Comparison of Ensemble Sentiment Analysis using Convolutional Neural Networks

[1]M. Aruna Safali, [2]Dr. Ch. Suneetha

[1]Research Scholar, Dept of Computer Science & Engg., Acharya Nagarjuna University, India.

[2]Associate Professor, RVR & JC College of Engineering, India.

[1]marunamtech@gmail.com, [2]suneethachittineni@gmail.com.

Abstract: Sentiment analysis is the fast-growing field in the present world and this is also the part of the data mining. Analysing the sentiments of various documents, texts and reviews can be done by the various application or tools. To improve the results for given inputs various NLP, AI and deep learning algorithms are implemented for the better results. The proposed sentiment analyzer is language independent i,e this can analyze any of the language called as multilingual sentiment analysers. Comparative results show the performance of the various algorithms.

Keywords: Artificial Intelligence, NLP, text mining.

I. INTRODUCTION

Various emotions are present in the form of text called sentiment analysis. SA classifies the various emotions such as positive, negative, stress and neutral. Sentiments are analyzed based on the sentences, tweets, reviews, and various documents. This can be also analyzed the opinion of the users. Text may in various formats i,e document format, text file format and concept format. Various classification techniques are implemented to get sentiment analysis for the various types of documents, sentences etc. To get the better sentiment analysis (SA) various NLP and deep learning techniques can be utilized for the better analysis.

In general, some subjective sentiment expressions are pointed out necessarily and this is explained by Wilson et al. [2]. For the short documents, it is known that there is no basic difference between document and sentence level classifications [3]. In many applications, it is important that document level or sentence level classification text does not supply needed information or opinions on various features of every entity. Based on the above level it is known that these are not efficient because of their nature. In this paper, the proposed level of sentiment aims to classify the sentiment based on the specific aspects of objects. Firstly in this, to find the objects and their features or aspects. Various opinions are given by the various opinion holders for the same object. For example, "The food in this restaurant is not good but the cost of each item is low compared with other restaurants". This is taken consideration by the proposed aspect.

In the last few years, many applications and enhancements are done on SA algorithms and techniques. The proposed Ensemble Feature Analysis Classifier (EFAC) is implemented and this will take two parameters into consideration I,e review rating and sentiment analysis for the given sentence which is done with artificial intelligence and with the convolutional neural network algorithm.

II. LITERATURE SURVEY

This section, describes various sentiment analysis techniques.

Bruce and Wiebe made an effort to manually tag sentences as subjective or objective by different judges and the resultant confusion matrix was analyzed [19]. 14 articles were randomly chosen and every non-compound sentence was tagged. Also a tag was attached to conjunct of every compound sentence. Authors then attempted to identify if pattern exists in agreement or disagreement between human judges. Authors observed that manual tagging suffered due drawback of biased nature of human beings during tagging phase.

Subrahmanian and Reforgiato graded sentiments by the combination of adjective, verb and adverb [11]. In contrast to the algorithms that extracted the sentiments using adjective - verb combination or adverb - adjective combination, the model was trained using adjective, verb and adverb combination. The opinion was drawn from eight combinations of positivity or negativity of adjective, verb and adverbs in the reviews.

Cai et al. stated that solution for sentiment analysis should include a sentiment classification scheme as well as a sentiment topic detection scheme [15]. The sentiment classification component measured the relative sentiment (on a positive/negative scale) expressed by the words. The sentiment topic detection component detected the most significant topics hidden behind each sentiment category

 DOI : 10.35291/2454-9150.2019.0005

using a combined Point-wise Mutual Information and word support metrics.

Twitter sentiment classification is rather different than movie and product reviews, mostly due to the language style and the shortness of the tweets, but is an increasingly popular topic [7, 8, 11, 12, 16]. The upside of the informal and slanglike language is the presence of emoticons or smileys that can be used as an indicator of the sentiment the author wants to express. Automatically assigning sentiment based on emoticons greatly facilitates the creation of training and testing corpuses. In [17] instead of classifying the sentiment of a tweet, the authors determine the general sentiment about a specified hash tag.

Natural Language Processing (SentiWords)

The next resource containing roughly 155,000 English words is SentiWords. Words are associated with a sentiment score included between -1 and 1. Words are in the form lemma#PoS and are aligned with WordNet lists that include adjectives, nouns, verbs and adverbs.

AFINN is a manually labeled by Finn Årup Nielsen in 2009–2011 list of English words rated for valence with an integer between minus five (negative) and plus five (positive). It is possible to try it on this site.

III. RESULTS AND DISCUSSIONS

Various machine learning techniques are implemented with artificial intelligence and other deep learning techniques. Our research is totally based on knowledge discovery and sentiment analysis by using various algorithms and techniques to get accurate results. Previously there are various supervised machine learning techniques show a better performance compare with the unsupervised lexicon based methods. According to the researchers, it is known that Support Vector Machines (SVM) has high accuracy than other algorithms. In this paper, the research is based on opinion mining, knowledge discovery and sentiment analysis. The performance of proposed algorithms and techniques are discussed.

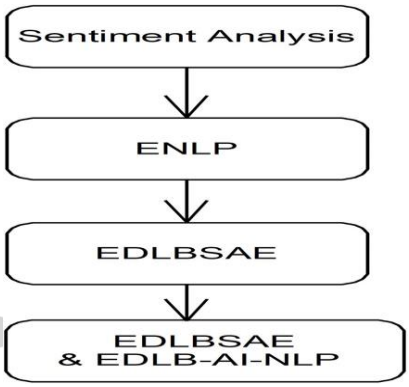

Figure: 1, Proposed Sentiment Analysis

Author	Dataset	Technique & Accuracy	Dis-advantage	Limitation
RuiXia, FengXu, JianfeiYu(2016)	Movie Review	PSDEE Approach based on Rule based and statistical method. (87.1%)	Polarity Shift Problem	Limited in Accuracy
Yanyan Zhao, Honglei Guo,Zhongsu and Ting Liu (2015)	Chinese Blog Review Dataset Dataset	Sentimentcompression technique before the aspect based sentiment analysis. Manual_comp_ssc = 88.95 Auto_comp_scc= 8 7.95	Challenge to syntactic parsers	Extractive compression technique fails to achieve accuracy
Ziang Lia, Wei Xu,Likuran Zhang(2014)	UIC(Unemployment Initial Claims) values between jan.2004 and Mar.2012 from the official website of the US Dept of Labor	Domain Ontology 81.7%	To improve the accuracy of unemployment rate prediction	Rate Prediction is Not Accurate
M. Aruna Safali, Dr. Ch. Suneetha [16]	Synthetic dataset	ENLP &AI 93.5%	To improve the result rate	Limited for text documents not for movie review

Table: 1, Various Methods used for Processing of Sentiment Analysis

IV. CONCLUSION

In this paper, various comparisons are explained ensemble approach which is merged with the features of the text mining approach, NLP technique, and artificial intelligence, deep-learning and artificial intelligence-based TSA prediction method that comprises of a stacked auto encoder (SAE) model that is used to learn generic linguistic and text semantic features and it is trained in a layer wise greedy fashion and Ensemble Feature Analysis Classifier to incorporate the new domain dimension within the rating and text-based sentiment analyzer.

REFERENCES

[1] R. Sagayam, A survey of text mining: Retrieval, extraction and indexing techniques, International Journal of Computational Engineering Research, vol. 2, no. 5, 2012.

 DOI : 10.35291/2454-9150.2019.0005

[2] N. Padhy, D. Mishra, R. Panigrahi., "The survey of data mining applications and feature scope," arXiv preprint arXiv:1211.5723, 2012.

[3] W. Fan, L. Wallace, S. Rich, and Z. Zhang, "Tapping the power of text mining," Communications of the ACM, vol. 49, no. 9, pp. 76–82, 2006.

[4] S. M. Weiss, N. Indurkhya, T. Zhang, and F. Damerau, Text mining: predictive methods for analyzing unstructured information. Springer Science and Business Media, 2010.

[5] S.-H. Liao, P.-H.Chu, and P.-Y. Hsiao, "Data mining techniques and applications–a decade review from 2000 to 2011," Expert Systems with Applications, vol. 39, no. 12, pp. 11 303–11 311, 2012.

[6] Oza, K.S., Naik, P.G., 2016. Prediction of online lectures popularity: a text mining approach. ProcediaComput. Sci. 92 (2016), 468–474.

[7] Sun, S., Luo, C., Chen, J., 2017. A review of natural language processing techniques for opinion mining systems. Inf. Fusion 36 (2017), 10–25.

[8] Muhammad, A., Wiratunga, N., Lothian, R., 2016. Contextual sentiment analysis for social media genres.Knowl.-Based Syst. 108, 92–101.

[9] Lee, N.Y., Kim, Y., Sang, Y., 2017. How do journalists leverage Twitter? Expressive and consumptive use of Twitter. Social Sci. J. 54 (2), 139–147.

[10] Corley, C.D., Cook, D.J., Mikler, A.R., Singh, K.P., 2010. Text and structural data mining of influenza mentions in web and social media. Int. J. Environ. Res. PublicHealth 7 (2), 596–615

[11] Liang-Chih Yu, Jheng-Long Wu, Pei-Chann Chang, Hsuan-Shou Chu, "Using a contextual entropy model to expand emotion words and their intensity for the sentiment classification of stock market news", Knowl-Based Syst, 41 (2013), pp. 89-97 Article Download PDF View Record in Scopus.

[12] Michael Hagenau, Michael Liebmann, Dirk Neumann. Automated news reading: stock price prediction based on financial news using context-capturing features. DecisSuppSyst; 2013.

[13] Xu Tao, QinkePeng, Yinzhao Cheng Identifying the semantic orientation of terms using S-HAL for sentiment analysis Knowl-Based Syst, 35 (2012), pp. 279-289 View Record in Scopus.

[14] Isa Maks, PiekVossen A lexicon model for deep sentiment analysis and opinion mining applications Decis Support Syst, 53 (2012), pp. 680-688 Article Download PDF View Record in Scopus.

[15] B. Pang, L. Lee Opinion mining and sentiment analysis Found Trends Inform Retriev, 2 (2008), pp. 1-135 CrossRefView Record in Scopus.

[16] M. ArunaSafali, Dr. Ch. Suneetha, "Ensemble Text Mining using NLP and AI Techniques", IJRECE VOL. 6 ISSUE 2 APR.-JUNE 2018.

[17] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up?: sentiment classification using machine learning techniques," Proceedings of the ACL-02 conference on Empirical methods in natural language processing, vol.10, 2002, pp. 79-86.

[18] Ji Fang and Bi Chen, "Incorporating Lexicon Knowledge into SVM Learning to Improve Sentiment Classification", In Proceedings of the Workshop on Sentiment Analysis where AI meets Psychology (SAAIP), pages 94–100, 2011.

[19] R. F. Bruce and J. M. Wiebe, "Recognizing Subjectivity: A Case Study in Manual Tagging," In Natural Language Engineering, ACM, vol. 05, issue 02, pp 187-205, 1999.

 DOI : 10.35291/2454-9150.2019.0005