ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 8, Issue 12, June 2019

Sentiment Analysis and Opinion Polarity Techniques in Customer Product Relationship: A Survey

Amit Kumar

Abstract*—Twitter, has more than 340 million monthly active users and over 540 million tweets per day. Machine learning is about forecasting the future based on the past information. At the present time, an establishment or association makes available a commercial service which essentials to acquire advice from consumer. The consumer review is significant to progress service for organization or establishment, which have together nearby opinion and undefended opinion. Support Vector Machine algorithm is a supervised machine learning procedure which can be applied for classification. Opinion mining is investigation of reviewer's attitudes, emotions as well as emotions concerning related to events, individuals, topics or issues. This paper represents the survey of different sentiment analysis methods. It also represents the features and limitations of different sentiment polarity techniques.

Keywords: **Social Media, Twitter, Sentiment analysis, Opinion mining, Sentiment Polarity, Machine learning.**

I. INTRODUCTION

With the speedy extension of corporation or association have additional services and goods online and increase consumer contentment. The supplier will read consumer assessment and other consumers who necessity to procedure amenities or goods will read assessment to express opinions on the facilities. The numeral of consumer assessment is accumulative or vast from website, ha blogs, opportunities and community broadcasting, which the amenities or invention is motivating. Therefore, countless consumers will read comment randomly which is tough to read altogether comments and create choice the facilities or goods. If consumer reads an insufficient review, consumer might acquire opinion assessment to be partiality. Therefore, opinion [1] excavating is a procedure of field area of info abstraction from text dispensation, which is assistance and several occasions to progress or change factor to business grind by this investigation. Investigation of judgment has equal of sentimentality from sentiment word and considered total of correspondence or collection with the benevolent of expression as optimistic or undesirable called sentiment polarity [2].

The problem is some customers given rating contrast with their comments. The other reviewers must read many comments and comprehensive the comments that are different from the rating. The challenge is in process of opinion mining or sentiment analysis that is unstructured and noisy data on website. The datasets used in OM are an essential problem in opinion review field.

Manuscript received: 25 May 2019
Manuscript received in revised form: 19 June 2019
Manuscript accepted: 06 July 2019
Manuscript Available online: 10 July 2019

The foremost issues that interfere with the user comfort and security are confidentiality breach, groups lacking opt-in options, disorder created out of various groups in which a consumer is a member of and struggle in managing group ideologies. The purpose of utmost of the groups in widely held social networking societies such as Twitter is misleading by uproars like spams and advertisements, and other clatters that obstruct with a group members interest.

The challenging task related with the groups is the administration of group guidelines. The furthermost widespread social networking site Twitter has group's members with over 100K. Consequently, it turns out to be difficult for the supervision to track the followers violating group guidelines. This represents that there occurs a need for a measure to classify the member posts based on suitability and group behavior. A scheme to control the entry of inappropriate messages within a group is very much necessary for the smooth operational of a social networking sites. These days, a company or organization makes available a business service which essentials to get opinion from customer. The customer assessment is essential to increase better service for company.

The remaining of the paper is summarized as follows. Section 2 represents background of sentiment analysis, sentiment polarity and machine learning. It also represents the types of machine learning algorithms. Section 3 provides literature survey. Section 4provides conclusions and future design.

II. SENTIMENT ANALYSIS AND POLARITY

Opinion Mining (OM) also termed as Sentiment Analysis (SA) [3] is the computational analysis of public's attitudes, opinions, and sentiments or opinions concerning an entity. The object can signify events, individuals, or topics. These subjects are furthermost likely to be concealed by analyses. The two expressions OM or SA are interchangeable and express a common meaning. Opinion Mining is an unending field of research in document mining field. Opinion mining will review different post of users and mine their opinion about related subjects. Opinion mining classification aims to categorize the opinion polarity of a tweet as negative, positive, or neutral.

The Clustering and natural language processing [4] procedure will be applied for opinion mining. A portion of opinion mining denotes using of natural language processing by suggested dissimilar method of dictionary for sentimentality analysis of text data as lexicon, corpus, and specific language dictionary.

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 8, Issue 12, June 2019

With the speedy extension of corporation or association have additional services and goods online and increase consumer contentment. The supplier will read consumer assessment and other consumers who necessity to procedure amenities or goods will read assessment to express opinions on the facilities. The numeral of consumer assessment is accumulative or vast from website, ha blogs, opportunities and community broadcasting, which the amenities or invention is motivating. Therefore, countless consumers will read comment randomly which is tough to read altogether comments and create choice the facilities or goods. If consumer reads insufficient reviews, consumer might acquire opinion assessment to be partiality. Therefore, opinion excavating is a procedure of field area of info abstraction from text dispensation, which is assistance and several occasions to progress or change factor to business grind by this investigation. Investigation of judgment has equal of sentimentality from sentiment word and considered total of correspondence or collection with the benevolent of expression as optimistic or undesirable called sentiment polarity.

Machine learning is a branch of computer science which works around building systems that have the ability to learn without being explicitly programmed for it.

Machine learning is about predicting the future based on the past. The difficulties faced by systems relying on hard-coded knowledge suggest that AI systems need the ability to acquire their own knowledge, by extracting patterns from raw data. This capability is known as machine learning. The introduction of machine learning enabled computers to tackle problems involving knowledge of the real world and make decisions that appear subjective. The performance of these simple machine learning algorithms depends heavily on the representation of the data they are given.

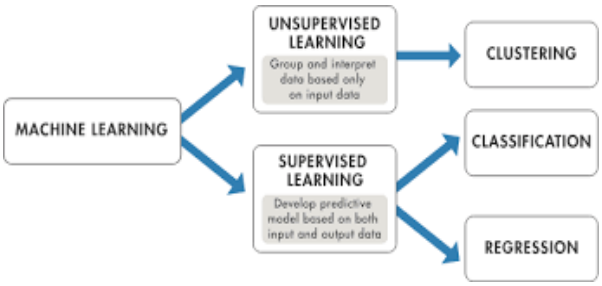

Fig. 1. Types of machine learning algorithm

The machine learning algorithm [5] are classified as supervised and unsupervised. Unsupervised learning is basically used as clustering techniques. Supervised learning can be used as regression and classification problems.

Supervised learning is the machine learning task of learning a function that maps an input to an output based on example input-output pairs. It is called supervised learning because the process of algorithm learning from the training dataset can be thought of as a teacher supervising the learning process. We know the correct answers; the algorithm iteratively makes predictions on the training data and is corrected by the teacher.

Unsupervised learning is the training of an artificial intelligence algorithm [6] using information that is neither classified nor labeled and allowing the algorithm to act on that information without guidance. In unsupervised learning, an AI system is presented with unlabeled, uncategorized data and the system's algorithms act on the data without prior training.

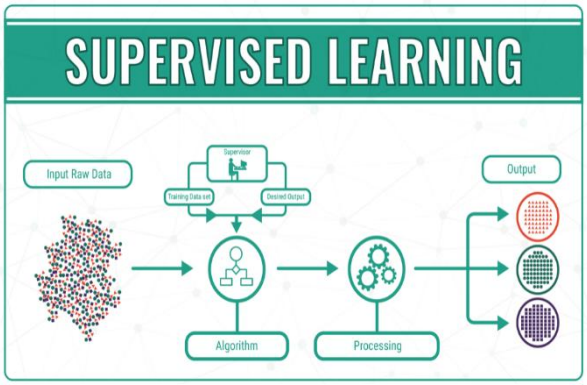

Fig. 2. Supervised Learning

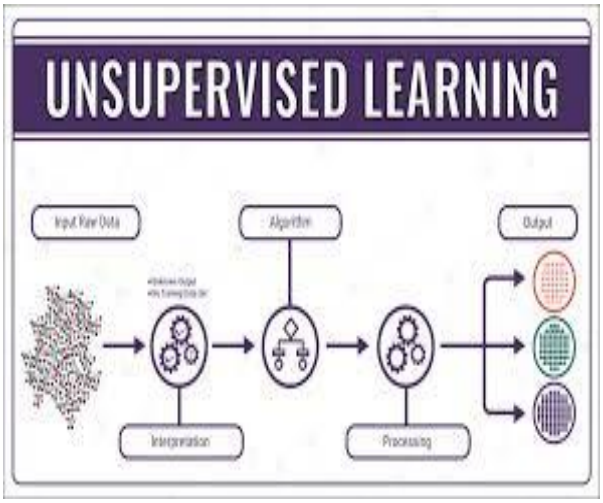

Fig. 3. Unsupervised learning

The figure 3 represents the working of unsupervised learning.

III. LITERATURE SURVEY

Wararat Songpan et.al.[7] Suggests the prediction rating and analysis from customer from hotel examinations who mentioned as open opinion with the help of probability's model as classifier. In further step, this classifier model has computed probability which represents value of style to provide the rating by applying naive Bayes procedures, which gives appropriately classifier to 93.47% as comparison with decision tree Methods. The suggested classifier models are applied in case study of consumer hotel review's in open posts comments for training dataset to categorized consumer comments as negative or positive.

The cleansing of data is performed by removing unused stop data and high frequency word selection by applying classifier model. The consumer reviews which can be

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 8, Issue 12, June 2019

negative as well as positive using data from training set and test set. The customer review's from hotel website agent service module for reservation system. The dataset from hotel checkout and checked in logs are arranged for classification. The existing literature proposes the analysis and prediction rating from customer reviews who commented as open opinion using probability's classifier model.

The classifier prototypes are mentioned case study of consumer reviews of hotel in open and in observations for training data to classify comments as negative or positive called opinion mining. In summation, this main classifier prototypical has premeditated the possibility that represents value of development to provide the rating by applying Naive Bayes procedures, which provides correctly categorized to 94.37% as compared with famous decision tree Techniques.

Su et. al [8]. Considers the probable of semantic expertise to address these enquiries. Afterward summarizing approaches to disambiguate and extract context information, the author present visualization procedures to discover the geospatial, lexical, and relational background of entities and topics referenced in these sources. The specimens stem as of the, The Climate Resilience Toolkit Media Watch on Climate Change, and the NOAA Media Watch—3 applications that combined environmental resources as of a widespread range of online sources. These schemes not only represents the value of as long as wide-ranging of information the public, but then again also have assisted to improve an innovative communication success metric that goes out there bipolar calculations of sentiment.

Kampset. el. [9] applied the Word Net dataset to search out the polarity of tokenized words. As they related a target text word to two key words (`bad' and `good') to search the lowest route distance related the pivot words and searched word the in the WordNet dataset hierarchy. The lowest path distance was transformed to an added total and this assessment was stored with the tokenizer word in the words dataset. The described accuracy level of this approach was 63.2%.

Littman et.el.[10] mapped the semantic relationship between the search final word and to every dataset word from the designated set of negative and positive words to a real number. By means of subtracting a word's relationship strength to a set of negative words as of its association strength to a group of positive words, an accuracy rate of 83% was accomplished.

SVMs were applied by Li and A. Li [11] as a sentiment polarity classifier. Dissimilar the binary classification difficulty, authors debated that expresser credibility and opinion subjectivity should also be considered into consideration. Authors suggested a framework that make available a condensed numeric summarization of sentiments on micro-blogs platforms. Authors extracted and identified the subjects mentioned in the opinions connected with the requests of users, and then categorized the opinions by using SVM. Authors worked on Twitter posts data for experiment. Author found out that the concern of user opinion subjectivity and credibility is necessary for accumulating micro-blog opinions. The proposed method proved that mechanism can effectually determine market intelligence (MI) for assistant decision-makers by instituting a monitoring method to track exterior opinions on dissimilar aspects of a business in actual time.

The proposed methodology used Thai customer review's hotels [12] from a website of hotel agent service, which service in hotel reservation directly. The process is started from collected data and preprocessing is cleaned data by removal stop words and using the high frequency of word which will be selected into attribute for using classifier model. The classifier model will be solve the text of customer review that is positive of negative from training data and test data which are train from behavior posting from customer of hotel service group. A part of opinion mining refers using of natural language processing by proposed different method of dictionary for sentiment analysis of text as corpus, lexicon and specific language dictionary [13]. The author tried to extract word from sentences for removal stop word or unnecessary word automatically. After reviewing literature we found some research gaps. The problem with most of the research is some customers give rating contrast with their comments. It is very difficult to find customers emotions related to comments. The other reviewers must read many comments and comments that are different from the rating. The research challenge is in process of opinion mining or sentiment analysis that is unstructured and noisy data on website. The data may contain fake review which are not considered in many research literatures. These challenges become obstacles in analyzing the accurate meaning of sentiments and detecting the suitable sentiment polarity.

Table 1: Data set S for data model

Set1 (14 words)	Positive: Best, good, better, excellent, Nice, very good, Beautiful Negative: Bad, not good, old, uncomfortable, expensive, unpleasant, problematic
Set2 (24 words)	Positive: Best, good, better, excellent, Nice, very good, Beautiful, safe, cheap, attractive, comfortable, popular Negative: Bad, not good, old, uncomfortable, expensive, unpleasant, problematic, troublesome, unsafe, inconvenient, costly, unfriendly
Set3 (40 words)	Positive: Best, good, better, excellent, Nice, very good, Beautiful, safe, cheap, attractive, comfortable, popular, new, thanks, special, enjoy, luxurious, not expensive, delicious, satisfied Negative: Bad, not good, old, uncomfortable, expensive, unpleasant, problematic, troublesome, unsafe, inconvenient, costly, unfriendly, not delicious, immortal, risky, unsatisfied, rare, not worth, not beautiful

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 8, Issue 12, June 2019

Table 2: API opinion mining

Sr. No	API	Meaning
1	google.cloud	Google cloud library
2	google.cloud.language	Google cloud natural language processing library
3	json	Library for JSON
4	requests	Facebook dataset requests library
5	facebook	Facebook SDK
6	collections	Library for collection
7	textblob	Library for sentiment analysis

Table 1 represents the dataset which can be used in sentiment analysis and sentiment polarity. Table 2 represents API which can be used to implement the sentiment analysis and sentiment polarity problems

IV. CONCLUSIONS

Twitter has currently become a platform for individuals and organizations that have a strong political, social, or economic concern in enhancing and maintaining their reputation. These days, a company or organization make available a business service which essentials to get opinion from customer. The customer assessment is essential to increase better service for company. Opinion mining provides these organizations the capability to monitor dissimilar social media web sites in real time. A portion of opinion mining denotes using of natural language processing by suggested dissimilar method of dictionary for sentimentality analysis of text data as lexicon, corpus, and specific language dictionary. Opinion mining is the process of automatically identifying whether a post segment contains opinionated or emotional content, and it can likewise determine the post's polarity. This paper provides the survey to customer review on different products with sentiment analysis and opinion polarity. It also represents different machine learning algorithms with advantages and disadvantages.

REFERENCES

[1] AranoScharil, David Hering, Wallter Rafeelsberger, Alexander Hubman-Haidvogeel, Ruslaan Kaamolov, Daniel Fishchl, Michal Foils, and Alberto Weichsebraun, "Semantic System and Visual Tools Applications to Support Environmental Communication", IEEE System Journal, VOL. 12, NO. 2, JUNE 2017, pp. 752-762.

[2] I. Qaarqaz, Araabic and R.M. Duwaii Sentiment Analysis using Supervised Classification. Proceeding of 2014 International Conference on Internet of Things and Cloud Computing. 2014, pp. 589-593.

[3] D.C. Londhee and B.V. Raaut, "Survey on opinion mining and summarization of user review on web", Journal of Information Technology and Computer Science, Volume. 4, 2014, pp. 1036-1040.

[4] S. Mohammed, Fiaidhi, O. Mohammed, T.H, Kim, S. Fong, Opinion Mining over Twitter and space: Classifying tweets programmacally using the R approach. Proceeding of the 7th International Conference on Digital Info Management, 2012, pp. 323-329.

[5] Athiroo U, and Saboo M. Thampl, "Linguistic Feature Constructed Filtering Mechanisms for Recommendation Post in a Social Networking Group", IEEE 18, pp. 4479-4494.

[6] S. 1. Wu, R.D. Chiang and Z.H. Ji, Development of a Chinese opinion mining system for application to Internet online forum, The Journal of Supercomputing, Springer US[Online],2016.

[7] Waraaraat Songpaan, The Analysis and Prediction of Customer Review Rating Using Opinion Mining, IEEE SEARA 2017, pp. 72-78.

[8] R. Zhang, Li, W, L. Lin, Yu and C. Sun, Opinion mining and sentiment analysis in social nets: A retweeting structure with aware approach. Proceeding of the 6th International Conferences on Cloud Computing and Utility, 2014, pp.880-885.

[9] Kampas, J., Marxi, M.K., Mokkeen, R. J.Using Word Nett to Measure Semantic Orientation of Adjective. LRREC 2004. Vol. IV, pp. 1215-1218.

[10] T.V. Le, H.S. Le, and V.T. Phaami, Aspect Analysis for Opinion Mining of Vietnamese Text. Proceeding of International Conference on Advance Application and Computing, 2015, pp.128-133.

[11] L.Liu, Z. Li, and Ci.Li, Analysis of customer satisfaction from China reviews by using opinion mining, Proceeding of the 7th IEEE International Conferences on Software Engineering and Service Science (ICESES). 2015, pp.105-109.

[12] Andrevskia, A., Berglar, S., Urseana, M.All Blogs Are Not Made Equals: Exploring Genre Differences in Sentiment Tagging of Blog. International Conference on Weblog and Social Media (ICSWM-2007), Booulder, CAO. 2007.

[13] Vandaana V. Chaudhary, Chitraa A. Dhawala and Sanjeev Mishra, "Sentiment Analysis Classification: A Brief Review", I J C T A, 9(23) 2016, pp. 447-454.