State-wise timeline prediction if Covid 3rd wave occurs in India

CS460-Project Proposal

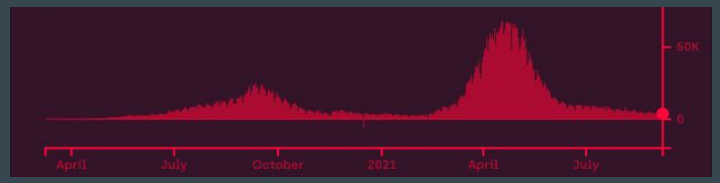
Nelson Kshetrimayum

Muhammed Jabir T

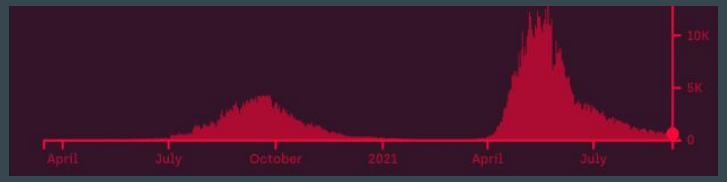
Objective

- To predict when a state will start seeing the affect. Take for example Maharashtra and Odisha:

  Maharashtra:



  Odisha:



- To get a forecast of the infections

Dataset

Covid19india.org has made the data publicly available.

- Testing data for each state
- Cases for each state
- Vaccination data

https://data.covid19india.org/

Approach Plan

- To study/find the correlation (spreading rate) to a neighbouring (nearby) state from a heavy infected (hotspot) state.
- Find dataset of other factors we can consider (such as population density).
- Start by implementing multiple linear regression. Use gradient boosting predictor, train using LightGBM. (This is not finalized)
- Autoregressive integrated moving average (ARIMA) model: statistical analysis model that uses time series data to either better understand the data set or to predict future trends.

Work division

Midway plan:

- Sorting/rearranging of the data to fit our use, making sure of proper labeling and merging the datasets that are going to be used. - Nelson
- Finding out the correlation between the datas (Vaccination efficiency, infectiousness of virus variant, etc) - Jabir
- Finding out the optimal method to implement on the problem and making our model. - Both

Later:

- To automate the input of daily data and improve our model.

References

- [Kaggle] Shreyas P J, 2021, *Covid forecasting using DL and statistical models.* - Link
- [International Health] Sherry M, Ashok K. P., Md. Arshad & Ubydul H, 2021, *Short-term forecasting of the Covid-19 outbreak in India.* - Link
- [Youtube] Bharani Akella, Great Learning, *Predicting COVID-19 With Machine Learning.* - Link
- [Nature] Yazeed Zoabi, Shira Deri-Rozov & Noam Shomron, 2021, *Machine learning-based prediction of COVID-19 diagnosis based on symptoms.* - Link

Thank you