

# Conception Analysis on User Tweets by Using Natural Language Processing Algorithm

Priyanka Asthana

*Dept. of Computer Science and Engineering  
R.B.S College Bichpuri Agra*

Submitted: 05-02-2022

Revised: 18-02-2022

Accepted: 20-02-2022

## ABSTRACT

Conception analysis of data or text available on social media sites in the form of blogs which gives the information about the people perspective of service providers like airtel, idea etc that they interested. We uses a Natural Language Processing algorithm (NLP) for conception analysis. The aim of this research paper is that to collect or extract the tweets from twitter in which some of them are positive and some are negative which shows people perspective about service providers and on each tweet we perform a conception analysis and calculated the score of positive and negative tweets and generate a histogram of these scores. The information which we gained through conception analysis is used to improve the services of service providers and helps them to take the appropriate decisions about services of ISP's.

## I. INTRODUCTION

Today's Conception analysis is used in wide variety of fields such as social networking sites where people share their views and opinions regarding different Internet Service Providers (ISP) which shows the attitude of a people towards the particular service which helps the service providers to judge the responses of people it is either positive or negative response and tries to improve the services. Conception analysis refers to use Of NLP, text analytics, computational linguistics and to identify, extract, quantify and subjective information [1].

**Problem Statement :** We use Natural Language Processing (NLP) algorithm in this Paper and analyze the performance of Internet Service Providers (ISP) on the basis of user's feedback that can be positive as well as Negative. We

collect the tweets from the Twitter as a feedback. This helps the Service providers to enhance the Performance of ISP. Because manually we have to check the reviews of ISP that is time taken so we can reduce the time and increase the performance of system.

This approach is better than manual approach. Manually companies face difficulties to check review of services because it is a voluminous data so extract the information from voluminous data is a very difficult task. So, this approach is better than the other approach and also beneficial for the companies.

Through these reviews company also check whether how many users are satisfied and how many users are not satisfied with the service. In Conception Analysis process which is shown in Fig.1.1 first tweets are fetched and data preprocessing is done and then stop words are removed with the help of Data Cleaning process

## II. RELATED WORK

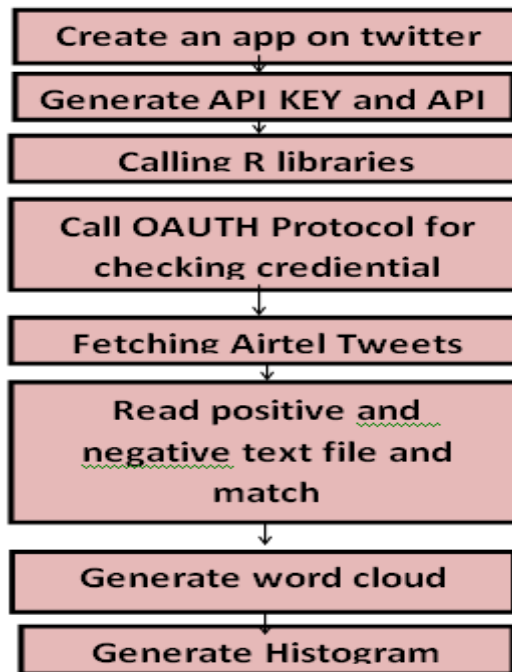
There are a large number of approaches that has been developed to date for classifying conceptions or polarities in English texts. Conception Analysis (B.Liu[2], Pang&Lee[3], Turney[4]) a research area that aims at detecting the authors sentiment, emotions or opinion about the events, topics or individuals expressed in text. Thelwalle [5] analyzed popular events in Twitter and showed that they are related to an increase in negative sentiment strength. Asur and Huberman [6] constructed a model based on tweet-rate about particular topics for predicting profit of movies before their release. They further showed how sentiment extracted from Twitter posts can improve the predictive power. Bollen, Mao[7], and Pepe[8] found that there is a association between the public mood (in terms of

tension, depression, anger, vigor, fatigue, and confusion) that emotion are expressed in Twitter posts and social, political, cultural, and economic events. One of the simplest ways to select the examples to be labeled is based on maximizing the expected informativeness of labeled examples. For example, the learner may find the examples with the highest uncertainty to be the most informative and request them to be labeled. Zhu et al.[9] applied (among other strategies) uncertainty sampling in two scenarios: (i) local uncertainty sampling which selects instances for labeling based only on a current batch of data from the data stream, and (ii) global uncertainty sampling which in the selection process takes into account also previous classifiers and thus forms a classifier ensemble. N. A. H. Haldar[10] the popularity of social sites like Twitter, which encourages users to exchange short messages, Twitter is being used by public for upgrade and feeling expression.

### III. PROPOSED METHOD

In this paper we have done Conception Analysis on “Airtel tweets” to fetch the tweets that are posted on the twitter. This helps to check the opinion of Airtel tweets and then it helps in decision making As per Fig. 3.1 these are the ways to fetch the tweets . In this paper we use R language to do Conception Analysis on ISP Tweets”. 1.Make an account on twitter and Then Create an app on Twitter.

- When we make it own app then api key and api secret are generated.
- Then call R libraries which are mandatory and then call Oauth Authentication Protocol and it check the authenticated users.
- Then Tweets are fetched and then positive and negative words is read in Lexicon Dictionary.
- When tweets are fetched then positive words match with positive and negative words match with negative.
- Then scoring of tweets are done Histogram and Wordcloud generated.



### IV. RESULTS & ANALYSIS

There are the following results when we perform Conception Analysis then determine the positive, negative and neutral tweets and then find the scores of sentiments. Then calculate score by using Score Functions i.e.

Score= Positive Tweets - Negative Tweets

We use Precision and Recall as the evaluation metric to measure our rating prediction performance.

**Accuracy:** Accuracy can be defined as the no. of correct cases divided by the total no. of cases.

Accuracy = (total correct items/ total no of items)/100

**Precision:** Precision can be defined as the fraction of retrieved instances or it can also be defined as the percentage of selected items are correct.

$$\text{Precision} = \frac{tp}{tp + fp}$$

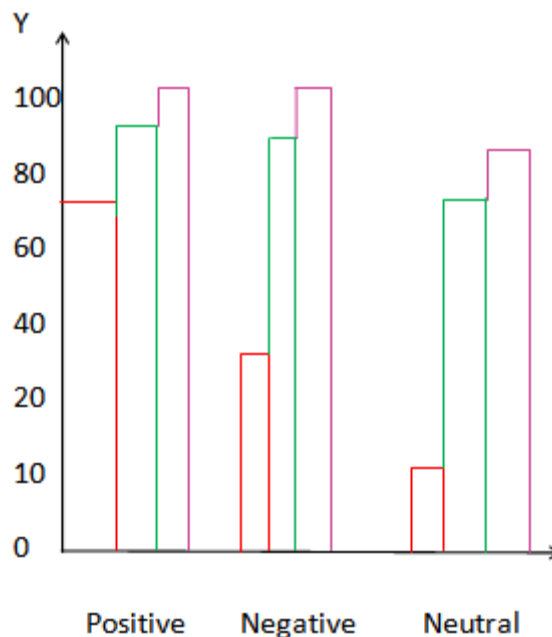
**Recall:** It can be defined as the percentage of correct item can be selected.

$$\text{Recall} = \frac{tp}{tp + fn}$$

Where tp is true positive, tn is true negative. fp is false positive and fn is false negative and tweets accuracy, precision and recall of the tweets and through these data company enhance the services and also analyze what is the repo of specified ISP and also analyze performance of the company then rating of companies also decided with the help of users post.

**Table 1: shows score of tweets**

Class	Accuracy	Precision	Recall
Positive	71.86%	93.3%	97.7%
Negative	28.13%	81.1%	92.7%
Neutral	6.67%	57.14%	80%

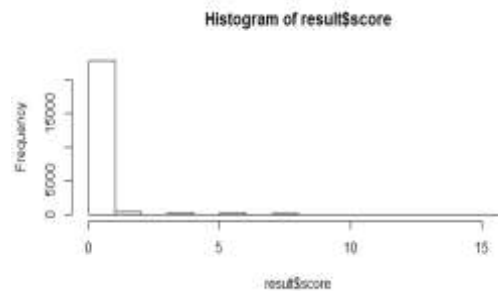


**Fig. 4.1: show accuracy, precision, recall of tweets**

X- axis shows the number of tweets and Y - axis shows the Accuracy, Precision and Recall of positive, negative and neutral tweets in Fig. 4.1 Red lineshows the Accuracy, Green line shows the Precision and pink line shows the Recall of positive, negative and neutral tweets. The Accuracy of positive tweets is higher than the negative then the value of Accuracy decreases rapidly. In case of Precision the value of precision also decreases the highest value of precision is at Positive tweets. In case of Recall the value of

recall continuously decreases. So in the three cases the value of Accuracy, Precision, Recall is highest for positive tweets.

Plot 1: Shows the scoring of tweets which is computed by function score. In this figure x- axis tells the score and y- axis tell the number of tweets and scoring of tweets. In Fig. 4.2 we analyze the score of tweets that is almost 0 at 5,10,15 that shows the negative tweets but at 0 which represent positive tweets.



**Fig. 4.2: Histogram Of Tweets**

PLOT 2 : frequently words when user post the tweets on Twitter and viewed tweets.



**Fig. 4.3: Wordcloud Of tweets**

In this word cloud based on the frequency or occurrence of words in tweets there is variation in size of words. It shows the word cloud of frequently used tweets below shows the distribution of positive tweets, negative tweets and neutral tweets which is shown in Fig. 4.3.

#### Naïve Baye's Classifier

Naive Baye's classifier are the most important algorithm for learning to classify the text document. Naive Baye's classifier basically used in learning algorithms. A Naive Baye's classifier is a program which predict a class value give in a set of attributes.

For each known class value:

1. Calculate probabilities for each attribute , conditional on the class value.
2. Use the product rule to obtain a joint conditional probabilities for the attributes.

**BAYE'S THEOREM:** Baye's theorem describe the probability of an event , based on priorknowledge of condition that might be related to the event.

$$P(H|E) = \frac{P(E|H).P(H)}{P(E)}$$

Meaningful data is given to the naïve baye's classifier and then naïve baye's classifier trained the meaningful data then machine learning algorithm is applied to the trained data then data is crawled to the classifier and then classifier will do prediction to the trained data and classify the object.

#### APPLICATION OF NAÏVE BAYES

- It also helps us to classify the new article about technology , politics, report.

- It also check piece of text expressing positive emotion, negative emotion.
- It is also used in face recognition .

#### V. CONCLUSIONS

These results shows that our conception analysis method is helpful is categorizing the tweets as feary, joyful, disgusting or broader terms either positive or negative. These kind of information may help companies to improve their quality or service for the products. This kind of feedback is easily accessible and needs money for feedback . The action can be taken and we may revisit the customers for their revised feedback . Predicting future trends , events, phenomenon is an interesting task , commonly connected to the analysis of public mood. Various studies indicates that analysis of such texts can be automated and can produce useful results. This dissertation presents the study whose main task is to investigate whether sentiment analysis of Twitter microblogging is a suitable datasource for predicting tweets for various services.

In order to perform the high quality sentiment analysis of twitter message which have specific nature , we conduct a series of experiment to determine the most suitable. Twitter conception analysis algorithm and the best text processing setting . The experiment showed that , by applying different combination of preprocessing settings.

## VI. FUTURE WORK

Substantial amount of work is left to be carried on, here we provide a beam of light in direction of possible future avenues of research.

The planned future work is concerned with, on the one hand, improving the methodology for collecting manual labeling, preprocessing and analysis of twitter messages, on the other hand, advancing the application of the study.

First, we plan to take special care on tweets and their manual labeling. Mainly, we intend to examine whether the approach to collecting the relevant tweet from Twitter API and composing a dataset for hand-labeling can be improved in terms of better preparing the queries for the Twitter API and additional processing of the collected data. We also plan to analyze the process of manual annotation and obtained hand labels

## REFERENCES

- [1]. Vryniotis, Vasilis (2013). The importance of Neutral Class in Sentiment Analysis.
- [2]. B. Liu, "Sentiment Analysis and Subjectivity," Handbook of Natural Language Processing, 2nd Edition, CRC Press, New York, 2010.
- [3]. Pang & Lee, L. (2008). Opinion mining and sentiment analysis. Foundations and trends in information retrieval, 2(1-2), 1-135.
- [4]. Zhu X, Turney P, Lemire D, Vellino A. Measuring academic influence: not all citations are equal. Journal of the Association for Information Science and Technology. 2015;66(2):408-27.
- [5]. Thalwaller et al (2010). Sentiment Analysis Area. Journal of the Association for Information Science and Technology.
- [6]. Asur and Huberman (2010): Predicting the future with Social Media at SitaramAsur Computing Lab.
- [7]. Bollen, Johan; Mao, Huina; Zeng, Xiaojun (28 February 2011). "Twitter mood predicts the stock market". Journal of Computational Science. 2 (1): 1-8. doi:10.1016/j.jocs.2010.12.007.
- [8]. Pepe, A., and Bollen, J. 2008. Between conjecture and memento: shaping a collective emotional perception of the future.
- [9]. Pepe, A., and Bollen, J. 2008. Between conjecture and memento: shaping a collective emotional perception of the future.
- [10]. Zhu, Y., Song, D., Tran, N.T., and Nguyen, N. (2007) The effects of the members of growth hormone family knockdown in zebrafish development. Gen Comp Endocrinol 150(3): 395-404.
- [11]. N. Azam, Jahiruddin, M. Abulaish and N. A. H. Haldar, "Twitter Data Mining of Events Analysis and Classification," 2015 Second International Conference on Soft Computing and Machine Intelligence (ISCMI), Hong Kong, 2015, pp. 79-83.