Different approaches to tackle COVID-19

I. Problem Statement

Evaluate the effect of a change in the country's approach and policies on the spread and containment of the virus and attempt to establish a realistic timeline for the longevity of the virus in the region while also considering factors like testing rate and availability of resources.

II. Description of Data Set

In response to the COVID-19 pandemic, the White House and a coalition of leading research groups have prepared the COVID-19 Open Research Dataset (CORD-19). CORD-19 is a resource of over 44,000 scholarly articles, including over 29,000 with full text, about COVID-19, SARS-CoV-2, and related coronaviruses. This freely available dataset is provided to the global research community to apply recent advances in natural language processing and other AI techniques to generate new insights in support of the ongoing fight against this infectious disease. There is a growing urgency for these approaches because of the rapid acceleration in new coronavirus literature, making it difficult for the medical research community to keep up. This data set will be used to extract keywords related to COVID - 19.

Also, we will be using the data repository for the 2019 Novel Coronavirus Visual Dashboard operated by the Johns Hopkins University Center for Systems Science and Engineering (JHU CSSE) to monitor and access the real-time data on a count of the number of deaths, recovered cases, country/region.

III. Implementation Plan

PART 1-- Using NLP, try to extract keywords from the journal and news articles present in the first data set.
PART 2-- Using dataset 2, build a prediction model for the number of COVID-19 cases in every region.
PART 3-- Link the outputs of the two parts and establish a relationship between country policies and the data curve observed for that region.

IV. Team Members and Task Allocation

1) Natural Language Processing for extraction of keywords -- **Yuvraj Shivtare**
2) Different Machine Learning Algorithms and their analysis and accuracy -- **Meetkumar Patel and Raghav Daga**

V. References

1. https://github.com/CSSEGISandData/COVID-19
2. https://www.kaggle.com/allen-institute-for-ai/CORD-19-research-challenge