A STUDY ON THE APPLICABILITY OF DIFFERENT SENTIMENT ANALYSIS TECHNIQUES

P.Radha¹, Dr. N.Sudha Bhuvaneswari²

¹P.Radha, Asst. Prof, Dept. of Computer Science, Sri Krishna Arts and Science College, Tamilnadu, India.
²Dr.N.Sudha Bhuvaneswari, Associate Professor, Dept. of Computer Science, Dr.G.R.Damodaran College of Science Tamilnadu, India.

---***---

Abstract - *In the world of modern business techniques driven by technology and science, an analysis of feedback on the product or service plays a major role. The method of analyzing and deciding whether the feedback provided by the customers about product, service or person as positive, negative or neutral is called as the process of Sentiment Analysis. Sentiment analysis is a major area of research the helps large business data analysts to gauge public sentiment, perform complex and difficult market research, to track the credibility of brands and goods, and understand consumer experiences. This paper aims in providing a vivid report on the applicability and usability of the different analysis methods that can be used for performing sentiment analysis.*

*Key Words***:   Lexicon, Sentiment, Opinion, Polarity, Entropy**

1.INTRODUCTION

Sentiment analysis can be otherwise called as text analysis which can measure the attitude of the customer towards the aspects of a service or product that is described or discussed in the text. Basically, sentiment analysis takes a piece of text that can be a sentence, a comment or a document, analysis it completely and then returns a score that measures how positive or negative the text is based on which inferences could be drawn. With huge volume of data sweeping into companies' database in the form of feedbacks and reviews it becomes very difficult to manually check what sentiment this information carries and hence the necessity for performing exclusive sentiment analysis that can be automated [1].

The method of using natural language processing, text analysis, and statistics to analyze consumer sentiment can be called as sentiment analysis. Companies that wanted to perform the best by understanding what their consumers feel, what they say go for this type of sentiment or opinion analysis. In tweets, comments or reviews where customers mention brand, consumer sentiment can be found. The field of understanding these emotions with software is Sentiment Analysis. Sentiment Analysis uses natural language processing, text and statistical analysis to extract, and identify the sentiment of words into positive, negative, or neutral categories.

2. APPLICABILITY OF SENTIMENT ANALYSIS

Sentiment analysis has been constantly growing reaching out different domains of business and this rapid growth has made this technique to create landmark in various business areas. Few important arena where sentiment analysis is efficiently used are: brand monitoring where it gives a full 360 view of the product or the brand that the company markets mostly through social media and this can be used by companies to mainly measure the launch of new products or services and to get a feedback on the campaigns conducted to capture the competitive market. Sentiment analysis also plays a major role in enhancing the customer service by prioritizing the customer calls and urgent mails, thereby satisfying the frustrated customers and building a good relationship.

Sentiment analysis can be used for market research and analysis where an analysis can be performed on the product, service and the reviews of the customer based on which inference can be drawn and this type of service is very helpful in identifying the pattern and thought of customers and their opinions. Other areas of application of Sentiment Analysis includes the political, social and psychology to analyze trends, ideological bias, opinions, gauge reactions [7].

3. CHALLENGES OF SENTIMENT ANALYSIS

Though Sentiment Analysis can offer potential solutions in the area of opinion mining, there are still many challenges faced by this technique. The foremost issue faced by this analysis is understanding how to classify the sentiment. Sentiment needs to be categorized appropriately depicting the nature and mood of people like happiness, sadness, excited, bored and so on.

Defining sentiments is a major task and in addition to this when making analysis of the sentences that humans speak there are different perspectives giving ideas and meanings in different dimensions and expression of opinion poses complex issues with sentences having sarcasm, irony and other implied meaning that could mislead the sentiment

that is shown in the sentence. Understanding the context is a major and thrust area of research that requires more exploration to understand the tone and the emotion that is embedded and this can be combined with previous analytical data to arrive at conclusions. The final challenge in the sentiment analysis is deriving at a model and finding out ways to train the model. Though there are lot of pre-defined models available, for a long run these models are not suitable and cannot withstand and hence it is a good idea to design and train our own model [8].

4. LITERATURE REVIEW

An extensive study of literature on Sentiment Analysis highlights the number of researches that are carried out in this particular domain and the general inference of this extensive study focuses that mostly for sentiment analysis the process of data mining is used basically. It is also understood there are different approaches available to draw inferences and as per literature no one such approach gives the best and optimum performance thereby paving way for more research in this area. Based on many literatures drawn few important literatures on sentiment analysis is drawn and presented in Table 1

Table -1: Types of Sentiment Analysis Algorithms

AUTHOR & METHOD NAME	ALGORITHM	ADVANTAGES	DISADVANTAGES
[2]Amolik et al Data Mining	Naïve Bayesian and Support Vector Machine	Naïve Bayesian and Support Vector Machine performs well and also provide higher accuracy	SVM algorithm is not suitable for large data sets
[3]Cho and Kang Data Mining	Support Vector Machine (SVM**)**	A strong performance in the Classifying the emotion behind the text	Long training time for large datasets
[5]Dos Santos and Gatti Deep Learning Method	Deep Convolutional Neural Network (DCNN)	Achieves a highest accuracy sentiment prediction	Not possible to analyze character-level representations for sentiment analysis
[9]Lakshmi et al Deep	Convolutional Neural Network (CNN) and K	improves the efficiency and	Different initial partitions can result in different final

| Learning Method	means	accuracy	clusters
[10]Le and Nguyen Machine Learning Method	Naive Bayes (NB) and Support Vector Machine (SVM)	highly effective and accurate on the analysis of feelings	Assumption of independent predictors
[12]Wan and Gao Machine Learning Method	ensemble sentiment classification	improve the overall accuracy in twitter sentiment classification	Only texts of the tweets are considered and other information like the users who tweet them, the times of the retweets
[6]Ghiassi et al Neural Network	Dynamic Artificial Neural Network (DAN2)	high degree of coverage over twitter corpus, improved sentiment classification accuracy	Not focused on key sentiment indicators on incoming tweets from users
[4]Coban et al Neural Network	Extreme Learning Machine (ELM)	better classification performance for twitter sentiment analysis	few hidden nodes employed would lead to under-fitting/over-fitting issues in pattern classification
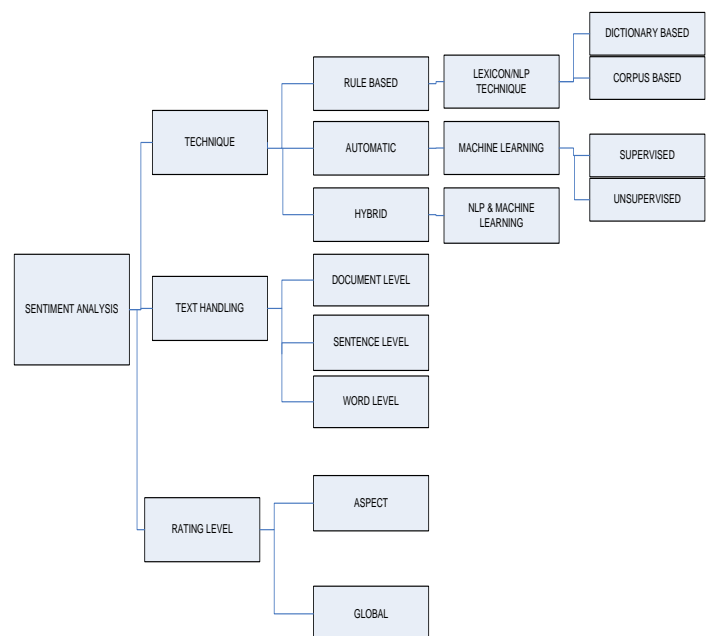
Fig -1: Classification of Sentiment Analysis

5. SENTIMENT ANALYSIS METHODOLOGY

The process of Sentiment Analysis has a set of steps to be followed to finally arrive at the inference that could be drawn from the analysis done. This is a five-stage process that followed properly could yield better approximation. The various steps involved in this process are:

Data Collection: This process includes collection of user data from social media and blogs in different formats, slangs and grammar and then organizing them. Manually doing this process will be lengthier and complex and almost impossible. Text analytics and natural language processing can be used for identifying and organizing the contents exactly.

Text Preparation: Trimming of the sentence by removing irrelevant contents and cleaning the data to be used for further data analysis [1].

Sentiment Detection: The extracted sentence from the user reviews and feedback is further examined to classify the sentence as subjective and objective. Subjective expressions highlight on views and opinions and they are retained and objective expressions that are factual are eliminated.

Sentiment Classification: From the retained subjective expressions, sentiments are identified and classified as positive, negative, like, dislike, good, bad etc.,

Output Generation: In sentiment analysis unstructured text is converted into meaningful data and the output inferred from the analysis can be projected as graphs and charts for better understanding [11].

6. COMPARATIVE STUDY ON SENTIMENT ANALYSIS TECHNIQUES

Techniques	Approaches	Features	Advantages	Limitations
Rule Based/Lexicon Based	1. Dictionary Based 2. Novel Machine Learning 3. Corpus Based 4. Ensemble	Manual Construction	Good coverage to accurately judge the sentiments	Lexicons restrict on the number of words that are classified, so only few sentiments could be identified.
Machine Learning	1. Bayesian Networks 2. Naive Bayes Classification 3. Max. Entropy 4. Neural Networks 5. Support Vector Machine	cy. Negations and extraction of opinions and new phrases	The potential to Adapt and produce models that are educated for specific reasons, and	Applicability of newly extracted data is very less taking into constraint the cost factors.
Hybrid	1. Machine Learning & Lexicon Based	Lexicons constructed using public resources and natural language and then sentiment processed using machine learning techniques	Learning and detecting of lexicon can be done at the concept level and not much changes or sensitiveness in the domain.	Very noisy reviews

7. CONCLUSIONS

This paper provides a comparative study on different types of sentiment analysis process and draws inferences based on the study carried out. Based on the extensive literature survey carried out inferences are drawn for the three major types of sentiment analysis: lexicon based, machine based and hybrid. According to the machine learning algorithm polarity of sentiments can be predicted based on trained and test data set. When considering the lexicon-based approach no training required for data mining. It uses a list of pre-defined words that are associated with a sentiment and this can be used to mine the data. The hybrid approach is a combination of machine learning and lexicon-based approach and this helps in optimizing the performance of sentiment analysis. These approaches can be applied in different domains of business and service mainly to identify the opinions of the customer and the reputation of the brand. The future study and research in this area is more focused towards handling social media data and identify the sentiments in which the process becomes much more complex due to the vast vocabulary of the language and the jargons that are being used while posting the messages.

REFERENCES

[1] Agarwal, B. Xie, I. Vovsha, O. Rambow, R. Passonneau, "Sentiment Analysis of Twitter Data", In Proceedings of the ACL 2011Workshop on Languages in Social Media,2011 , pp. 30-38

[2] Amolik, N. Jivane, M. Bhandari and M. Venkatesan, Twitter sentiment analysis of movie reviews using machine learning techniques, International Journal of Engineering and Technology, Vol.7, No.6, Pp.1-7, 2016.

[3] S.H. Cho and H.B. Kang, Text sentiment classification for SNS-based marketing using domain sentiment dictionary, IEEE International Conference on Consumer Electronics (ICCE), Pp.717-718, 2012.

[4] O. Coban, B.M. Ozyildirim and S.A. Ozel, An empirical study of the extreme learning machine for Twitter sentiment analysis, International Journal of Intelligent Systems and Applications in Engineering, Vol.6, No.3, Pp.178-184, 2018.

[5] C. Dos Santos and M. Gatti, Deep convolutional neural networks for sentiment analysis of short texts, Proceedings of cooling, the 25th International Conference on Computational Linguistics: Technical Papers, Pp. 69-78, 2014.

[6] M. Ghiassi, J. Skinner and D. Zimbra, Twitter brand sentiment analysis: A hybrid system using n-gram analysis and dynamic artificial neural network, Expert Systems with applications, Vol.40, No.16, Pp.6266-6282, 2013.

[7] Hassan, A., Abbasi, A., Zeng, D. 2013. Twitter Sentiment Analysis: A Bootstrap Ensemble Framework, Proceedings of the ASE/IEEE International Conference on Social Computing, pp. 357-364

[8] Hatzivassiloglou, V. & McKeown, K.R. 1997. Predicting the semantic orientation of adjectives. In Proceedings of the 8th conference on European chapter of the association for computational linguistics Madrid, Spain, pp.174-181.

[9] B.S. Lakshmi, P.S. Raj and R.R. Vikram, Sentiment Analysis Using Deep Learning Technique CNN with KMeans, International journal of pure and applied mathematics, Vol.114, Pp.47-57, 2017.

[10] B. Le and H. Nguyen, Twitter sentiment analysis using machine learning techniques, Advanced Computational Methods for Knowledge Engineering, Pp.279-289, 2015.

[11] Wan, X.. "A Comparative Study of Cross-Lingual SentimentClassification". In Proceedings of the 2012 IEEE/WIC/ACMInternational Joint Conferences on Web Intelligence and Intelligent Agent Technology-Volume 01 (pp. 24-31). IEEE Computer Society.2012

[12] Y. Wan and Q. Gao, An ensemble sentiment classification system of twitter data for airline services analysis, IEEE International Conference on Data Mining Workshop (ICDMW), Pp.1318-1325, 2015.