A STUDY ON APPLICATIONS AND CONCEPTS OF SENTIMENT ANALYSIS

P. Uma[1*], A. Aloysius[2]

[1]Dept. Computer Science, MRK College of Arts and Science, Kattumannarkoil, India
[2] Dept. Computer Science, St. Joseph's College (Autonomous), Trichy, India

*Corresponding Author: Umasathiya87@gmail.com

Abstract-At present, the stage of Internet has changed the way that the people express their views and opinions in Social Media. Millions of people are using social network sites like Facebook, Twitter, Google Plus, etc. to express their emotions, opinion and share views about their daily lives. Through the online communities, sellers get an interactive media where consumers inform and influence others by online forums. Social media is generating a large volume of sentiment rich data especially in the form of tweets, status updates, blog posts, comments, reviews, etc. Sentiment analysis has become a very popular field of research and lot of researchers have explored this field but still, there are many issues as sentiment analysis processes text-based unstructured data. The dictionary-based approach takes less processing time than supervised learning approach but accuracy is not up to the mark. Supervised learning approach provides better accuracy and it is found that sentiment classifiers are severely dependent on domains or topics. In this research paper, applications of sentiment analysis, types of approaches present, and evaluation metrics' involved in sentiment analysis are discussed briefly.

Keywords-Data Mining, Sentiment Analysis (SA), Emotion Detection (ED), Opinion Mining, Social Media Network, Applications, Evaluation Metrics.

I. INTRODUCTION

Sentiment analysis is one of the wide areas of research. Sentiment analysis is considered for its improvement in techniques and for classification approaches. It takes part in opinion mining and text summarization, which helps not only in social network sites but also supported in Politics, Business intelligence, Public activities, etc [11]. Sentiment analysis can be done on different approaches to machine learning that use the training set of data. So it is said to be a supervised approach in the present scenario. The second one is Lexicon based approach. In this, the trained data set as like machine learning is used but it has pre-defined dictionary related to words that mostly used in this field. The next to Lexicon based approach is hybrid, which is the combination of both machine learning as well as lexicon-based approach [19]. Sentiment Analysis is widely used in research areas like Spam detection and Neural Network. The neural network is the current interesting area among all.

II. APPLICATIONS OF SENTIMENT ANALYSIS

A. Emotion Detection
Sentiment analysis performs NLP task to discover opinion about an item to avoid ambiguities in a different type of opinion in sentiment and emotion. Emotion detection is a kind of Sentiment Analysis (SA) task in which SA is mainly used for positive and negative opinion but Emotion Detection (ED) is identified and detected from the various texts. It can be implemented by using machine learning but most probably used a lexicon based approach [1].

B. Building Resources
The main task of building resources is to create lexicon, dictionary, and corpora in which expressions are based according to the polarity. Building resources is not a task of SA but it can improve ED as well as SA [9]. The main function of this category is multi granularity ambiguities of word and difference in their opinion in text and gesture.

C. Transfer Learning
It extracts knowledge to improve learning process in the target domain, for example, it transfers knowledge to Wikipedia document from English to the Arabic language which is in the use of enhancing text mining task like sentimental analysis, text classification, part-of-speech tagging name entity recognition. In sentiment analysis, this learning process is mainly applied to transfer sentimental classification in one domain to another [6].

D. Business and organizations
Social media provides an opportunity for businesses by giving a platform to connect with their customers in the

name of advertisement. People mostly depend upon user-generated content online to a great extent for decision-making [11]. For the establishment and smooth functioning of a business, it is very important to find out what the users think about their services for their product and service benchmarking. User's opinion is the basic criterion for the improvement of the quality of products and enhancement of the services [17].

E. Individuals
A customer making a purchase knowing the opinion of other users can play an influential role in his decision making. The Internet provides a large number of user-generated reviews of numerous products. E-commerce websites such as amazon.com (product reviews), zomato.com (restaurant reviews), reviewcentre.com have millions of reviews by customers which are useful for individuals [5].

F. Detection of Hatred
It is possible to scan the social media and blog posts to detect the arrogant words to find out the disagreement or anger over a particular issue. Monitoring the emails or tweets can mine the hatred language used which can prove to be harmful [10].

III. RECENT EMERGING FIELDS OF SENTIMENT ANALYSIS

Applications that use Reviews from Websites
Today Internet has a large collection of reviews and feedbacks on almost everything. This includes product reviews, feedbacks on political issues, comments about services, etc. Thus there is a need for a sentiment analysis system that can extract sentiments about a particular product or services. It will help us to automate the provision of feedback or rating for the given product, item, etc. This would serve the needs of both the users and the vendors.

Applications as a Sub-component Technology
A sentiment predictor system can be helpful in recommender systems as well. The recommender system will not recommend items that receive a lot of negative feedback or fewer ratings. In online communication, it comes across abusive languages and other negative elements [9]. These can be detected simply by identifying a highly negative sentiment and correspondingly taking action against it.

Applications in Business Intelligence
It has been observed that people nowadays tend to look upon reviews of products which are available online before they buy them. In many businesses, the online opinion decides the success or failure of their product [12]. Thus, Sentiment Analysis plays an important role in businesses. Businesses also wish to extract sentiment from the online reviews in order to improve their products and in turn their reputation and help in customer satisfaction.

Applications across Domains
Recent researches in sociology and other fields like medical, sports have also been promoted by Sentiment Analysis that show trends in human emotions, especially on social media [10].

Applications in Smart Homes
Smart homes are supposed to be the technology of the future. In future entire homes would be networked and people would be able to control any part of the home using a tablet device. Recently there has been a lot of research going on the Internet of Things (IoT). Sentiment Analysis would also find its way in IoT. Like for example, based on the current sentiment or emotion of the user, the home could alter its ambiance to create a soothing and peaceful environment. It can also be used in trend prediction [6]. By tracking public views, important data regarding sales trends and customer satisfaction can be extracted.
The field the Sentiment Analysis usage stand tall are

- Opinions in the social and geopolitical context
- Predicting stock prices based on opinions that people have about the companies and resources
- Determine areas of a product that needs to be improved by summarizing product reviews to see what parts of the product are generally considered good or bad by users
- Customer Preference

IV. SENTIMENT ANALYSIS TECHNIQUES

Sentiment analysis can be defined as a process that automates mining of attitudes, views, opinions, and emotions from speech, text, tweets, and database sources through Natural Language Processing. It involves classifying opinions from text into categories like positive, negative, or neutral [20]. It is also referred to as subjectivity analysis, opinion mining, and appraisal extraction.

Sentiment Classification techniques are divided into three kinds of approaches such as Machine Learning approach, Hybrid approach and Lexicon based approach [7]. All these approaches are explained in the following sections.
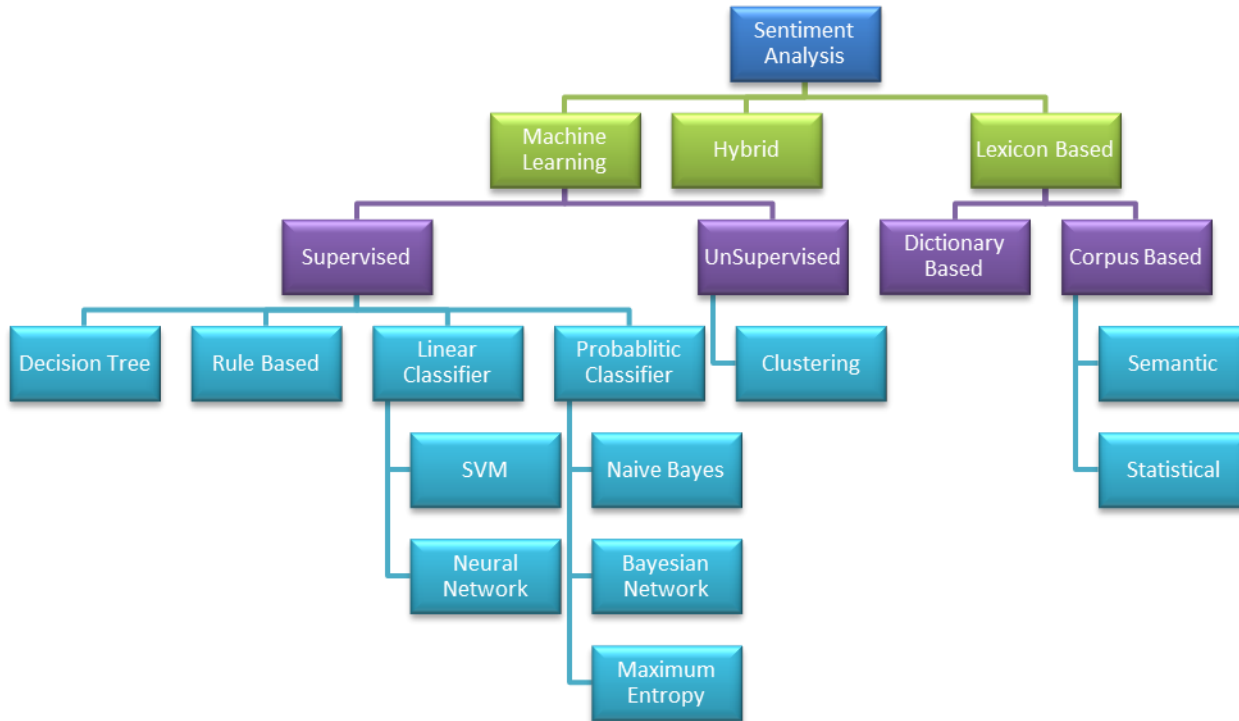
Figure 1: Sentiment Classification Technique

- **Machine Learning Approach**

Machine learning based approach uses classification technique to classify text into classes. There are mainly two types of machine learning techniques such as supervised and unsupervised methods for sentiment classification [19]. Figure 1 provides the route map of all the approaches in seperate category. The success of both these learning methods mainly depends on the selection and extraction of the specific set of features used to detect sentiment.

➢ *Unsupervised learning*
    Unsupervised learning does not consist of a category and they do not provide with the correct targets at all and therefore rely on clustering [15].

➢ *Supervised learning*
    Supervised learning is based on the labeled dataset and thus the labels are provided to the model during the process [4]. These labeled datasets are trained to get meaningful outputs when encountered during decision-making.

- **Hybrid approach**

The hybrid approach is a combination of machine learning and lexicon based approaches. The hybrid technique is sentiment lexicon constructed using product reviews for initial sentiment analysis [21]. These sentiment analysis reviews are features in machine learning method. A hybrid approach is much faster than both of two approaches, in this approach the sentiment symbol detection is facts and

detection of sentiment is measured at a conceptual level and lesser sensitivity to change the domain. There is the only limitation in this approach its noisy reviews [2].

- **Lexicon Based Approach**

Lexicon approach depends on finding opinion lexicon which analyzes the sentiment of text. As mentioned in figure 1, this approach has two methods such as Dictionary based and Corpus-based [13].

➢ *Dictionary Based approach*
    The main strategy of dictionary-based approach is by working on manually created a set of opinion that are repeated and then finds their synonyms and antonyms by iterations and saves these words in seed list. These iterations repeated until no synonyms and antonyms are found. After that it manually removes and corrects errors [11]. The limitation of dictionary approaches is its low applicability.

➢ *Corpus-Based approach*
    Corpus approach improves the limitations and helps to improve the finding opinion in particular area or orientation [16].

V. LEVELS OF SENTIMENT ANALYSIS

There are four levels of Sentiment analysis which is shown in figure 2.

Word Level

Word sentiment classification use mostly adjectives as features. The two methods of automatically annotating sentiment at the word level are:

- ▪ Dictionary-Based Approaches
- ▪ Corpus-Based Approaches.

Both are Lexicon based sentimental analysis techniques, which use word-level analysis effectively [18].

Sentence level

In sentence level approaches, every sentence is considered as an entity and analysis approaches are applied to individual sentence, then their result is summarized to provide the overall result of the document. Generally, it deals with tagging individual sentences with their respective sentiment polarities. Sentence level sentiment classification classifies sentence into positive, negative or neutral class [8].

Document level

In this approach, the whole document is considered as a single entity and the analysis approach is applied to the whole document. It deals with tagging individual documents with their sentiment. In Document level, the whole document is classified either into positive or negative class.

Feature level

It deals with labeling each word with their sentiment and also identifying the entity towards which the sentiment is directed. Aspect or Feature level sentiment classification concerns with identifying and extracting product features from the source data. Techniques like dependency parser and discourse structures are used in this.
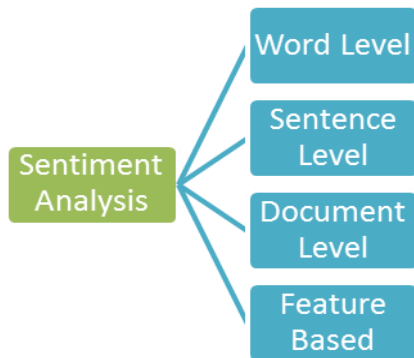

Figure 2: Levels of Sentiment Analysis

VI. PERFORMANCE EVALUATION METRICS OF SENTIMENT CLASSIFICATION

The performance of sentiment classification can be evaluated by using four indexes calculated by the following equations [3]:

Confusion Matrix [19]:

CONFUSION MATRIX	Predicted Positive	Predicted Negative
Actual Positive	TP	FN
Actual Negative	FP	TN

Precision - Precision is the fraction of retrieved instances that are relevant

Precision= TP / TP + FP

Recall - Recall is the fraction of relevant instances that are retrieved

Recall= TP / TP + FN

F1 score - These two measures are sometimes used together in the F1 score (also F-score or F-measure) which is a measure of test accuracy.

F1= 2 x Precision x Recall / (Precision + Recall)

In which TP, FN, FP, and TN refer respectively to the number of true positive instances, the number of false negative instances, the number of false positive instances and the number of true negative instances [14].

Accuracy= (TP + TN) / (TP + TN + FP + FN)

VII. CONCLUSION

In this research work, sentiment analysis are discussed to cover techniques, applications, recent techniques and approaches that promise to directly enable opining mining information seeking system. This paper provide a detailed study of existing techniques for opinion mining including machine learning and lexicon-based approaches, together with cross-domain and cross-lingual methods.These applications and techniques are aided to apply in the sentiment analysis and get the powerful reflection for it. These techniques are extracted from the mathematical and statistical model, which help to practice quickly, understand and decision making for the analysis. Also studied the effects of various approaches on sentimental analysis classifier and conclude that more the cleaner data, more accurate results can be obtained based on classifier.

REFERENCES

[1]. AkriviKrouska, Christos Troussas, Maria Virvou, "Comparative Evaluation of Algorithms for Sentiment Analysis over Social Networking Services", Journal of Universal Computer Science, Volume 23, Issue 8, 2017, Pages 755-768.

[2]. Amreen Shaikh, Madhuri Rao, "Survey on Sentiment Analysis", International Conference on Emanations in Modern Technology and Engineering, Volume 5, Issue 3, 2017, Pages 164 – 167.

[3]. Ankita Gupta, JyotikaPruthi, Neha Sahi, "Sentiment Analysis of Tweets using Machine Learning Approach", International Journal of Computer Science and Mobile Computing, Volume 6, Issue 4, 2017, Pages 444 – 458.

[4]. Anu Sharma and Savleen Kaur, "Techniques for Sentiment Analysis survey", I J C T A, Volume 41, Issue 9, 2016, Pages 355 - 360.

[5]. Bing Liu. "Sentiment Analysis and Opinion Mining", Morgan & Claypool Publishers, May 2012.

[6]. Brian Keith, Exequiel Fuentes, and Claudio Meneses. A Hybrid Approach for Sentiment Analysis Applied to Paper Reviews. In Proceedings of ACM SIGKDD Conference, Halifax, Nova Scotia, Canada, 2017, Pages 1 – 10.

[7]. Edison M., A. Aloysius, "Concepts and Methods of Sentiment Analysis on Big Data", International Journal of Innovative Research in Science, Engineering and Technology, Volume 5, Issue 9, 2016, Pages 16288 - 16296.

[8]. Edison M., A. Flora, D. Vinnarasi, A. Aloysius, "Sentiment Analysis: A Review", International Journal of Engineering Technology and Computer Research (IJETCR), Volume 5, Issue 4, 2017, Page 26 - 30.

[9]. Govindarajan M., Romina M., "A Survey of Classification Methods and Applications for Sentiment Analysis", International Journal of Engineering and Science (IJES), Volume 2, Issue 12, 2013, Pages 11-15.

[10]. Jianbo Yuan, Quanzeng You and JieboLuo, "Sentiment Analysis Using Social Multimedia", Springer International Publishing Switzerland, Multimedia Data Mining and Analytics, 2015, Pages 31 - 59.

[11]. Mahima Gaur and JyotikaPruthi, "A Survey on Sentiment Analysis and Opinion mining", International Journal of Current Engineering and Technology, Volume 7, Issue 2, 2017, Pages 444 - 446.

[12]. MrunaliOmprakashThakare, MrudulaNimbarte, "User Tracking Using Tweet Segmentation And Word Recognition", International Conference on Electronics, Communication and Aerospace Technology, 2017, Pages 664 - 668.

[13]. Praveen Kumar, Umesh Chandra Jaiswal, "A Comparative Study on Sentiment Analysis and Opinion Mining", International Journal of Engineering and Technology (IJET), Volume 8, Issue 2, 2016, Pages 938 - 943.

[14]. Purti Jain, ShrutiKohli, "Big Data Analysis, Algorithms, and Applications: A Survey", International Journal of Emerging Trends in Engineering Research (IJETER), Volume 3, Issue 1, 2015, Pages 19 – 24.

[15]. Revathi S., N Rajkumar, "A Survey on Sentiment Analysis for Web-Based Data by Using Various Machine Learning Techniques", International Journal of Engineering Trends and Applications (IJETA) – Volume 1, Issue 2, 2014, Pages 1 - 5.

[16]. Shailesh Kumar Yadhav, "Sentiment Analysis and Classification: A Survey", International Journal of Advanced Research in Computer Science and Management Studies, Volume 3, Issue 3, 2015, Pages 113 - 121.

[17]. Tapan Kumar Hazra, YashikaBihani, Sneha Mishra, "Sentiment Learning Using Twitter Ideograms", IEEE, 2017, Pages 115 - 120.

[18]. Vidisha M. Pradhan, Jay Vala, PremBalani, "A Survey on Sentiment Analysis Algorithms for Opinion Mining", International Journal of Computer Applications (0975 – 8887) Volume 133, Issue 9, 2016, Pages 7 - 11.

[19]. Vishal A. Kharde, S.S. Sonawane, "Sentiment Analysis of Twitter Data: A Survey of Techniques", International Journal of Computer Applications, Volume 139, Issue 11, 2016, Pages 5 - 15.

[20]. Walaa Medhat, Ahmed Hassan, HodaKorashy, "Sentiment analysis algorithms and applications: A survey", Ain Shams Engineering Journal, Volume 5, 2014, Pages 1093 – 1113.

[21]. Yakshi Sharma, VeenuMangat, Mandeep Kaur, "Sentiment Analysis & Opinion Mining", International Journal of Soft Computing and Artificial Intelligence, Volume 3, Issue 1, 2015, Pages 59 - 62.