machines usually work better than traditional ML tools because they t oo learn the feature extraction. Deep Learning performs better than other techniques if the size of the data is large. Deep Learning techniques require to have high end infrastructure to train in a period of time. When there is not enough knowledge of understanding the domain for feature observation, deep Learning techniques works better and we can have less worry about feature engineering. Deep Learning out perform for complex problems such as natural language processing, image classification and speech recognition. Deep learning methods has multi layer processing with less time and better accuracy performance. Sub sampling layers give better output, by using CNN and auto-encoders. With the increase number of auto encoders, the accuracy and the clarity increase. Similarly increase number of sub sampling too gives the better result.

II. OVERVIEW OF DEEP LEARNING

Neural network is impacted by human brain and has various neurons that make a magnificent network. Deep learning is capable of providing training to supervised and unsupervised methods. Deep learning has various networks such as Convolutional Neural Networks, Recurrent Neural Networks like Long Short Term Memory, Gated Recurrent Unit, Bidirectional Long Short Term Memory and many more. Neural networks are very advantageous in generating text, representation of word, sentence classification and modeling Sentence.

III. CONVOLUTIONAL NEURAL NETWORK

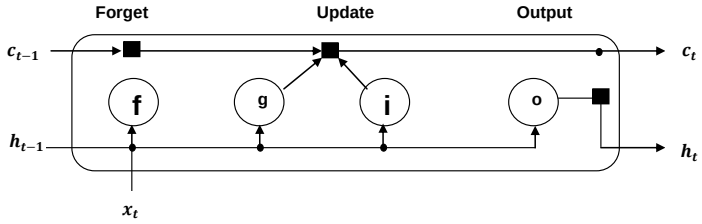
Convolutional neural network (CNN) is a neural network, comprising of convolution layers, pooling layers and fully-connected layers. Convolution layers and pooling layers are associated and then again the output of the layer is the next layer input. Convolution layer is an extraction layer, which extract the features through filters and create feature map which is calculated by convolution function, then output to pooling layer. Pooling layer is a mapping layer, which samples the features produced by the convolution layer and the local features. After a few convolution and pooling layers, there may be at least one fully-connected layers perform better.

By combining the pre-trained word feature generated using Global Vector (GloVe) polarity features and input, the feature sets into a deep CNN [1]. It avoids error and improves classification performance. CNN efficiently get the sentiment information in the sentence, retain the order of the word and reduce the problem of data sparseness. [11,13]

IV. LONG SHORT TERM MEMORY

Long Short Term Memory is a unit of Recurrent Neural Network. Recurrent Neural Network's and Feed-Forward Neural Networks are both named as the method they pass the information. In a Feed-Forward neural network, the information moves straight from the input layer, through hidden layers and to the output layer in only one direction. It also have no memory of input they received previously and so they are bad in assuming what's comes next. But in a RNN it takes the information of the current input and also the inputs it learned received previously. RNN has only a short-term memory. When with used with a LSTM they have a long-term memory.

LSTM has additional "forget" gates over the simple RNN. Its unique feature enables to overwhelm both the exploding and vanishing gradient problem.



Input activation	$a_t = \tanh(W_a . x_t + U_a . out_{t-1} + b_a)$ ---(1)
Input gate	$i_t = \sigma(W_i . x_t + U_i . out_{t-1} + b_i)$ ---(2)
Forget gate	$f_t = \sigma(W_f . x_t + U_f . out_{t-1} + b_f)$ ---(3)
Output gate	$o_t = \sigma(W_o . x_t + U_o . out_{t-1} + b_o)$ ---(4)

The LSTM network takes three inputs. X_t is the input, h_t denotes the hidden state, c_t denotes the cell state, h_{t-1} denotes the output from the preceding LSTM unit and C_{t-1} is the "memory" of the preceding unit, which is the most important input, f is the forget gate, g is the memory cell, i is the input gate and o is the output gate.

Deep learning-based recurrent neural networks (RNNs) is to classify short messages whether it is secure and non-insecure [2,10,12]. The Short messages are preprocessed,

features are extracted through word2vec and the short messages are classified via RNN by achieving the accuracy of 92.7% which is better than SVM.

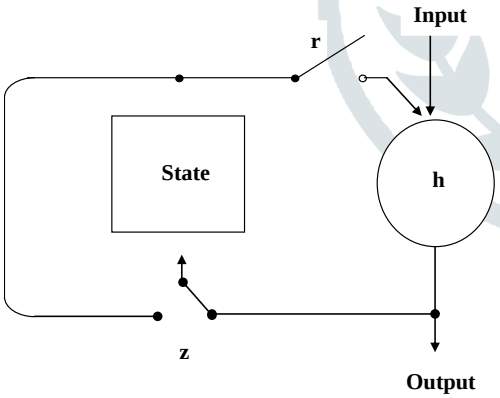
Sentiment classification based on Recurrent neural network (RNN)[3] choose the method of constructing a vector for each word in a sentence, then with RNN they train sentence vectors of different length, so that the sentence contain both word sequence and semantic features. Finally they used softmax regression classifier to predict sentiment orientation for each sentence. A basic LSTM model does not capture target information. So target dependent LSTM has significant big improvement over LSTM when target information is taken. By including target connection LSTM we obtain best performance in term of classifying accuracy. [4]

In traditional neural network, inputs are independent with each other. In Natural Language Processing it is important to know previous word to predict the next word in a sentence. In this work [5,11,13] input are taken as word embeddings and send to a convolutional Neural Network to bring out the high level features. The extracted output is given to a LSTM Recurrent Neural Network model and followed by a Classifier layer.

V. GATED RECURRENT UNIT

There are a number of differences in LSTM that are utilized today. One such variation of the LSTM is the Gated Recurrent Unit (GRU). It combines the input and forget gates into a “update gate”. Thus, it reveals the hidden information without any control.

The working of GRU is as follows:



Update Gate	$z_t = \sigma_g(W_z . x_t + U_z . h_{t-1} + b_z)$	---(5)
Reset Gate	$r_t = \sigma_g (W_r . x_t + U_r . h_{t-1} + b_r)$	---(6)
Output	$h_t = z_t . h_{t-1} + (1 - z_t) . \sigma_h(W_h . x_t + U_h (r_t . h_{t-1} + b_h))$	---(7)

Where x is the input for step t, z is the update gate which is the combination of new memory and previous memory, r is the reset gate which determines the proportion of new word and previous information in generating new memory and h is the new memory generated at step t. σ_g and σ_h is the activation functions.

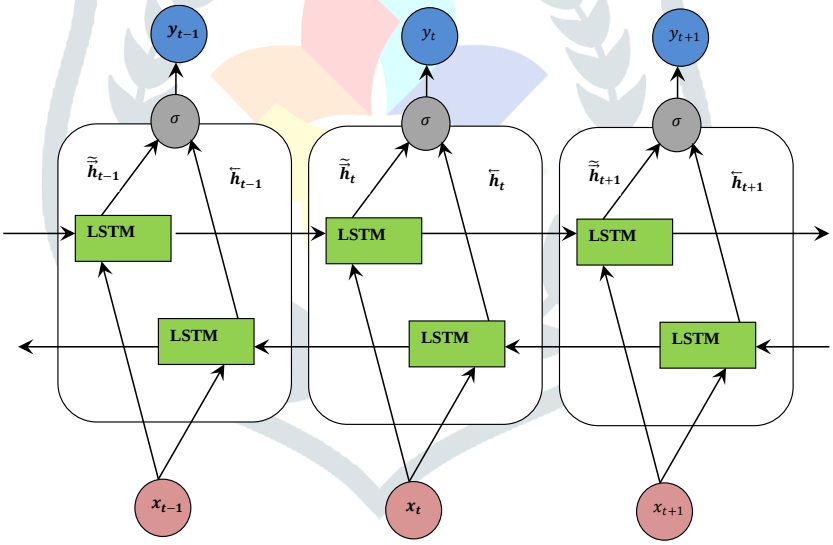
The CNN and RNN architecture in which the coarse-grained features produced by CNN and long-distance dependencies acquired via RNN for sentiment analysis of short texts. Then we apply LSTM and Gated Recurrent Unit for getting good results. [6,10]

Comprehensive Attention Recurrent Neural Networks was developed to store previous, next and local contexts of any position in a sequence [7]. To access the past and future information bidirectional recurrent neural networks is used, where a convolution layer utilized to capture local information. To further improve the effectiveness long short-term memory and gated recurrent unit were used. Another important feature of this model is that without any intervention of human it can be trained end-to-end. So the implementation was very easy.

The work [8] modeled sentence representation with convolution neural network or long short-term memory. Then the meaning of sentences and their relations are encoded in a document with gated recurrent neural network. Gated recurrent neural network outperformed recurrent neural network in modeling the document for sentiment classification.

VI. BIDIRECTIONAL LSTM – CONDITIONAL RANDOM FIELD

The LSTM used in the Bidirectional LSTM has two gates (an input gate i_t, an output gate o_t) and a cell activation vectors c_t. Bidirectional LSTM uses two LSTMs to learn each token of the sequence based on both the past and the future context of the token. The working of Bidirectional-LSTM is as follows:



As another commonly used sequence model, conditional random fields (CRF) is a discriminative undirected probabilistic graphical model, which represents a single log-linear distributions over structured outputs as a function of a particular observation input sequence.CRF describe the conditional probability of a set of output values.

Bidirectional LSTM CRF has more complicated hidden units than other DNN approaches. CRF with features gives comparable performance to LSTM, but lower performance than more complex DNN models.

Bidirectional LSTM give higher accuracy compared to LSTM since they can learn each token, in view of both previous and next content of the token [9, 14]. While LSTM only

use past context of the token, conditional random field (CRF) define the conditional probability of a set of output values. DCNN is used for classification of sentiment.

VII. CONCLUSION

In this paper, we present the survey of existing techniques for textual sentiment analysis with deep learning approaches. The deep learning techniques like Convolution Neural Network, Long Short Term Memory, Gated Recurrent Unit and Bidirectional Long Short Term Memory with Conditional Random Field have been expressed to classify polarity and have attained great success in sentiment analysis. Convolution neural network once trained their predictions was very fast and works best for more data points. By using Long Short Term Memory on top on Convolution network we can effectively reduce the number of Convolution layers to capture long term dependencies. In Gated Recurrent Unit training was faster than Long Short Term Memory and have performance high with less complexity. It is an efficient recurrent neural network than Long Short Term Memory. Bidirectional Long Short Term Memory with Conditional Random Field gives better performance compared to Long Short Term Memory because they can learn each token of the sequence based on both the past and the future context of the token, while Long Short Term Memory only use the past context of the token. Finally the performance of the deep learning techniques especially Bidirectional Long Short Term Memory with Conditional Random Field is impressive. Hence, the researchers can use deep learning techniques in the future to classify the polarity of text sentiment with high speed and accuracy.

REFERENCES

[1] Zhao Jianqiang , GuiXiaolin and Zhang Xuejun," Deep Convolution Neural Networks for Twitter Sentiment Analysis", IEEE Access, 2018, Volume: 6,Pages: 23253 – 23260.

[2] Jin Wang, Liang-Chih Yu, K. Robert Lai and Xuejie Zhang, "Dimensional Sentiment Analysis Using a Regional CNN-LSTM Model", Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, pages 225–230, Berlin, Germany, August 7-12, 2016.

[3] Zhang Yangsen, Jiang Yuru and Tong Yixuan, "Study of Sentiment Classification for Chinese Microblog Based on Recurrent Neural Network",Chinese Journal of Electronics Vol.25, No.4, July 2016.

[4] Duyu Tang, Bing Qin,Xiaocheng Fengand Ting Liu, "Effective LSTMs for Target-Dependent Sentiment Classification", arXiv:1512.01100v2 [cs.CL] 29 Sep 2016.

[5] Abdalraouf Hassan and Ausif Mahmood, "Convolutional Recurrent Deep Learning Model for Sentence Classification", IEEE Access,2018 , Volume: 6,Pages: 13949 – 13957.

[6] XingyouWang, Weijie Jiang and Zhiyong Luo, "Combination of Convolutional and Recurrent Neural Network for Sentiment Analysis of Short Texts", Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers, pages 2428–2437, Osaka, Japan, December 11-17 2016.

[7] Yong Zhang, Meng JooEr, Rajasekar Venkatesan, Ning Wangand MahardhikaPratama, “Sentiment Classification Using Comprehensive Attention Recurrent Models”, 2016 International Joint Conference on Neural Networks (IJCNN).

[8] Duyu Tang, Bing Qin and Ting Liu, ”Document Modeling with Gated Recurrent Neural Networkfor Sentiment Classification”, Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, pages 1422–1432,Lisbon, Portugal, 17-21 September 2015. c 2015 Association for Computational Linguistics.

[9] Tao Chen, Ruifeng Xu, Yulan He and Xuan Wang, ”Improving sentiment analysis via sentence type classification using BiLSTM-CRF and CNN”, Elsevier Volume 72, 15 April 2017, Pages 221-230.

[10]Sadam Al-Azani and El-Sayed M. El-Alfy, ”Emojis-Based Sentiment Classification of Arabic Microblogs Using Deep Recurrent Neural Networks”, March 2018 International Conference on Computing Sciences and Engineering (ICCSE).

[11]Qiongxia Huang, Riqing Chen, Xianghan Zheng and Zhenxin Dong, “Deep Sentiment Representation Based on CNN and LSTM”, 2017 International Conference on Green Informatics (ICGI).

[12]AdyanMarendraRamadhani and Hong Soon Goo, ”Twitter Sentiment Analysis using Deep Learning Methods”,2017 7th International Annual Engineering Seminar (InAES), Yogyakarta, Indonesia.

[13] Jin Wang, Liang-Chih Yu, K. Robert Lai and Xuejie Zhang, “Dimensional Sentiment Analysis Using a Regional CNN-LSTM Model”, Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, pages 225–230, Berlin, Germany, August 7-12, 2016.

[14] Muhammad Asif Khan, Tayyab Naveed, Elmaam Yagoub and Guojin Zhu, Donghua University, China and GIFT University, Pakistan, ”Bidirectional Long Short-Term Memory (BILSTM) with Conditional Random Fields (CRF) for Knowledge Named Entity Recognition in Online Judges (OJS)”, International Journal on Natural Language Computing (IJNLC) Vol.7, No.4, August 2018.