IBM HACK CHALLENGE PROJECT REPORT 1.0
Team Name:TheMinions
Problem Statement:COVID-19 Tweet Visualisation Web Application

AHAN BOSE
PRITHWIJIT BANERJEE
VEDICA RAO

1.**Introduction**:

1.1 Overview

Sentiment Analysis (Opinion Mining) is an extremely powerful tool for predictive modelling and analysis. Sentiment Analysis relies on advanced text mining techniques and then performs Natural Language Processing on the texts to predict the latent sentiment or opinion of the text in a given context like analysing movie reviews, restaurant ratings or in this case government response against the backdrop of the coronavirus pandemic.

1.2 Purpose

The purpose of this project is to perform sentiment analysis on streamed tweets based on keyword search for visual and graphical analysis of public opinion on government measure. We also saved the incoming tweets in a database. Based on the analysed data, local authorities would be able to enact measures in accordance to the needs and desires of the public and enable better understanding of requirements and difficulties faced by the people during the pandemic thereby making better public policies.

2.**Literature Survey**
2.1. Existing Problems

India has a large and dynamic population. Therefore the authorities find it hard to reach out to the masses and often fail to truly understand their needs. Social media in general and twitter in particular connects millions of users provides a powerful means for the analysts to perform sentiment analysis on tweets and understand their response to government measures with regard to COVID 19 and lockdown in India.

More often than not people express themselves on twitter using images attached to the tweet and analysis of the image adds to the sentiment of the tweet and perception of the user on the subject.

Public perception can vary geographically and because of the pandemic it is essential to understand how the needs of the public are varying on a location basis in India.

2.2 Proposed Solution

2.2.1 Sentiment Visualisation Dashboard

- Getting access to twitter API:

Connect to twitter API by creating a developer account and requesting credentials for access.

- Setting up IBM cloud services:

Setting up an IBM cloud account in and creating Node Red Cloud Foundry App. Making all necessary Watson service connections with the Node Red example Watson Visual Recognition.

- Building Node Red Flow for line chart:

Using Node Red Flow Editor we design a visualisation dashboard. The dashboard connects to the twitter API using twitter stream node and streams tweets. Sentiment analysis is performed on any language of tweet using multi-lang sentiment node and subsequently displayed on the dashboard in terms of line charts.The incoming tweets are connected and stored in a cloudant database.

- Building Node Red Flow for time sentiment analysis:

In this flow we stream tweets, perform sentiment analysis and average the sentiment scores over 20, 100 and 1000 Tweets and continuously on a time series plot each line plot indicating a category of average tweets.

- Building Node Red Flow for Image Analysis:

Here, we use Watson Visual Recognition, to analyse images attached to tweets and display the results on the dashboard along with a gauge to display the sentiment score.

- Building Node Red Flow for Twitter Map:

We plot the tweets along with their sentiment on the world map.The tweets are localised mainly to India. A legend of the plotted tweets and sentiments is also displayed.

The inputs to this flow are taken based on keyword search from the user.

- Social Media Dashboard and Lockdown Analysis

The streamed tweets are saved in a database and then queried using ElephantSQL to return the pie chart representing sentiment ratio, top locations of tweets and average sentiment per day.

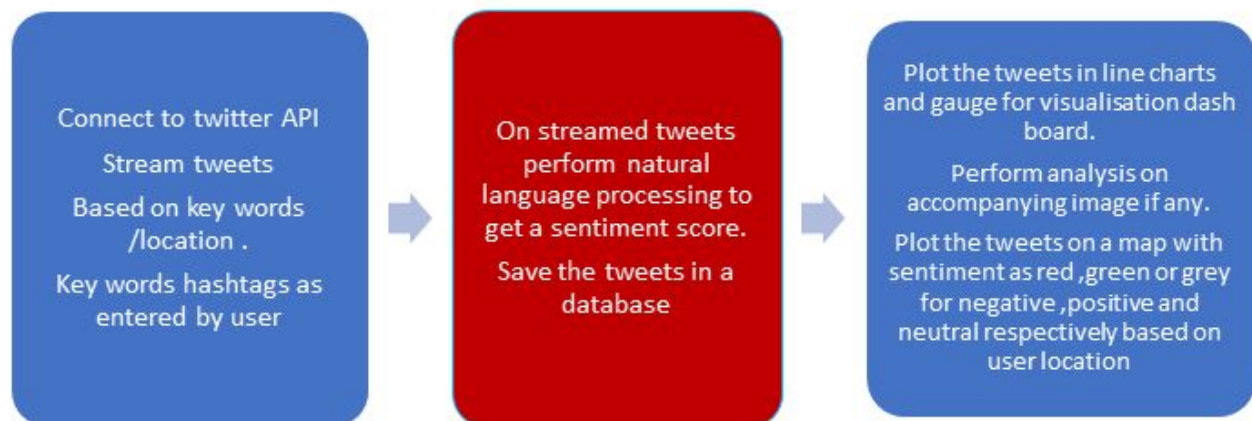## 2.2.2 User Interface Design

User interface designed using HTML and CSS.
CSS animations/Perspective used in the design process.The landing page has a title section, a menu section(sentiment visualization dashboard, news updates, statistics and chatbot), health guidelines and COVID-19 centers.

- Sentiment Visualization Dashboard - As explained in 2.2.1
- News Updates - Twitter handles of relevant and important organisations like the WHO, Ministry of Health and Family Welfare(MoHFW) , IRCTC , AAI etc. The latest tweets from these twitter handles designed with a masonry layout are automatically updated on the web page so that users can stay updated about the latest COVID-19 news.
- Statistics - Latest COVID-19 related data embedded on dynamic charts and maps for comprehensive analysis. Charts and Maps imported from Our World in Data.

- Chatbot - COVID-19 assistant created using IBM Watson Assistant.The chatbot is trained to answer important covid-19 queries like country wise covid-19 statistics, general CDC health guidelines and other relevant information.
- Health guidelines - relevant youtube videos from World Health Organisation embedded for users to view and follow.
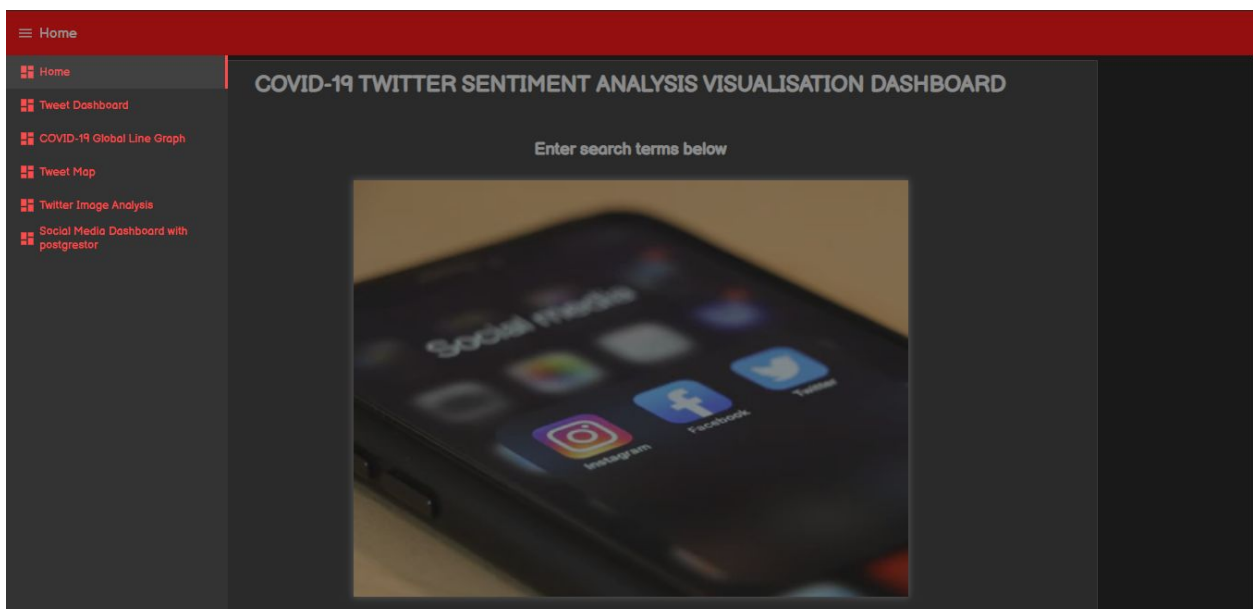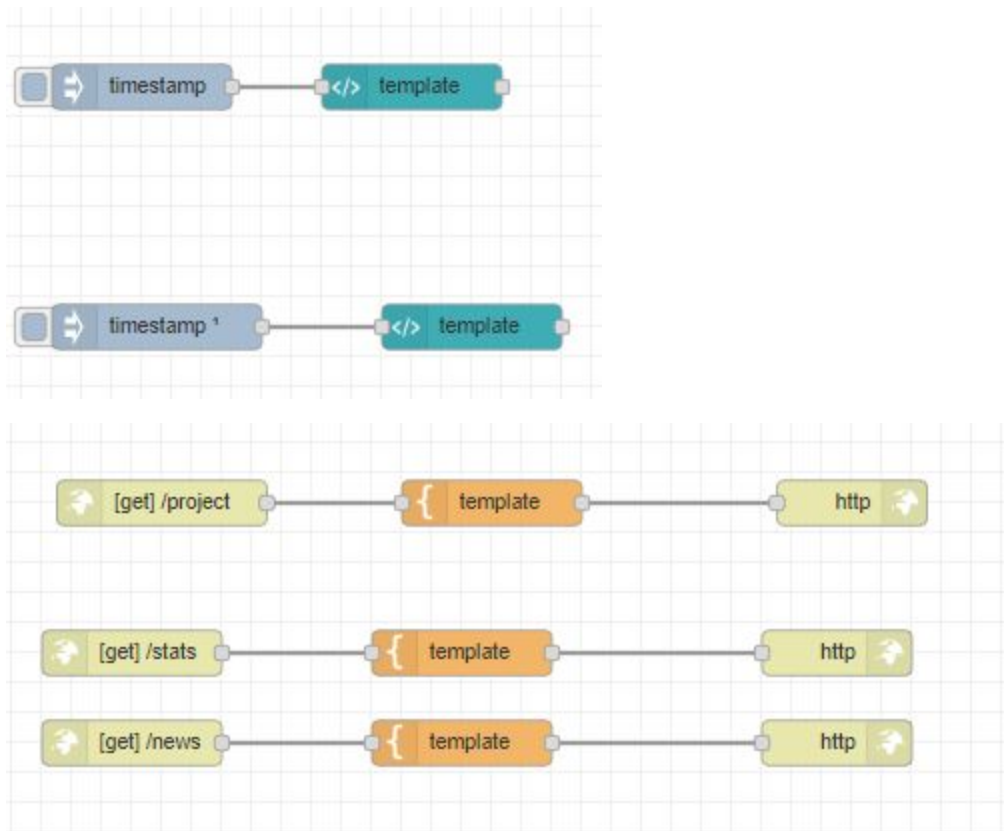- COVID-19 centres - data collected from Karnataka government and plotted using Google My Maps. The map shows all COVID-19 centres in Karnataka.

3. **Theoretical Analysis**

3.1 Block Diagram(Sentiment Visualization Dashboard)

Connect to twitter API

Stream tweets

Based on key words /location .

Key words hashtags as entered by user

On streamed tweets perform natural language processing to get a sentiment score.

Save the tweets in a database

Plot the tweets in line charts and gauge for visualisation dash board.

Perform analysis on accompanying image if any.

Plot the tweets on a map with sentiment as red ,green or grey for negative ,positive and neutral respectively based on user location
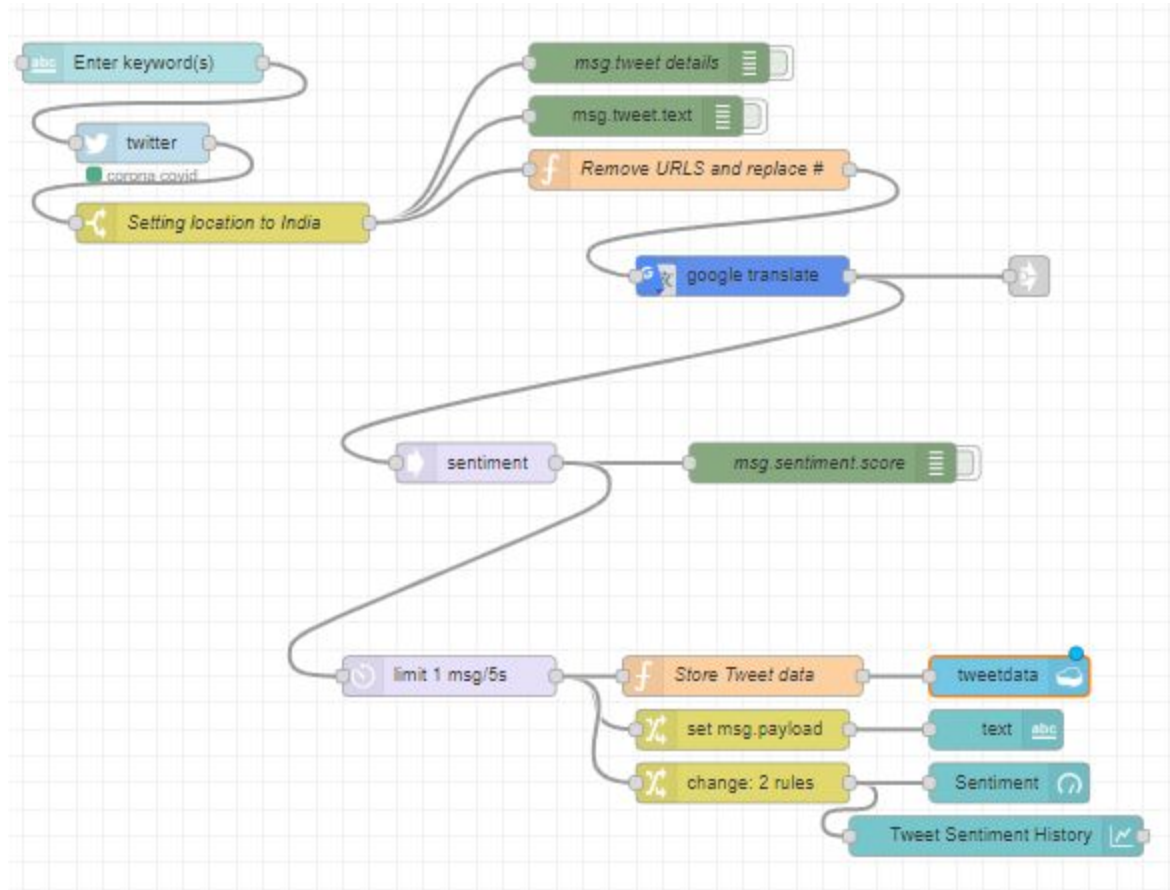
3.2 Software Designing
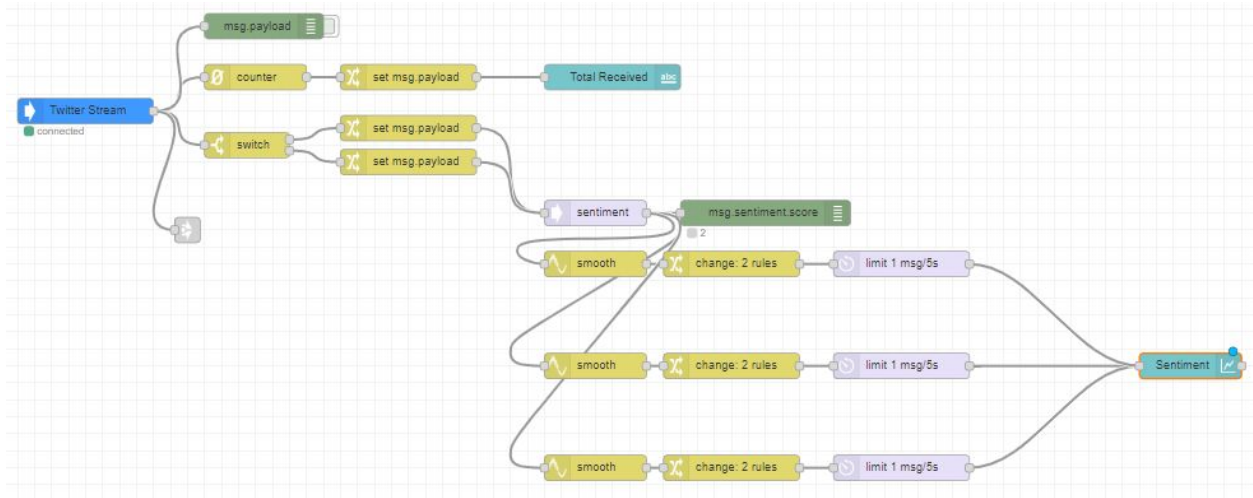
Flow 1 - Welcome/Home Page

Home page displays the title of the project and the user enters their keyword/hashtags for querying tweets for analysis.
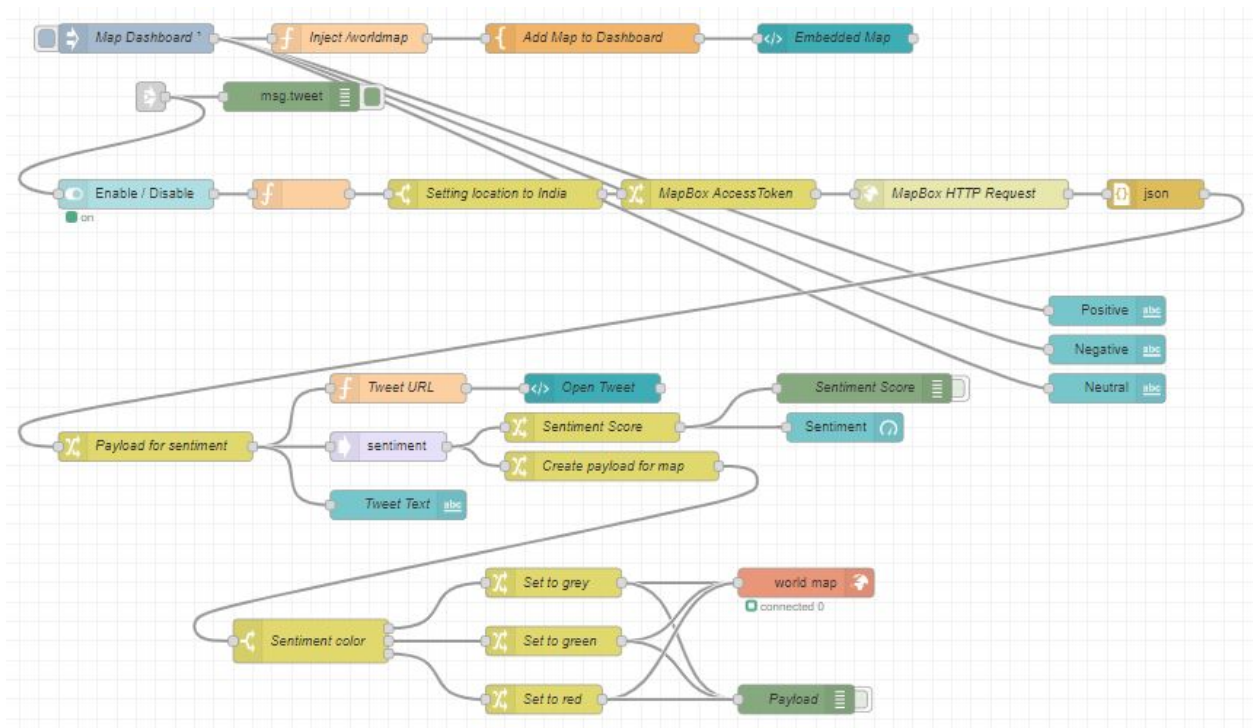
Flow 2 - Twitter Dashboard

Saving tweets in a cloudant database and displaying the sentiment of the same on line charts and gauge for visualisation.
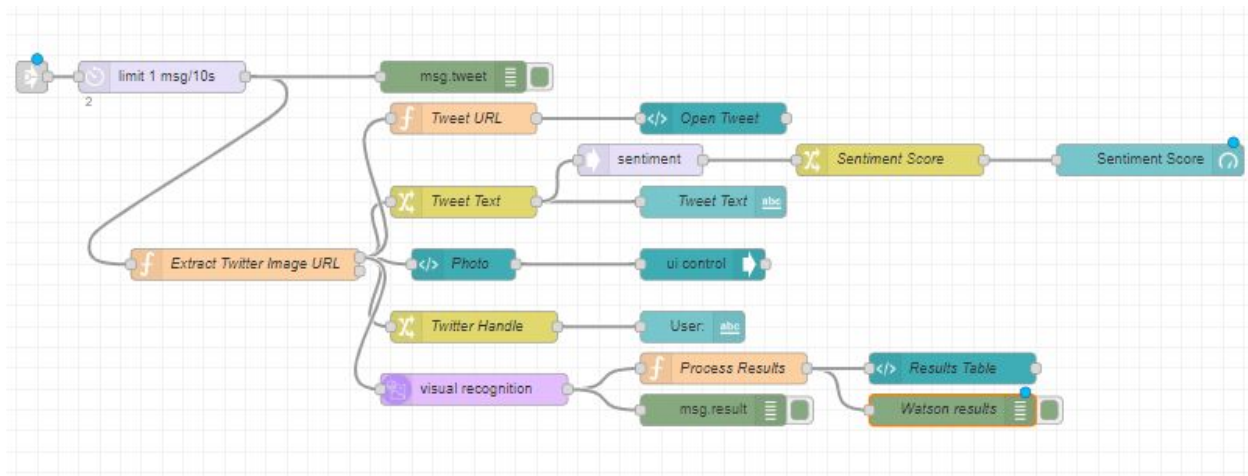
Flow 3 - COVID-19 Line graph

Line Chart of averaged tweets over 20,100 and 1000 tweets category displayed.

Flow 4 - Map Tweets

This flow accepts keywords/hashtags from the users and performs sentiment analysis. The location of the tweets are plotted on a world map using node red contrib world map node and Mapbox API.

Flow 5 - Twitter Image Analysis

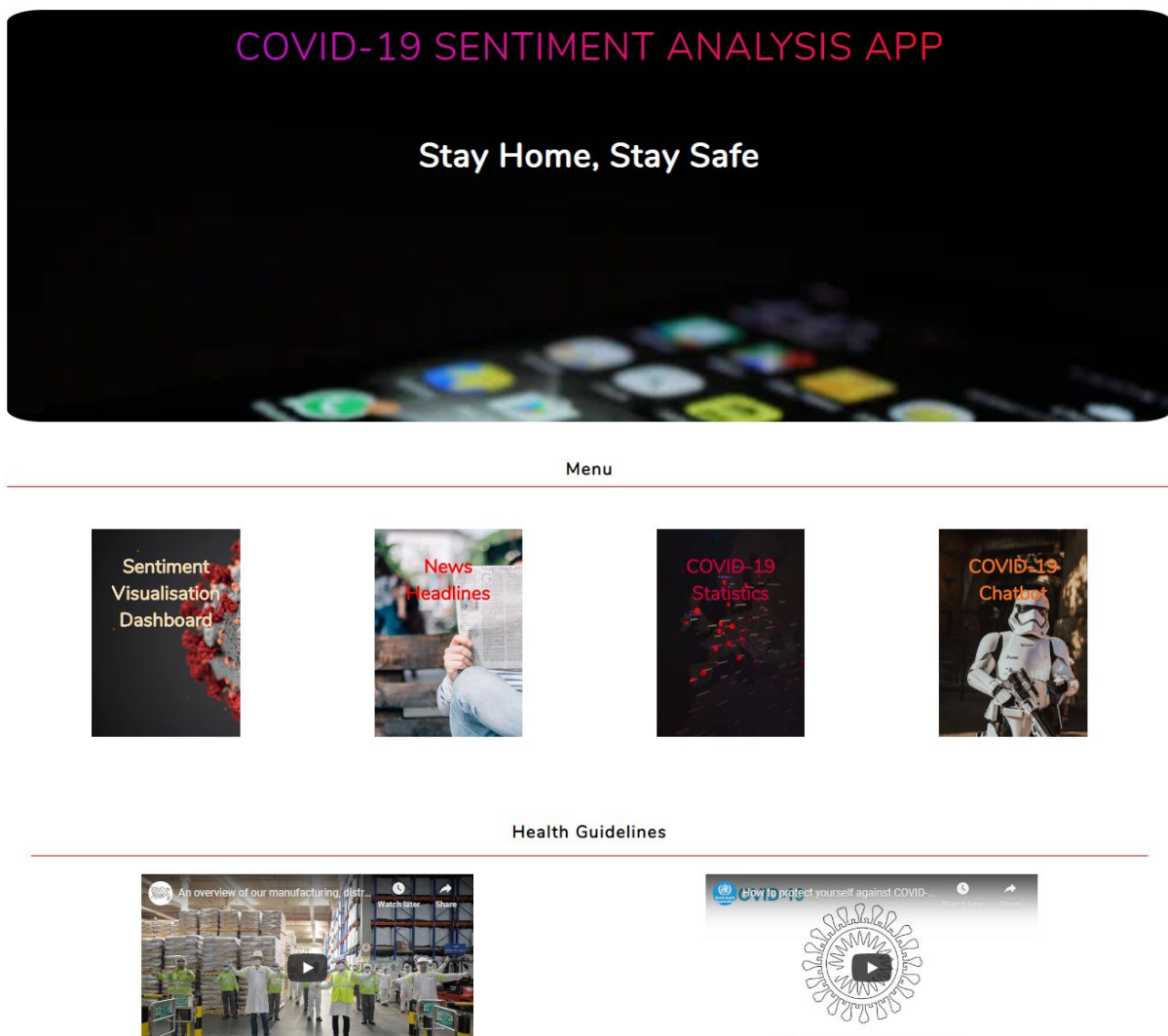User based input, Watson Visual Recognition can analyse the images attached to the tweets.

Flow 6 - Social Media Dashboard and Lockdown Analysis
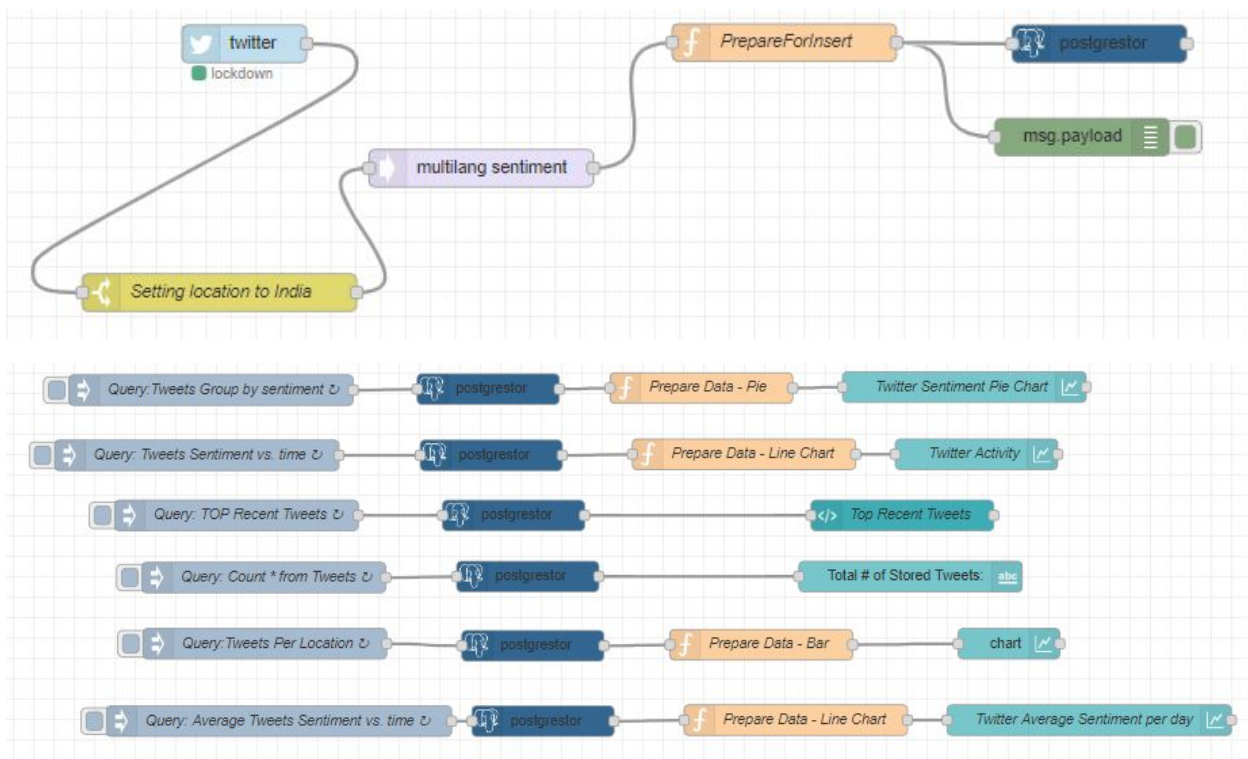
Streamed tweets are saved in postgresql and queried using ElephantSQL. Top tweets location wise, pie chart of sentiment ratio and average sentiment per day are displayed on the dashboard.
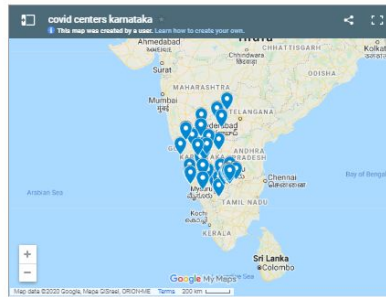
3.3. User Interface Design
A.Home Page

COVID-19 Centres

B. Sentiment Visualisation Dashboard

Shown in section 5(Result)

C. News Headlines

D. COVID-19 Statistics

Cumulative confirmed COVID-19 cases, Jul 10, 2020
The number of confirmed cases is lower than the number of total cases. The main reason for this is limited testing.

Source: European CDC – Situation Update Worldwide – Last updated 10 July, 11:22 (London time)

OurWorldInData.org/coronavirus • CC BY

Dec 31, 2019   Jul 10, 2020

CHART   MAP   TABLE   SOURCES   DOWNLOAD

Stay-at-home requirements during the COVID-19 pandemic, Jul 10, 2020

E. COVID-19 Chatbot

Build your own assistant using
IBM Watson Assistant

Assistant preview

Hello, I'm the COVID Crisis Communication Bot ready to answer your questions about COVID-19. How can I help you?

Should I wear a mask?

It is recommended that people who are well wear a facemask to protect themselves from respiratory illnesses, including COVID-19. A facemask should be used by people who have COVID-19 and are showing symptoms. This is to protect others from the risk of getting infected. The use of facemasks also is crucial for health workers and other people who are taking care of someone infected with COVID-19 in close settings (at home or in a health care facility).

Type something...

→

4. **Flowchart**
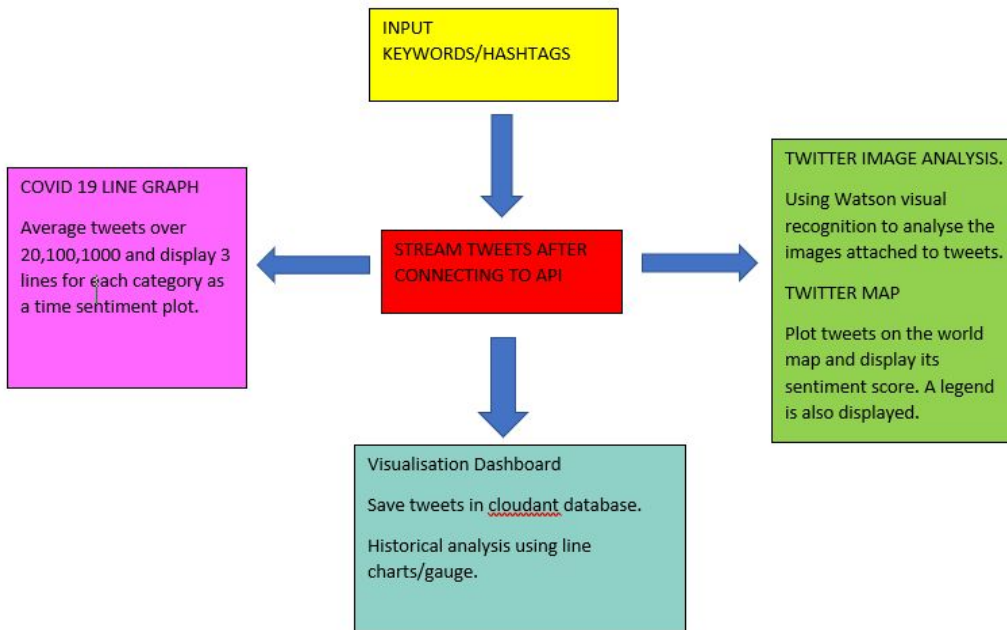
5.Result
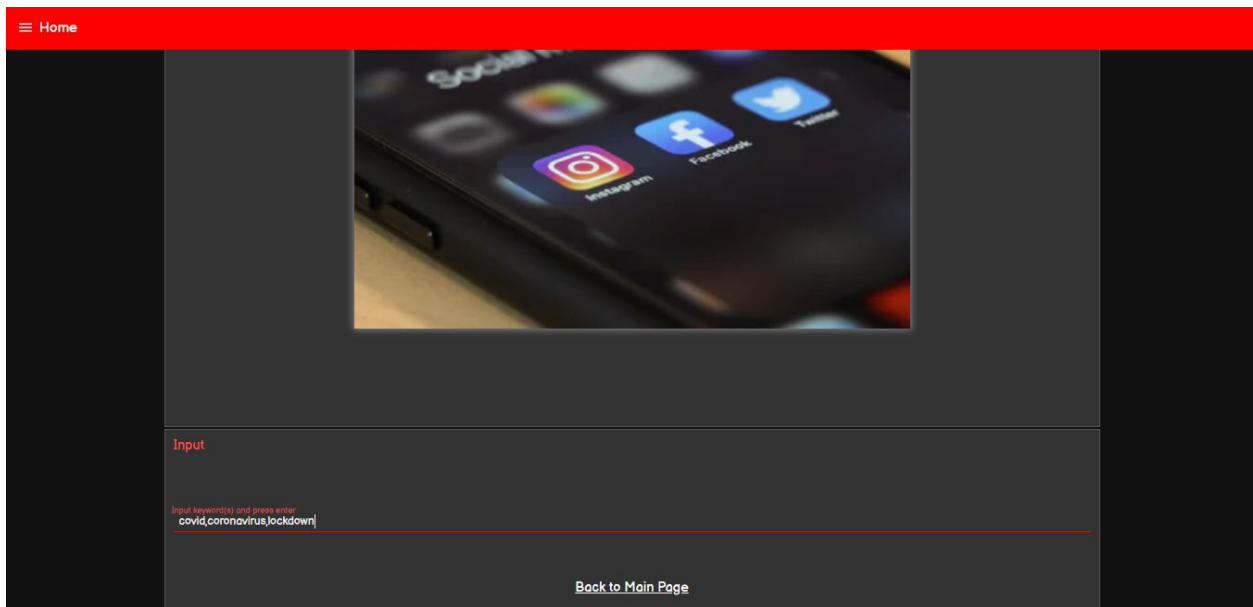
HOME TAB
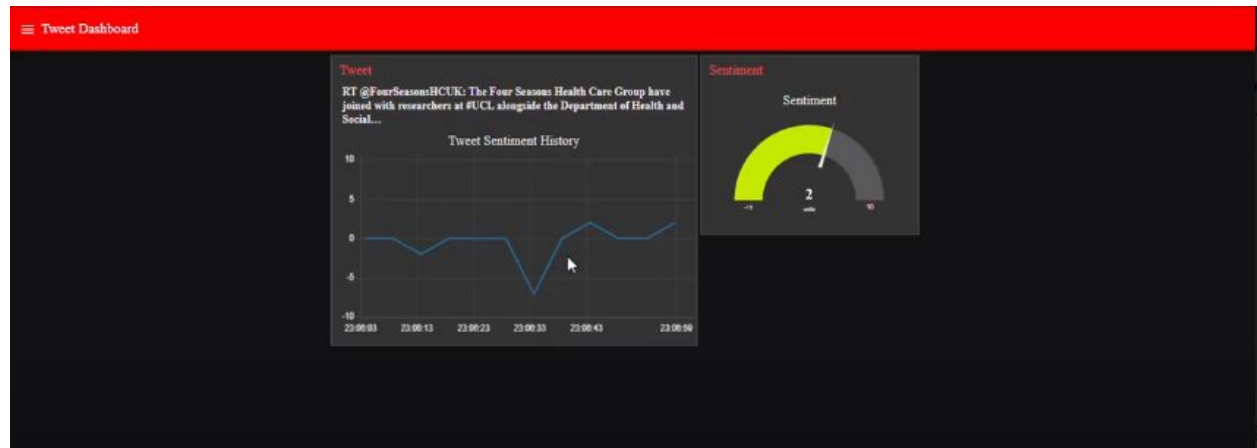
TWEET DASHBOARD
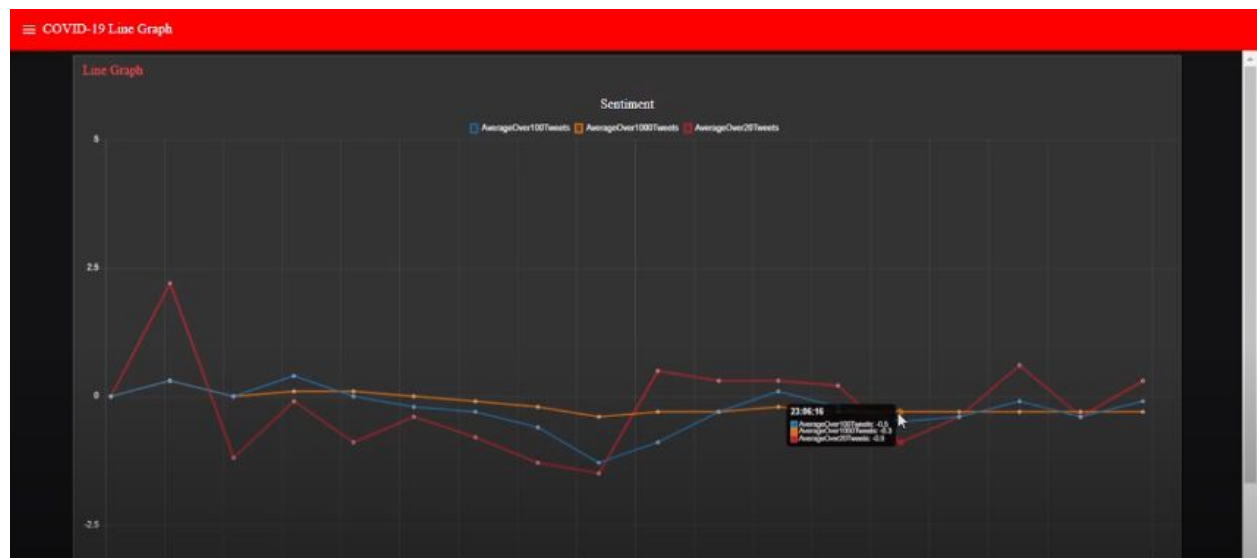
COVID-19 GLOBAL LINE GRAPH   (averages tweets over 20,100 and 1000)

Global Tweets

Enable / Disable

RT @SanjaySDutt: Drawing inspiration from Soniaji &
@RahulGandhi Ji, @nsui led by I/c @guptar & Pres.
@Neerajkundan hve initiated #SpeakUpF…

SandeepKhasa_

Open Tweet

Sentiment

2
units

-10   10

Global Tweet Map

Legend

Positive

Negative

Neutral

Leaflet | Tiles © Esri

Global Tweets

Enable / Disable

RT @XpressHyderabad: COVID numbers in districts other than
AihikS GHMC, RR and Medchal have been increasing. Watch this video
to know the present…

Open Tweet

Sentiment

0
units

-10   10

Global Tweet Map

1281572108602978304

place : Bengaluru, India
sender : oorvsss
text : Meanwhile Covid-19 - #gocoronago
https://t.co/UwUFhkqNF2
sentiment : 0
lat, lon : 12.96991, 77.59796

Lat: 7.8271 Lon: 75.2490

Legend

Positive

Negative

Neutral

Leaflet | Tiles © Esri

TWEET MAP

TWITTER IMAGE ANALYSIS

SOCIAL MEDIA DASHBOARD AND LOCKDOWN ANALYSIS

6.Advantages and Disadvantages

The advantages of our model are that it provides a visualisation with respect to location,time and also provides an analysis of any images in the tweets in a single dashboard.This allows for a comprehensive review of government measures and for enacting future measures in accordance to the needs of the public.

Due to the wide economic gap in India, many labourers particularly in the unorganised sector, do not have access to internet and their demands and difficulties during this pandemic would remain unaccounted in our model.

Also because our model is reliant on the twitter API, it is also restrained by the limitations that come with its usage.

7.Application

Our model can be used by data analysts to predict the general sentiment of people towards any government action.

Another application of our model(since our project allows for the user to input any keywords or hashtags)is to gauge and analyse the sentiment and its variation on their topic of choice.

8.Conclusion

The COVID-19 Twitter Sentiment Analysis Visualisation Dashboard is made using Node Red Flow editor.
We have displayed line charts for time analysis, gauge for sentiment score, visual recognition of images attached to tweets and plotting streamed tweets on a map with sentiment score and accompanying legend.
We have also trained a COVID-19 chatbot, displayed twitter news updates, statistics, and COVID-19 centers in Karnataka.

9.Future Scope

Future Scope of our project that is, improvements that can be made, are being able to display a variation of sentiment over a period of months-since from the

beginning of the pandemic till date.Twitter streaming limitations if overcome, could provide a visualisation with a much longer timeline.

Seeing the variation over a timeline of months analysed along with newscasts will give a better analysis.

10.Bibliography

1. **https://nodered.org/docs/getting-started/ibmcloud**
2. **https://cognitiveclass.ai/courses/node-red-basics-to-bots**
3. **https://flows.nodered.org/node/node-red-node-watson**
4. **https://flows.nodered.org/node/node-red-contrib-maps**
5. **https://dev.to/poojamakes/why-you-should-be-using-node-red-right-now-1n1l**
6. **https://ourworldindata.org/coronavirus**
7. **https://unsplash.com/**
8. **https://www.google.com/maps/d/u/0/**