Twitter Sentiment Analysis

Parth D Patel[1], Yash R Patel[2], Kishan Patel[3], Sharvil Patel[4]

[1]parth.p2@ahduni.edu.in

[2]yash.p3@ahduni.edu.in

[3]kishan.p1@ahduni.edu.in

[4]sharvil.p@ahduni.edu.in

***Abstract*— This is an aim to create a model for companies that can deliver better experiences to their customers across the world on the basis of thoughts, beliefs, and mindset of customers themselves through their tweets. To analyze Twitter data to perform algorithms and predict better results through Machine Learning models to evaluate a company or brand in real-time on the basis of the past datasets and for changes in their business models, detect anomalies with alert and public engagement.**

***Keywords*— Sentiment Analysis | Feature Extraction | Classification | Trend Analysis | Regular Expression | Naive Bayes Classification**

I. Introduction

This project addresses the problem of sentiment analysis in twitter; that is classifying tweets according to the sentiment expressed in them: positive, negative or neutral. Analysing the public sentiment is important for many applications such as firms trying to find out the response of their products in the market, predicting political elections and predicting socioeconomic phenomena like stock exchange and then making relevant changes in their business models and strategies. The aim of this project is to develop a functional classifier for accurate and automatic sentiment classification of an unknown tweet stream and letting the brands make it easier to know about the trends going on in the real market through this wonderful micro-blogging platform, Twitter.

II. Literature Survey

Sentiment analysis is learning of people's emotions, views, attitude, and opinions. It is also called opinion mining. Sentiment analysis identifies the sentiment articulated in a text then analyzes it. Sentiment Analysis is a type of Natural Language Processing (NLP) for tracking the mood of the public through a particular product. There is a wide range of applications of Sentiment Analysis in real life usage and one such we will be describing here in detail along with the process, models, and inferences.

Twitter Sentiment Analysis is a very known topic in this field due to its wide range of applications in various fields such as research, politics, trend recognition, personality tracking, brand improvement, etc. Through a wide range of research in this field from past few years, it has obtained a huge amount of attention for its application in this field and a lot of articles and projects have been developed. The scale to which this can be implemented is now reaching new heights and we aim to do one such research project involving the techniques of machine learning that can be useful for real life applications at a large scale for companies, brands, individuals as well as governments.

III. Implementation

First we understood the basic sentiments of the customer/ public (tweet which are considered in our case) which a brand needs to monitor and take action on. First of all, only binary classification is done. But, for the case of tweets in this micro-blogging platform, the corpus/sentiments are divided into three categories - negative ,neutral and positive. For a broader and more accurate classification for a larger dataset, it can be even divided into 5 categories as extremely positive, positive, neutral, negative and extremely negative. Initially, we have used the sample data of relatively small dataset of tweets for trying our initial machine learning model. Then, after exploring the Twitter Developer account and its relevant resources, we found out about the use case of Twitter API and found the required tweetS data regarding different brands and companies from different locations (states and countries).

In the medial part, we need to pre-process the data, to make it model application ready, It involves various factors, like use of special characters such as "@" and "#" in tweets which are very popular. We need to check about other factors too sucha s stopwords, POS tagging, noun phrase extraction, spelling correction, parsing and use of multiple n-gram models for efficient deployment and application of model with an aim to achieve as maxim accuracy as possible.

When we extract the tweets we check for mentions ,hashtags for our brand. Also, if available locations and geographical segmentation factors are taken into consideration for more details of the dataset and then generating trends and results as per it. Then using those tweets which extract the features of that tweet (tokenization and data cleaning). Then, using the trained model we use a bag of predefined words or lexicons. Then using that model we classify the test data (live tweets). And using python libraries we aim to plot the pie chart of sentiments.

IV. Results

This project is to develop a functional classifier for accurate and automatic sentiment classification of an unknown tweet his project is to develop a functional classifier for accurate and automatic sentiment classification of an unknown tweet

And using that we plot a pie chart to show the percentage of each sentiment.For further better visual representation, we aim to deploy the live streaming of dataset through applications, tools and technologies such as Plotly, Dash where run time data of live tweets can be fetched from the official data of Twitter API.

V. Conclusion

In this paper about the sentence-level sentiment analysis. Users are getting information from the online whether it may be a positive polarity or negative polarity or neutral. Finally they come under specific polarity from the reviews or views or opinions from the online like social media. Sentences may text or emotions or emoticons or expressions. Even though emotions look like a diagram, it represents some textual information.

VI. References

1] A. Go, R. Bhayani, and L. Huang, "Twitter Sentiment Classification using Distant Supervision," *Stanford.edu*. [Online]. Available: https://cs.stanford.edu/people/alecmgo/papers/TwitterDistantSupervision09.pdf. [Accessed: 17-Mar-2021].

2] L. Jiang, M. Yu, M. Zhou, X. Liu, and T. Zhao, "Target-dependent twitter sentiment classification," *Aclweb.org*. [Online]. Available: https://www.aclweb.org/anthology/P11-1016.pdf. [Accessed: 17-Mar-2021].

3] I. G. Councill, R. Mc Donald, L. V. Google, and 76 Ninth Avenue, "What's great and what's not: Learning to classify the scope of negation for improved sentiment analysis," *Googleusercontent.com*. [Online]. Available: http://static.googleusercontent.com/media/research.google.com/en//pubs/archive/36744.pdf. [Accessed: 17-Mar-2021].

4] "Twitter Sentiment Analysis using Python - GeeksforGeeks," *Geeksforgeeks.org*, 24-Jan-2017. [Online]. Available: https://www.geeksforgeeks.org/twitter-sentiment-analysis-using-python/. [Accessed: 17-Mar-2021].

5] "For Academics - Sentiment140 - A Twitter Sentiment Analysis Tool," *Sentiment140.com*. [Online]. Available: http://help.sentiment140.com/for-students. [Accessed: 17-Mar-2021].

6] A. J. Schoubye, "Descriptions, truth value intuitions, and questions," *Linguist. Philos.*, vol. 32, no. 6, pp. 583–617, 2009.

7] B. Pang and L. Lee, "Opinion mining and sentiment analysis," *Found. Trends® Inf. Retr.*, vol. 2, no. 1–2, pp. 1–135, 2008.

8] M. Afham, "Twitter Sentiment Analysis using NLTK, Python," *Towards Data Science*, 25-Sep-2019. [Online]. Available: https://towardsdatascience.com/twitter-sentiment-analysis-classification-using-nltk-python-fa912578614c. [Accessed: 17-Mar-2021].

9] *Iitk.ac.in*. [Online]. Available: https://cse.iitk.ac.in/users/cs365/2015/_submissions/ajaysi/slides.pdf. [Accessed: 17-Mar-2021].