International Journal for Research in Engineering Application & Management (IJREAM)
ISSN : 2454-9150 Vol-06, Issue-10, JAN 2021

Context Sensitive Lexicon Features and Hybrid CNN-LSTM for Aspect based Sentiment Analysis

[1]M.Ramesh, Research Scholar, Department of CSE, Acharya Nagarjuna University, Guntur, Andhra Pradesh, India. ramesh.welcome@gmail.com

[2]Dr. Bandla Srinivasa Rao, Professor, Department of CSE, NVR College of Engineering and Technology, Tenali, Andhra Pradesh, India. sreenibandla@gmail.com

Abstract - Sentiment analysis aims to distinguish the direction of opinion in a portion of text. Manually analyzing all data is extremely unproductive and contains biased or fake reviews. Challenges are existence of related lexical features but dissimilar sentiments, different style of writing but identical sentiment, lexicons is not adequate for sentiment analysis and reviews may not be authentic. We propose a context sensitive lexicon-based method to confine contextual semantics polarity and extremely relevant aspects are confined using Hybrid CNN-LSTM to provide which will progress the accuracy of aspect based sentiment analysis. The experiments were done with the SemEval 2016 dataset consisting of consumer electronics like laptops and restaurants.

Keywords - Aspect-based, sentiment analysis, context sensitive, Hybrid CNN-LSTM, consumer reviews, polarity

I. INTRODUCTION

Sentiment Analysis make use of Natural Language Processing techniques to mine and classify emotions, evaluations, opinions, and attitudes associated to services, organizations, products, people, events and subjects articulated in open text. Aspect Based Sentiment analysis (ABSA) is Sentiment towards an aspect or opinion goal or feature. Like Polarities and strengths in sentiment attributes of words can serve to offer a word-level establishment for analyzing the sentiment of sentences as well as documents. We explore an effective method to use sentiment lexicon features [1].In polarity identification for a given set of aspect expressions contained by a sentence, resolved whether the polarity of every aspect expression is neutral,positive, negative or conflict i.e., mutually positive and negative.

More complicated sentence-level features such as the counts of positive as well as negative words, their whole strength, with the highest strength, etc, include also been exploited (Kim and Hovy, 2004; Wilson etal., 2005; Agarwal et al., 2011). Such lexicon features have been revealed highly efficient, leading to the finest accuracies in the SemEval shared task (Mohammad et al., 2013).In sentiment analysis there are aspect term extraction which is used for discovery of product's characteristics illustrated in textual system [2], aspect term sentiment opinion for detection of the sentiment polarity typically positive, negative or neutral and related to every aspect, and aspect aggregation is for not always present, executes the grouping of recognized aspects. And aspect term sentiment inference task consists of two main approaches the lexicon-based approaches, and machine learning ones.

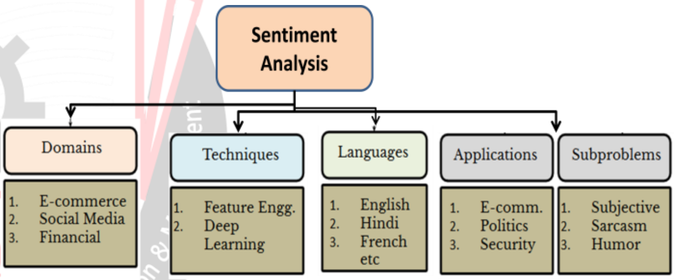


Figure 1: A broader view of Sentiment Analysis (SA)

We inspect a method that can potentially deal with the issues like bag-of-word models do not unambiguously hold semantic compositionality (Polanyi and Zaenen, 2006), by using a recurrent neural network (RNN) to confine context dependent semantic work effects above sentences and word sense differences (Devitt and Ahmad, 2007). The key proposal is to utilize a bi-directional long short term memory (LSTM) (Hochreiter and Schmidhuber, 1997; Graves et al., 2013) and CNN model to confine global syntactic need and semantic information, relied on which the weight of every sentiment word mutually with a sentence-level sentiment predisposition score are predicted.

Efficient and effective LSTMs for Target-Dependent Sentiment Classification [Tang et al. 2016] are Long Short-Term Memory (LSTM) which models the semantic illustration of a sentence exclusive of considering the target word being estimated. Whereas Target Dependent Long Short Term Memory (TD-LSTM) broaden LSTM by allowing for the objective word. And Target Connection

 DOI : 10.35291/2454-9150.2021.0001

Long Short-Term Memory (TC-LSTM) is a semantic related of objective with its context words is included [3]. Simple LSTM models the semantic illustration of a sentence with no considering the objective word being evaluated as no bias between the any two instances.

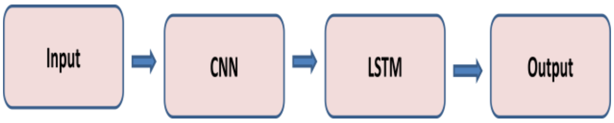


Figure 2: Architecture of Convolutional Neural Network and Long Short-Term Memory Network (CNN-LSTM)

II. RELATED WORK

A research study done by Citius [4] demonstrate that Naïve Bayes gave elevated routines in the procedure of sentiment analysis on Tweets in english.The Aspect Category discovery job can be consideration of as related to article classification task, which has a massive consideration of outstanding literature. Particularly investigating into categorization of reviews, (Kiritchenko et al., 2014) illustrated state-of-art presentation; by means of motivating linguistic along with lexicon features [5].And the author [5] also used an internal entity classification system to locate labels for Outside Term (O) in addition to Aspect Term (T).

Various studies explore whether the enclosure of an ontology progress results. In [6], the focal point is on a knowledge-based procedure that complements typical machine learning algorithms. The authors of [6] augment the sentiment analysis using area ontology information. To make simpler the sentiment study, author intended an aspect as well as sentiment association model (ASUM) [7] that determines join up of aspect as well as sentiment within a sentence, below the statement that single aspect with one sentiment be able to produce every words in every sentence.

Aspect-based assessment is the most consequential application in opinion mining, as well as researchers are in receipt of more concerned in product aspect extraction; nevertheless, more intricate algorithms are needed to deal with this issue correctly with bigger corpora [8]. Deep neural network methods together implement feature extraction and categorization for document classification [9]. CNN newly accomplished extremely performance on NLP tasks.

Table 1.Synopsis of deep learning based sentiment analysis.

Year	**Study**	**Research Work**	**Method**	**Dataset**	**Remarks**
2019	Abid et al. [10]	Sentiment analysis using recent recurrent variations	CNN,RNN	Social media - Twitter	Domain specific word entrenching
2019	Wu et al.[11]	Sentiment analysis by means of variation auto encoder	LSTM, Bi-LSTM	Social media - Facebook	Encoding, sentiment forecast, with decoding
2018	Li et al.[12]	How textual eminence of online reviews influence categorization act	LSTM, and CNN	Movie reviews	Influence of two significant textual features, specifically the word count with review understandable
2018	Jangid et al[13]	Economic sentiment analysis	CNN, LSTM, RNN	Economic tweets	Aspect-based sentiment analysis

III. THE PROPOSED HYBRID MODEL BY CNN AND LSTM:

Our proposed method is hybrid approach of CNN and LSTM based Aspect Sentiment Classification.

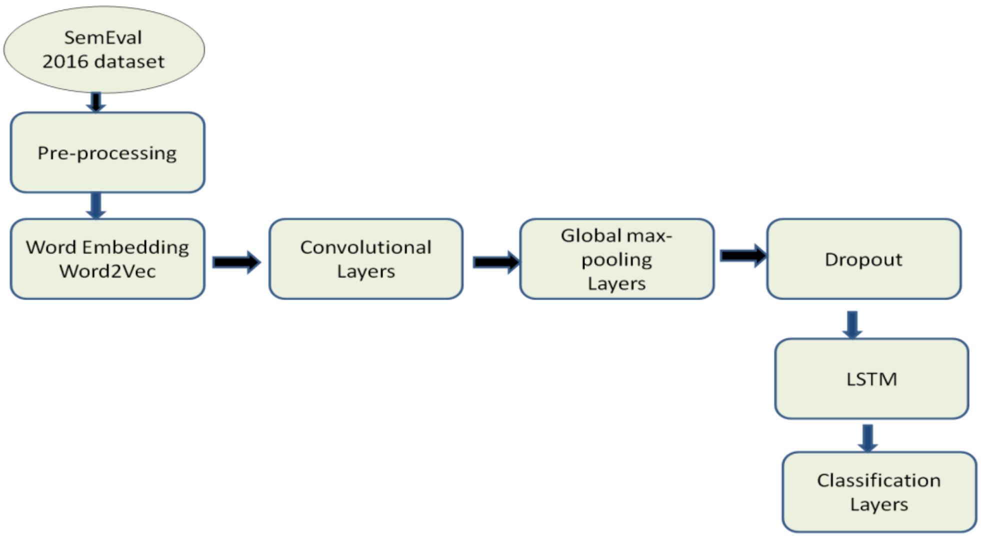


Figure 3: Proposed methodology Hybrid CNN-LSTM Model

 DOI : 10.35291/2454-9150.2021.0001

The practice of sentiment analysis is conferred below. Data cleaning with feature extraction were achieved in the preprocessing phases.

3.1 Preprocessing:

To find the sentiment of a review and to find negative comments or to find fake reviews about products a sequence of a text of indefinite length has to be transformed into a category of text. From a Machine Learning point of view, these are basically the same problem with presently the target labels altering and nonentity else[9]. With that said, the adding together of business acquaintance can help build these models more robust and integrate while preprocessing the data for text classification.

Review id:LPT1 (Laptop)
"The So called laptop Runs to Slow and I hate it! Do not buy it! It is the worst laptop ever."

For this preprocessing should match the preprocessing that was used prior to training the word embedding and get clear of is the unusual characters in text data.Next step is to clean the numbers and remove misspells & contractions

3.2 Word Embedding

After datasets be cleansed, sentences be divided into individual words, which be returned to their support and at this position; sentences were transformed into vectors of incessant real numbers acknowledged as feature vectors by means of means of word embedding. Word embeddings present a way to utilize an efficient, intense representation in which related words have a related encoding. An embedding is a opaque vector of floating point values. As an alternative of indicating the values for the embedding manually, they are trainable parameters .That is vectorizing the text and accumulate their embedding for upcoming analysis[10]. The function embed is the embed layer that is initialized by means of arbitrary weights and which will be trained the embed for every word in the training datasets.

3.3 Convolutional Neural Networks (CNN)

The CNN is a exceptional kind of neural network as well as engaged from the domain of image processing. Nevertheless, CNN representation has been efficiently used in text categorization. In CNN form, a division of key in to its prior layers is associated with a convolutional layer and CNN layers are called feature map. The CNN form makes use of polling layer to decrease the computational complication. The polling methods in CNN decrease the output dimension of one stack layers to subsequently in such a move that significant information is conserved. There are several polling techniques existing, though, max-polling is frequently used in which pooling window encloses max value component. The flattened layer is utilized to provide for the result of polling layer as well as maps it to subsequently layers. The concluding layer in CNN characteristically is completely connected. Figure 4 illustrates the fundamental architecture of CNN.

$$c_i := \sum_{k,j} (\mathbf{S}_{[i:i+h]})_{k,j} \cdot \mathbf{F}_{k,j}$$

A joining of every vector in a sentence makes a feature vector c ∈ Rn−h+1. The vectors be concerned then combined above all m filters into a feature map matrix C ∈ Rm×(n−h+1). The filters are trained throughout the training stage of the neural network.

3.4 Global max pooling layers

The result of the convolutional layer is accepted from side to side a non-linear activation function, prior to toward the inside a pooling laye[11]. The final combined vector elements by captivating the greatest over a predetermined place of non-overlapping intervals. Due to the exclusive structure of global pooling layers where the pool shapes equal the input shapes.

Global max pooling = normal max pooling layer through pool size equals to the dimension of the input (minus filter size + 1, to be accurate).

$$h_t = f_w(h_t - 1, x_t)$$

The above equation represents broad RNN representation where *ht* is the novel state at time *t*, fw is a function through *vwv* factor,*ht*-1is an previous state and *xt* is input vector at time *t.*

$$h_t = \tanh(W_{hh}h_1 + W_{xh}x_t)$$

3.5 Long Short Term Memory networks (LSTM)

LSTM which is a unique type of Recurrent Neural Network (RNN). In RNN, the weight connected with the input of the previous state (w1) and weight associated with output for the previous state are multiplied. Then passed to the Tanh function to

 DOI : 10.35291/2454-9150.2021.0001

obtain the new state. But RNN endures from a fading gradient problem that is very important changes in the weights that do not assist the model learn[12]. To defeat this LSTM was introduced. The arrangement of LSTM is chain like as well as it is alike to RNN, though, LSTM utilizes three gates to control and conserve information into each node state[13]. The details of LSTM gates along with cells are provided in the following equations.

$$\text{Input Gate } In_t = \sigma(W_{in}.[hs_t-1], x_t + b_{in}) \qquad (1)$$

$$\text{Memory Cell } C_t = \tanh(W_c.[hs_t-1], x_t + b_c) \qquad (2)$$

$$\text{Forget Gate } f_t = \sigma(W_f.[h_{st}-1], x_t + b_f) \qquad (3)$$

$$\text{Output Gate } f_. = \sigma(W_o.[h_{st}-1], x_t + b_.) \qquad (4)$$

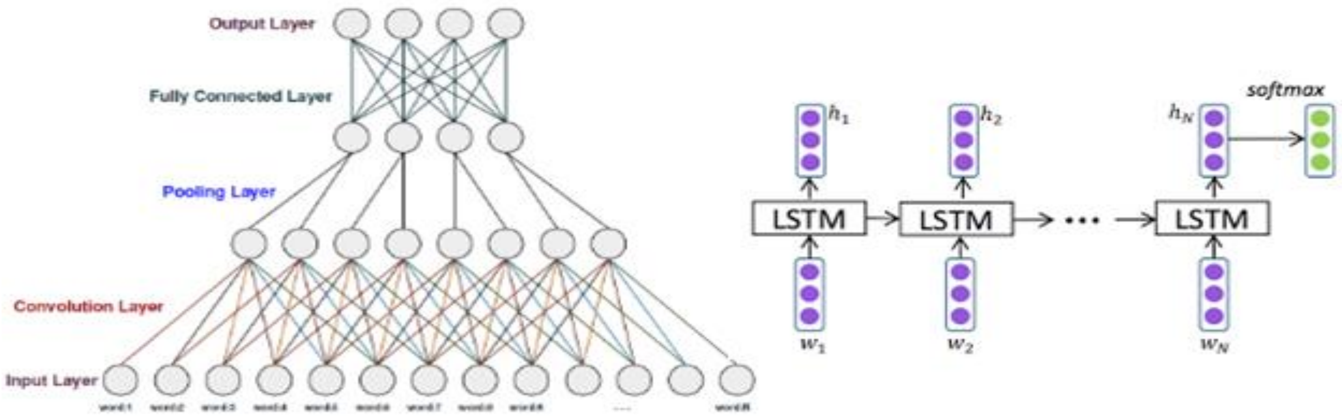


Figure 4:General architectures of CNN and LSTM

Dropout Techniques

We apply dropout method as it avoids form from over fitting. It drops the unrelated information from the network which does not give in further processing to improve the presentation of model.

Dense Layer

Dense layer has been used in the anticipated model. It connects every input with each output by means of weights.

SoftMax

It is a function that is frequently used in the last layer of the neural network. It gets the standard of the random results in to 0,1 outline.

3.6 Combined CNN and LSTM

In this learning, a combined routine was developed to routinely identify sentiments using laptop, restaurant reviews. The arrangement of this design was intended by merging CNN and LSTM networks, where the CNN is utilized to mine difficult features from reviews sentences given by customers, and where as LSTM is used as a classifier.

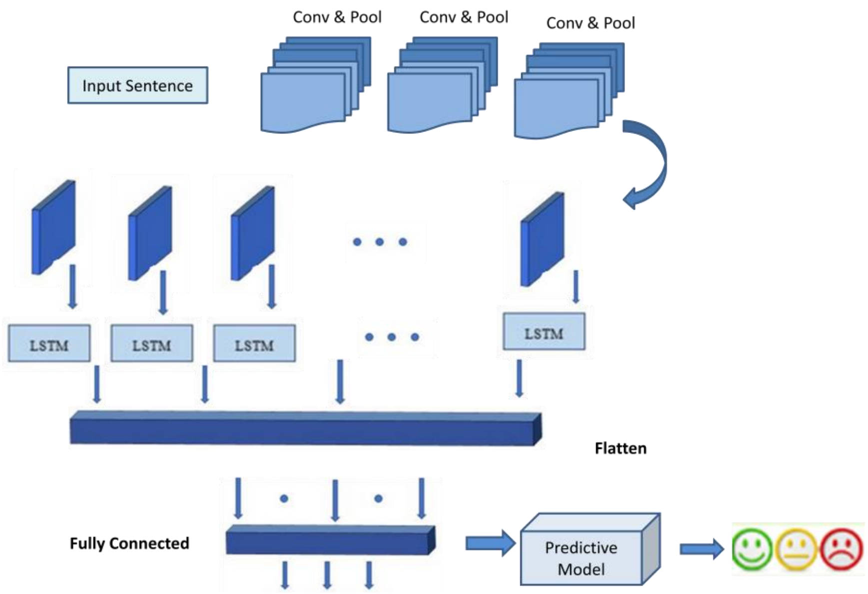


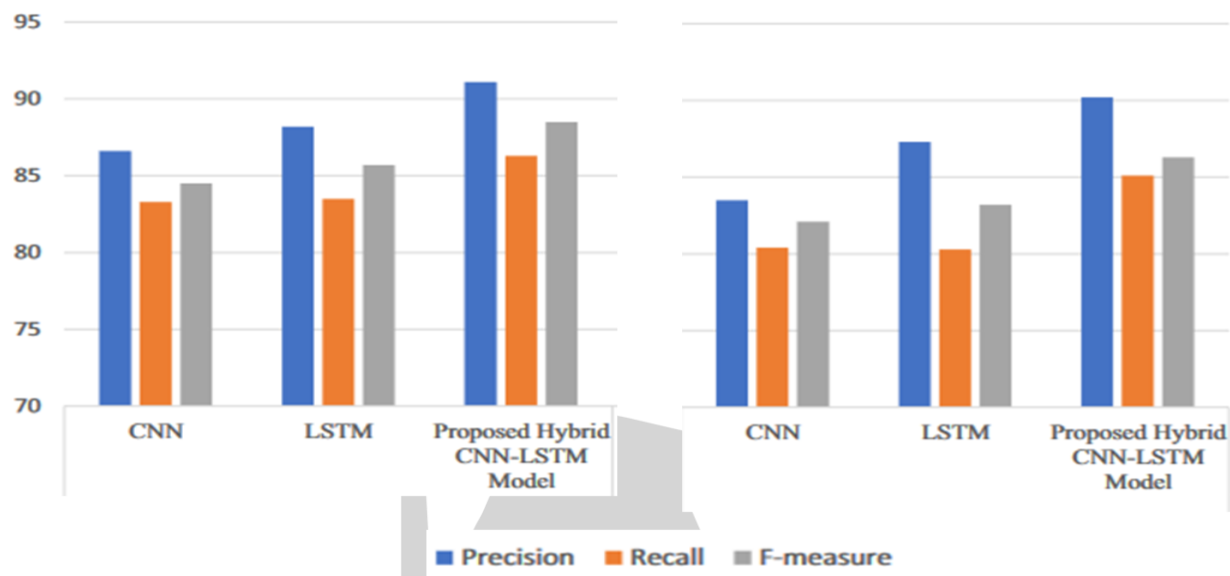
Figure 5: Illustration of the proposed method hybrid CNN-LSTM model

 DOI : 10.35291/2454-9150.2021.0001

International Journal for Research in Engineering Application & Management (IJREAM)
ISSN : 2454-9150 Vol-06, Issue-10, JAN 2021

IV. EXPERIMENTS & RESULTS

We executed two regular deep learning representations CNN, LSTM and proposed method Hybrid CNN LSTM Model on laptop, restaurant reviews. Execute many experiments on SemEval 2016 dataset sentiment analysis dataset to effort a straight similarity with viable techniques. We applied planned hybrid Cmethod NN LSTM model on reviews dataset. We used word2vc method to compose the words as vector space with word2vec utilize skip gram as well as bag-of-words method to alter the words in vector depiction.



(a) Laptop Reviews (b) Restaurant Reviews

Figure 6: Performance Anlysis of Hybrid CNN-LSTM representation with CNN & LSTM with respect to Precision, Recall and F-measure

The primary overviews of our outcome on dataset are that the proposed the hybrid model enhances the f-measure up to 4-8% when evaluated with CNN as well as LSTM separately. Hybrid model employ 10convolutional layers to mine neighboring information in a proficient way. The conventional CNN architectures are recognized to be capturing the lexical as well as structural features very fine. We include this idea into our effort to learn sentiment embed vector. Empirically successful method of pioneering sentiment lexicon features to state-of-the-art LSTM form for sentiment analysis. We performed experiment on both laptop and restaurant reviews dataset and compared their results among traditional models CNN and LSTM, and shown the f-measure, precision as well as recall. The results illustrate that performance of proposed hybrid model and improved accuracy is better than baseline algorithms CNN and LSTM.

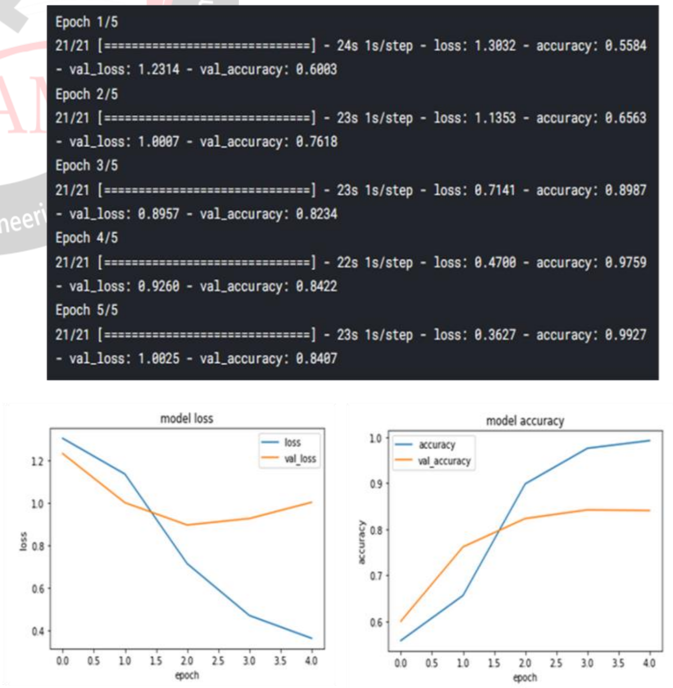


Figure 7: Model for 5 epochs with batch size 256 and visualization of accuracy

V. CONCLUSION

We may wrap up that the approaches used to build and combine lexicons were efficient; given the obvious

 DOI : 10.35291/2454-9150.2021.0001

improvement in the outcome of the application of the techniques using these lexicons .The LSTM representation is competent to confine long-term reliance among word sequences. CNN assist to find out how to mine features from the data. Nevertheless, it as well necessitates many convolution layers to incarcerate the long-term need, incarcerate needs becomes poorer with the augment of input progression of extent in a neural network. Essentially, it directs towards a extremely deep layer of convolution neural networks. We build a model Hybrid CNN-LSTM representation for sentiment analysis. The build model executed extremely well on two standard reviews datasets as measure up to particular individual CNN and LSTM baseline models in terms of accuracy. The implemented method Hybrid CNN-LSTM model accomplishes 91% accuracy when compared to CNN and LSTM baseline models.

REFERENCES

[1] Jiahua Chen, Shuai Wang, Sahisnu Mazumder, Bing Liu, A Knowledge-Driven Approach to Classifying Object and Attribute Coreferences in Opinion Mining, Findings of the Association for Computational Linguistics: EMNLP 2020, pages 1616–1626November 16 - 20, 2020.c©2020 Association for Computational Linguistics.

[2] l-Smadi M, Talafha B, Al-Ayyoub M, Jararweh Y (2018) Using long short-term memory deep neuralnetworks for aspect-based sentiment analysis of Arabic reviews. International Journal of Machine Learningand Cybernetics, 1-13.

[3] Cheng Z, Ying D, Lei Z, Mohan K (2018) Aspect-aware latent factor model: Rating prediction with ratingsand reviews. In: Proceedings of theWorld Wide Web Conference on World Wide Web, pp. 639-648.

[4] P. Gamallo and M. Garcia, “Citius: A naïve-bayes strategy for sentiment analysis on english tweets,” in Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014) (Association for Computational Linguistics and Dublin City University, Dublin, Ireland, 2014), pp. 171–175.

[5] Svetlana Kiritchenko, Xiaodan Zhu, Colin Cherry, Saif Mohammad, Detecting Aspects and Sentiment in Customer Reviews, Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), pages 437–442,Dublin, Ireland, August 23-24, 2014.

[6] K. Schouten, F. Frasincar, and F. de Jong.Ontology-Enhanced Aspect-Based Sentiment Analysis.InProceedings of the 17th International Conference onWeb Engineering (ICWE 2017), volume 10360, pages302–3320. Springer, 2017.

[7] Y. Jo, Aspect and sentiment unification model for online review analysis, Proc. Fourth ACM Int. Conf. Web Search Data Min. (2011), pp. 815-82.

[8] O. Gross, A. Doucet, and H. Toivonen, "Document summarization based on word associations," in Proc. of SIGIR'14, Gold Coast, Queensland, Australia, 2014, pp. 1023--1026.

[9] Kim, Y., Convolutional neural networks for sentence classification. In Proceedings of the 2014 Conference on Empirical Methods in Natual Language Processing (EMNLP). Pages 1746-1751, Doha, Qatar, October 2014, Association for Computational linguisitcs.

[10] Abid, F.; Alam, M.; Yasir, M.; Li, C.J. Sentiment analysis through recurrent variants latterly on convolutional neural network of Twitter. Future Gener. Comput. Syst.2019,95, 292–308.

[11] Wu, C.; Wu, F.; Wu, S.; Yuan, Z.; Liu, J.; Huang, Y. Semi-supervised dimensional sentiment analysis with variational autoencoder.Knowl. Based Syst.2019,165, 30–39.

[12] Li, L.; Goh, T.-T.; Jin, D. How textual quality of online reviews affect classification performance: A case of deep learning sentiment analysis.Neural Comput. Appl.2018, 1–29.

[13] Ashok Tholusuri, Manish Anumala, BhagyarajMalapolu, G. Jaya Lakshmi, Sentiment Analysis using LSTM, International Journal of Engineering and Advanced Technology (IJEAT)ISSN: 2249 –8958, Volume-8, Issue-6S3, September 2019.

 DOI : 10.35291/2454-9150.2021.0001