## **COVID-19 Tweets and understanding popular sentiments**

Coronavirus disease (COVID-19) , an infectious disease caused by the novel virus, has resulted in a global pandemic. There has been a global lockdown to reduce the spread of viruses. With quarantines and isolations people are feeling a heightened level of anxiety is something that we hear commonly on the news. In this study we aim to study the tweets of people during this pandemic and understand the public sentiments. More specifically:

1. Understand how people feel towards this quarantine
2. Whether there are shifts in sentiments? If so, what stressors/ factors contribute to this shift?
3. What are some most talked about things/activities during this time?

Dataset: <https://www.kaggle.com/lopezbec/covid19-tweets-dataset>

## **Fake news detection**

The recent growth of fake news on social media has a great impact on society. There is a lot of information from disparate sources among various users around the world and developing a technique to detect fake news has become a necessity in this era of social media. The aim of this project is to develop a model that can detect fake news. In order to achieve this Twitter data has been used from Kaggle. The dataset consists of two categories of labelled data True News and Fake News. Essentially this will be a supervised classification task.

Dataset:<https://www.kaggle.com/clmentbisaillon/fake-and-real-news-dataset>

## **Movie recommendation system and predicting the success of a movie**

This dataset consists of a large movie data combined with ratings. Analysing different features of a movie e.g cast, genre, language and ratings we can suggest users movies that they might like given they have a preference for a given movie. We could also combine these factors and the budget of the movie to see if we can predict which movie can be successful.

Dataset: <https://www.kaggle.com/rounakbanik/the-movies-dataset?select=ratings.csv>

## **3.**