# Spread & Peak Prediction of Covid-19 using ANN and Regression (Workshop Paper)

Anupam Prakash, Piyush Sharma, Indrajeet Kumar Sinha and Upendra Pratap Singh
Machine Learning and Optimization Lab, Department of Information Technology
Indian Institute of Information Technology Allahabad, Prayagraj
Email: {ihm2016002, ism2016007, pcl2016004, rsi2017001}@iiita.ac.in

***Abstract*—Covid-19, caused by Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) virus, has presented tough times for countries all over the world with number of cases and casualties running in millions. While virologists and doctors have spent sleepless nights to come up with a potent vaccine, the work life of government personnel including administrative staffs, hospital employees etc. has not been any easier. Amidst this turmoil, the common question crossing every mind is concerned with the statistics about this infection including expected number of infections, peak prediction etc. We try to answer these questions by analyzing the time series data of Covid-19 infections for certain hard-hit countries and states in India. A series of machine and deep learning models have been built to capture the infection distribution so that these models could predict the fate of this infection in the near future. We also make an attempt to predict the time when active cases would cease to increase.**

**Keywords: Severe Acute-Respiratory Syndrome (SARS), Exploratory analysis, Growth rate, Active cases, Regression**

## I. Introduction

Coronavirus disease (COVID-19) is an infectious disease caused by a newly discovered severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The first case of the disease was identified in Wuhan, the capital of the Hubei province of China in November, 2019. The disease has now spread globally and has been declared as the 2019–20 coronavirus pandemic. As of 6 June 2020, more than 6.79 million cases have been reported across 188 countries and territories, resulting in more than 396,000 deaths. In statistics, exploratory analysis is an approach in which we analyze data sets to summarize their main characteristics using visual methods. It refers to the critical process of performing initial investigations on data. It is extremely beneficial when one wants to discover patterns, spot anomalies, test hypothesis and check assumptions with the help of summary statistics and graphical representations. In this paper, we visualize the Covid-19 spread datasets for different hard-hit countries as well as Indian states and use machine and deep learning models to explore how the spread is going to span out in the near future.

Recently, a lot of efforts have been made to model Covid-19 data using SIER based regression models [1] [2] [3]. Authors in [4] have conducted detailed analysis of the infection data using time-dependent SIER model with an objective to answer questions like if it is possible to contain COVID-19 and if it could be contained, what would be its peak value. The authors have also considered herd immunity in their analysis. Further, [5] quantitatively compare existing SIER based models trained for deterministic and stochastic predictions. In the Indian scenario, works like [6] [7] [8] [9] [10] [11] have been reported in the literature. In these works, authors have proposed a variety of machine and deep learning models for predicting the spread of infection. In addition, these models also answer questions like what would be the peak value and how the outbreak trend would actually look like. [12] extends the above works to not only predict the spread of the infection but also goes on to suggest preventive measures depending upon the severity of the infection it predicts. Another section of works considers the COVID data as strictly a time series data and train only time series models for predicting the future course of infections [13] [14].

Effective machine learning models have also been proposed for other countries data in [15] [16] [17] [18] [19] [20]. Rigorous attempts to model country-wise trends in the infections and corresponding peak predictions give readers a comprehensive understanding of how things are actually shaping for the future worldwide [21] [22]. [23] provides an online platform with latest developments pertaining to Covid-19 as well as its infection statistics. In contrast to the above models where comparatively complex models have been proposed to model Covid-19 infections, we, in our work, try to work with relatively simpler models. The motivation remains to be able to perform a


DOI 10.1109/BigMM50055.2020.00062

rigorous exploratory analysis of the data at hand and to be able to predict the infection spread and its growth using lesser number of model parameters.

The rest of the paper is structured as follows - Section 2 discusses our proposed methodology, Section 3 contains the Experiment Setup details while Results are discussed in Section 4. We conclude our work in Section 5 and provide to the readers, our plans for the future. We acknowledge the guidance and support in Section 6.

## II. Proposed Methodology

In our experiments, we have considered the data from Indian states as well as other hard-hit countries across the world. Firstly, we have divided the Indian States into 3 categories depending upon the severity with which they have been affected. We call these categories, top-5 most affected, top-5 moderately affected (middle) and top-5 least affected states. This categorization is the result of our bold assumption that states belonging to a particular category have similar spread pattern and hence demand similar preventive measures from the administration and health-care staff. Apart from this, when we merge similarly-hit states, the predictions tend to be more smoothed. Our proposed methodology could be put in three different heads - exploratory analysis (for modeling infection-spread), modeling of active cases and peak predictions.

### A. Exploratory Analysis

To begin with, we plot the total number of confirmed and active cases against the day number for each category/country separately . We then model the spread of infections using a linear regression model with number of days and infection count as $\{(x, y)\}$ pair with regularization parameter for Ridge regression set to $\alpha = 0.05$; for testing, last $15$ days data has been used. We then extend our linear regression model to polynomial regression; in our experiments, we have used cross-validation to choose the degree of polynomial to be $4$.

We also model the infection spread using an artificial neural network and compare the model predictions with the ground truth. For this, we initialize a neural network with 3 hidden layers, with each layer containing $30$ neurons; the network is set to be trained for $1000$ epochs. The squared loss objective function was minimized using Adam optimizer with an initial learning rate set to $1.5$.

While building the above models, we have discarded the initial data as the initial data does not give a trustworthy picture of how the infection tends to spread.

### B. Modeling Active cases

Since the number of confirmed cases is an increasing function, we use number of active cases to model and predict the fate of this pandemic. The number of active cases is calculated by subtracting the sum of recovered and deceased cases from total confirmed cases i.e. Total active cases = Total confirmed cases - (total recovered cases + total deaths). In our next set of experiments, we model the number of active cases against the day number so that we can predict the day when these active cases actually go down to 0 i.e. day when this infection ceases to exist anymore. We model the active cases for our different datasets and from the modeling curve obtained, we estimate the day when active cases reach zero (or nearly zero). Consistent with the modeling so far, we have ignored the initial data.

### C. Peak Prediction

Further, we conduct experiments to figure out when this pandemic would be on its peak. For this, we have proposed growth-ratio that represents whether active cases have increased or decreased from previous day results.

Mathematically, the growth-ratio for the $i^{th}$ day is computed as:

$$gr(i) = \frac{\text{Number of Active Cases on day i}}{\text{Number of Active Cases on day i-1}} \quad (1)$$

Using growth-ratio for the $i^{th}$ day, active cases for next day could be found out using,

$$ActiveCases_{(i+1)^{th}} = ActiveCases_{(i)^{th}} + ActiveCases_{(i)^{th}} \times gr(i)$$

Active cases have been purposely chosen since they may decrease after a while (due to recovered or deceased patients) as against the total confirmed cases which is a monotonically increasing function. Since total active cases are expected to decrease after a while, we can always find out the day when peak occurs. As per the prevailing practices, initial data is not used for modeling. We switched to polynomial regression of degree $4$ for modeling growth-ratio since an ANN was over-fitting the data.

From the modeling curve, we find the day when growth-ratio changes its sign from positive to negative; the corresponding day marks the peak of the infections. Though, the peak is expected to be on the day when growth-ratio turns its slope from positive to negative, we waited for growth ratio to take a considerable negative value before we announce the peak-day. This also resonates with the fact that the first signs of growth-ratio changing signs cannot be fully trusted; this change must be consolidated before we infer anything out of it.

## III. Experiment Setup

### A. Dataset Description and Exploratory Analysis

The datasets for our experiments have been collected from two sources - for India, we have web-scrapped the official website of Ministry of Health and Family Welfare, Government of India using BeautifulSoup [24]. For other countries, we downloaded the datasets available at Github [25]. In the Indian scenario, we have collected data from January 30, 2020 while for USA, Italy and Spain, the data was available from January 22, 2020, January 31, 2020 and February 01, 2020 respectively. The end date for all the datasets is June 06, 2020.

### B. Data pre-processing

For the sake of simplicity, we have converted the 'dates' to 'number of days since the first case was reported' with 'date of first occurrence' to be taken as day 1. As the number of cases were growing exponentially, we have converted these values to logarithmic scale for better interpretation and modeling. The infection count for the first few days had to be discarded since this data was not the representative of actual spread (limited testing kits were available in the initial days). For testing our models, we have taken the data from last 15 days.

### C. Exploratory Analysis

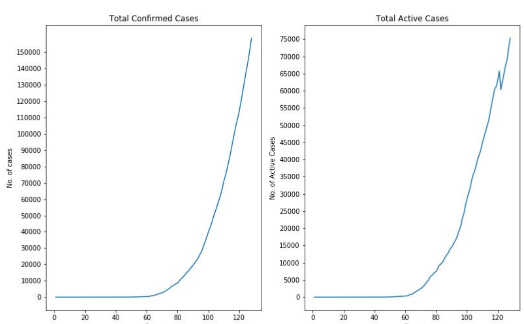


Figure 1. Infection spread for the Top-5 most affected States in India.

Figure 1 plots the total infection spread for the Top-5 worst-hit States in India. The left hand graph captures the total confirmed cases (day-wise) while the graph on the right hand side plots the total active cases (day-wise). Figure 2, on the other hand, reports the total confirmed cases and total active cases respectively (from left to right) for top-5 moderately infected States in India. Figure 3, finally plots the infection statistics of Top-5 least affected Indian states. In Figure 1, the steep rise in infections for Top-5 worst affected States can be clearly seen in the form of a near-exponential curve.

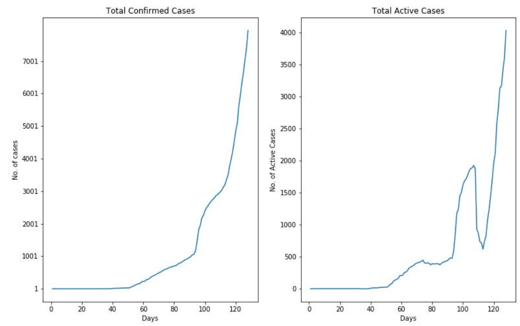


Figure 2. Number of cases reported in the Top-5 moderately affected States in India

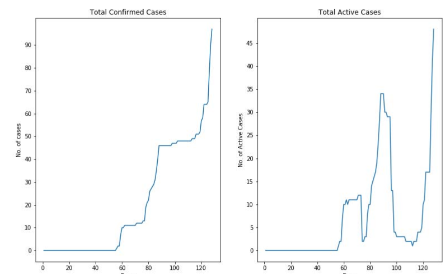


Figure 3. Total confirmed cases and total confirmed active cases for top-5 least affected Indian states

We also present the infection counts in Italy, Spain and USA (see Figure 4, Figure 5 and Figure 6 respectively).

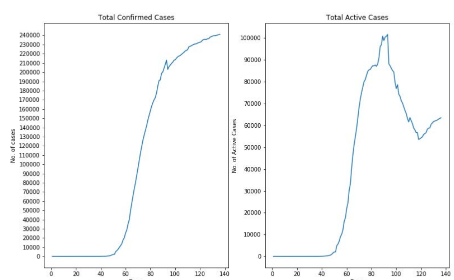


Figure 4. Covid-19 infections rise in Spain in terms of number of total confirmed cases and active cases.

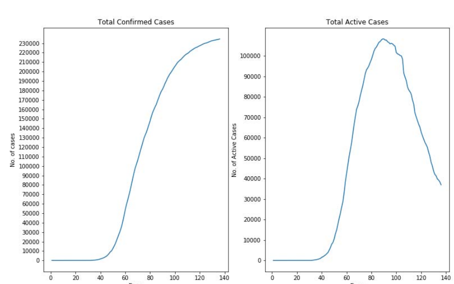


Figure 5. Infection count for Italy (total confirmed cases and total active cases)

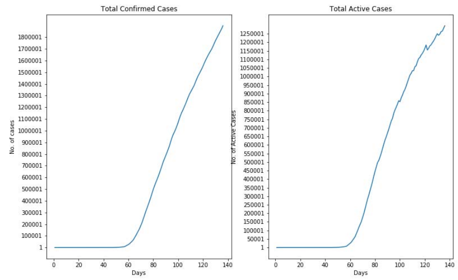


Figure 6. Infection count for USA (total confirmed cases and total active cases)

# IV. Results and Discussion

## A. Using Linear Regression

Figure 7-12 capture the predictions made by our linear regressor model on different Indian states and countries data. For better representation of the infection count, we have considered logarithmic scale as well as real number scale. As is evident, even a simple model like linear regression with added regularization term is able to capture the data distribution for Covid-19 infections. For Italy, Spain and USA, it can be seen that the model predictions are a bit deviant from the actual counts; this was not unexpected since we chose a very simply model (lesser number of model parameters) to model the infection spread.

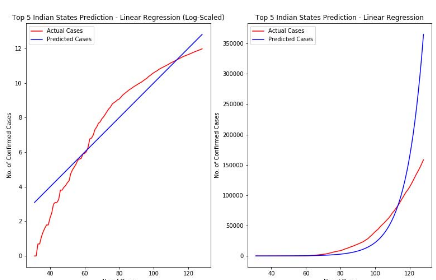


Figure 7. Predictions made by the Linear Regression model for top-5 most affected Indian States. The graph captures infection count (on logarithmic as well as real number scale) against the day number

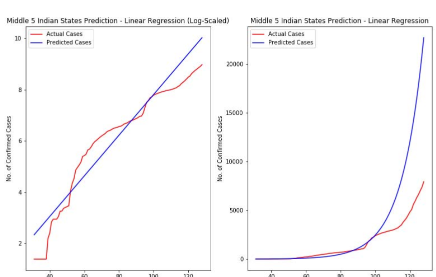


Figure 8. Linear Regression model predictions for top-5 moderately affected Indian States. Infection count (on logarithmic and real number scale) against the day number

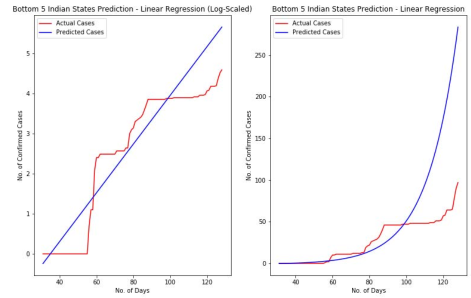


Figure 9. Linear Regressor predictions on top-5 least affected Indian States against the day number

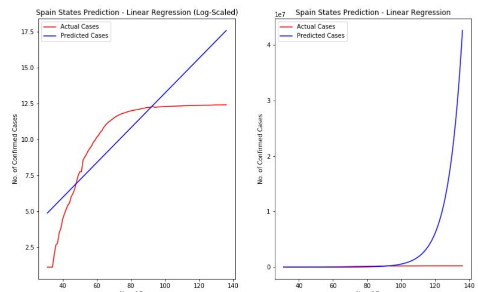


Figure 10. Linear Regressor predictions on infections that took place in Spain

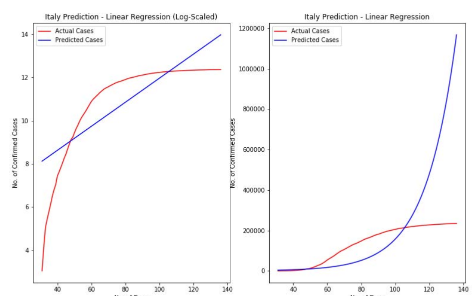


Figure 11. Linear Regressor predictions on infections that took place in Italy and comparison with the ground truth

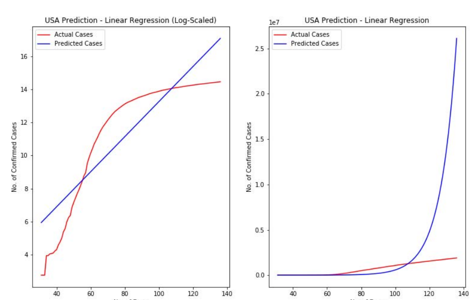


Figure 12. Linear Regressor predictions on infections that took place in US

### B. Using Polynomial Regression

The modeling results have been depicted in Figures 13-18. It can be seen that our polynomial regressor predictions follow closely the ground truths for the Indian scenario.

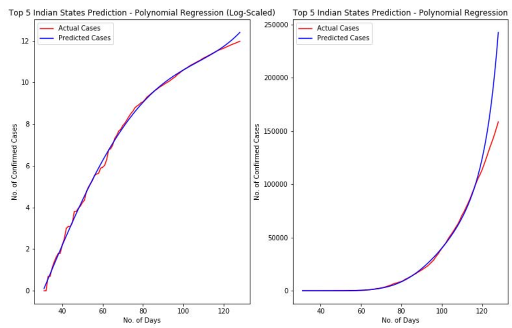


Figure 13. Predictions from polynomial regression compared against ground truths on top-5 most affected states in India. The left hand side of the graph captures the number of confirmed cases on logarithmic scale while the one on the right hand side plots the number of confirmed cases on the real number scale

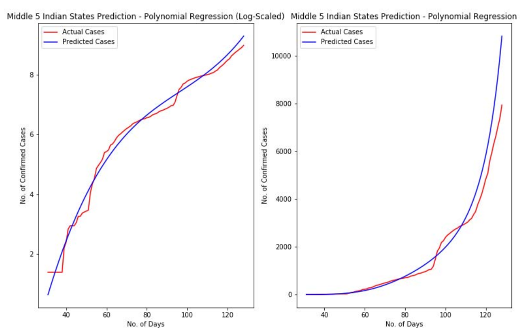


Figure 14. Polynomial regressor predictions and its comparison with the ground truth on top-5 moderately affected Indian States on log and real number scale

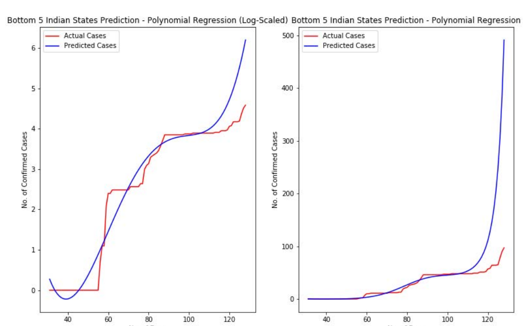


Figure 15. Results of the polynomial regressor along with ground truths for top-5 least affected Indian States. Total number of confirmed cases in log and real number scale plotted against number of days since the first infection was confirmed

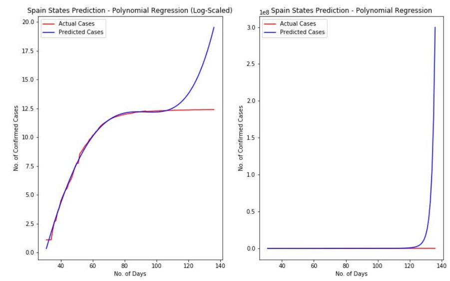


Figure 16. Results of the polynomial regressor compared against actual number of confirmed cases on Spain data.

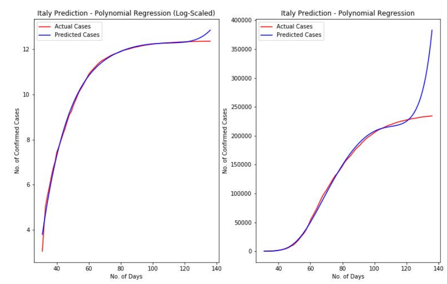


Figure 17. Results of the polynomial regressor compared against actual number of confirmed cases on Italy data.

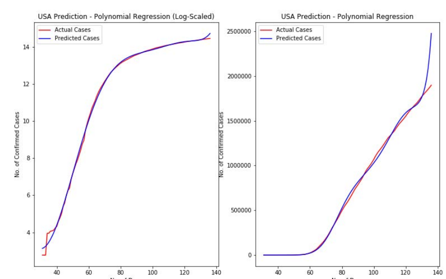


Figure 18. Results of the polynomial regressor compared against actual number of confirmed cases on US data.

By comparing the results obtained using linear regression with fourth degree polynomial regression, we can clearly see that the predictions of polynomial regression follow closely the ground truth. However, the predictions made by polynomial regression were distant from the actual number of cases on Spain and Italy data. We believe that this is due to decreased number of cases (for Spain and Italy) in the last 15 days of the time frame that we have used in our experiments.

### C. Using Artificial Neural Networks

Figure 19-24 capture the predictions made by the ANN model and its deviation with the ground-truth. For a better comparison of model predictions with the ground-truth, we have considered real as well as logarithmic scales.

We can see that the modeling and subsequent predictions done by our ANN model have been fairly accurate; particularly for the Top-5 most affected Indian states. For the Top-5 least affected States, there has been some deviation while Top-5 moderately affected States have shown average modeling performance. This may be due to the fact that the least affected states didn't have many cases initially and hence the spread was not necessarily following any pattern. But once we have considerable number of cases, ANN is able to model the pattern and the accuracy of its prediction improves as evident from the graph of Top-5 most affected states. By comparing the predictions made by the ANN model with actual number of confirmed cases, it is evident that ANN does far better modeling of infection data; particularly for the Indian states that have been worst affected by the pandemic.

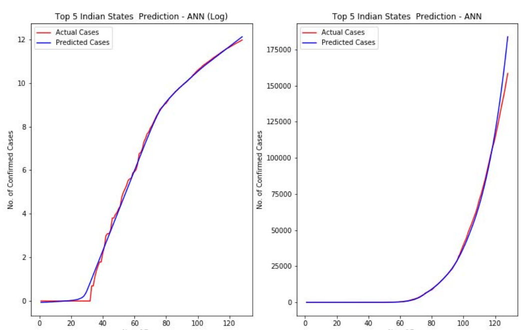


Figure 19. Prediction of infection spread ( of total number of confirmed cases) using ANN against the actual count for top-5 severely affected Indian states on log and real number scales.

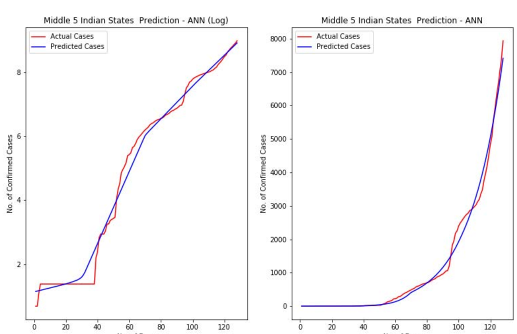


Figure 20. Predictions for top-5 moderately affected Indian states using ANN model (on linear as well as logarithmic scale). The predictions when compared against the actual counts show a fair amount of agreement.

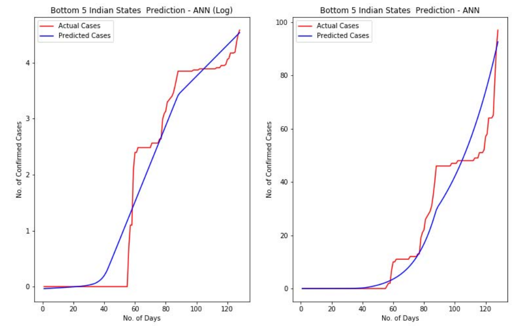


Figure 21. ANN predictions and comparison with ground truth for top-5 least affected Indian States on log as well as real number scale.

As we can observe from the graph of Top-5 least affected states, the actual spread is quiet uneven and the curve has many inflection points. This can be associated with reasons like low rates of testing initially, arrivals of labours and students from different states and other external factors.

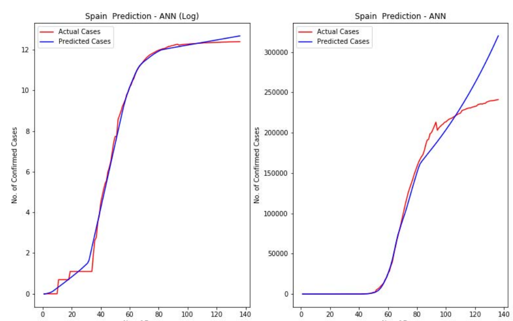


Figure 22. ANN predictions on Spain data. Left hand side predicts the number of confirmed cases on log scale while the graph on the right hand side makes predictions on real number scale.

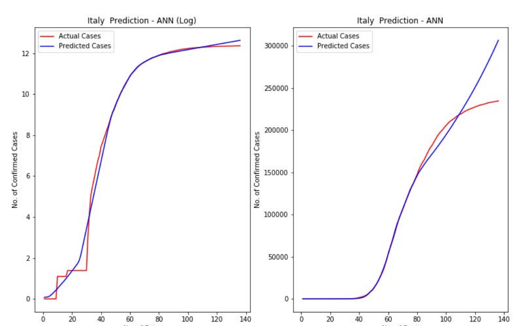


Figure 23. Predictions by ANN model on Italy data (on both the scales)

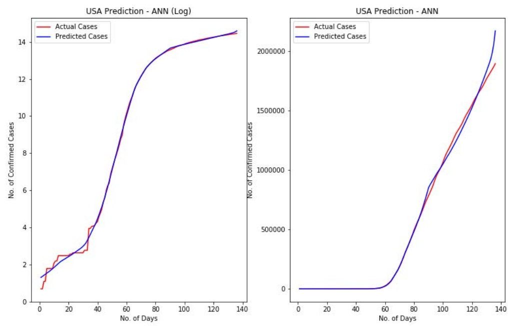


Figure 24. Predictions by ANN model on US data (on both the scales)

Table I
COMPARING THE NUMBER OF CONFIRMED CASES AS PREDICTED BY THE ANN MODEL AGAINST ACTUAL NUMBER OF CONFIRMED CASES AS ON JUNE 06, 2020.

| Dataset | Actual cases | Predicted cases |
|---|---|---|
| Top 5 Indian States | 158,498 | 183,867 |
| Middle 5 Indian States | 7,936 | 7,538 |
| Bottom 5 Indian States | 97 | 92 |

Again the predictions for Italy and Spain show some deviations. We believe that this deviation is due to the following reasons - one, the cases had started to flatten by the end of the time frame that we have chosen; secondly, our model does not take into account external factors such as dynamics in lockdown restrictions and upgrades in the health-care facilities.

### D. Modeling Active Cases

We model the active cases using polynomial regression. We experimented with ANN as well but it was over-fitting on the data.

The results obtained for Actives Cases prediction for Spain is shown in Fig 25.

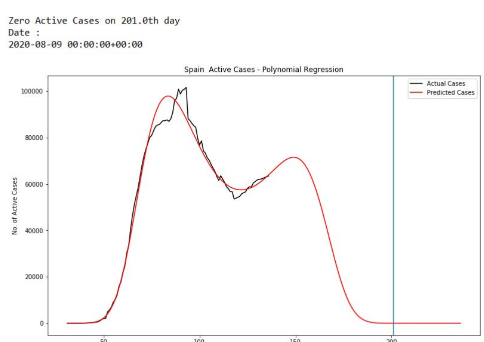


Figure 25. Modeling of total number of active cases on Spain data using Polynomial regression. The modeling curve obtained is then used to determine when these active cases would become zero.

Figure 25 depicts the curve fitting on the number of active cases for Spain data using polynomial regression. We use this curve to find out the day number when the active cases will actually go down to zero. From the figure, the vertical line marks the required day which for this dataset is 201 i.e. August 09, 2020. Similarly, the day when active cases go down for Italy data was 189 i.e. July 28, 2020 as shown in Figure 26. Figure 27, on the other hand, fit the polynomial regressor model on US data in order to estimate the day when active cases would go down to zero. The day number when active cases cease to exist is 183 i.e. July 20, 2020; this can be seen by the vertical blue line on the curve.

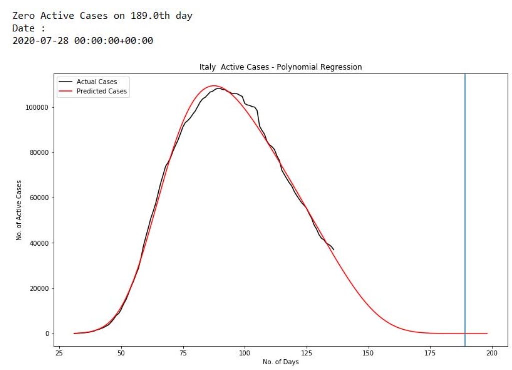


Figure 26. Modeling the active cases for Italy data using our polynomial regression model. Day when active cases go to zero can easily be read from the graph.

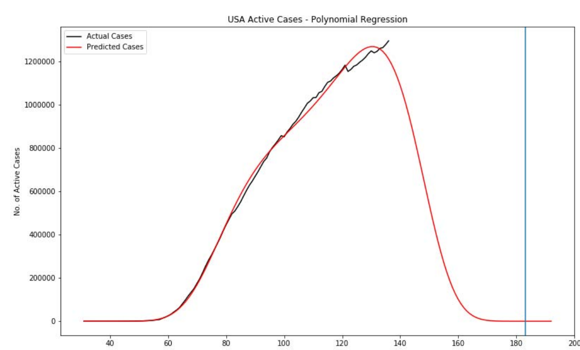


Figure 27. Modeling the active cases for US data using our polynomial regression model. Day when active cases go to zero can easily be read from the graph.

From the two graphs modeling the active cases for Italy and Spain, it has to be observed that we were actually able to predict the day when these cases go down to zero. This could be attributed to the fact that these countries had started to show a fall in the number of active cases early on and the curve we fitted could be extended to read the required day.

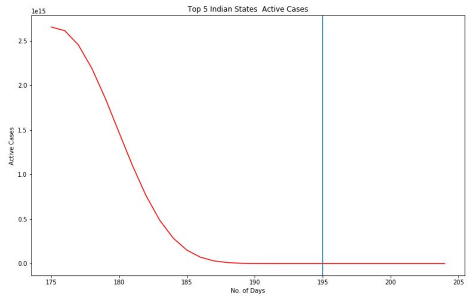


Figure 28. Modeling the active cases for top-5 most severely affected Indian states. Vertical blue line on the curve marks the day when active cases cease to exist anymore.

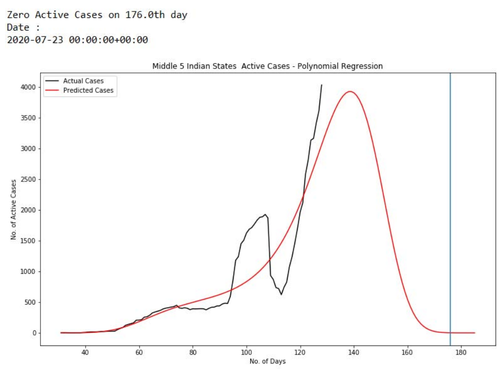


Figure 29. Modeling the active cases for top-5 moderately affected Indian states and subsequent determination of the day when these active cases would go down to zero using polynomial regression model.

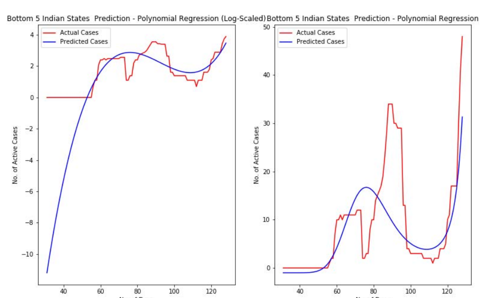


Figure 30. Modeling the active cases for top-5 least affected Indian states.

Our model predicts that for top-5 moderately affected Indian states, the day when active cases would be zero is 176 i.e. on July 23, 2020. It is to be understood that the day number that is being predicted for any dataset is actually the day from its first confirmed case; so it is dataset specific.

Likewise, for top-5 most severely affected states, the day when active cases go down to zero is estimated to be 208 i.e. August 11, 2020. The number of active cases for top-5 least affected states continued to rise; as a consequence we could not determine the day when these cases would actually go down to zero. These trends are captured in Figures 28-30.

### *E. Growth Ratio and Peak Prediction*

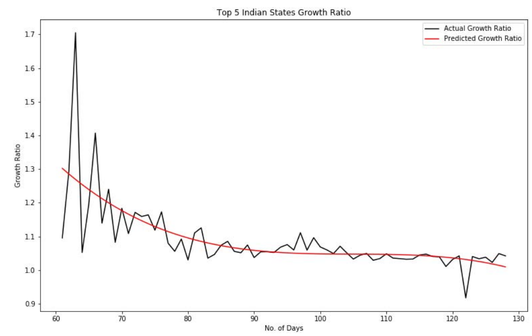


Figure 31. Growth-ratio modeling for top-5 most affected Indian states using degree 4 polynomial regression

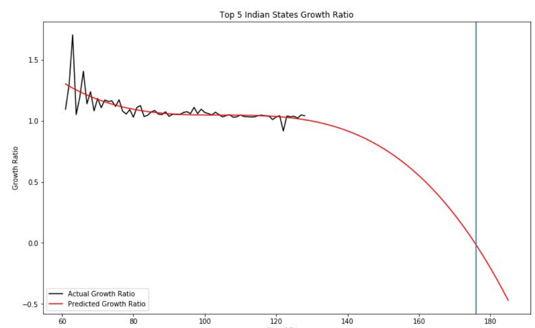


Figure 32. Peak prediction for top-5 most affected Indian states. From the curve, we can expect the peak to occur on $176^{th}$ day i.e. on July 23, 2020.

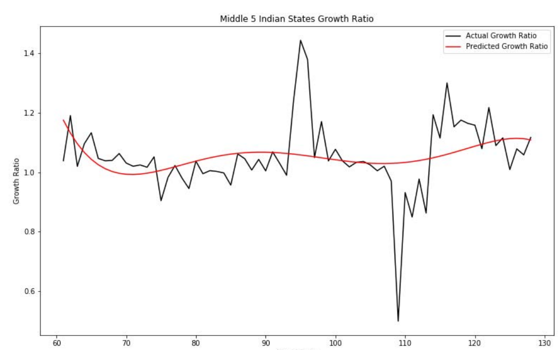


Figure 33. Growth-ratio modeling for top-5 moderately affected Indian states. The growth-ratio in this case was quite erratic, however, our model tries its best to approximate the trend.

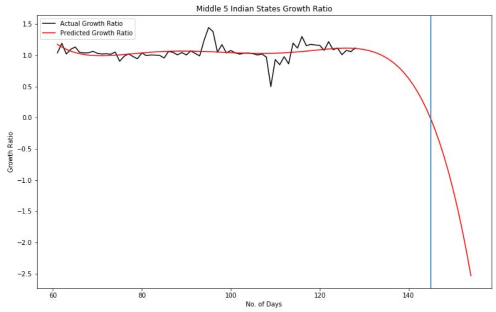


Figure 34. Peak prediction for top-5 moderately affected Indian states. We may expect the peak to be occurring on $145^{th}$ day i.e. on June 22, 2020.

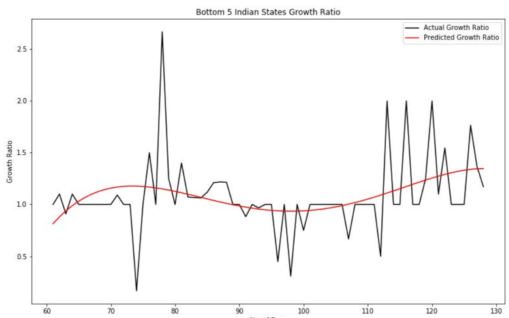


Figure 35. Growth-ratio for top-5 least affected Indian states using polynomial regressor

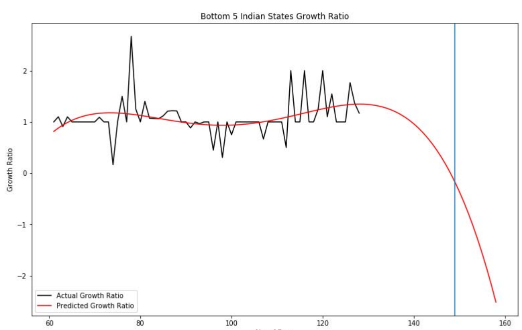


Figure 36. Peak prediction of the infection for top-5 least affected states in India is expected to be $149^{th}$ day i.e. June 26, 2020.

Figure 31-36 capture the growth-ratio modeling and peak prediction on Indian states for the three categories. The modeling and subsequent prediction has been done using a degree 4 polynomial regressor model. We figure out the day when the growth ratio changes its sign from positive to negative; the corresponding day (vertical blue line on the graph) is expected to witness peak value of the infection. Figures 37-42 capture the same trends for US, Italy and Spain data in the same order.

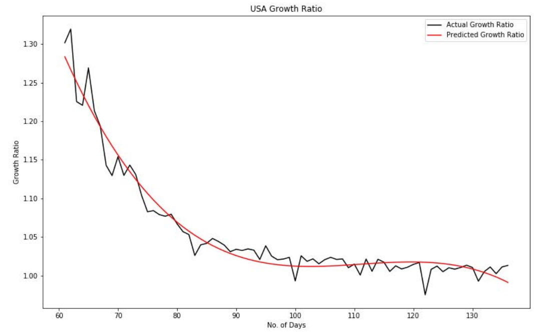


Figure 37. Growth-ratio modeling for the USA data.

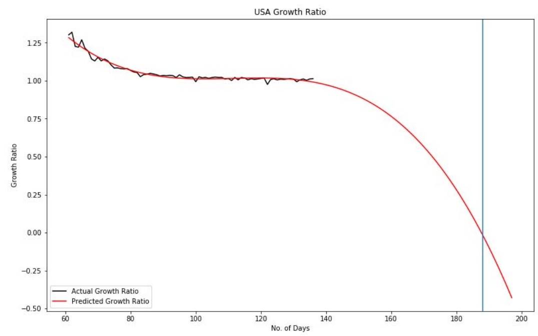


Figure 38. Peak prediction for USA data. Peak is predicted on $188^{th}$ day i.e. On July 27, 2020

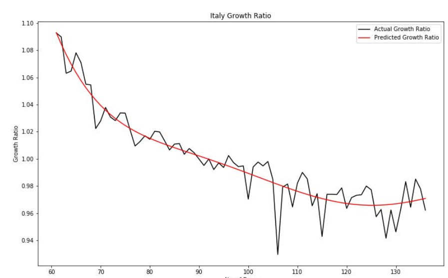


Figure 39. Growth-ratio modeling for Italy data.

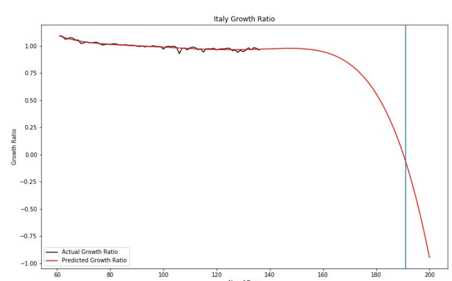


Figure 40. Peak prediction for Italy data. Peak is predicted on day 191 i.e. On July 30, 2020

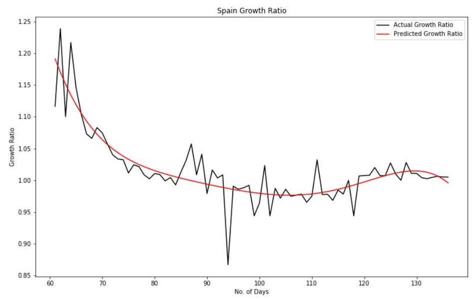


Figure 41. Growth-ratio modeling for Spain data.

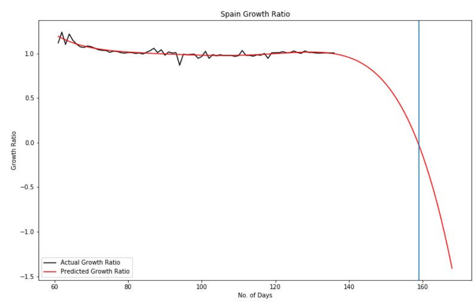


Figure 42. Peak prediction for Spain data. Peak is predicted on day 159 i.e. On June 28, 2020

## V. Conclusion

In this paper, we have build prediction models to analyze the global spread of Covid-19 pandemic in Indian as well as worldwide scenario. We started with simple models like regression and eventually went up to ANNs. Next, we also create a growth-ratio metric to predict peak as well as the end of this pandemic. Further, for the sake of simplicity, we have ignored factors like severity of lockdown imposed, health-care facilities available etc. and have relied only on the infection statistic for our modeling. In future, we plan to use differentially private time series models to model the infection data.

## VI. Acknowledgement

We express our sincere regards to Krishna Pratap Singh, Associate Professor in the Department of Information Technology, IIIT Allahabad for his consistent guidance and support. This work was completed as an assignment in the machine learning course offered by the Department.

## References

[1] G. Pandey, P. Chaudhary, R. Gupta, and S. Pal, "Seir and regression model based covid-19 outbreak predictions in india," *arXiv preprint arXiv:2004.00958*, 2020.

[2] Z. Yang, Z. Zeng, K. Wang, S.-S. Wong, W. Liang, M. Zanin, P. Liu, X. Cao, Z. Gao, Z. Mai *et al.*, "Modified seir and ai prediction of the epidemics trend of covid-19 in china under public health interventions," *Journal of Thoracic Disease*, vol. 12, no. 3, p. 165, 2020.

[3] J. Wangping, H. Ke, S. Yang, C. Wenzhe, W. Shengshu, Y. Shanshan, W. Jianwei, K. Fuyin, T. Penggang, L. Jing *et al.*, "Extended sir prediction of the epidemics trend of covid-19 in italy and compared with hunan, china," *Frontiers in medicine*, vol. 7, p. 169, 2020.

[4] Y.-C. Chen, P.-E. Lu, C.-S. Chang, and T.-H. Liu, "A time-dependent sir model for covid-19 with undetectable infected persons," *arXiv preprint arXiv:2003.00122*, 2020.

[5] B. M. Ndiaye, L. Tendeng, and D. Seck, "Comparative prediction of confirmed cases with covid-19 pandemic by machine learning, deterministic and stochastic sir models," *arXiv preprint arXiv:2004.13489*, 2020.

[6] R. Sujath, J. M. Chatterjee, and A. E. Hassanien, "A machine learning forecasting model for covid-19 pandemic in india," *Stochastic Environmental Research and Risk Assessment*, p. 1, 2020.

[7] S. Sengupta and S. Mugde, "Covid-19 pandemic data analysis and forecasting using machine learning algorithms," *medRxiv*, 2020.

[8] P. N. Mahalle, N. P. Sable, N. P. Mahalle, and G. R. Shinde, "Predictive analytics of covid-19 using information, communication and technologies," 2020.

[9] N. S. Punn, S. K. Sonbhadra, and S. Agarwal, "Covid-19 epidemic analysis using machine learning and deep learning algorithms," *medRxiv*, 2020.

[10] S. Tiwari, S. Kumar, and K. Guleria, "Outbreak trends of coronavirus disease–2019 in india: A prediction," *Disaster medicine and public health preparedness*, pp. 1–6, 2020.

[11] R. O. Ogundokun and J. B. Awotunde, "Machine learning prediction for covid 19 pandemic in india," *medRxiv*, 2020.

[12] A. Tomar and N. Gupta, "Prediction for the spread of covid-19 in india and effectiveness of preventive measures," *Science of The Total Environment*, p. 138762, 2020.

[13] F. M. Khan and R. Gupta, "Arima and nar based prediction model for time series analysis of covid-19 cases in india," *Journal of Safety Science and Resilience*, 2020.

[14] R. Salgotra, M. Gandomi, and A. H. Gandomi, "Time series analysis and forecast of the covid-19 pandemic in india using genetic programming," *Chaos, Solitons & Fractals*, p. 109945, 2020.

[15] F. Rustam, A. A. Reshi, A. Mehmood, S. Ullah, B. On, W. Aslam, and G. S. Choi, "Covid-19 future forecasting using supervised machine learning models," *IEEE Access*, 2020.

[16] L. Jia, K. Li, Y. Jiang, X. Guo *et al.*, "Prediction and analysis of coronavirus disease 2019," *arXiv preprint arXiv:2003.05447*, 2020.

[17] S. Tuli, S. Tuli, R. Tuli, and S. S. Gill, "Predicting the growth and trend of covid-19 pandemic using machine learning and cloud computing," *Internet of Things*, p. 100222, 2020.

[18] T. Chakraborty and I. Ghosh, "Real-time forecasts and risk assessment of novel coronavirus (covid-19) cases: A data-driven analysis," *Chaos, Solitons & Fractals*, p. 109850, 2020.

[19] S. F. Ardabili, A. Mosavi, P. Ghamisi, F. Ferdinand, A. R. Varkonyi-Koczy, U. Reuter, T. Rabczuk, and P. M. Atkinson, "Covid-19 outbreak prediction with machine learning," *Available at SSRN 3580188*, 2020.

[20] Z. Hu, Q. Ge, L. Jin, and M. Xiong, "Artificial intelligence forecasting of covid-19 in china," *arXiv preprint arXiv:2002.07112*, 2020.

[21] D. Yadav, H. Maheshwari, and U. Chandra, "Outbreak prediction of covid-19 in most susceptible countries," *Global Journal of Environmental Science and Management*, vol. 6, no. Special Issue (Covid-19), pp. 11–20, 2020.

[22] R. Pal, A. A. Sekh, S. Kar, and D. K. Prasad, "Neural network based country wise risk prediction of covid-19," *arXiv preprint arXiv:2004.00959*, 2020.

[23] F. B. Hamzah, C. Lau, H. Nazri, D. Ligot, G. Lee, C. Tan *et al.*, "Coronatracker: worldwide covid-19 outbreak data analysis and prediction," *Bull World Health Organ*, vol. 1, p. 32, 2020.

[24] G. of India, "Ministry of health and family welfare." [Online]. Available: https://www.mohfw.gov.in/

[25] "Novel coronavirus 2019 time series data on cases." [Online]. Available: https://github.com/datasets/covid-19