

A MULTI-PRONGED APPROACH TO PREDICT REPRODUCTION NUMBER (R_0) OF SARS-CoV2 TRANSMISSION IN INDIA

Mrinmoy Roy*,
Associate Professor
PES University
Bangalore, India
mroy2612@gmail.com,
0000-0002-4646-9301

Gopal Krishna Patra
Chief Scientist
CSIR – 4PI
Bangalore, India
gkpatra@csir4pi.in

Abstract— The sudden rapid outbreak of the coronavirus has posed significant challenges to all communities. Researchers and doctors have struggled to fully comprehend all aspects associated with SARS-CoV2, such as its transmission rate, due to the virus's unique nature and the emergence of new strains. Making it incredibly tough to simulate accurately. Even as more data and information become available, these characteristics remain dynamic, making it impossible to quantify them. The lack of good models and sufficient data on confirmed cases and deaths at the start of the pandemic made it difficult to anticipate future cases and plan for key hospital resource needs and effective mitigation efforts. Because of the coronavirus's unfamiliarity and novelty, contradictory messages were transmitted about the outbreak's potential severity, the importance of mask-wearing and social distancing, and other related precautions, resulting in ad hoc, unplanned lockdowns and a rise in SARS-CoV2 cases. Supply chains and hospitals around the world were unable to keep up with the demand for ventilators, personal protective equipment (PPE), and intensive care unit (ICU) beds as doctors and scientists began to better understand SARS-CoV2's destructive and disruptive nature, causing critically ill patients to be turned away. The world's strongest and largest economies came to their knees, and resources have been stressed out due to the speed of coverage of the pandemic and the rate of transmission of the disease. Seeing the magnitude of growth in a number of cases also its consequences on the health administration of the country, an accurate prediction seems to be a very important requirement to control this in future and be prepared for the same. In this research, Mathematical, Statistical and LSTM modelling and identification techniques are applied in developing a prediction system to forecast the spread of SARS-CoV2 in India. When any Epidemic or Pandemic breaks out there is a need for immediate measures in exponential cases as it's evident that developing a vaccine would take some time, so non-therapeutic interventions such as social distancing or contact tracing have benefitted to delay the rising of SARS-CoV2 cases. Through this research, we are proposing a multi-pronged approach for bracketing the SARS-CoV2 cases in India. We found ARIMA model grossly underestimates, SEIRD model grossly overestimates, the Polynomial has shown promises for a long-term prediction (~20 days), The changing degree of polynomials for best fit, Deep learning models seems to have immense potential to bracket the number of cases especially for short-term (~10 days). A Comprehensive platform (most likely a hybrid one with DL, Statistical and Mathematical) is very useful to Sharpen the bracket as much as possible and use of clustering algorithms to identify vulnerable districts to

potentially become SARS-CoV2 hotspots. Estimate changing R_0 (basic reproduction number) from data using Deep Learning, influenced by lockdown and unlockdown patterns and design downscaling approaches. Anaconda environment with Python, Jupyter Notebook and Google Collab ML platforms are used as simulation tools for the algorithms to test and predict Indian SARS-CoV2 disease data.

Keywords— *Multipronged Approach, SARS-CoV2, ARIMA, SEIRD, Deep Learning*

I. INTRODUCTION

COVID-19 outbreak has brought disaster and emergency all over the world. Infected cases are rapidly increasing every day. India being second after US in terms of total cases till now have 34.285 million COVID-19 cases and out of those 0.15 million active cases as per the data updated on Nov 1, 2021 (www.covid19india.org). World's strongest and largest economies came to their knees, resources have stressed out due to the speed of coverage of the pandemic and rate of transmission of the disease. Seeing the magnitude of growth in number of cases also its consequences on the health administration of the country, an accurate prediction seems to be very important requirement to control this in future and be prepared for the same. The following are few prediction models have exercised by some researchers. **Edward et al., (2022)** formed **CoVCom9** which is a modified SEIR Compartmental model consists of 9 compartments describing SARS-CoV-2 transmission dynamics in Ghana. Their team carried out a detailed mathematical analysis and derived basic reproduction number R_0

Tuli et al., (2020) deployed Machine Learning (ML) and Cloud computing technology very effectively to keep a track of the disease, predicting growth of the epidemic and suggestive strategies and policies for health administration to manage the spread. An improved mathematical model is applied here to analyze and predict the growth of the epidemic. Worldwide lots of AI, ML-based algorithmic model is applied to predict the potential threat of COVID-19. But this study used iterative weighting for fitting Generalized Inverse Weibull distribution, this can be better fit to develop a prediction framework.

Tomar et al., (2020) used data-driven estimation and prediction methods of AI viz. long short-term memory (LSTM) and curve fitting to find out the number of COVID-19 cases in India 30 days prior and effective preventive measures to be taken to manage the same. The predicted number of positive cases and recovered cases obtained by the proposed method is accurate

within a certain range and is a beneficial tool for administrators and health officials.

Pandey et al., (2020) analyzed outbreak of the disease in India for a certain date and prediction of number of cases done for the next 2 weeks. The SEIR model and Regression models are used for predictions based on the data collected from John Hopkins University repository in the time period of 30th January 2020 to 30th March 2020. The evaluation of performance of the proposed model using RMSLE achieved 1.52 for SEIR model and 1.75 for the regression model. The RMSLE error rate between SEIR model and Regression model was found to be 2.01. Even the value of R_0 which determines the spread of the disease was calculated to be 2.02. The predicted number of cases may rise between 5000-6000 in the next two weeks of time according to the above method.

Sarkar et al., (2020) proposed a mathematical model that specifically predicts the COVID-19 dynamics in 17 provinces of India and overall. According to this study an estimated pandemic life cycle along with the real data or history to date has shown which reveals the predicted inflection point and ending phase of COVID 19. Actually the proposed model monitors the dynamics of six compartments, namely susceptible (S), asymptomatic (A), recovered (R), infected (I), isolated infected (I_q) and quarantined susceptible (S_q), collectively expressed $SARII_qS_q$. Also a sensitivity analysis is conducted to determine the robustness of model predictions to parameter values and the sensitive parameters are estimated from the real data on the COVID-19 pandemic in India. Their result shows by reducing contact rate between uninfected and infected individuals through quarantine the susceptible individuals, leads effective reduction in proportionate number of cases. The proposed model simulation too demonstrates COVID 19 pandemic eradication possible by combining the restrictive social distancing and contact tracing. As their predictions are based on real data with reasonable assumptions, whereas the accurate course of epidemic heavily depends on how and when quarantine, isolation and precautionary measures are enforced.

Rajesh Ranjan, (2020) compared Indian COVID-19 data with several countries as well as the important states in the US struck with a major outbreak, and he also found that the basic reproduction number R_0 for India is in the range of 1.4-3.9 which was as per expectation. Also growth rate of infections in India is almost near to that of Washington and California. Based on Exponential and classic susceptible-infected-recovered (SIR) model estimation India is

going to enter an equilibrium by May end 2020 with final epidemic size of 13,000 approximately. But the above estimation will not be valid if India acquires community transmission stage. There is also an assessment of social distancing, assumption of no community transmission with impact on increase in number of cases by comparing data from different geographical locations.

Rajesh et al., (2020) also used a very simple and effective mathematical model SIR(D) to predict the future of the epidemic in India by using the existing data. Even they also estimated the effect of lock-down/social isolation through a time-dependent coefficient of the model. The said model studied with realistic parameters set reveals the epidemic will be at its peak around the month end of June or the first week of July with almost 108 Indians most likely being infected if the lock-down is relaxed after May 3, 2020. Even it also estimates one-third of total number predicted population will be infected if it is considered that people only in the red zones (approximately one-third of India's population) are susceptible to the infection.

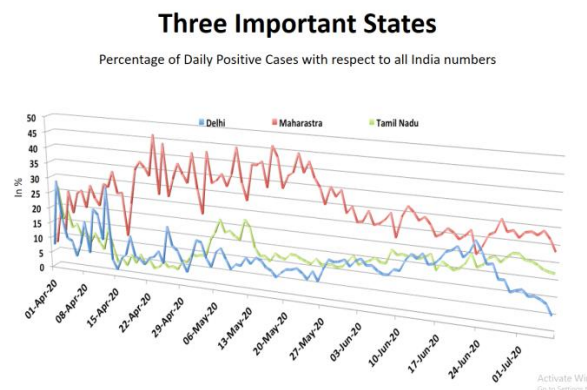
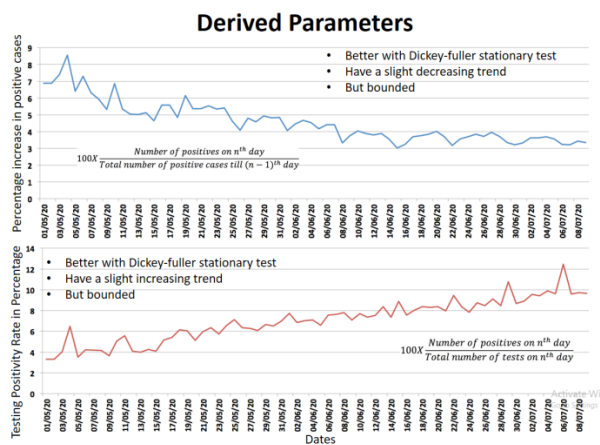
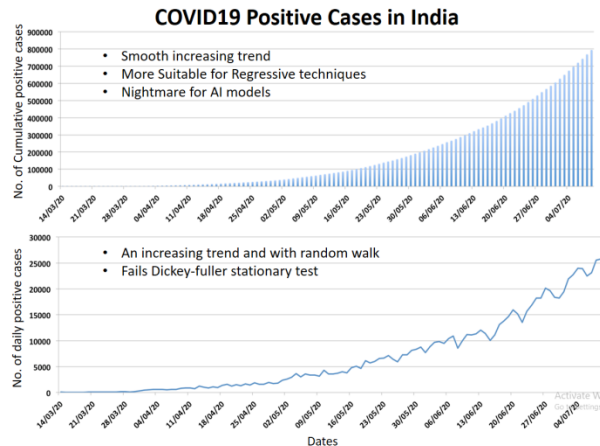
Ray et al., (2020) used a Bayesian extension of the Susceptible-Infected-Removed (eSIR) model designed for intervention forecasting to study the short- and long-term impact of an initial 21-day lockdown on the total number of COVID-19 infections in India compared to other less severe non-pharmaceutical interventions. The comparison based on effects of hypothetical timeline of lockdown on reducing the number of active and new infections. They found if lockdown implemented correctly, can reduce the total number of cases on short term basis and India can get time to prepare its healthcare and disease-monitoring system. Their analysis shows we need to have some measures of suppression in place after the lockdown for increased benefit (as measured by reduction in the number of cases). A longer lockdown between 42–56 days is preferable to substantially “flatten the curve” when compared to 21–28 days of lockdown. Their models focus solely on projecting the number of COVID-19 infections and, thus, inform policymakers about one aspect of this multi-faceted decision-making problem.

II. OBJECTIVES

With all these constraints and uncertainties, we still will have to provide a practically usable estimations

with a practically implementable lead time.
 Comprehensive platform (most likely a hybrid one with DL, Statistical and Mathematical)
 Sharpen the bracket as much as possible
 Use of clustering algorithms to identify vulnerable districts to potentially become covid19 hotspots.
 Estimate changing r_0 (basic reproduction number) from data using Deep Learning

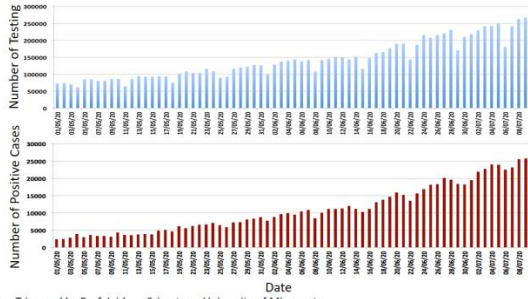
III. METHODOLOGY



LSTM (LONG SHORT TERM MEMORY)

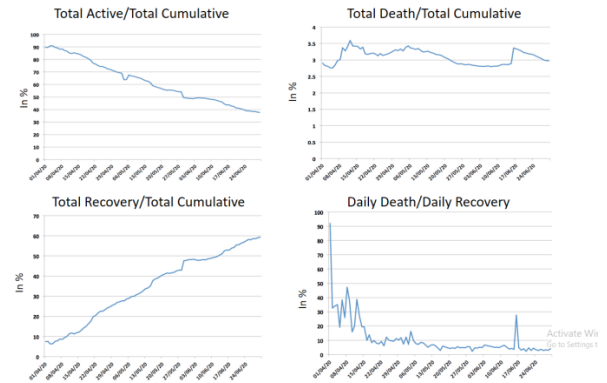
- VANILA LSTM:** A Vanilla LSTM is an LSTM model that has a single hidden layer of LSTM units, and an output layer used to make a prediction.
- Stacked LSTM:** Multiple hidden LSTM layers can be stacked one on top of another in what is referred to as a Stacked LSTM model.
- Bidirectional LSTM:** On some sequence prediction problems, it can be beneficial to allow the LSTM model to learn the input sequence both forward and backwards and concatenate both interpretations.
- CNN LSTM:** A CNN model is used in a hybrid model with an LSTM backend where the CNN is used to interpret subsequences of input and provided as a sequence to an LSTM model to interpret.
- Conv LSTM:** A type of LSTM related to the CNN-LSTM, where the convolutional reading of input is built directly into each LSTM unit.
- Encoder-Decoder LSTM:** Encoder reads the input sequence and encodes it into a fixed-length vector, and decoder decodes the fixed-length vector and outputting the predicted sequence. Here VANILA LSTM is used for both encoding and decoding.

Importance of Testing to Positivity Rate

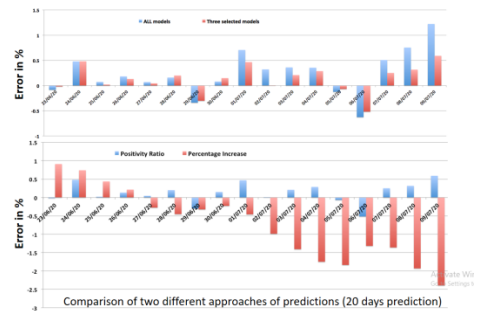


- Triggered by Prof. Jaideep Srivastava, University of Minnesota
- Testing to positivity rate is the least influenced by (though still affected by whom do we test) external interventions.
- From the practical aspect it make sense as the authorities do have ideas about the number of tests that are likely to be done well in advance.

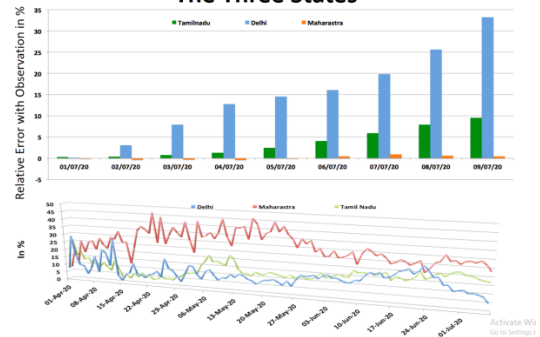
Important Ratios in Percentage



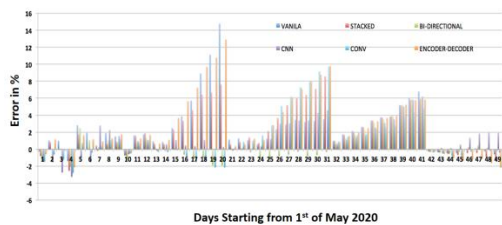
Predicted Total number of Covid19 cases in India with Testing to Positives ratio as the predicting parameter



The Three States

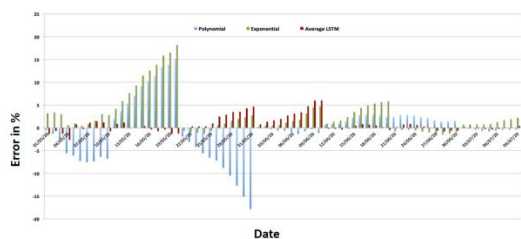


Analysis of Different Flavors of LSTM



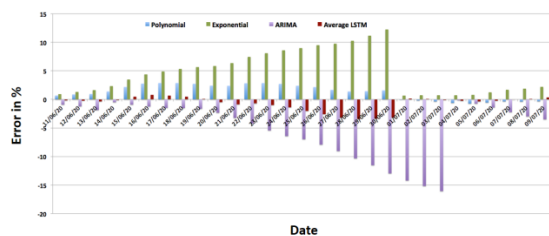
We are still evolving and improving our prediction capability

Regression Models and LSTM



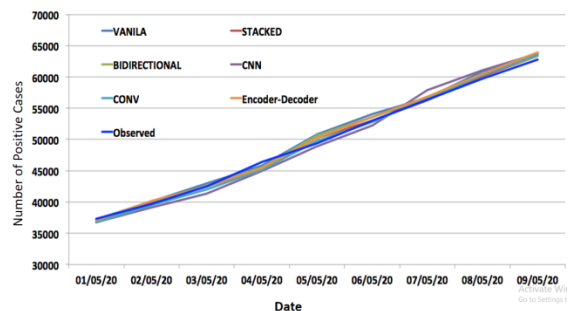
Relative error in percentage in prediction of cumulative Covid19 cases in India using the regression (polynomial and exponential) and the deep learning models (variants of LSTM) at 4PI. The predictions are done as a block of 10 days starting from 1st May 2020 (i.e. 1st May, 11th May and so on for the next 10 days).

Regression Models and LSTM

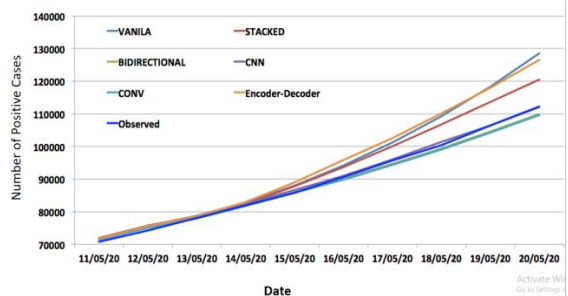


Relative error in percentage in prediction of cumulative Covid19 cases in India using the regression (polynomial and exponential) and the deep learning models (six variants of LSTM) at 4PI. The predictions are done 20 days in advance starting from 11th June 2020 till 9th July 2020.

Does a Bound Exist?

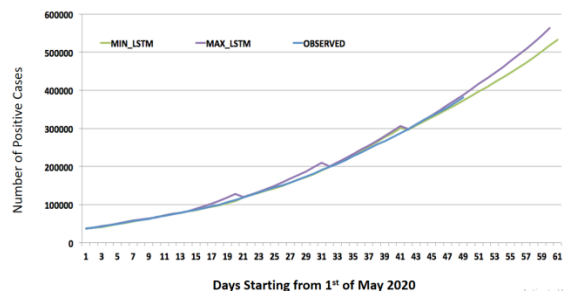


Does a Bound Exist?



USEFULNESS TO CAPACITY BUILDING

Does a Bound Exist?



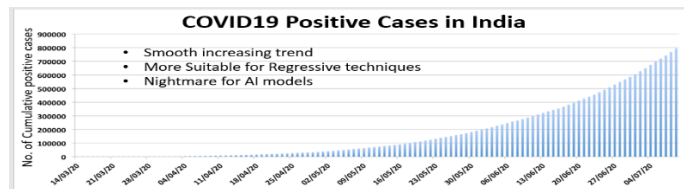
	Confirmed	Recovered	Deceased	Other	Tested
count	79689.000000	79689.000000	79689.000000	79689.000000	3.865400e+04
mean	1319.981528	874.223105	30.012612	0.278232	2.510640e+04
std	7230.458336	5390.172805	239.222477	7.171635	6.525639e+04
min	-3.000000	-67.000000	-11.000000	-3.000000	1.000000e+00
25%	31.000000	11.000000	0.000000	0.000000	3.113000e+03
50%	143.000000	81.000000	1.000000	0.000000	1.002750e+04
75%	562.000000	359.000000	7.000000	0.000000	2.247475e+04
max	162527.000000	146588.000000	7442.000000	305.000000	1.443004e+06

IV. MAIN FINDINGS

General analysis

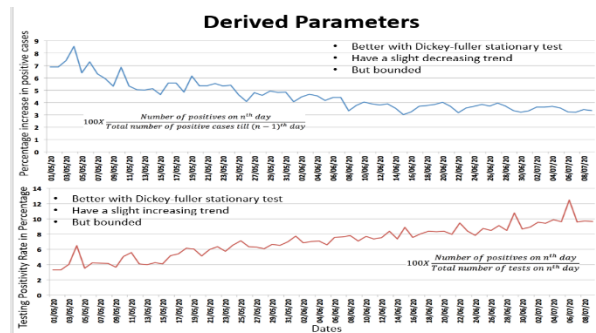
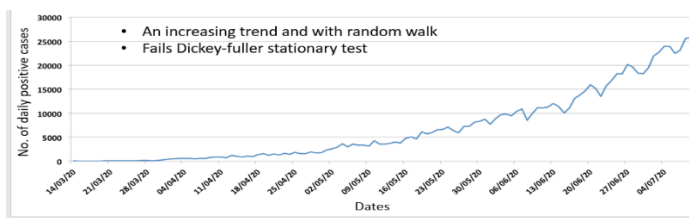
The following table depicts basic statistics about the data of districts containing confirmed cases deaths etc.

The following graph shows the trend of cumulative cases in India starting from 26th April upto August 2020

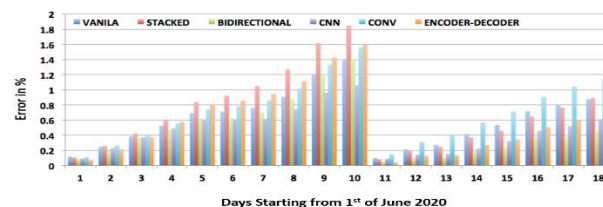


The two graphs here show the predicted trend of increasing cases in India upto July

2020 using both Regressive techniques and random walk .



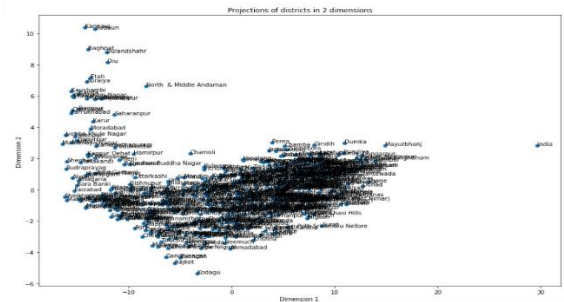
Analysis of Different Flavors of LSTM



Dimensionality Reduction on the Demographic data of Indian districts

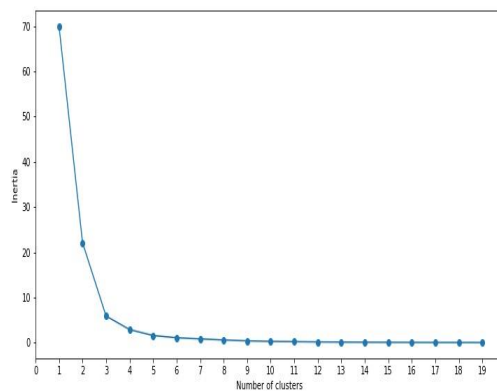
The below graph shows the distribution of various Indian districts based on their demographic parameters taken from the 2011 census data that includes attributes such as population , working population , organizational sectors etc. Of all these parameters , the data was scaled and

subjected to dimensionality reduction using Principal Component Analysis to visualize the districts in a two dimensional vector space domain . This shall be a vital in the further work of combined clustering of Indian states and districts based on the severity of COVID-19 and their demographic conditions .



We have worked on clustering of the districts, grouping the districts based on the similarity of the population, hospital facilities, rise in COVID cases and various other parameters. The clustering is done by means of **K-Means and Agglomerative clustering algorithms** with initial transformation and scaling of data for better fit to the model .

Also, we have considered various parameters, such as hospital facilities

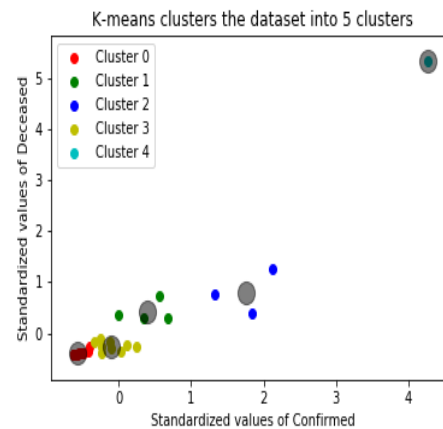


The first figure shows the optimal number of clustering required when 2 parameters are considered, i.e. Confirmed and Deceased (till 20th August). The

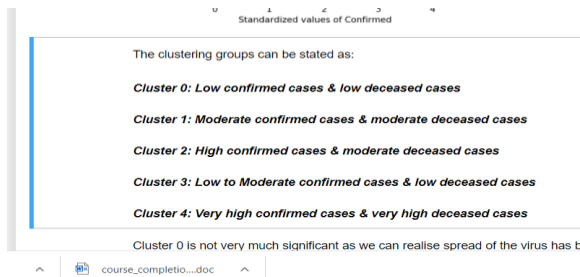
available in different states. The rise in cases and the lack of facilities in some places are grouped. Considering various other factors involving the cases we have built a model to predict the change in number of cases in future.

ARIMA Modelling

Firstly, the needed clustering algorithm for different states resulted in 5 clusters.



second figure shows the actual clustering positions in a standardized scale. The clusters can be described as follows:



So, it would be best to examine **Cluster 4 and Cluster 1 states** as their proportion of deceased to confirmed is evidently high. Each state has around 25 or more districts and an appropriate ARIMA model was fitted to all of them and their fit was checked by Ljung-Box

test. Some of the predicted values by ARIMA modelling standing on '24-08-2020' for important districts like Mumbai, Chennai and Bangalore are enlisted below.

These are the cumulative number of deceased cases.

	Mumbai	Chennai	Bangalore Urban
25/8	7466	2596	1669
26/8	7511	2614	1670
27/8	7557	2632	1670
28/8	7604	2649	1671
29/8	7651	2666	1672
30/8	7699	2683	1672
31/8	7747	2700	1673
1/9	7796	2717	1673
2/9	7845	2734	1674
3/9	7895	2750	1674

SEIRD Modelling

Regarding the SEIRD model, we came to some fine-tuned values for the parameters Infection rate(beta), incubation period(alpha), $\gamma=1/(\text{no of days taken})$

to recover) and frac = fraction of the infected that are dead.

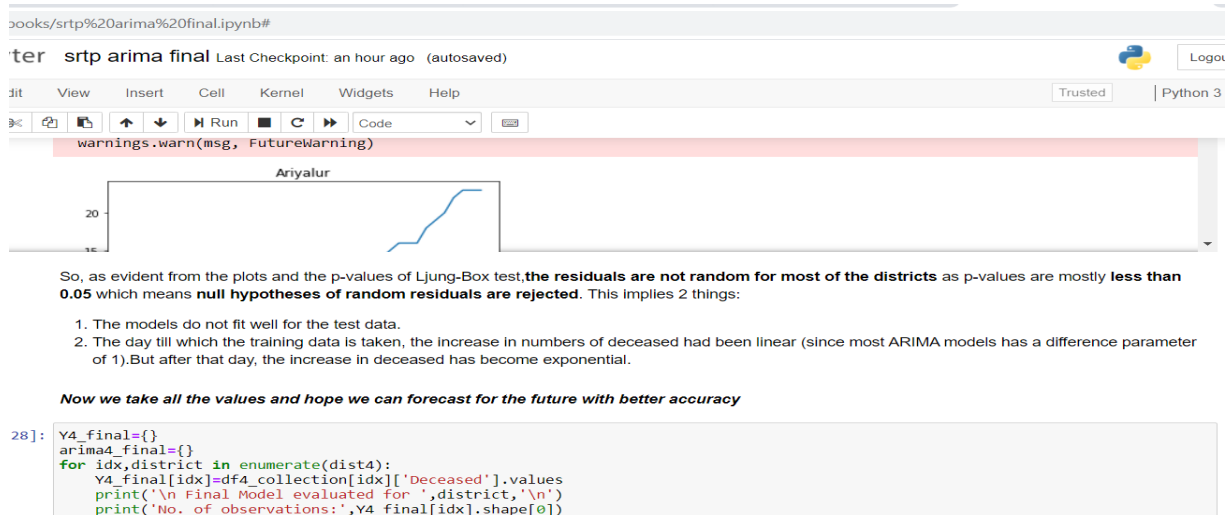
Beta	Alpha	Gamma	frac
0.1555	5.5	0.0833	0.0244

With these values, we can estimate the current values of deceased and recovered. All lower values are underestimated and all upper values are overestimated.

V. DISCUSSION

From ARIMA modelling, it is evident that the split in the dataset to take training and test dataset is very poor. The deceased curve rises steeply after the training data ends, which means covid-19 effect has started after that time of split. (The

data started from 26th April and our proportion of trained to tested was 2:1 i.e. training data for 81 days). The models change from test data to when all the data points are taken.

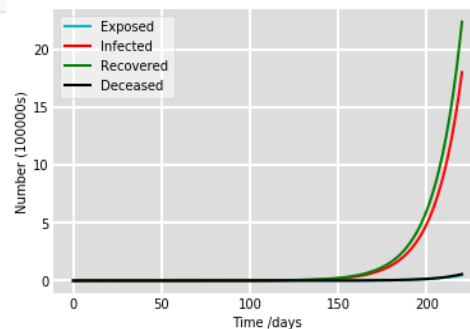
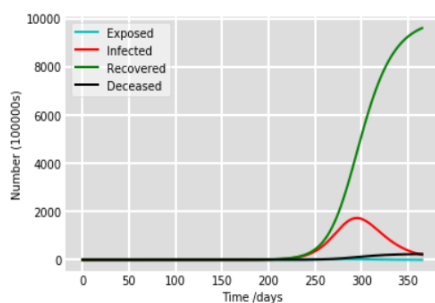


The models would have predicted better if we had more available data points and lockdowns had been exercised in India for a long period.

From SEIRD modelling, it can be concluded that taking the incidence rate constant over time is making the model very sensitive. In real world, incidence rate

decreases with time as our bodies start developing antibodies against the virus. By our model, a very blatant estimate of peak of deceased cases can be estimated which is around **300 days since the onset i.e. 30th January, 2020.**

```
for spine in ('top', 'right', 'bottom', 'left'):
    ax.spines[spine].set_visible(False)
plt.show()
```



Susceptible curve is left out as the value is very high for India compared to the other values and the graph was not being clear due to massive scale difference.

VI. FUTURE PERSPECTIVES

- i) The Reproduction number (R_0) is fixed in this case and its value is **1.8667**. Future estimates will always be overestimates, so the government and hospitals authorities can accordingly figure out their policies for the next few days.
- ii) Most important future perspective would be to regress the Beta parameter on time and then fit it in the SEIR model. Possible models for beta can be **logistic, rational decreasing function and other such functions**. Beta is by nature a decreasing function and it should ideally tend towards 1 over time.
- iii) Reducing the error of prediction as much as possible
- iv) Design downscaling approaches

VII. CONCLUSION

In this research, Mathematical, Statistical and LSTM modeling and identification techniques are applied in developing a prediction system to forecast the spread of COVID-19 in India. When any Epidemic or Pandemic breaks out there is a need of immediate measure in exponential cases as developing vaccine would take some time, so non-therapeutic interventions such as social distancing or contact tracing have benefitted to delay the rising of COVID-19 cases. Through this research we are proposing a multi-pronged approach for bracketing the COVID-19 cases in India. We are right now just at the end of 3rd wave and 4th wave is predicted by lots of studies to be happening in the month of June 2022 so it is in the learning phase due to Changing Govt. guidelines from time to time, reporting policies by different states Time to time reconciliations (pollutes the most), Political Interventions, geographically varying lockdowns and unlock, varying priorities (healthcare vs Economic), Lack of sufficient testing capability, Varying lags in results declaration. We found ARIMA model grossly underestimates, Exponential grossly overestimates, Polynomial has shown promises for a long-term prediction (~20 days), The changing degree of polynomial for best fit, Deep learning models seems to have immense potential to bracket the number of cases especially for short-term (~10 days). A Comprehensive platform (most likely a hybrid one with DL, Statistical and Mathematical) is very useful to Sharpen the bracket as much as possible, Use of clustering algorithms to identify vulnerable districts to potentially become covid19 hotspots. Estimate changing r_0 (basic reproduction number) from data using Deep Learning, Influenced by lockdown and un-lockdown patterns and design downscaling approaches. Python, Jupyter Notebook and Google Collab is used as simulation tool for the

algorithms. It's important to remember that these frameworks all have comparable, if not identical, components; the crucial difference is how we structured them and from what perspective, or lens, we examined them. These frameworks have the advantage of providing viewpoints on topics where there has been a lot of attention and where there is a gap.

References:

- Optimization in the Context of COVID-19 Prediction and Control: A Literature Review ELIZABETH JORDAN¹, DELIA E. SHIN¹, SURBHI LEEKHA², AND SHAPOUR AZARM¹
- Gupta, H.; Kumar, S.; Yadav, D.; Verma, O.P.; Sharma, T.K.Ahn, C.W.; Lee, J.-H. Data Analytics and Mathematical Modeling for Simulating the Dynamics of COVID-19 Epidemic--A Case Study of India. *Electronics* 2021, 10, 127. <https://doi.org/10.3390/electronics10020127>
- Alali, Y., Harrou, F., & Sun, Y. (2021). Optimized Gaussian Process Regression by Bayesian Optimization to Forecast COVID-19 Spread in India and Brazil: A Comparative Study. 2021 International Conference on ICT for Smart Society (ICISS). doi:10.1109/iciss53185.2021.9532501
- Yadav SK, Singh G, Sarin N, Singh S, Gupta R. Optimization of man power deployment for Covid-19 screening in a tertiary care hospital: A study of utility of queuing analysis. *Disaster Med Public Health Prep*. doi: <https://doi.org/10.1017/dmp.2021.228>.
- Gandhi R. T., Lynch J. B., del Rio C. Mild or moderate COVID-19. *N Engl J Med*. 2020. 10.1056/NEJMcp2009249. [Google Scholar](#)
- Rafiq D., Batool A., Bazaz M. A. Three months of COVID-19: a systematic review and meta-analysis. *Rev Med Virol*. 2020. doi:10.1002/rmv.2113. [Google Scholar](#)
- E. Mahase **Coronavirus: COVID-19 has killed more people than SARS and MERS combined, despite lower case fatality rate** *BMJ*, 368 (2020) [Google Scholar](#)
- C. Huang **Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China** *Lancet*, 395 (2020), pp. 497-506 [ArticleDownload PDFView Record in ScopusGoogle Scholar](#), <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>. [Google Scholar](#)
- J.T. Wu, K. Leung, G.M. Leung **Nowcasting and forecasting the potential domestic and international spread of the 2019-ncov outbreak originating in wuhan, china: a modelling study** *Lancet*, 395 (2020), pp. 689-697 [ArticleDownload PDFView Record in ScopusGoogle Scholar](#)
- Kucharski A.J., et al. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. *Lancet Infect Disease*. 2020. 10.1016/S1473-3099(20)30144-4. [Google Scholar](#)
- T. Kuniya **Prediction of the epidemic peak of coronavirus disease in japan** *J Clin Med*, 9 (2020), p. 789 [CrossRefGoogle Scholar](#)
- A.M. Almeshal, A.I. Almazrouee, M.R. Alenizi, S.N. Alhajer **Forecasting the spread of COVID-19 in kuwait using compartmental and logistic regression models** *Appl Sci*, 10 (10) (2020), p. 3402, [10.3390/app10103402](https://doi.org/10.3390/app10103402) [CrossRefGoogle Scholar](#)
- M. Mandal, S. Jana, S.K. Nandi, A. Khatua, S. Adak, T.K. Kar **A model based study on the dynamics of COVID-19: prediction and control** *Chaos Solitons Fractals* (2020), [10.1016/j.chaos.2020.109889](https://doi.org/10.1016/j.chaos.2020.109889) [Google Scholar](#)
- F. Ndairou, I. Area, J.J. Nieto, D.F.M. Torres **Mathematical modeling of COVID-19 transmission dynamics with a case study of Wuhan** *Chaos Solitons Fractals* (2020), [10.1016/j.chaos.2020.109846](https://doi.org/10.1016/j.chaos.2020.109846) [Google Scholar](#)
- Tomar, N. Gupta **Prediction for the spread of COVID-19 in India and effectiveness of preventive measures** *Sci Total Environ*, 728 (2020), p. 138762, [10.1016/j.scitotenv.2020.138762](https://doi.org/10.1016/j.scitotenv.2020.138762) [ArticleDownload PDFGoogle Scholar](#)
- Liu D., Clemente L., Poirier C., Ding X., Chinazzi M., Davis J.T., Vespignani A., Santillana M. A machine learning methodology for real-time forecasting of the 2019–2020 COVID-19 outbreak using internet searches, news alerts, and estimates from mechanistic models. 2020. <https://arxiv.org/abs/2004.04019> [Google Scholar](#)
- H. Yonar **Modeling and forecasting for the number of cases of the COVID-19 pandemic with the curve estimation models, the Box-Jenkins and exponential smoothing methods** *Eurasian J Med Oncol*, 4 (2) (2020), pp. 160-165, [10.14744/ejmo.2020.28273](https://doi.org/10.14744/ejmo.2020.28273) [View Record in ScopusGoogle Scholar](#)

M.A. Al-qaness, A.A. Ewees, H. Fan, M. Abd El Aziz **Optimization method for forecasting confirmed cases of COVID-19 in China** *J Clin Med*, 9 (2020), p. 674 [CrossRefGoogle Scholar](#)

B.F. Maier, D. Brockmann, *Effective containment explains sub-exponential growth in confirmed cases of recent COVID-19 outbreak in mainland china*, *arXiv preprint: 2002.07572* (2020).

Y. Li, M. Liang, X. Yin, X. Liu, M. Hao, Z. Hu, Y. Wang, L. Jin, *Covid-19 epidemic outside china: 34 founders and exponential growth*, *medRxiv* (2020).

M. Raygoza, *Covid-19, exponential growth, and the power of showing up in social solidarity: The math behind the virus* (2020).

Y. Bai, Z. Jin, *Prediction of SARS epidemic by BP neural networks with online prediction strategy*, *Chaos Soliton. Fractal*. 26 (2) (2005) 559–569.

Y.-H. Hsieh, J.-Y. Lee, H.-L. Chang, *Sars epidemiology modeling*, *Emerging Infect. Dis.* 10 (6) (2004) 1165.

D. Lai, *Monitoring the SARS epidemic in china: a time series analysis*, *J. Data Sci.* 3 (3) (2005) 279–293.

W. Wang, S. Ruan, *Simulating the SARS outbreak in beijing with limited data*, *J. Theor. Biol.* 227 (3) (2004) 369–379.

D. Smith, L. Moore, *The SIR model for spread of disease: the differential equation model*, *Loci*. (originally Convergence.) <https://www.maa.org/press/periodicals/Loci/Joma/The-Model-for-Spread-of-Disease-the-Differential-Equation-Model> (2004).

F.R.S. De Gusmao, E.M.M. Ortega, G.M. Cordeiro, *The generalized inverse weibull distribution*, *Statist. Papers* 52 (3) (2011) 591–619.

J.J. Moré, *The levenberg-marquardt algorithm: implementation and theory*, in: *Numerical analysis*, Springer, 1978, pp. 105–116.

P. Pulla, *Covid-19: India imposes lockdown for 21 days and cases rise*, 2020.

O. Analytica, *Japan'S partial COVID-19 lockdown may be insufficient*, *Emerald Expert Briefings* (oxan-es).

J. Thornton, *Covid-19: A&e visits in england fall by 25% in week after lockdown*, 2020. [Y. Zhang, B. Jiang, J. Yuan, Y. Tao, *The impact of social distancing and epicenter lockdown on the COVID-19 epidemic in mainland China: a data-driven SEIQR model study*, *medRxiv* (2020).

R. Niehus, P. Martinez de Salazar Munoz, A. Taylor, M. Lipsitch, *Quantifying bias of COVID-19 prevalence and severity estimates in Wuhan, China that depend on reported cases in international travelers* (2020).

J.T. Wu, K. Leung, G.M. Leung, *Nowcasting and forecasting the potential domestic and international spread of the 2019-ncov outbreak originating in wuhan, china: a modelling study*, *The Lancet* 395 (10225) (2020) 689–697.

Covid-19 science: Why testing is so important | american heart association, (<https://www.heart.org/en/news/2020/04/02/covid-19-science-why-testing-is-so-important>). (Accessed on 05/07/2020).

R. Mancini, S. Tuli, T. Cucinotta, R. Buyya, *iGateLink: A Gateway library for linking IoT, edge, fog and cloud computing environments*, *Proc. Int. Conf. Intell. Cloud Comput.* (2019).

X. Tang, C. Wu, X. Li, Y. Song, X. Yao, X. Wu, Y. Duan, H. Zhang, Y. Wang, Z. Qian, et al., *On the origin and continuing evolution of SARS-cov-2*, *Natl. Sci. Rev.* (2020).

D. Benvenuto, M. Giovanetti, L. Vassallo, S. Angeletti, M. Ciccozzi, *Application of the ARIMA model on the COVID-2019 epidemic dataset*, *Data Brief* (2020) 105340. [38] *There's an unlikely beneficiary of coronavirus: the planets*, link: <https://edition.cnn.com/2020/03/16/asia/china-pollution-coronavirus-hnk-intl/index.html>, [online accessed].

H.S. Maghdid, K.Z. Ghafoor, A.S. Sadiq, K. Curran, K. Rabie, *A Novel AI-enabled Framework to Diagnose Coronavirus COVID 19 using Smartphone Embedded Sensors: Design Study*, *arXiv preprint: 2003.07434* (2020). [40] S. Wang, B. Kang, J. Ma, X. Zeng, M. Xiao, J. Guo