**Problem statement:**

**Title: Sentiment Analysis of COVID-19 Tweets – Visualization Dashboard**

**Description:** The sentiment analysis of Indians after the extension of lockdown announcements to be analyzed with the relevant #tags on twitter and build a predictive analytics model to understand the behavior of people if the lockdown is further extended.

Also develop a dashboard with visualization of people reaction to the govt announcements on lockdown extension

**Expected Solutions:** *Develop a twitter sentiment analysis model to understand the following*

*1. Get to know people’s sentiment towards the epidemic*

*2. Understand the sentiments of people on govt. decision to extend the lockdown*

**Technologies & Tools**

Python 2 or 3, IBM Watson Studio, IBM Cloud for Deployment, Any Web frameworks

**Reference**

1.https://www.kaggle.com/smid80/coronavirus-covid19-tweets

2. https://ieee-dataport.org/open-access/corona-virus-covid-19-tweets-dataset

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

Solution Description:

**Work Flow:**

**Method 1:**

1. From the tweets available in the dataset, we will pick the keywords and classify them as positive or negative words . this will be done by an ML classification algorithm, as the classification will be divided into two categories we use random forest classifier or could go with naive bayes algorithm
2. Further it would be cateqorize  based upon how they affected those people who tweeted.

**Method 2:**

1. We'll be using natural language processing to extract words by extracting data from our own twitter developer account and classify them by using our own modified algorithm by simply changing either maths used behind or changing their output by other changes which we could analyse while studying regarding previous implemented algorithm's research papers. And it would be added advantage for us if we are provided with mentor for this new algorithm development task.
2. Rather than visualizing final outcome on jupyter notebook we will visualize that on our newly made website which would be either made using django framework or using react(will depend on time left while modifying the algorithm)

**Timeline:**

|  |  |
| --- | --- |
| Work to be done | Tentative time required |
| Research Paper Analysis for getting to know what sort of previous algorithms has been implemented regarding twitter sentiment analysis and try to modify it for our convenience and more optimised solution | 65% and would be in parallel to other work |
| Dataset preprocessing | 3% |
| Model selection | 3% |
| Model optimization | 3% |
| Dashboard creation | 10% |
| Content on dashboard(  psychological analysis  - economic analysis as per people thinking  - health analysis  - lifestyle analysis  -happiness analysis  ) | 3% |
| Integrating with backend and in turn with ibm cloud | 10% |
| Other Tasks | 3% |

#### Novelty/ Uniqueness:

#### 1st uniqueness:

If we design an algorithm it would be unique in itself.

#### 2nd uniqueness:

We would be analyzing all sort of

- psychological analysis

- economic analysis as per people thinking

- health analysis

- lifestyle analysis

-happiness analysis

Business/Social Impact:

Social Impact:

1. As its outcome would be  people's mental health analysis which would in turn help government to take further measures regarding lockdown

2. As Mental health would be known so it will help to to prepare people for post lockdown situations

Business Impact:

1. It will let people know potential of local brands

2. After knowing mindset of local people business could flourish accordingly post lockdown and government could plan economic strategies accordingly.

Technology Stack:

Python  3,

IBM Watson Studio,

IBM Cloud for Deployment,

React or Django (depending on time left after improving the algorithm to be used in this approach.)

Scope of work

1. The findings of the study might be helpful in identifying the knowledge and attitudes of Covid-19 among all age groups
2. The findings of the study might be helpful to future researchers.
3. The knowledge attitudes of youths about covid-19 will be assessed. The prevention measures followed by the people will be assessed.
4. Best part of this would be a new algorithm would be devised.