Neural Network Based Context Sensitive Sentiment Analysis

S.Suruthi, M.Pradeeba, A.Sumaiya
Pondicherry Engineering College
Pudhucherry, India

J.I Sheeba
Pondicherry Engineering College
Pudhucherry, India

Abstract: Social media communication is evolving more in these days. Social networking site is being rapidly increased in recent years, which provides platform to connect people all over the world and share their interests. The conversation and the posts available in social media are unstructured in nature. So sentiment analysis will be a challenging work in this platform. These analyses are mostly performed in machine learning techniques which are less accurate than neural network methodologies. This paper is based on sentiment classification using Competitive layer neural networks and classifies the polarity of a given text whether the expressed opinion in the text is positive or negative or neutral. It determines the overall topic of the given text. Context independent sentences and implicit meaning in the text are also considered in polarity classification.

Keywords: Sentiment analysis, neural network, data mining, implicit meaning, soft computing.

1. INTRODUCTION

Sentiment analysis also known as emotion mining refers to the identification of the emotion in the given text. Sentiment analysis determines the attitude of a speaker or a writer. Sentiment analysis of short texts such as online political debate post is challenging because of the limited contextual information which normally contain. Measuring public opinions emerge as a challenging task [1]. This type of posts are increasingly used to determine consumer sentiment towards a brand. The existing literature on sentiment analysis uses various methods that are used in many text classification problems. Sentiment classification has received considerable attention in the natural language processing research community due to its many useful applications such as online product review classification and opinion summarization[2][3].

Sentiment analysis has become a mainstream research field since the early 2000s. Its impact can be seen in many practical applications, ranging from analyzing product reviews to predicting sales and stock markets using social media monitoring [4][5]. The users' opinions are mostly extracted either on a certain polarity scale, or binary (positive, negative); various levels of granularity are also taken into account, e.g., document-level, sentence-level, or aspect-based sentiment [6].

Techniques for sentiment analysis can be broadly categorized into two classes of approaches. The first class involves the application of a sentiment lexicon of opinion-related positive or negative terms to evaluate text in an unsupervised fashion [7]. The second class of approaches utilize a textual feature representation coupled with machine learning algorithms to derive the relationship between features of the text segment and the opinions expressed in the writing in a supervised fashion [8]. Models based upon such supervised techniques require a large training set of instances complete with class labels, to calibrate the models and tune the relevant parameters. Labels can be assigned to training instances manually through human evaluation of the text, or resources with explicitly defined ratings (such as the number of stars assigned in movie and product reviews) can be leveraged.

2. RELATED WORKS

Sentiment topic models for social emotion mining presents two sentiment topic models called Multi-label Supervised Topic Model (MSTM) and Sentiment Latent Topic Model (SLTM).Both MSTM and SLTM can be applied to the tasks of social emotion classification and generating social emotion lexicons. Both MSTM and SLTM allow to distinguish between different affective senses of the same word, and to discover meaningful topics evoking social emotions [9].

In Sentiment analysis in Czech social media using supervised machine learning describes in-depth research on machine learning methods for sentiment analysis of Czech social media. It evaluates state-of-the-art supervised machine learning methods for sentiment analysis. It also explores different pre-processing techniques and employ various features and classifiers. It significantly outperformed the baseline in three-class classification and achieved an F-measure of 0.69 using a combination of features and preprocessing techniques [10].

In Bi-view semi-supervised active learning for cross-lingual sentiment classification proposed a new model based on bi-view classification by combining active learning and semi-supervised co-training approaches in order to reduce the human labeling effort in cross-lingual sentiment classification. Both labeled and unlabeled data are represented in the source and target languages using a machine translation service to create two different views of data [11].

Twitter brand sentiment analysis: A hybrid system using n-gram analysis and dynamic artificial neural network makes several contributions to twitter sentiment analysis, demonstrated through application on a corpus of tweets related to the Justin Bieber brand. Using this approach, they state that ''more than 80%'' of tweets contain no sentiment. Their process resulted in a more accurate estimation of sentiment in experimentation on the Justin Bieber Twitter corpus [12]. A neural network based approach for sentiment classification in the blogosphere proposed a neural network based approach which combines the advantages of the Machine language techniques and the Information Retrieval techniques. The back-propagation neural network has been selected as the basic learner. This method uses the results of the Semantic Orientation indexes as the inputs for the BPN. The experimental results indicate that method can efficiently increase the performance of sentiment classification [13].

Document-level sentiment classification aims to automate the task of classifying a textual review, which is given on a single topic whether it is positive or negative. Except for some unbalanced data contexts, their experiments indicated that ANN produce superior or at least comparable results to SVM's. ANN outperformed SVM by a statistically significant difference, even on the context of unbalanced data [14].

To overcome all the problems of the existing techniques the proposed framework has been introduced here.

3. THE PROPOSED FRAMEWORK

Figure 1 shows the overall architecture of the proposed work. The input data is taken from social media. In the first step, the input data is preprocessed that means stop words are removed. After preprocessing, feature extraction is done. In the feature Extraction, features like Noun, Adjective and Verbs will be extracted using QTag tool. All the words are tagged based on the POS Tagger. The occurrence of words are calculated in the frequency calculation step. The output of feature extraction has been given as input to the competitive layer neural network. Finally the positive, negative and neutral words are classified using competitive layer neural network. The topic of the given input is also identified in the proposed framework.
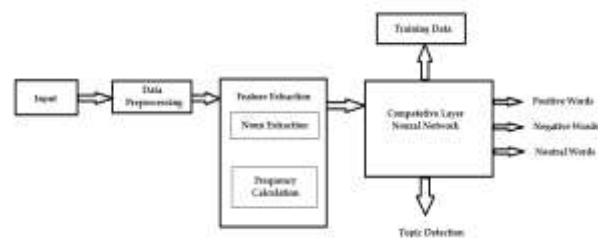

Figure. 1 Architecture of proposed framework

In the proposed framework in order to analyze the polarity of the given input the following steps are carried out.

- Input
- Data preprocessing
- Feature Extraction
- Competitive layer neural network

3.1 Input

The input data can be a conversation or comments or other types of unstructured documents for which the polarity is to be determined. This distribution contains the data used in the publication "Recognizing Stances in Ideological On-line Debates." There are 6 directories, one for each debate domain. There are totally 84 debates and 3921 posts are available in this dataset [16].

3.2 Data preprocessing

Data preprocessing is an important step in the data mining process. First, the reduced feature set decreases the computing demands for the classifier, and, second, removing irrelevant features can lead to better classification accuracy [10]. Quality decisions must be based on quality data. Quality data is obtained through preprocessing. First step is to remove the stop words from the dataset. The stop words add little semantic value to a sentence. For example "to", "I", "has", "the", "be", "or", etc. are the stop words. Stop words bloat memory space and processing time without providing any extra value.

3.3 Feature Extraction

3.3.1 Noun Adjective and Verb Extraction

Feature extraction is done to further reduce the data for achieving maximum efficiency in the output. All the sentiment words will be mostly bundled in the noun, adjective and verbal part of a sentence. So, the noun, verbs and the adjectives are extracted. Q tag tool is used to tag the sentence. This tagged words are parsed and only Noun Verbs and Adjectives are extracted.

3.3.2 Frequency Extraction

The frequency extraction involves extracting the occurrence count of the words in the data. This is used to find the overall topic of the given data. Only those words whose counts are above the threshold set will be considered.

Those words that are noun, adjective and verb and those words that are above the given threshold will be sent as input to the neural network. Feature extraction is done in order to reduce the processing time in the neural net.

3.4 Competitive layer Neural Network

It is a type of unsupervised neural network. Unsupervised networks are trained by letting the network continually adjust itself to new updates [15]. Competitive layer neural network recognizes and group similar input vectors. Some data whose individual meaning changes when seen as a whole sentence. Even in such context independent situation sentiment analysis are made to find a positive or negative output. Neural networks are organized in layers which is made up of a number of interconnected nodes. These nodes are also called as neurons. Input layer communicates with one or more hidden layers where actual processing is done in hidden layer and it is connected to the output.
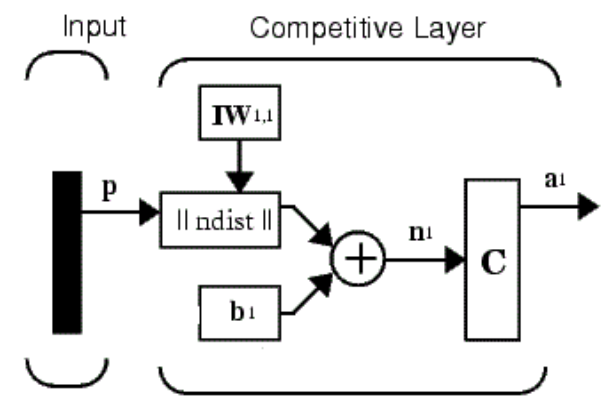

Figure. 2 Competitive layer neural network architecture.

The ‖ dist ‖ box in the figure 2 accepts the input vector p and the input weight matrix IW and gives a vector having elements S1. They are the negative of the distances between the input vector and vectors $iIW_{1,1}$ formed from the rows of the input weight matrix. Then n1 is the net input which is computed by adding the negative distance and the bias value. The competitive transfer function accepts a net input vector and returns 1 for all neurons that win the competition and 0 for those neurons that does not win the competition [15].

3.4.1 Kohonen learning rule:

The weights of the winning neuron are adjusted with the help of Kohonen learning rule. If the ith neuron wins, all the elements of the ith row of the input weight matrix are adjusted as shown below.

$$iW1,1(q)=iIW1,1(q-1)+\alpha(p(q)-iIW1,1(q-1))$$

The Kohonen rule will make the weights of a neuron to learn an input vector [15].

Thus, the neuron whose weight is closest to the input vector is changed such that it is even closer. Thus the winning neuron is more likely to win the competition next time and less likely to win when a very different input vector is presented. As more inputs are given, each neuron that is closest to a group of input vectors adjusts its weight vector toward the respective input vectors. Thus, the competitive network learns to categorize the input vectors it sees.

3.4.2 Bias leaning rule:

One of the limitations of competitive networks is that some neurons will not be allocated. Some neuron weight may start very far away from input vectors and never win the competition, no matter how long the training is made. These neurons, is called as dead neurons, does not perform any function [15].

In order to stop this, biasing is given to the neurons. A positive bias, added to the negative distance, makes a distant neuron to win. To do this an average of neuron outputs is taken. It is equivalent to the percentages of times each output is 1[15]. This average is used to update, so that frequently active neurons bias value become smaller, and infrequently active neurons bias value become larger.

Infrequently active neurons bias value increase and so the input space to those neurons respond increases. As the input space increases, the infrequently active neuron responds and moves toward input vectors. Thus the neuron responds to the same number of vectors as other neurons.

4. CONCLUSION

The proposed framework mainly used for finding the polarity of the text using competitive layer neural network whether it is positive or negative or neural. It also considers implicit meaning and context independent data in the polarity classification and the overall topic of the given data is identified. This proposed framework will give better accuracy when compared to the other techniques.

5. ACKNOWLEDGMENTS

6. REFERENCES

[1] E. Cambria, B. Schuller, Y. Xia, & C. Havasi, (2013) Knowledge-based approaches to concept-level sentiment analysis: new avenues in opinion mining and sentiment analysis, IEEE Intel. Syst. Vol.28, pp. 15–21.

[2] Kang, H., Yoo, S. J., & Han, D. (2012). Senti-lexicon and improved Naïve Bayes algorithms for sentiment analysis of restaurant reviews. Expert Systems with Applications, Vol. 39(5), pp. 6000-6010.

[3] Ku, L. W., Liang, Y. T., & Chen, H. H. (2006, March). Opinion Extraction, Summarization and Tracking in News and Blog Corpora. In AAAI spring symposium: Computational approaches to analyzing weblogs Vol. 100107.

[4] Stepanov, E. A., & Riccardi, G. (2011, December). Detecting general opinions from customer surveys., 2011 IEEE 11th International Conference on Data Mining Workshops (ICDMW) pp. 115-122

[5] Yu, L. C., Wu, J. L., Chang, P. C., & Chu, H. S. (2013). Using a contextual entropy model to expand emotion words

and their intensity for the sentiment classification of stock market news. Knowledge-Based Systems, Vol. 41, pp. 89-97.

[6] Sadegh, M., Ibrahim, R., & Othman, Z. A. (2012). Opinion mining and sentiment analysis: A survey. International Journal of Computers & Technology, Vol. 2(3), pp. 171-178.

[7] Turney, P. D. (2002, July). Thumbs up or thumbs down?: Semantic orientation applied to unsupervised classification of reviews. In Proceedings of the 40th annual meeting on association for computational linguistics pp. 417-424.

[8] Pang, B., Lee, L., & Vaithyanathan, S. (2002, July). Thumbs up?: sentiment classification using machine learning techniques. In Proceedings of the ACL-02 conference on Empirical methods in natural language processing, Vol. 10, pp. 79-86.

[9] Rao, Y., Li, Q., Mao, X., &Wenyin, L. (2014). Sentiment topic models for social emotion mining. Information Sciences, Vol. 266, pp. 90-100.

[10]Habernal, I., Ptácek, T., & Steinberger, J. (2013, June). Sentiment analysis in Czech social media using supervised machine learning. In Proceedings of the 4th workshop on computational approaches to subjectivity, sentiment and social media analysis pp. 65-74.

[11]Hajmohammadi, M. S., Ibrahim, R., &Selamat, A. (2014). Bi-view semi-supervised active learning for cross-lingual sentiment classification. Information Processing & Management, Vol. 50(5), pp.718-732.

[12]Ghiassi, M., Skinner, J., &Zimbra, D. (2013). Twitter brand sentiment analysis: A hybrid system using n-gram analysis and dynamic artificial neural network.Expert Systems with applications,Vol. 40(16), pp. 6266-6282.

[13]Chen, L. S., Liu, C. H., & Chiu, H. J. (2011). A neural network based approach for sentiment classification in the blogosphere. Journal of Informetrics, Vol.5(2), pp.313-322.

[14]Moraes, R., Valiati, J. F., &Neto, W. P. G. (2013). Document-level sentiment classification: An empirical comparisonbetween SVM and ANN. Expert Systems with Applications, Vol.40 (2), pp.621-633.

[15] http://in.mathworks.com/help/nnet/competitive-layer.html

[16] http://www.aclweb.org/anthology/W10-0214.