Master of Technology
Department of Computer Science and Engineering
Indian Institute of Technology, Kanpur

COVID-19 Analysis and Prediction

Group No. : 27

Mani Kant Kumar (20111030) mani20@iitk.ac.in

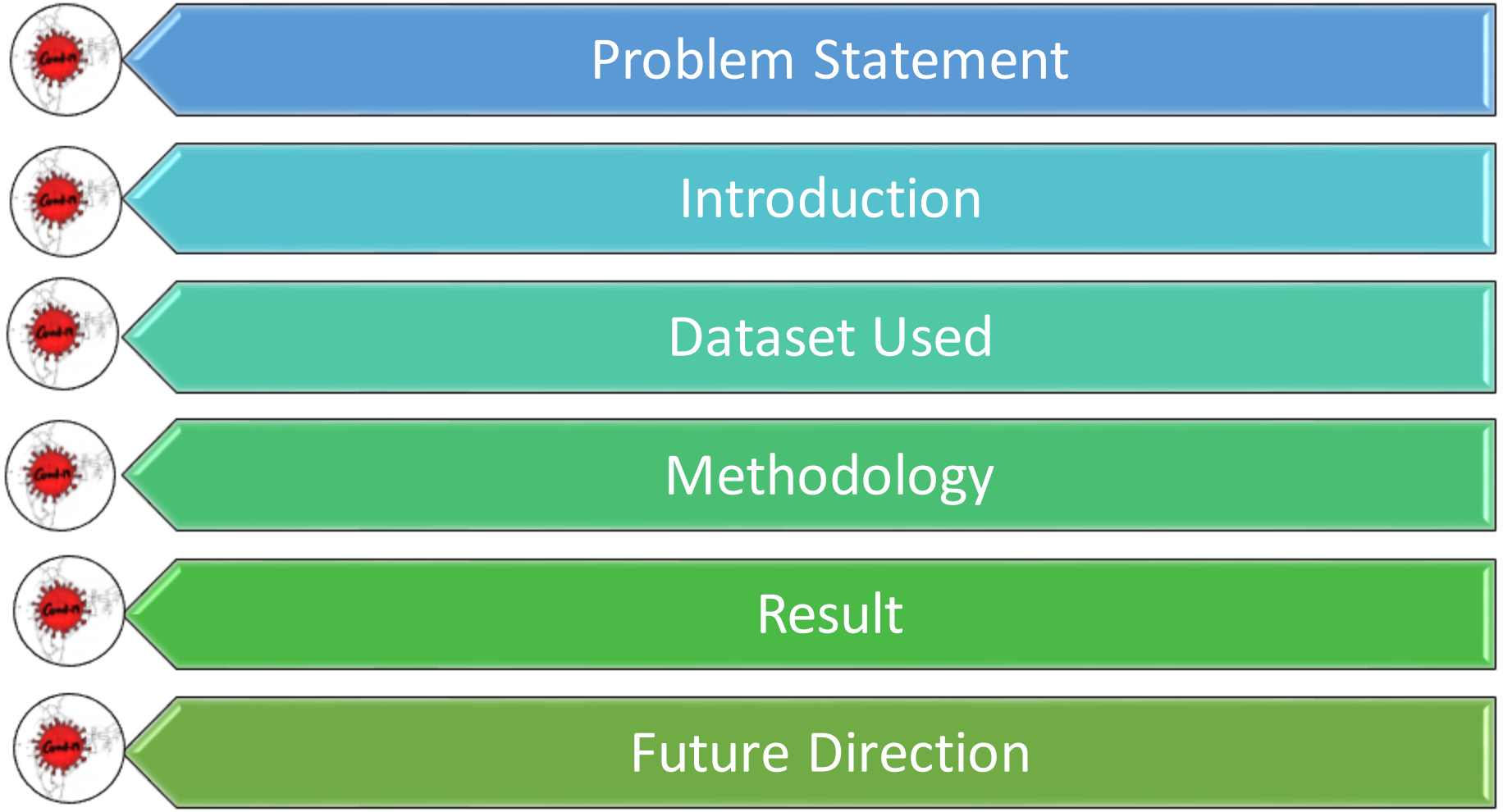
Rohit Raj (20111051) rohitjha20@iitk.ac.in

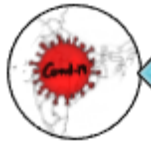
Shruti Wasnik (20111062) gdshruti20@iitk.ac.in

Sonam Tshering (20111064) somsring20@iitk.ac.in

Project Instructor:
Associate Prof. Arnab Bhattacharya

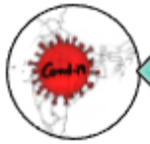
Content


Problem Statement
Introduction
Dataset Used
Methodology
Result
Future Direction



Problem Statement

- **Aim**: To process covid-19 cases for analysis and future prediction.
- Analyze covid-19 cases in India
  - Range of date of cases
  - Searching cases with filter options
  - Active cases
  - Recovered cases
  - Deceased cases
  - Hot-spots and cold-spots
  - Mortality rate and its factor
  - Future Prediction

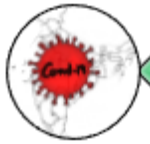


Benefits of Our Analysis

- The analysis could be used for finding the hot-spots, so that more attention can be provided to those regions
- The mortality rate by age data can be used to analyze the age group which are more prone to the spread. Special medical facilities and preventive measures can be provided to them.
- The predicted cases can be used by the hospitals to arrange extra vacant beds and medical equipment in advance, such that treatment can be provided to everyone.

Dataset Used

- coivd19india.org (Main Source)
- Installation of Libraries: seaborn, plotly, sklearn, matplotlib
- Data are available in different forms like:
  - District-wise, daily, time series, state wise and raw data.
  - district_wise.csv
  - Day-wise data from data-all.json.
  - Patient-wise data from raw_data.json files.
  - states_india.geojson

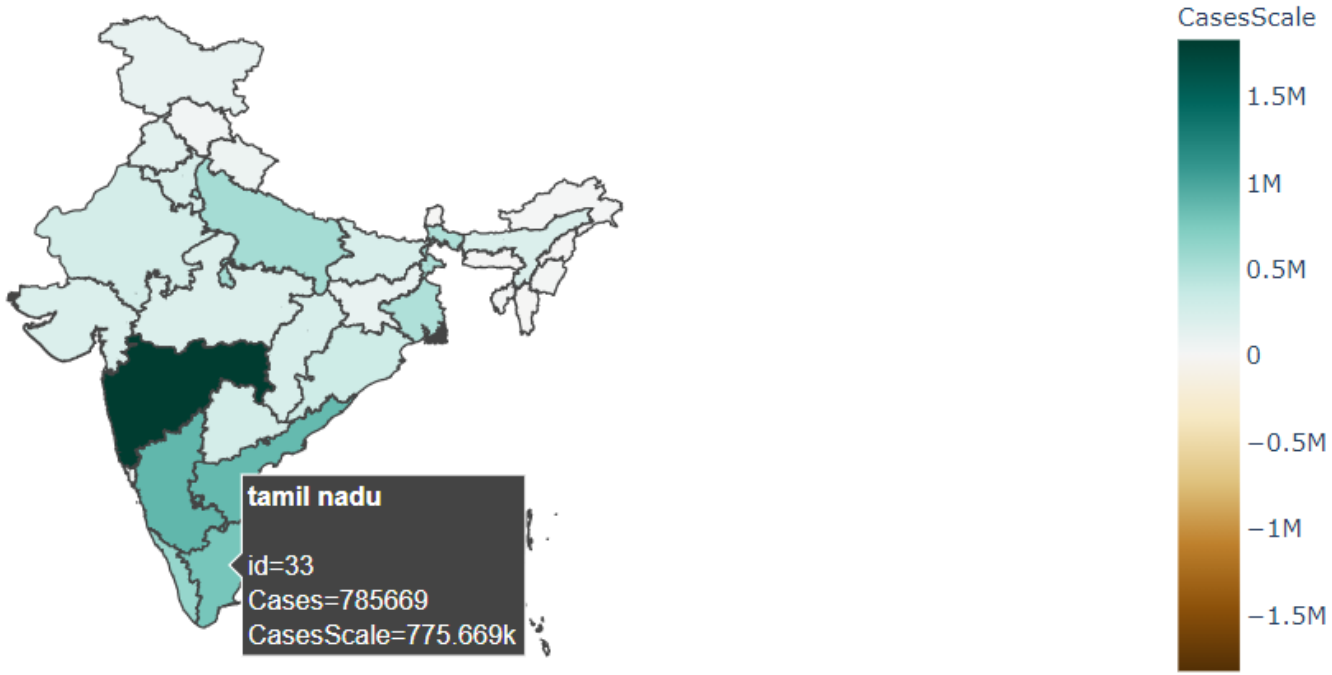


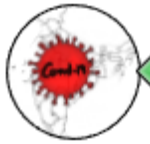
Methodology

- **Pre-processing**
  - Data cleaning, handling missing data and data integration.
- **Visualization**
  - Presenting our analysis and prediction of data by means of graphs, charts and maps.
  - Python ipywidgets to get user input.
  - Matplotlib.pyplot, seaborn and plotly to plot analyzed data
- **Prediction**
  - Tested our ideology on models
    - Linear Regression ,and Polynomial Regression.
    - Our prediction model is based on Linear Model (Best Accuracy)

Results

- Current trend of COVID-19 cases in Indian States





Results cont…

Hotspots

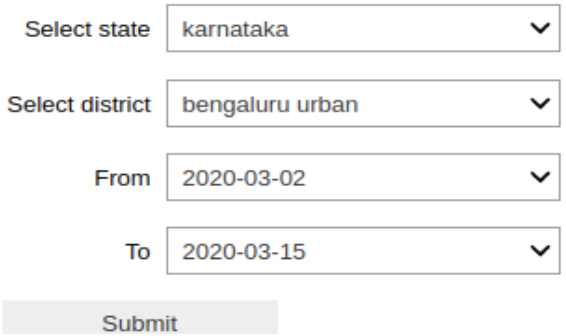
state	cases	deceased	recovered
maharashtra	1834519	47257	1708862
karnataka	889857	11815	853279
andhra pradesh	869578	6989	857002
tamil nadu	785669	11738	763834
kerala	625299	2355	561532

Coldspots

state	cases	deceased	recovered
dadra and nagar haveli and daman and diu	3334	2	3289
mizoram	3912	6	3690
andaman and nicobar islands	4742	61	4611
sikkim	5147	111	4587
ladakh	8712	120	7734

Results cont...

- Analysis of covid-19 cases with filter options.
  - From 2nd March till date
  - Particular range of date
    - All state
    - Particular state
    - All district in a state
    - Particular district

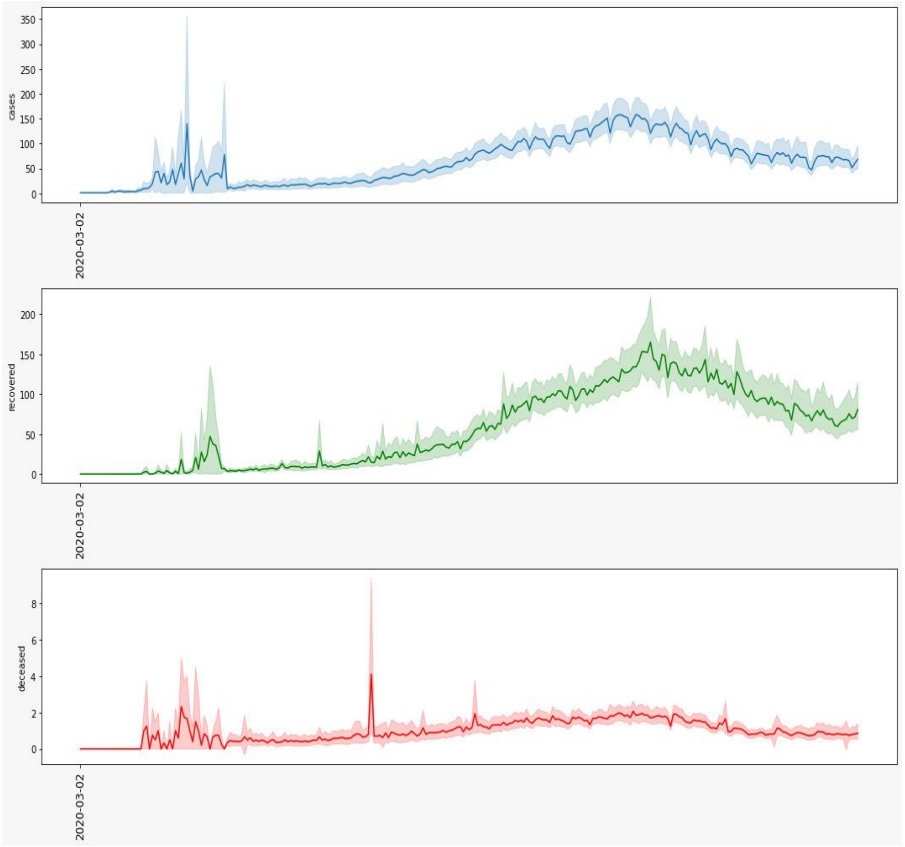


Sorry! no cases were found for the given details. Please try with different entry.

Total Cases : 9585717

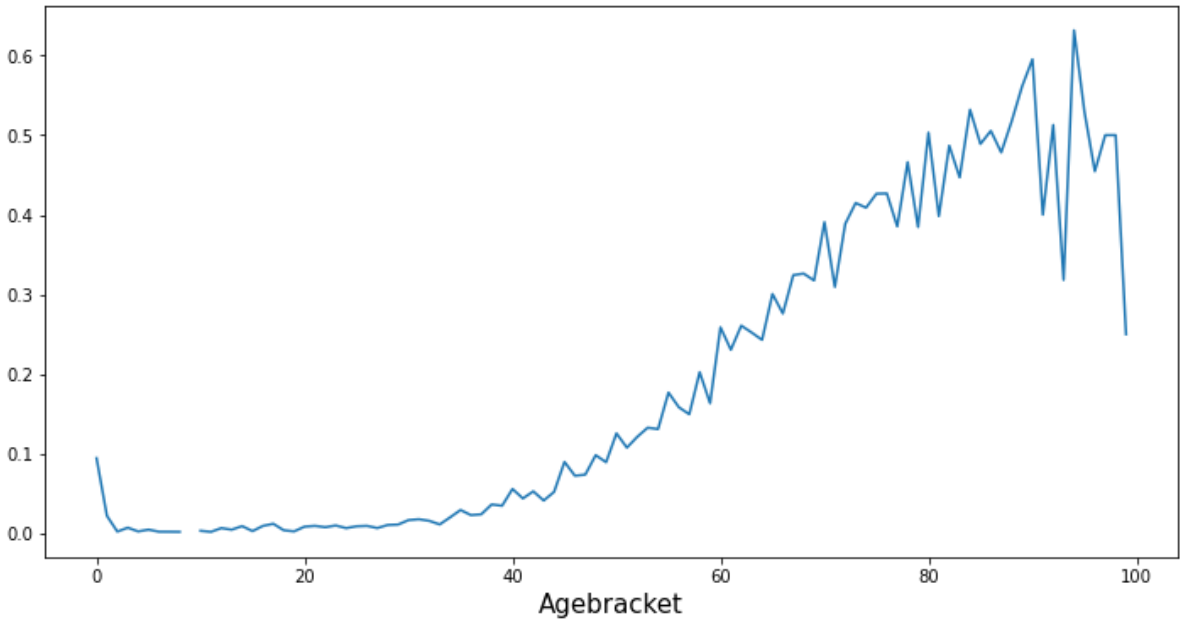
Recovered : 9053787

Deceased : 138944

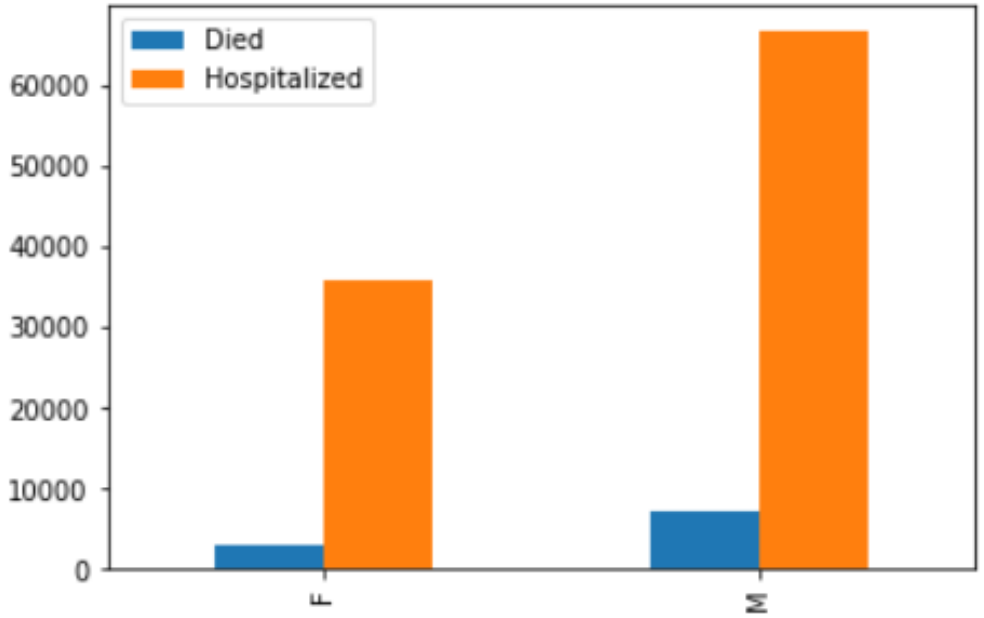


Results Cont...

- Age Bracket prone to death



- Gender wise Death vs Hospitalised

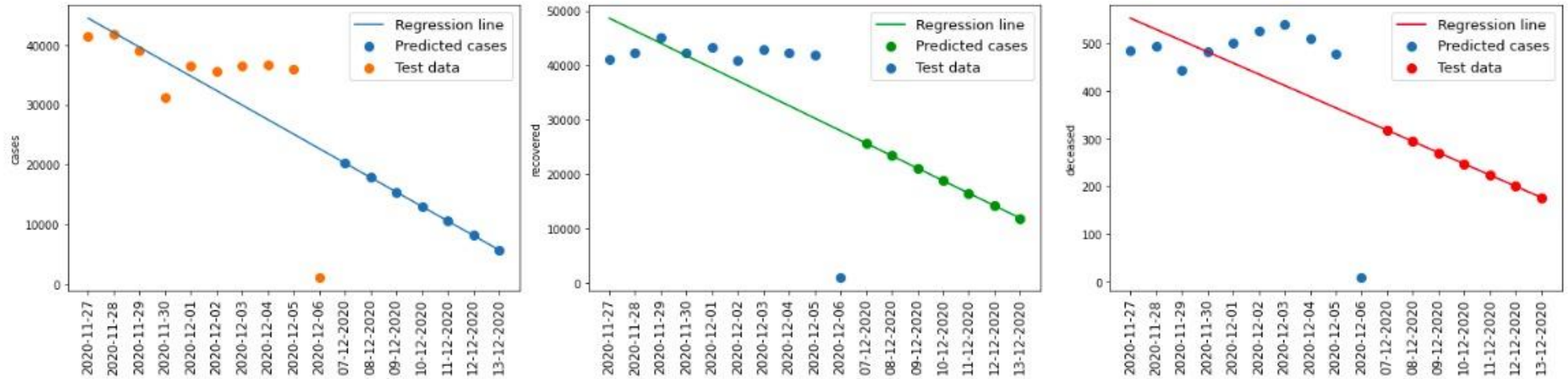


Results Cont...

- Prediction evaluation
  - Model is trained based on 10 previous consecutive days and tested on next 7 consecutive days.

Predictions for state : ALL and district : ALL

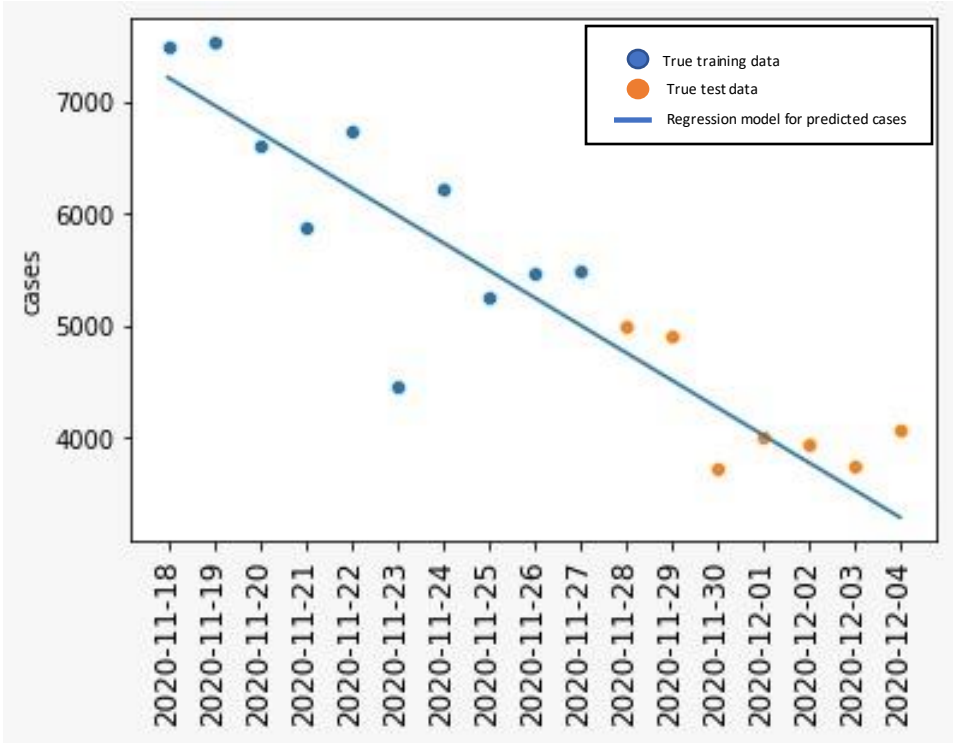
Prediction graphs

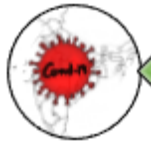


Results Cont…

- Predicted cases

	date	predicted cases	actual cases
0	2020-11-28	4761	4998
1	2020-11-29	4515	4906
2	2020-11-30	4269	3726
3	2020-12-01	4023	4006
4	2020-12-02	3777	3944
5	2020-12-03	3530	3734
6	2020-12-04	3284	4067





Future Direction

- Our work only includes cases of India, it can be applied for other nations as well.
- We could predict for more days if we enhance our project.
- We actually planned to make web-interface to interact with users which is pushed for future for now.
- We plan to represent the hostpots and coldspots at district-level.

Thank you