

Analysis of Emotion using Machine Learning on Social Media Platform

Sovit Nayak

Student, B.Tech- Electronics and Communication
Engineering Department of Electronics
SRM Institute Of Science And Technology
Kattankulathur, India

Jugal Prasad Dutta

Student, B.Tech- Computer Science Engineering
Department of Computing
SRM Institute Of Science And Technology
Kattankulathur, India

ABSTRACT

In this rapid, non-stop world everything and everyone goes on dramatically and enthusiastically. And people don't usually express themselves to their loved ones, this creates a sense of unreliability which leads to harm. Instead of speaking to their near and dear ones, they usually express themselves using social media. Usually, a lot of time people don't even know that they are depressed or have anxiety. Which leads to unreasonable actions like suicide? However, if we could know how they feel and somewhat solve their dilemma by consulting doctors and help groups. Here we are using machine learning algorithms to analyze their thoughts using polarity and subjectivity. **Polarity** determines emotions expressed in a sentence, emotions are closely related to sentiments. The strength of sentiment or opinion is typically linked to the intensity of certain emotions, e.g., joy and anger. And **Subjectivity** expresses some personal feelings, views, or beliefs. We would be able to find out if their views are positive or negative, further we could diagnose their mental condition and act accordingly.

General Terms

Natural Language Processing, Sentimental Analysis

Keywords

Depression, Anxiety, Sentimental Analysis, NLP, Polarity, Subjectivity

1. INTRODUCTION

Emotions are the most integral part of any human being to bring a balance in work and personal life. If one doesn't have a grip on his/her emotions its effects his/her life along with their closed one's. And as emotions are closely related to sentiment, it's required to have an observation on sentiment to get hold of one's emotion.

This paper proposes a detailed Sentimental Analysis using the help of Natural Language Processing. In the era of Internet, mostly surveys have been found people expressing their thoughts openly over various social platforms, some with identity and some being anonymous. Here we worked on Machine Learning Algorithms to analyze individual persons thought through their posts, comments etc. over social platform. These Algorithms are then utilized for Natural Language Processing through Polarity and Subjectivity.

1.1 Natural Language Processing

To describe Natural Language Processing in simple words, it is the ability of a computer program to understand human language as it is spoken and written. Just as humans have

different sensors -- such as ears to hear and eyes to see -- computers have programs to read and microphones to collect audio, thus with NLP a vast improvement in communication between human and computer is observed. And just as humans have a brain to process that input, computers have a program to process their respective inputs. Natural Language Processing basically revolves around two processes, data preprocessing and algorithm development.

1.2 Polarity

It simply means emotions expressed in a sentence. As we mentioned above, Emotions are closely related to sentiments. The strength of a sentiment or opinion is typically linked to the intensity of certain emotions, e.g., joy and anger. It is the expression that determines the sentimental aspect of an opinion. In textual data, the result of sentiment analysis can be determined for each entity in the sentence, document or sentence. The sentiment polarity can be determined as positive, negative and neutral. Polarity is basically a float which ranges from -1 to 1 where 1 implies positive statement and -1 implies a negative statement. The *sentiment* function of text blob returns two properties, that is, polarity, and subjectivity.

1.3 Subjectivity

Subjectivity analysis recognizes the contextual polarity of opinions, attitudes, emotions, feelings etc. Subjective sentences generally refer to personal opinion, emotion or judgment. As the basic task, it classifies a text as subjective or objective. While subjective text expresses opinions about an object or issue using sentiment expressions, objective text describes an object or issue considering their facts. It is also a float which lies in the range of [0, 1]. This paper proposes collection of different words into a word cloud by using subjective patterns for subjectivity analysis in different social platform domains.

2. SENTIMENTAL ANALYSIS

It is a part of Natural Language Processing that deals with text analysis, computational linguistics, and biometrics and benefits to systematically identify, withdraw emotions within text or speech. It also helps in analyzing a sentence and helps us determine whether the connotations are positive, negative, or neutral.

Sentiment analysis models focus on polarity (*positive, negative, neutral*) but also on feelings and emotions (*angry, happy, sad, etc.*), urgency (*urgent, not urgent*), and intentions as well.

There are different types of Sentimental analysis classified into some of the following:-

A) Fine-Grained Sentimental analysis

This deals with precisions in determining polarity. For example, think about how you went to a restaurant and rate their service by using stars or points. This usually contains words that give a sense of the following.

- Very positive = 5 stars /points
- Positive = 4 stars/points
- Neutral = 3 stars/points
- Negative = 2 stars/points
- Very negative = 1 star/point

B) Emotion detection

This deals with the sensibility and emotional condition of the person expressing their views which includes anger, frustration, sadness, happiness, and even depression.

On the other hand, it can also determine intense feelings like (egg: murder, kidnap, extortion, drugs deals, etc.)

The main challenge in doing all of these is extracting data in real-time and real emotions and views should be expressed from the other side as well, i.e. the user.

2.1 Sentimental analysis of data

The process of obtaining the sentiments from the tweets consists of several steps explained down below:

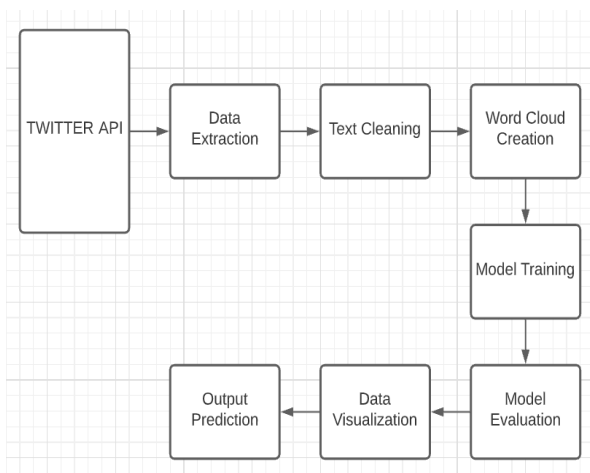


Fig. 1

1) Extraction of Data: This is the first step in the entire process of sentimental analysis wherein data is collected from websites like Twitter using the API key and Token key. This extraction includes all data like tweets like dislike and comments.

2) Text Cleaning: After the data is extracted from the source, then these datasets are cleaned from emoticons, URLs, and redundant words which would eventually impact the overall emotions of the extracted tweets. Because if a negative word is repeated uncontrollably in one particular tweet it would give result as more negative but in reality, the word has been repeated.

3) Performing Sentimental Analysis: At this stage, the dataset is ready for classification. Now Natural Language Processing is used for semantic reasoning, where all the words in the sentence are examined and will be categorized into positive, negative, and neutral according to their subjectivity and objectivity which is differentiated into connotations such as good, bad, or like, dislike. Subjective sentences generally

refer to personal opinion, emotion, or judgment whereas objective refers to factual information. After all, Polarity is used to assign percentage scores for each word.

4) Output: Once the algorithm is complete, by generating meaningful data from the obtained raw datasets, which we've gotten from Twitter API. The denouement can be viewed using data visualization in the form of Bar graphs, Pie charts and Time Series, etc. To evaluate the percentage of likes and dislikes we can use a pie chart. Similarly, for positive, negative, and neutral thought percentage calculations we can use bar graphs. If the whole procedure is taken into account for its entirety, we can observe that the last step is of paramount importance as this lets us view the work being implemented in the flesh. We can furthermore create live flamboyant dashboards using Tableau or Power BI to make the end result look more appealing

3. SYSTEM ARCHITECTURE

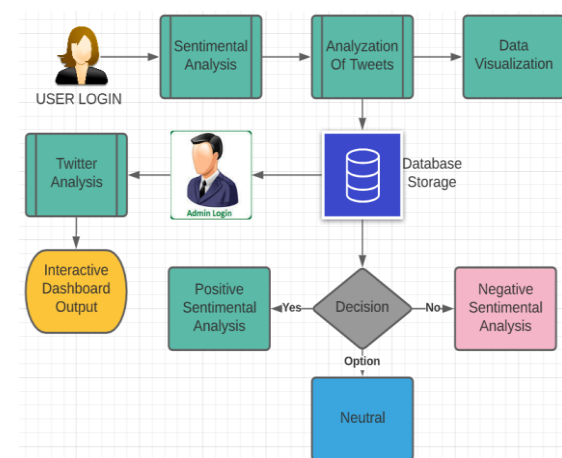


Fig. 2

The role of the admin is here to perform the analysis from the data input from different users by collecting credentials, likes, dislikes, and re-tweets, new tweets. Sentimental analysis is performed on the user data and is checked for any words pointing towards depression and anxiety. Database plays a major role in storing user information which can be extracted at any given time. Real-time tweets and credential is collected from the user and a specific dataset is formed. Along with those texts, analysis is performed, which is evaluated using a graph that is further stored in the database. We can also use filters to separate between positive, negative, and neutral tweets for which specific keyword is set. The positive keyword is set for those tweets that indicate depression or anxiety. Similarly, Negative keywords indicate nothing related to depression and anxiety, and finally, the neutral keyword is for those tweets where the differentiation could not be made between positive and negative tweets. Or if the tweet contains both positive and negative tweets the level of positivity or negativity is evaluated and specified.

4. OUTPUT GRAPHS

Here we have analyzed “Bill Gates” last 200 tweets and categorized it into Positive, Negative and Neutral based on the below parameters i.e Subjectivity and Polarity.

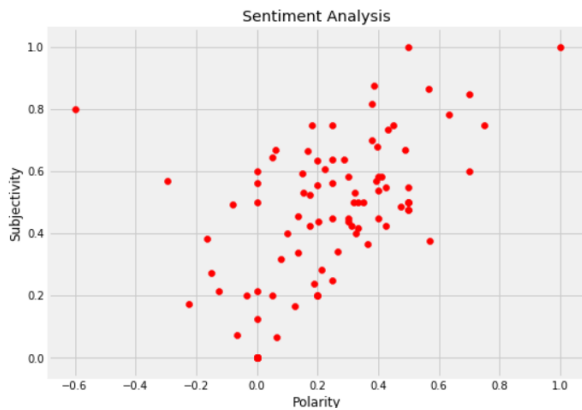


Fig. 3

After analyzing the tweets we have found that 71% of the tweets are Positive, 19% of the tweets are Neutral and approximately 10% of the tweets are Negative, based on the graphs given below in Fig. 4

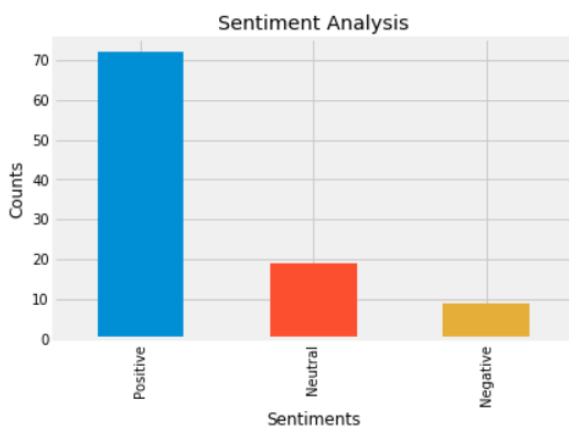


Fig. 4

5. CONCLUSION & FUTURE SCOPE

The main purpose of doing analysis is to recognizing the positivity and negativity in anonymous groups through polarity and subjectivity. This works by extracting the anonymous tweets and comments and grouping into different words to form a word cloud. It is later processed by natural language processing through different Subjectivity and Polarity algorithms, which further gives output as range of negativity and positivity of individual tweets. In long run it can be beneficial to both the business leading sectors and individual anonymous for their personal growth and development with analyzing their state of mind and taking required therapeutics.

Being in 21th century, where technology develops with a snap of finger, Sentimental Analysis plays vital role in bringing human race to next level upgradation. Along with the algorithm proposed in this paper, we can not only detect in which state human lies but also a scheming solution for their betterment of health .In the near future we can improve this prototype by implementing it in different languages so that we can have a wide range of users. Not only twitter, we can also apply these model algorithm in variety of social media

platforms, which would increase the number of users eventually.

6. REFERENCES

- [1] Luciano Barbosa and Julian Feng. "Robust sentiment detection on twitter from biased and noisy data." Proceedings of the 23rd international conference on computational linguistics: posters. Association for Computational Linguistics, 2010.
- [2] Sahayak V, Shete V & Pathan A (2015). "Sentiment analysis on twitter data." International Journal of InnovativeResearch in Advanced Engineering (IJRAE), 2(1), 178-183.
- [3] Gupta B, Negi M, Vishwakarma K, Rawat G & Badhani P(2017). "Study of Twitter sentiment analysis using machine Learning algorithms on Python." International Journal of Computer Applications, 165(9) 0975-8887.
- [4] Mangain N, Mehta E, Mittal A & Bhatt G (2016, March). "Sentiment analysis of top colleges in India using Twitter data." In Computational Techniques, in Information and Communication Technologies (ICCTICT), 2016 International Conference on (pp. 525-530). IEEE.
- [5] Liu, B., Sentiment Analysis and Opinion Mining. Morgan & Claypool Publishers,2012.
- [6] Nisha, J., & Dr.E. Kirubakaran, "M-Learning sentiment analysis with Data MiningTechniques", International Journal of Computer Science and Telecommunications, Volume 3, Issue 8, 2012.
- [7] Schukla, A., "Sentiment analysis of document based on annotation", CORR Journal, Vol. abs/1111.1648, 2011.
- [8] Kasper, W. & Vela, M., "Sentiment analysis for hotel reviews", proceedings of the computational linguistics-applications, Jachranka Conference, 2011.
- [9] Zhang, L., Hua, K., Wang, H., and Qian, G., "Sentiments reviews for mobile devices products", The 11th International Conference on Mobile Systems and Pervasive Computing (MobiSPC-2014) ,procedia computer science, Volume 34,2014.
- [10] Godbole, N., Srinivasaiah, M., and Skiena, S., "Large-Scale Sentiment Analysis for News and Blogs", ICWSM'2007 Boulder, Colorado, USA, 2007
- [11] Kumar, A., & Sebastian, T. M., "Sentiment Analysis On Twitter". IJCSI International Journal of Computer Science, Issues 9(3):372–378, 2012.
- [12] Apoorv Agarwal, Fadi Biadisy, and Kathleen R. Mckeown. "Contextual phrase-level polarity analysis using lexical affect scoring and syntactic n-grams." Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics. Association for Computational Linguistics, 2009.
- [13] Soo-Min Kim and Eduard Hovy. "Determining the sentiment of opinions." Proceedings of the 20th international conference on Computational Linguistics. Association for Computational Linguistics, 2004.