

Sentiment Analysis of COVID-19 Tweets – Visualization Dashboard

Index

Chapt. No	Chapt. Name	Page No.
1	Introduction 1.1 Overview 1.2 Purpose	3
2	Literature Survey 2.1 Existing Problem 2.2 Proposed solution	4
3	Experimental Investigations	5
4	Flowchart	6
5	Result	7
6	Advantages and Disadvantages	8
7	Applications	9
8	Conclusion	10
9	Future Scope	11
10	Bibilography Appendix A. Source Code B. Video and source code	12

1. Introduction

1.1 Overview

On 11th March 2020, World Health Organization announced COVID-19 outbreak as a pandemic. Starting from China, this virus has infected and killed thousands of people from Italy, Spain, USA, Iran and other European countries as well.

While this pandemic has continued to affect the lives of millions, a number of countries have resorted to complete lockdown. During this lockdown, people have taken social networks to express their feelings and find a way to calm themselves down.

In this Project, We have used sentiment140 dataset, to analyse sentiments of people across the India.

The dataset have been pre-processed, and then used for building the predictive model for sentiment analysis.

As a conclusion, We can now address Live people's behavior on the lockdown if it is further extended with the help of Visualization Dashboard.

1.2 Purpose

The purpose of this project is to understand the behavior of people if the lockdown is further extended and to develop a dashboard with visualization of people reaction to the Government announcements on lockdown extension.

And to analyse how the citizens are dealing with pandemic ,their feelings , emotions etc.

2.Literature Survey

2.1 Existing Problem

As India is a population rich country. If Government takes any decision regarding any circumstances. In order to get sentiments from people the Government used to depend highly on News Channels, Social Media, etc.

The Limitation of News Channels is that they cannot cover whole country. Social Media can cover it but Social Media is such a huge it becomes very difficult to Government to track every people on Social Media.

If negative sentiments of the people increases beyond a certain limit then it becomes very difficult to handle the sentiments of huge population with respect to government decision.

2.2 Proposed solution

Inorder to solve above problem, We have build predictive system to understand the sentiments of people from Live Tweets. For building the model we have used sentiment140 dataset. It contains 1,600,000 tweets extracted using the Twitter API. The tweets have been annotated (0 = negative, 2 = neutral, 4 = positive)

According to the creators of the dataset:

"Our approach was unique because our training data was automatically created, as opposed to having humans manual annotate tweets. In our approach, we assume that any tweet with positive emoticons, like :), were positive, and tweets with negative emoticons, like :(, were negative. We used the Twitter Search API to collect these tweets by using keyword search"

So, with the help of visualization of these Live sentiments we get **prior notice** of increasing negative sentiments curve.

3.Experimental investigation

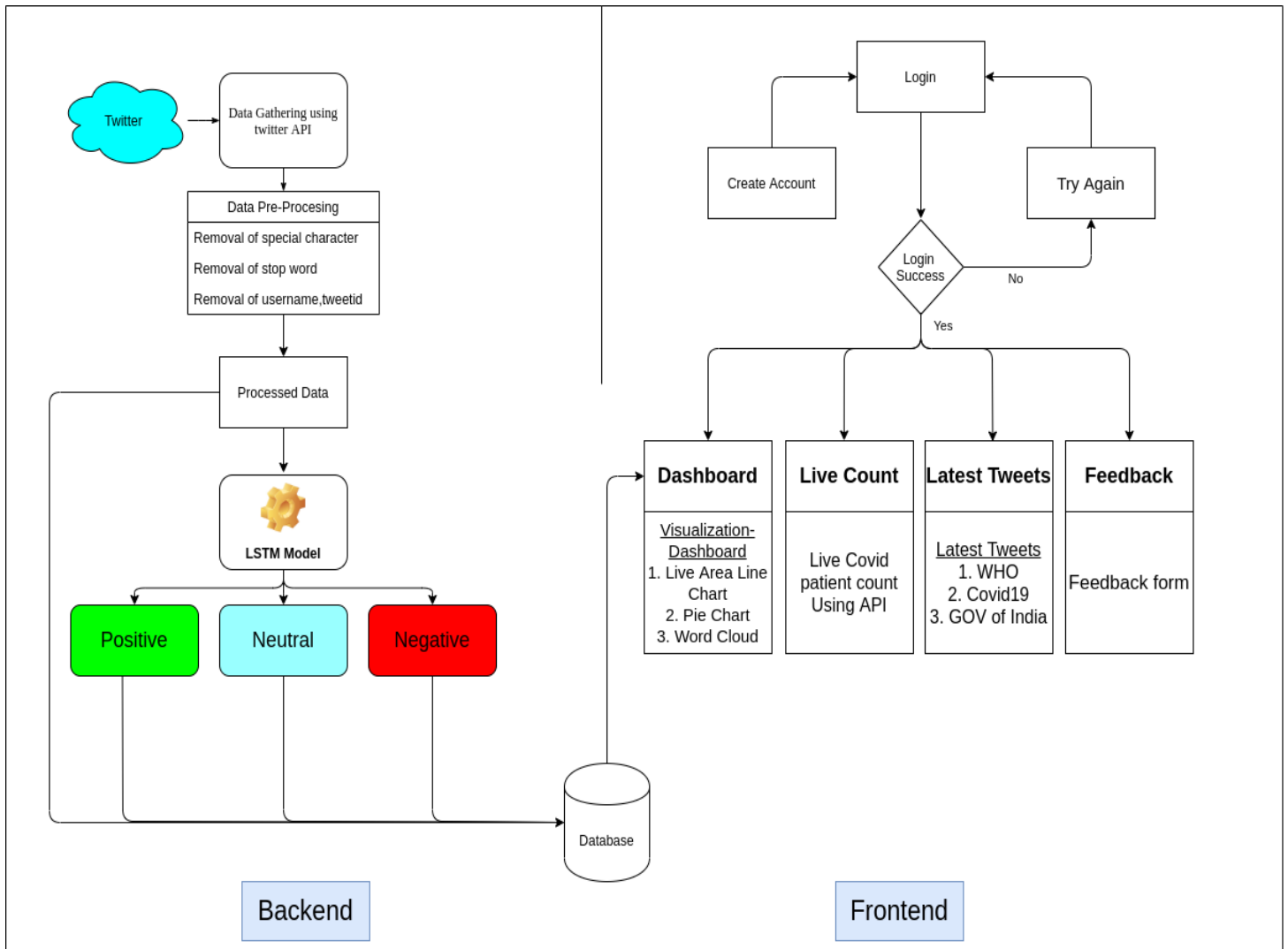
Initially, we started experimenting with probabilistic model that is by using MultinomialNB but it failed to give expected results because,

1. MultinomialNB assumes that your features (the attributes of the vectors we produced by using Bag of Words/Tf-Idf) are independent of one another.
2. The features generated by this representation is highly sparse and more biased towards the most common words.
3. It disregards the order/grammar of the text and thus loses the context in which a word is being used.

Then, we started experimenting with RNN (LSTM architecture), it gave pretty expected results because,

1. These networks read the data sequentially, while keeping a "memory" of what they read previously.
2. It is a discriminative model. It tries to figure out what the differences are between your class labels, in order to perform the classification.
3. The textual data is represented by using Word Embeddings, which made the feature representation dense.

4.Flowchart



5. Result

As a result, we have built a system that can predict sentiments from Live tweets across India. Visualization of these sentiments is shown through 3 Line Plots (Positive, Neutral, Negative).

We have also added Pie Chart, representing the count of Positive, Neutral and Negative for better understanding the sentiments.

In Addition, we have added Word Cloud, A Word Cloud is an image made of words. The size of a word shows how important/frequent word it is.

At end, we have also added Live Latest Tweets tab, which contains tweets from GovtofIndia, WHO, etc for quick knowledge of latest tweets from them.

6. Advantages and Disadvantages

● Advantages

1. The Model is Robust, in such way that it can also predict sentiments other than covid19 tweets.
2. Live Count of sentiments viz. Positive, Neutral, Negative plotted with three different Line Plots makes easy to understand
3. Pie Chart makes easy to understand the sentiments throughout the day
4. From Word Cloud, we easily get to know about the most Important/Frequent words
5. With the help Live Latest Tweets Tab, We directly get know about the latest tweets from Government, WHO, etc

● Disadvantages

1. The Model will only predict sentiments of English Language
2. The Model can predict sentiments in only three different categories i.e. Positive, Neutral and Negative

7.Application

The main Application of this project is to analyse Live behaviour of people as well as to understand the people reaction on Government decision in any type of pandemic situation.

As the Model is Robust, it can be used on any other application of Sentiment Prediction.

8.Conclusion

We can now address Live people's behavior on the lockdown if it is further extended with the help of Visualization Dashboard.

From this project, we can say that people's reactions vary day to day from posting their feelings on Twitter.

By analysing the Visualizations, outbreak of COVID-19 shows us how People, Government and Media Agencies broadcast the situations.

Thus, this Project, presented valuable informational about people's sentiment, which can be used to develop much needed Motivational Solutions and Strategies to Counter the rapid spread COVID-19.

9.Future scope

1. Currently, We have developed system which can analyse the Live tweets and predicts sentiment of people across the India through Twitter but this can be extended to other social media platform as well as across the Globe.
2. This system is not only based for Coronavirus health issue but it can also predict sentiment for the future similar cases as well as on any other Sentiment Prediction.
3. In addition, We can also add threshold value for negative sentiments if negative sentiments goes above this threshold value then we should get Alert Message.
4. We can also add 3 Pie Charts, representing the count of Positive, Negative as well as Neutral for better understanding the sentiments of Current day, Previous day as well as of past Seventh day.

10.Bibilography

Reference Links:

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7152888/>

<https://flask.palletsprojects.com/en/1.1.x/>

<https://ieee-dataport.org/open-access/coronavirus-covid-19-tweets-dataset>

Source code:

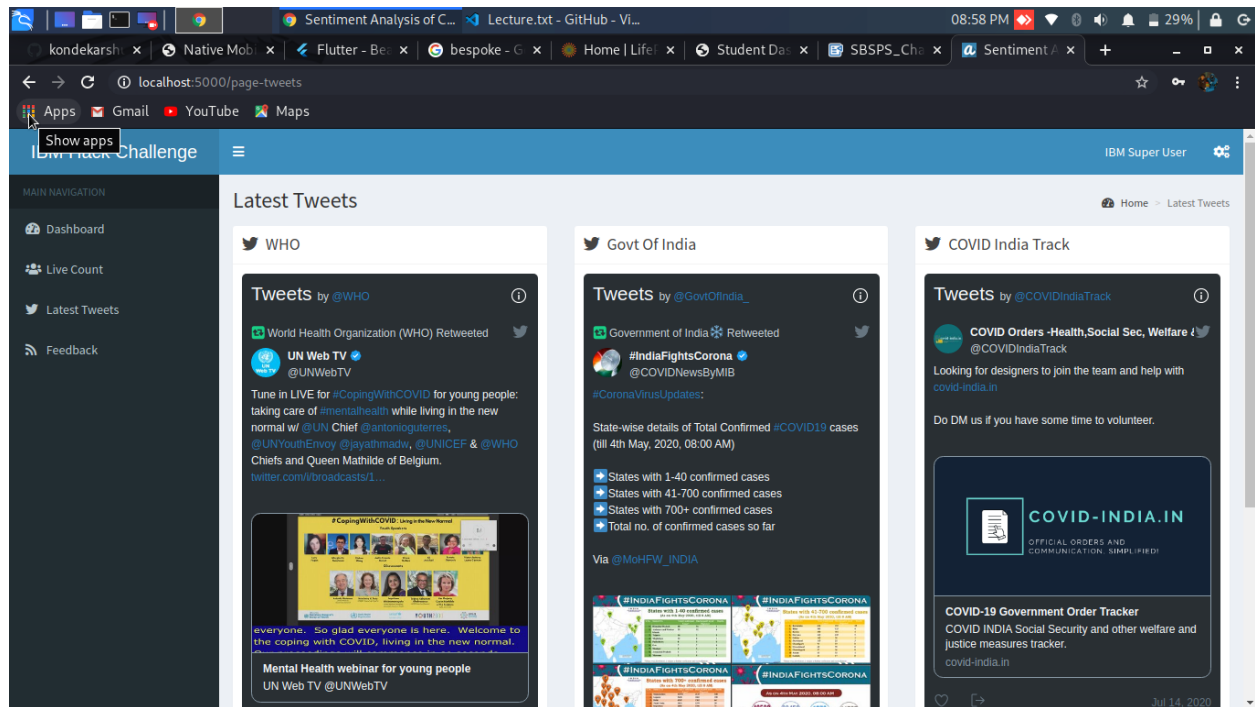
- <https://smartpracticeschool.github.io/SBSPS-Challenge-1399-Sentiment-Analysis-of-COVID-19-Tweets-Visualization-Dashboard/>

Video :



<https://youtu.be/p57Bs-Q2iMQ>

[illegible]



Team Profile:

1. Shubham Kondekar
2. Pranali Yangandul
3. Sarthak Phatate
4. Nilesh Chilka