Sentiment Analysis by using Recurrent Neural Network

Alpna Patel[1] and Arvind Kumar Tiwari[2]

Abstract—Sentiment analysis is the process of emotion extraction and opinion mining from given text. This research paper gives the detailed overview of different feature selection methods, sentiment classification techniques and deep learning approaches for sentiment analysis. The feature selection methods include n-grams, stop words and negation handling. This paper also discusses about various sentiment classification techniques named as machine learning based approach and lexicon based approach. There is various classification algorithms such as SVM, Maximum Entropy and Naïve Bayes used for sentiment classification. In this paper we also discuss about deep learning models such as RNN, CNN and LSTM which is used for sentiment analysis. There are various application of sentiment analysis in decision making, prediction and business application.

Keywords: Sentiment Analysis, Deep Learning, Sentiment Classification, Machine Learning

I. INTRODUCTION

Sentiment analysis refers as opinion mining or subjectivity analysis for a given task [1]. Sentiment analysis is the combination of emotions, opinion and sentiments towards the given topic. It is the procedure of extracting opinion from particular task. Sentiment analysis is the powerful tool for specific event such as movies reviews and general elections [2]. It shows the user's view for different commercial sites such as Flipkart, Amazon, Myntra etc.

Sentiment analysis focuses on user's likes and dislikes to the prediction of particular products review, box office performance and general elections. Sentiment is polarity based analysis. Polarity can be further divided into three parts positive, negative and neutral.

The aim of sentiment analysis is naturally determine the public view whether it is positive, negative or neutral [3]. We can also apply sentiment analysis on commercial websites, audios, videos, social networking sites and images [4].

Now a day's opinion extraction has become a important field of research because generally people are usually show their opinion on social networking sites and commercial sites. There are some important notations related to sentiment analysis have given below:

1. *Subjective:* Gold is awesome movie.
2. *Objective:* Reema Kagti is the director of Gold.
3. *Polarity:* Polarity can be divided into three parts on the basis of opinion.
   a. *Positive:* I love natural beauty.
   b. *Negative:* This is disgusting.
   c. *Neutral:* I usually go outside in the evening.
4. *Sentiment level:* There are various level of sentiment analysis such as document level, sentence level and phrase level.

II. METHODOLOGY

We focus on opinion mining for classification task. The main objective of classification is related to sentiment analysis such as box office performance, analysis of blogs, web texts and general elections. We extract most relevant feature from sentences for sentiment analysis. We trained different models by using different classification techniques and different set of features.

Feature Selection is a primary task to extract the feature from sentences for sentiment classification. In order to select the feature, feature vector can be divided into two categories such binary feature and numerical feature which represents the frequency occurrences. Here the following texts feature:

- *N- grams*: It can be defined as n term in texts. If one can take one (unigram) or more than one word (bigram) at a time accordingly.
- *Stop words*: In article 'a' and 'the', in preposition 'in', 'near', 'beside' and in pronouns 'he', 'she', 'it' are stop words.
- *POS tagging*: It is related to parts of speech such as noun, pronoun, adjective, adverb etc. In a sentence adverb and objective holds the maximum sentiments.
- *Stemming*: It is defined as the procedure of wrap out the suffixes and prefixes from text. For a example: 'writing', 'written' can be 'write'.

[1,2]*Department of Computer Science & Engineering*
Kamla Nehru Institute of Technology
Sultanpur, India
E-mail: [1]1703alpanapatel@gmail.com
[2]arvind@knit.ac.in

- *Discourse Feature*: A sentence may contain multiple sentiments which can be represented using discourse based connectors.
- *Negation*: It is related to negation word 'not'. For an example: Dosa is not good. In this sentence 'good' is the positive word but 'not' inverts the meaning of whole sentence.

III. Techniques of Sentiment Classification

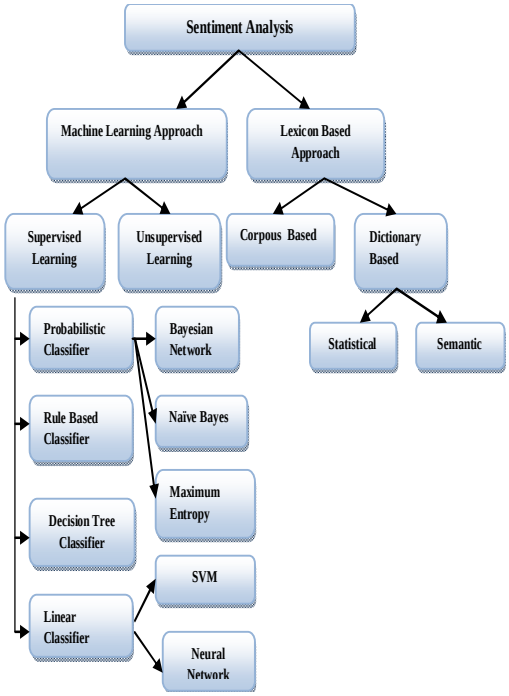

Figure. 1: Techniques of sentiment classification

A. *Lexicon Based Approach*

The objective of lexicon based approach is to find the opinion lexicon from given text. It is used for calculating sentiments. It focuses on counting the number of positive words and negative words in a given sentence. If the sentence gives more positive opinion then the sentence assigned to a positive score and if the sentence contains more negative words then it assigned to a negative score. If the sentence represents the equal number of positive words and negative words then it is assigned to a neutral score. Opinion lexicon is a tool for it [5]. There are two type of lexicon based approach

1. *Dictionary based approach*: It is the collection of opinion words which are collected manually. Opinion words help to form seed list. When we found new words (synonyms & antonyms) during sentiment classification, then it is added to seed list.
2. *Corpus based approach*: It is the collection of large amount of words based on specific topic. It helps to expand seed list [6]. It uses two basic approach named as statistical and semantic [7].

B. *Machine Learning Based Approach*

There are generally two type of approach such as supervised learning method and unsupervised learning approach used to train and test datasets. Supervised learning method uses well labeled data to train the machine. Each class contains different kind of features which labeled with it. Supervised learning method further divided into four classification methods named as Probabilistic classifier, Decision tree classifier, Linear classifier, Rule based classifier**.** There are different kinds of algorithms such as such as SVM [11], Neural Network [10], Maximum Entropy [9], and Naïve bayes [8] are used for classification.

IV. Deep Learning Based Models

A. *Recurrent Neural Network*

RNN is one of the deep learning approaches which are used for sentiment analysis. It produces the output on the basis of previous computation by using sequential information. Previously, traditional neural network uses independent inputs which are unfit for some task in Natural Language Processing. For an example: word prediction in a given sentence. RNN is efficient model for sentiment analysis. RNN uses memory cell that capable to capture information about long sequences, shown in fig. 2.

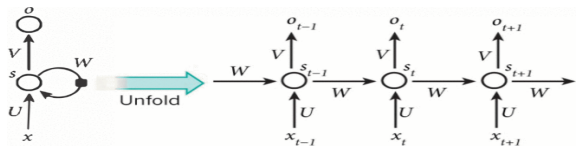

Figure. 2: RNN framework

The given equation shows the basic formula for RNN:

$$a_t = f(h_{t-1}, x_t) \quad (1)$$

Where a_t represents the output from previous node, activation function f is tanh function and x_t denotes the input sequences $(x_0, x_1, x_2, \ldots\ldots, x_t)$. Recursive neural networks represent the deep tree structure which is able to grab the semantic of texts. This process is time consuming task that is the disadvantage of Recursive neural network [22]. Recurrent neural network has improved the time complexity. In RNN model, analysis is done word by word and it is time taking process [12].

It is able to learn recent words compare to earlier words that leads the vanishing gradient problem. RNN uses LSTM to solve the problem of vanishing gradient.

B. *LSTM*

LSTM (Long Short Term Memory) is the part of RNN which is used to learn long-range dependencies for text sequences. LSTM contains memory blocks which also known as gates to control the text flow. The memory blocks contain three gates named as input gate; forget gate and output gate to control the flow of information [13], which is shown in fig. **3.**

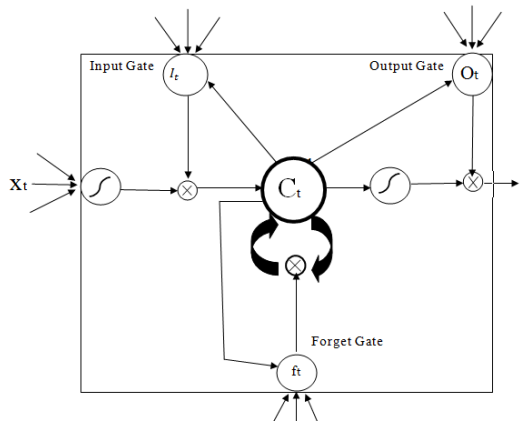

Figure 3: Long short term memory

C. *Convolutional Neural Network*

CNN is one of the deep learning models that were used in classification of sentence as opinion extraction and text tokenization [14]. CNN perform very well for Natural Language Processing. It preserves 2-D spatial orientation in texts, computer vision and images. CNN uses three layers such as convolutional layer, pooling layer and fully connected layer to perform the task.

D. *Word Embedding*

Deep learning models need word embedding as a input variable in NLP [21]. Word embedding is the process language modeling that converts texts into real numbers (e.g., word "fat"-(…., 0.18…0.26….0.35...)).In word embedding technique, each word is represented using dimension. **Word2Vec** is the prediction model which computes word embeddings from sentences. It includes CBOW (Continuous Bag of words) model and Skip- Gram model for feature learning.

V. Dataset Description

To evaluate the accuracy and error analysis of our model, we used IMDB movie reviews dataset for sentiment analysis [23]. This dataset contains movie reviews along with their associated binary sentiment polarity labels. It is intended to serve as a benchmark for sentiment classification. The core dataset contains 50,000 reviews split evenly into 25k train and 25k test sets. The overall distribution of labels is balanced (25k pos and 25k neg). We also include an additional 50,000 unlabeled documents for unsupervised learning. We used 20% of labeled training sets as validation set.

VI. Result and Discussion

In this paper to perform sentiment analysis, we used IMDB movie review dataset. Anaconda open source tool is used for python language to perform machine learning task. This paper used jupyter notebook for deep learning and Keras sequential model to implement Recurrent Neural Network. RNN model is used to evaluate the performance.

To achieve the sentiment analysis, split the dataset into two label; label 0 for negative review and label 1 for positive review. Deep learning methods are used to train data. After training the data we able to identify valid data for test classifier. The shape of train data is 15000, valid data shape 10000 and test data shape 25000.

In order to find unique word, NLTK uses different algorithm to remove punctuation, stop words and spaces from texts. The Sklearn uses different techniques such text, tfidfvectorizor for feature extraction by using data preprocessing techniques. The NLTK use porter stemmer and snowball stemmer algorithms for stemming the words. The word 'disappointed' stemmed as disappoint.

In order to get unique word, the NLTK used different feature selection methods such as stop word removing, negation handling, POS tagging and stemming.

The given graph shows the model training evaluation on dataset:

Figure 4: Model training evolution

A. *Performance Measure*

There are several parameters that is used to measure the performance. It is measured by using confusion matrix. Confusion matrix also known as error matrix. It includes TP (true positive), FN (false negative), TN (true negative), and FP (false positive) on test data.

Table 1: Confusion matrix

	Label 1 (Predicted)	**Label 2 (Predicted)**
Label 1 (Actual)	TN	FP
Label 2 (Actual)	FN	TP

With the help of these parameters, we calculate accuracy, recall, specificity etc.

- *Accuracy*: It is used for evaluating the performance of models by using binary classification. It is calculated by confusion matrix.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$$

- *Recall*: It is also known as sensitivity and it is calculated by using formula:

$$\text{Recall} = \frac{TP}{TP+TN}$$

- *Specificity:* It is calculated by using confusion matrix.

$$\text{Specificity} = \frac{TN}{TN+FP}$$

- *Precision*: It is also known as positive predictive values. It is calculated by using formula:

$$\text{Precision} = \frac{TP}{TP+FP}$$

- *F-measure*: It gives the measure on test's accuracy that is used both precision and recall. F-measure or F score calculated by using formula:

$$\text{F-measure} = \frac{2*precision*recall}{precision+recall}$$

- *MCC*: MCC stands for mathew correlation coefficient that is used for binary classification. It varies between ranges from -1 to +1. It can be calculated by using given formula:

$$\text{MCC} = \frac{TP*TN-FP*FN}{[(TP+FP)*(FN+TN)*(FP+TN)*(TP+FN)]^{1/2}}$$

The result achieved from test data is given in table II:

Table II: Evaluation of classifiers

Classifiers	Accuracy	Recall	Specificity	Precision	F-measure	MCC
RNN	0.8742	0.8717	0.8765	0.8753	0.8734	0.7484

In this paper, RNN model gives the accuracy 87.42% on IMDB movie review dataset. In order to evaluate the performance, we achieved; recall 87.17%, specificity 87.65%, precision 87.53%, f- measure 87.34% and MCC 74.14% value.

VII. CONCLUSION

Sentiment analysis is the process of extracting sentiments, attitude, emotions and opinion in given sentence. This research paper covered different sentiment classification techniques. Sentiment classification uses two approaches such as machine learning based approach and lexicon based approach for classification. Machine learning based approach is more suitable than lexicon based approach. This survey paper concludes various deep learning models for sentiment analysis. CNN and RNN are most popular deep learning algorithms for opinion mining. Recurrent Neural Network model produces the output on the basis of previous computation by using sequential information. This paper concludes that RNN gives the accuracy 87.42% on movie review dataset. Presently, sentiment analysis is the popular field of research.

REFERENCES

[1] Pang B, Lee L. , "Opinion mining and sentiment analysis" FoundTrends Inform Retriev:1–135, 2008.

[2] B. Heredia, T. M. Khoshgoftaar, J. Prusa, and M. Crawford, Cross-Domain Sentiment Analysis: An Empirical Investigation, 2016 IEEE 17th Int. Conf. Inf. Reuse Integr., pp. 160165, 2016.

[3] F. Luo, C. Li, and Z. Cao, Affective-feature-based sentiment analysis using SVM classifier, 2016 IEEE 20th Int. Conf. Comput. Support. Coop. Work Des., pp. 276281, 2016.

[4] Medhat, Walaa, Ahmed Hassan, and Hoda Korashy. "Sentiment analysis algorithms and applications: A survey" Ain Shams Engineering Journal 5.4:1093-1113, 2014.

[5] Kang Hanhoon, Yoo Seong Joon, Han Dongil., "Senti-lexicon and improved Naı¨ve Bayes algorithms for sentiment analysis of restaurant reviews", Expert Syst Appl ,39:6000–10, 2012.

[6] Keshtkar Fazel, Inkpen Diana., "A bootstraping method for extracting paraphrases of emotion expressions from texts" Comput Intell;vol. 0, 2012.

[7] Medhat, Walaa, Ahmed Hassan, and Hoda Korashy. "Sentiment analysis algorithms and applications: A survey" Ain Shams Engineering Journal 5.4 :1093-1113, 2014.

[8] Kang Hanhoon, Yoo Seong Joon, Han Dongil., "Senti-lexicon and improved Naı¨ve Bayes algorithms for sentiment analysis of restaurant reviews", Expert Syst Appl ,39:6000–10, 2012.

[9] Duric Adnan, Song Fei., "Feature selection for sentiment analysis based on content and syntax models", Decis Support Syst,53:704–11, 2012.

[10] Moraes Rodrigo, Valiati Joa˜o Francisco, Gavia˜o Neto Wilson P.,"Document-level sentiment classification: an empirical comparison between SVM and ANN", Expert Syst Appl ,40:621–33, 2013.

[11] Rui Huaxia, Liu Yizao, Whinston Andrew., "Whose and what chatter matters? The effect of tweets on movie sales", Decis Support Syst 2013.

[12] Socher, R., et al. Parsing natural scenes and natural language with recursive neural networks. in Proceedings of the 28th international conference on machine learning (ICML-Il). 2011.

[13] Senior, A. (n.d.). Long Short-Term Memory Recurrent Neural Network Architectures for Large Scale Acoustic Modeling Has. http://doi.org/arXiv:1402.1128

[14] Senior, A. (n.d.). Long Short-Term Memory Recurrent Neural Network Architectures for Large Scale Acoustic Modeling Has. http://doi.org/arXiv:1402.1128

[15] Kaur, H., Mangat, V., & Nidhi. (2017). A survey of sentiment analysis techniques. Proceedings of the International Conference on IoT in Social, Mobile, Analytics and Cloud, I-SMAC 2017, 921–925. https://doi.org/10.1109/I-SMAC.2017.8058315

[16] Yelena, M. (2009). Sentiment Analysis: An Overview Comprehensive Exam Paper. Computer Science Department, (May).

[17] Boiy, E., & Moens, M. F. (2009). A machine learning approach to sentiment analysis in multilingual web texts. Information Retrieval, 12(5), 526–558. https://doi.org/10.1007/s10791-008-9070-z

[18] Hassan, A., & Mahmood, A. (2017). Deep Learning approach for sentiment analysis of short texts. 2017 3rd International Conference on Control, Automation and Robotics, ICCAR 2017, 705–710. https://doi.org/10.1109/ICCAR.2017.7942788

[19] Yin, W., Kann, K., Yu, M., & Schütze, H. (2017). Comparative Study of CNN and RNN for Natural Language Processing, (July). https://doi.org/10.14569/IJACSA.2017.080657

[20] Zvarevashe, K., & Olugbara, O. O. (2018). A framework for sentiment analysis with opinion mining of hotel reviews. 2018 Conference on Information Communications Technology and Society, ICTAS 2018 - Proceedings, 1–4. https://doi.org/10.1109/ICTAS.2018.8368746

[21] Collobert R, Weston J, Bottou L, Karlen M, Kavukcuoglu K, and Kuksa P. Natural language processing (almost) from scratch. Journal of Machine Learning Research, 2011.

[22] Socher, R., et al. Parsing natural scenes and natural language with recursive neural networks. in Proceedings of the 28th international conference on machine learning (ICML-Il). 2011.

[23] https://www.kaggle.com/desiredewaele/sentiment-analysis-on-imdb-reviews/data