Approaches and Applications of Sentiment Analysis on Users Data

Sachin Balvir[1], Santosh Sahu[2], Jaynat Rohankar[3]

[1]Assistant Professor, Department of Information Technology, DMIETR, Wardha, India

[2]Research Scholar, Department of Computer Science Engineering, VNIT, Nagpur, India

[3]Assistant Professor, Department of Computer Science Engineering, TGPCET, Nagpur, India

ABSTRACT:

Every day the user's content on social networking sites, blogs and other e-commerce linked sites is growing drastically. This content can be prepared and come out with various outcomes will be very much helpful. At present, collecting reviews and suggestions of users from the sites is trending. The reviews and suggestions of customers are very essential for any industry. Therefore there is a great requirement of some techniques by which we can identify the unseen data from reviews and suggestions. Sentiment Analysis is process of extracting meaningful data from user's reviews and suggestions. Many people share their reviews and suggestions on websites, blogs data or information on social networking sites and on web forums. Such type of data gives meaningful information about the community that what they are thinking. Without this information one cannot find out whether there is any need for change. This information can be process and manipulate to come up with some results on which one can take decisions. This paper focuses on different types of sentiment analysis performed, techniques of sentiment analysis and various applications of sentiment analysis practice in real world.

Keywords - Sentiment analysis, Machine Learning, Applications of Sentiment Analysis

1. INTRODUCTION

The World Wide Web is a popular space where people express their individual opinions and emotions. It also includes an influencing aspect of life, with implication for marketing and business purpose. Sentiment analysis is a technology that helps to extract the information about the user's feedback. Sentiment analysis is a classification of emotions such as positive, negative and neutral within users review using text analysis techniques.

Sentiment analysis model identify duality within a text e.g. a positive, negative or neutral opinion whether it is a whole document, paragraph, sentence or word. For example, in business it is essential to understanding people's emotions. Automatically analyzes customer review to listen attentively to their customers and product and services to meet their needs. One of our customers used sentiment analysis to analyze 1,495+ reviews about their product, and identify that customers were happy about their pricing but complained a lot about their customer service [2, 3].

Focusing on the Social Media activities is an excellent approach to calculate customer's reliability and observance of their sentiment towards brands or products. Social Media are the subsequently reasonable marketing field. At present, Facebook has taken first place in the digital marketing. After facebook, twitter is the next popular place for marketing. [5]

This paper discussed about, What sentiment analysis is? , How sentiment analysis works?, and How to detect emotions with in text?. It includes different techniques for performing a sentimental analysis. In first part of the paper discuss various types and approaches of sentiment analysis. At last deliberate about the application of sentiment analysis in real world.

2. LITERATURE REVIEWS

The sentiment analysis can be drawn from financial news. The prepared sentiment rules are used to determine the polarity of sentence, the polarity may be positive or negative. The ratio of positivity with respect to negativity is used for the determination of sentiment value [6, 9]. Based on reviews on movies, the polarity i.e. positivity and negativity about the particular movie can be drawn. For this purpose machine learning and semantic orientation approach can be adapted to review a movie for evaluation. As the available information is always mixed type, it is the challenge for review mining [7].

Sentiment analysis is also used for the classification of Web forum opinions written in multiple languages. The usefulness of stylistic and syntactic attributes is evaluated for sentiment classification of English and Arabic content. The entropy weighted genetic algorithm (EWGA) is designed to improve performance and get a better assessment of key features. The features and techniques are evaluated on a benchmark movie review dataset and U.S. and Middle Eastern Web forum postings [10].

The author constructs the framework by employing hierarchical attention network and creating sentiment-level embedding. Also build up the attention system by allowing it to receive both lexicon level and sentiment-level information to improve the effectiveness of the model. The experiments are performed on IMDB movie reviews [11]. Depth models such as a Convolutional Neural Networks (CNNs) model joint with SVM text sentiment analysis improves the correctness of text sentiment classification effectively [12].

Opinion mining is parallel to emotion mining with having some difference in the concept. Opinion mining determines the thoughts of a writer towards a subject whereas emotion mining is concerned with detecting and classifying writers' emotions in the direction of events or topics [14]. The sentiments are extracted from the paragraph given by the customer as a feedback for the company's progress [15].

3. TYPES OF SENTIMENT ANALYSIS

Sentiment analysis has various types, from models that focus on polarity positive, negative and neutral to detect feelings and emotions such as sad, happy, angry etc.[1] The most popular types of sentimental analysis are as follows:-

3.1 Fine-grained Sentiment Analysis

In Fine- grained Sentiment Analysis it analyzes the sentence by parts. We apply this type on a sub-sentence level. That means it identify a target of a sentiment. It might consider expanding polarity categories to include:

- Very positive
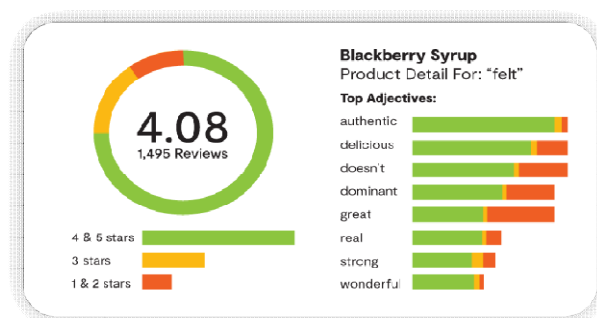- Positive
- Neutral
- Negative
- Very negative

Figure 1. Example of analysis

3.2 Emotion Detection Analysis:

Emotion detection sentiment analysis aims to detect emotions such as anger, sadness, happiness, frustration, surprise and so on. A lexicon is a list of words and emotions they convey. Lexicon or complex machine learning algorithms are mostly used in much emotion detection system.

3.4 Aspect- Base Sentiment Analysis:

It is a wording analysis technique that breaks down text into component of product or service, and then allocates every one aspect on a sentiment level such as positive, negative or neutral [16].

For example, "The battery life of this mobile is too short", aspect-base analysis able to determine that the sentence express a negative opinion about feature battery life.

3.5 Multilingual Sentiment Analysis:

Multilingual sentiment analysis is a mostly studied Natural Language Processing (NLP) task   where the goal is to determine evaluation, opinion, and emotions of users towards product, an entity or a service that they are reviewing.

4. APPROACH FOR SENTIMENT ANALYSIS

Sentiment analysis works on various Natural Language Processing (NLP) and algorithms. Sentimental analysis classification techniques are separated into two parts as Machine Learning (ML) and Lexicon base approach.
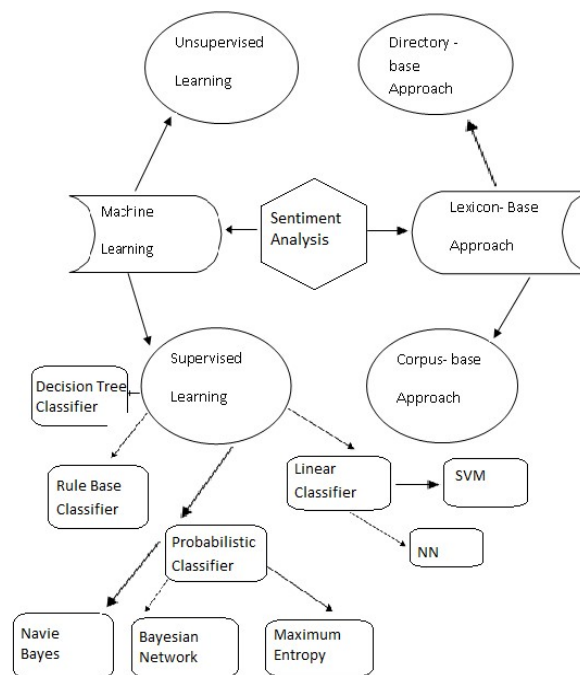

Figure 2. Sentiment Classification Technique

A.   Lexicon-based Approach:

The Lexicon-based Approach is depending on sentimental Lexicon. Lexicon includes as an important indicator for opinion word which is review of users called as sentiment. Lexicons divided into a two approach, dictionary approach and corpus based approach.

i. Dictionary approach:

Dictionary approach is an arrangement of sentimental word (e.g. good, bad, excellent and so on) is gathered with known instruction and this process done manually. It is a computational approach to measuring feeling or emotion that a text convey to the reader. Sentiment has a binary classification positive or negative; it can be extended to different dimensions like joy, sadness, fear, anger, etc.

ii. Corpus based approach:

Good for domain-specific application (e.g. micro blogs, reviews). Good for sentiment and aspect extraction. Lexicon expansion given a seed list of general-purpose sentiment words, discover other sentiment words from a domain corps.

B. Machine Learning Approach for Sentiment Analysis:

Machine Learning Approach's have been used for sentimental analysis. It has a two types supervised-based learning and unsupervised- based learning. The bag-of-words (BoW) is commonly used for sentiment analysis. BoW method is independence of words and ignores the importance of word and ignores the importance of semantic information in the text.

Supervised-based Learning:

Supervised learning technique is broadcast upon the presence of labeled Dataset. Majority of practical machine learning uses supervised learning. Following fig shows the process of supervised-based learning. [4]
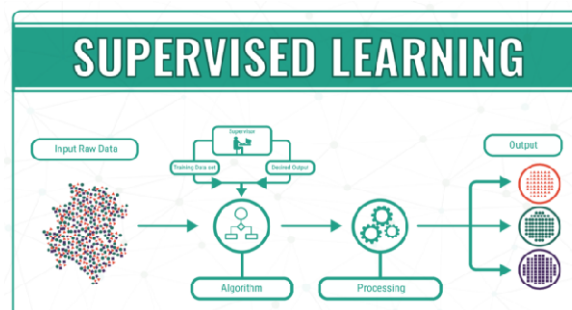

Figure 3. Supervised-based learning

i. Decision Tree Support:

Decision tree support classifier gives different leveled deterioration of the preparation space in which a condition on the credit esteem is used to separate the information. The information division space is done recursively until the leaf hubs contain base quantities of records which are utilized with the end goal of characterization. The choice tree execution in content characterization have a tendency to minimal minor departure standard bundles for instance as ID3.

ii. Linear Classifiers:

Linear classifiers sort various types of classifiers, among them Support Vector Machine (SVM) which is a kind of classifier that attempt to choose awesome direct separators between different classes and another technique is neural system. SVM is supervised learning model which support vector machines work on of decision plans.

Example, the concept of linear Support Vector Machine (SVM) in the objects either belongs to Black class or Blue class. Isolated line determines the choice. On the right hand side of the boundary, all objects are BLUE and to the left hand side of the boundary, all circles are BLACK.
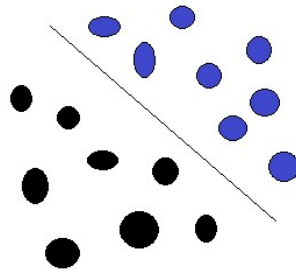
Figure 4. Example of linear SVM

A new white circle will be classified as BLUE if it fall to the right hand side of the boundary   or classified as BLACK If it falls to the left hand side.
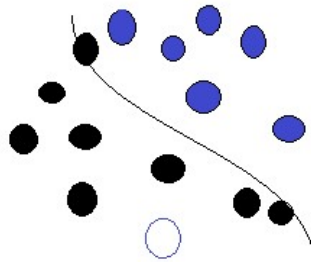
Figure 5. Example of Hyper plane SVM

iii. Rule Base Approach:

Normally, a rule-based system uses a set of manufactured rules to help detect   subjective, polarity or subject of an opinion. Manufactured rule may include different techniques developed in computational linguistics, are as follows:

- Stemming- stemming is the process of reducing word. Stemming is important in   Natural Language Understanding (NLU) and Natural Language Processing (NLP).
- Tokenization- This is process of replacing sensitive data with uniquely identification symbols that retain the critical information about the data.
- Part-of-speech-tagging-
- Parsing-

iv. Probabilistic Classifiers:

Probabilistic classifier use combine of models for the combination. Mixing demonstrates expect that every class is apportioning of the come together. Each combine model fragment is generative model that gives the like hood of exploratory a exacting term for that segment. The Probabilistic Classifier follows the following algorithms.

Naïve Bayes Algorithm:

Naïve Bayes Algorithm is fast and Strength forward classification algorithm and it is suitable for   large amount of data. Nave Bayes is used in various applications such as text classification recommendation system spam filtering and sentimental analysis [13]. It uses the probability for prediction of known dataset.

$$P(a\backslash B) = \frac{P(B\backslash a) * P(a)}{P(B)}$$

- P (a): Probability of hypothesis becomes true. This called as the prior probability of a.
- P (B): Probability of the Data. This called as prior probability.
- P (a|B): Probability the hypothesis a given the data B.
- P (B|a): Probability of data B given that hypothesis was true.

Bayesian Network:

Bayesian Network is one of the type classifier similar as a Naïve Bayes but this classifier does not show any relationship between attribute (i.e. conditional independent) at all.
Example,
 P (attribute (a)/parent (p))

- If attribute of parent (p) is true then the attribute (a) is also true

This algorithm gives a joint probability as solution on attribute dependency based on some sort of condition. Joint probability gives two types are Direct Acyclic Graph (DAG) and Conditional Probability Table (CPT).

- DAG- Ti is a graph which is contains a nodes and connectivity between them. It is not suitable for large data set.

Maximum Entropy Classifier (ME):

Maximum Entropy Classifier is a type of probabilistic classifier which has a position with the class of exponential models. It is relies on Principle of Maximum Entropy (ME) and from of each one of the models that fit preparing information, pick the greeter entropy.

5. APPLICATIONS OF SENTIMENT ANALYSIS

The Sentiment analysis can be used in various applications. It helps in reviewing movies, reviewing product, identifying sentiments regarding political leaders etc. The various applications where sentiment analysis helps are shown in table below.

Applications	Description
Automobiles	Cars/Two-Wheels: Product, brands, features, pre-owned vehicles.
Fashion	Accessories, Apparel, Outlets, Brands
Online Services	Brands, Factories, Services, Prices, Product.
Personal Finance	Finance institution , Loan ,cards, CAs
Books	Shops, Titles, Authors, Exchange
Electronics	Cell-phones, Computers, AC, TVs, Cameras
Business	Brand Reputation, online advertising, Consumer voice
Politics	Classification of politicians position, voting advice applications
Predictions	predicting election outcome or market trend from sentiment

Table 1. Sentiment Analysis Applications

6. CONCLUSION

Sentimental Analysis Technique helps in identifying emotional and attitude situation of peoples. People's emotions can be expressed in positive or negative ways. By applying sentiment analysis, any company can make a strategic change in their decision strategy for business growth. Such type of analysis is very important for all types of business. This paper discuss about technique which is important during performing Sentimental Analysis. Most of the researchers used the techniques for their research for sentiment analysis is mentioned here. Analysis performed on peoples view point about any product can work as feedback to company to know the quality and drawbacks of their product. This paper also gives point of view of the unique applications and types of sentimental analysis.

REFERENCES

[1] M. Joshi, P. Prajapati, A. Shekh, V. Vala, "A Survey on Sentiment Analysis" International journal of computer application, April 2017.

[2] F. Neri, C. Aliprandi, F. Capaci, M. Cuodras, T. Bay,"Sentiment Analysis on Social Media", ACM International Conference on Advances in Social Network Analysis and Mining, 2012 .

[3] S. Xiong, "Improving Twitter Sentiment Classification via Multi-Level Sentimental-Enriched Word Embeddings", Association for the Advancement of Artificial Intelligence, 2017.

[4] Walaa Meddhat , Ahmed Hassan ,Hoda Korashy "Sentiment analysis algorithms and applications: A survey", Ain Shams Engineering Journal (2014) 5, 1093–1113, Ain Sham University, Egypt. 19 April 2014.

[5] Bogdan Batrinc, Philip C. Treleaven "Social media analytics: a survey of techniques, tools and platforms", Springer, DOI 10.1007/s00146-014-0549-4, Published on 26 July 2014.

[6] Tan Li Im, Phang Wai San, Chain Kim On, Patricia Anthony, "Rule-based Sentiment Analysis for Financial News", IEEE International Conference on Systems, Man, and Cybernetics, 2015 IEEE

[7] Chaovalit, P., Zhou, L.:. "Movie review mining: A comparison between supervised and unsupervised classification approaches", in Proceedings of the Hawaii International Conference on System Sciences (HICSS), 2005.

[8] Camillo, F., Neri, F.: "Sentiment Mining: Automatic linguistic analysis and natural language understanding in an e-CRM systems", EURISBIS 2009, Proceedings, F.Mola et al. (Eds), TILAPIA, p. 136, Cagliari (IT), 01/06/2009.

[9] P. Tyagi and R.C. Tripathi, "A Review towards the Sentiment Analysis Techniques for the Analysis of Twitter Data", 2nd International Conference On Advanced Computing and Software Engineering, (ICACSE-2019)

[10] A. Abbasi, H. Chen, and A. Salem, "Sentiment Analysis in Multiple Languages: Feature Selection for Opinion Classification in Web Forums", ACM Transactions on Information Systems, Vol. 26, No. 3, Article 12, Publication date: June 2008

[11] X. Xu and P. Li, "Knowledge-oriented Sentiment-level Embedding for Sentiment Classification", ACAI '19, December 20–22, 2019, Sanya, China © 2019 Association for Computing Machinery. ACM ISBN 978-1-4503-7261-9/19/12

[12] Y. Chen and Z. Zhang, "Research on text sentiment analysis based on CNNs and SVM", 13th IEEE Conference on Industrial Electronics and Applications (ICIEA), 2018

[13] S. Kunal, A. Saha, A. Varma, V. Tiwari, " Textual Dissection Of Live Twitter Reviews Using Naive Bayes", International Conference on Computational Intelligence and Data Science (ICCIDS 2018), Procedia Computer Science 132 (2018) 307–313, Science Direct, 2018

[14] A. Yadollahi, A. G. Shahraki, and O. R. Zaiane, "Current State of Text Sentiment Analysis from Opinion to Emotion Mining", ACM Computing Surveys, Vol. 50, No. 2, Article 25, Publication date: May 2017

[15] B. Nandi, M. Ghanti, S. Paul, "Text Based Sentiment Analysis", Proceedings of the International Conference on Inventive Computing and Informatics (ICICI 2017), IEEE Xplore Compliant - Part Number: CFP17L34-ART, ISBN: 978-1-5386-4031-9

[16] https://monkeylearn.com/sentiment-analysis