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

CS460 – Midway Presentation

• • •

Nelson Kshetrimayum

Muhammed Jabir T

Analysis

1. Forecasting of COVID-19 in India Using ARIMA Model

- Used python package 'pmdarima', auto_arima function to achieve the relevant values for p, d and q.
- Goodness of fit - R-squared value of 0.994 suggests this be a good model.
- Forecast of cases of the following 10 days.
- Results indicate that the predictions are within the range of 95% confidence interval.

2. Analysis and prediction of COVID-19 trajectory: A machine learning approach

- Build predictive models that can predict the number of positive cases. Regression-based (Nonlinear Regression), Decision tree-based, and Random forest-based are tested and compared for effectiveness.

- The data were split into 80% training for developing the model and 20% testing to validate the effectiveness of the models..

- Mean Absolute Percentage Error -
  - Nonlinear regression - 0.24%
  - Decision tree - 0.18%
  - Random forest - 0.02%

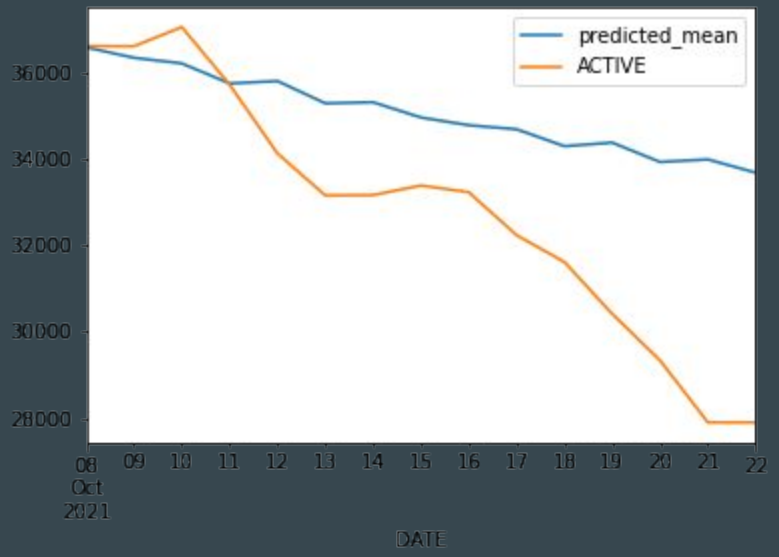
Using ARIMA to forecast for the following 15 days

- From the data we have collected from *covid19india* and following what has been done in one of our reference *Short-term forecasting of the Covid-19 outbreak in India*, we have been able to generate a prediction for the following 15 days.

- For this we have used ARIMA, specifically *auto_arima* function from *pmdarima* library is used to find the best model and the using *ARIMA* from *statsmodels* we have trained and tested.

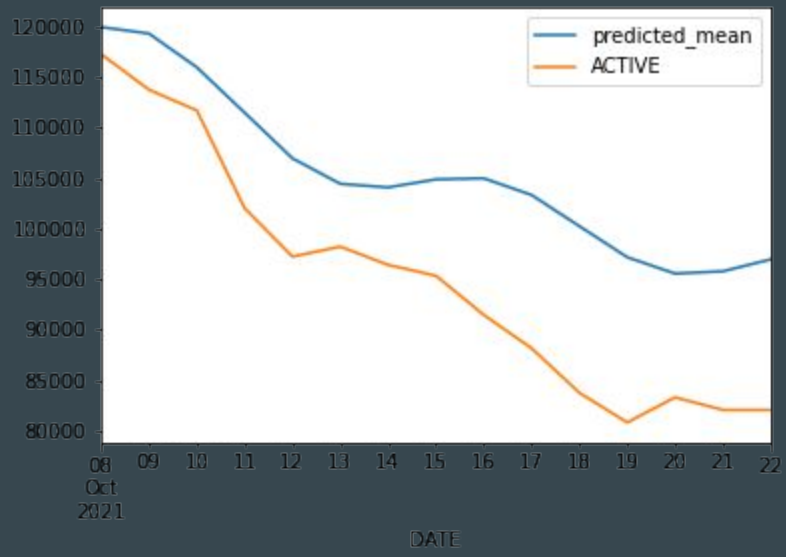
- First we tried on the dataset for India and checked if we can get satisfactory results then proceeded the same procedure with state datas.

What we got: Active cases prediction



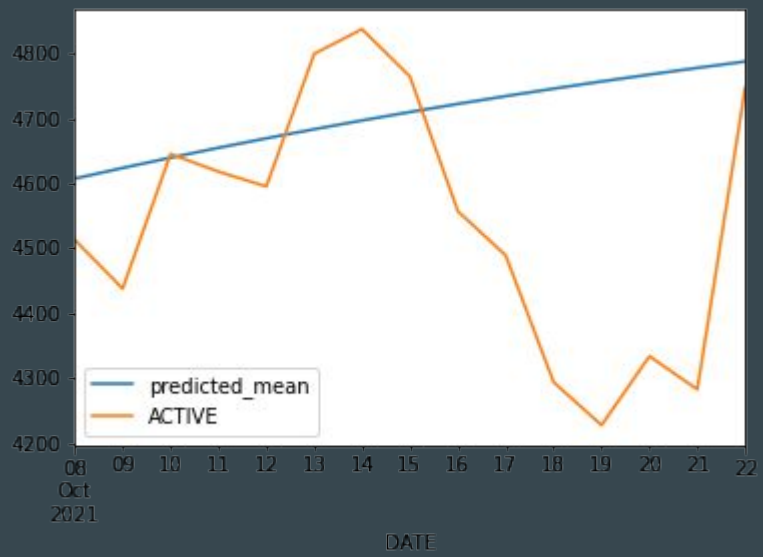
Maharashtra

R-sq - 0.2281
MSE - 6453065.06



Kerala

R-sq - 0.0538
MSE - 128921666.67



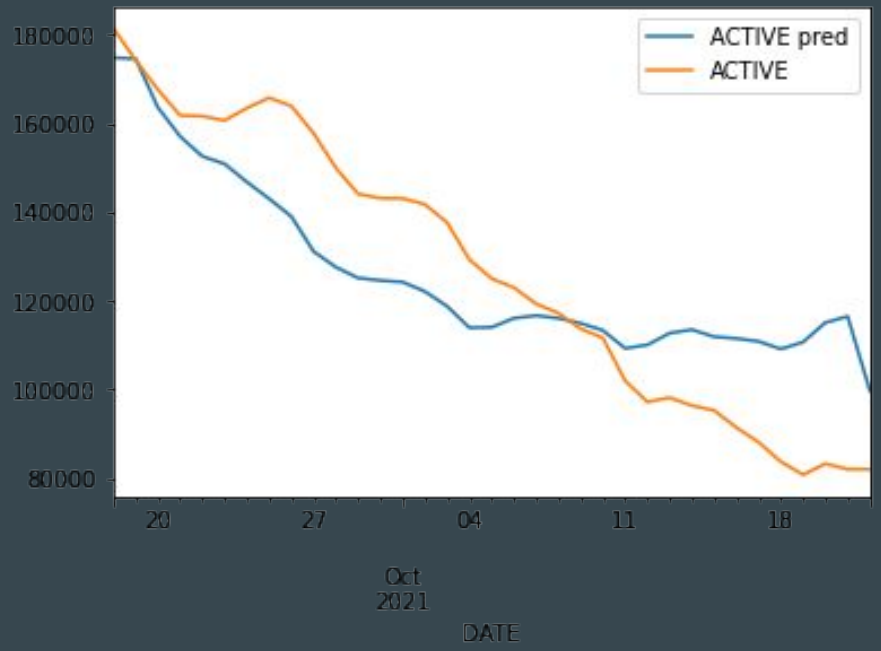
Odisha

R-sq - -0.98
MSE - 72690.91

Using VARMAX (Vector Autoregression Moving-Average with Exogenous Regressors)

- **Endogenous variables**: *Input variables that are influenced by other variables in the system and on which the output variable depends.*
- **Exogenous variables**: *also called covariates, can be thought of as parallel input sequences that have observations at the same time steps as the original series*

- **Endogenous vectors**: Active cases, recovered, deaths
  - Vaccination data?
- **Exogenous vectors: Active testing data**



Integrating COVID-19 compartmental models and DL models

An idea proposal

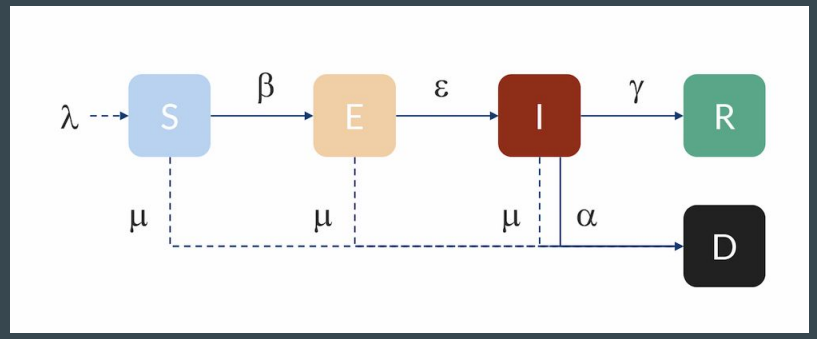
• • •

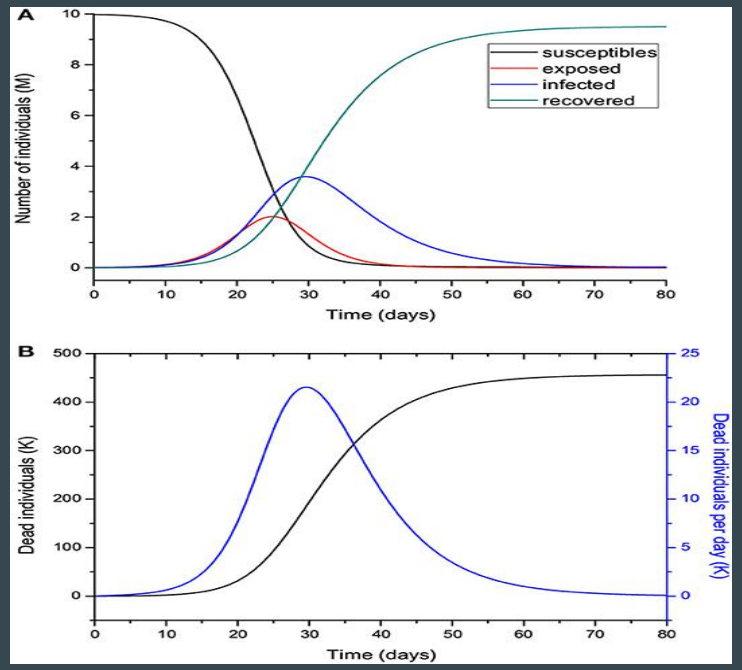
Why neural networks for COVID-19 modelling?

- Neural network is good at describing the characteristics of complex system with strong nonlinearity and difficult to be expressed by precise mathematical model
- adaptive ability
- COVID-19 dynamics within populations is a complex system:
  - random mutation of the virus
  - complexity of social mobility.
- LSTM layer: a type of Recurrent Neural Network (RNN) that utilizes learning from concurrent data.

Compartmental modeling

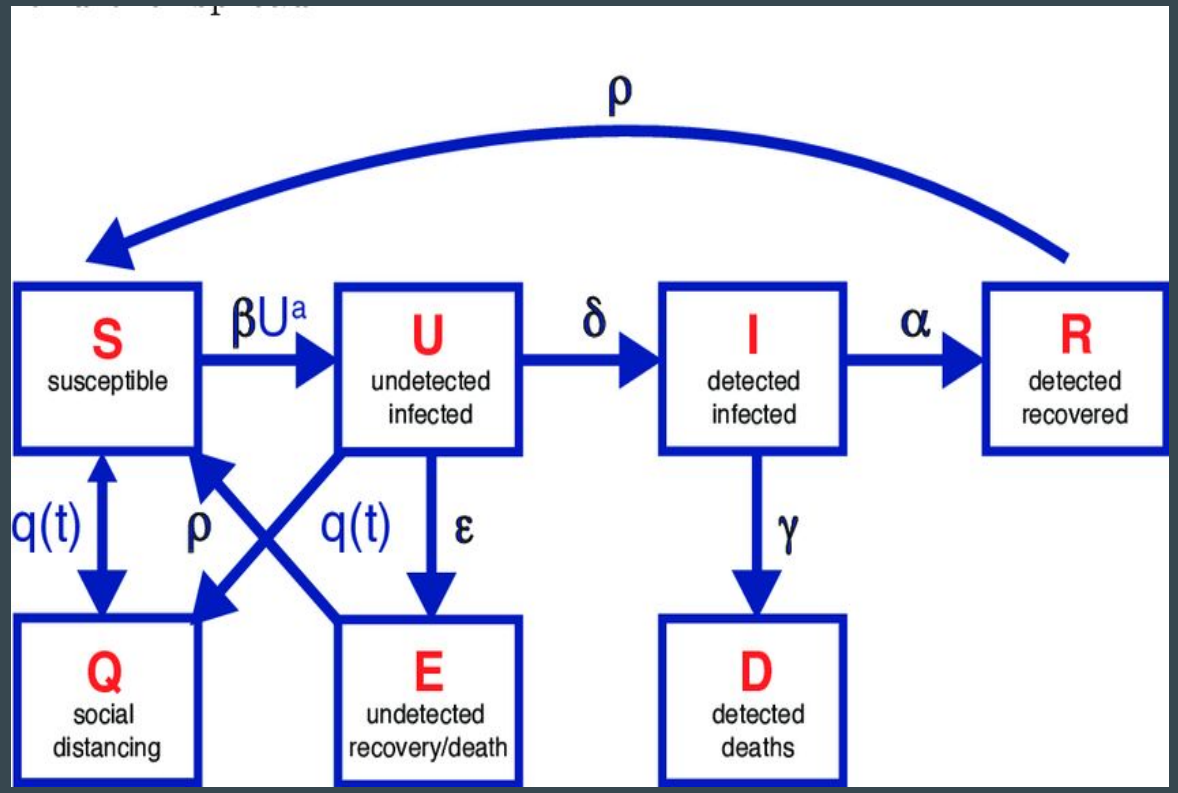
- SEIR:

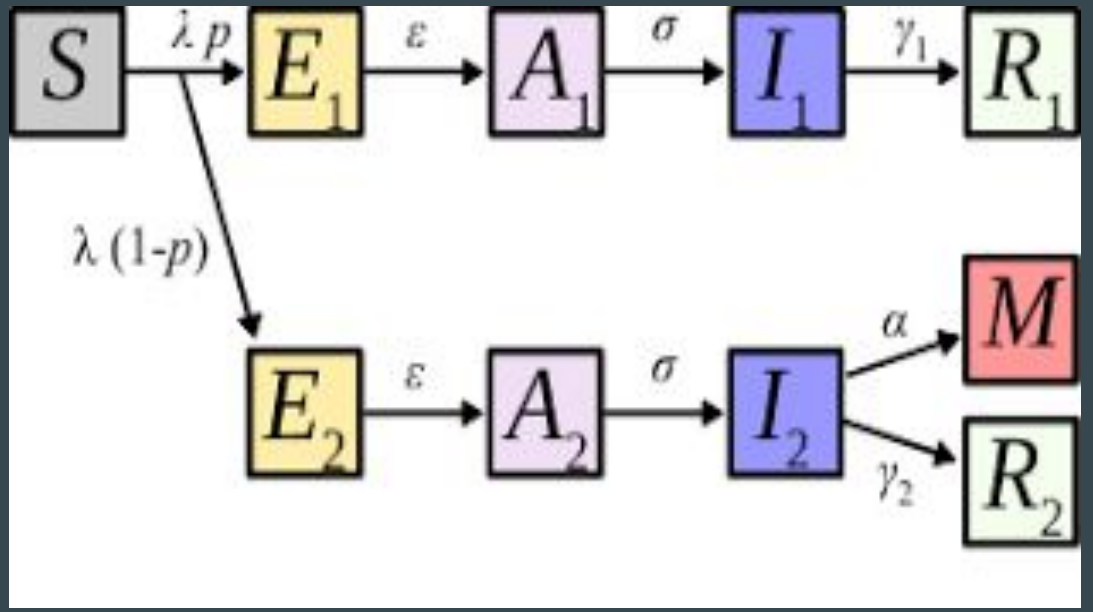




$$\frac{dS}{dt} = -\beta S \frac{I}{N}$$
$$\frac{dE}{dt} = \beta S \frac{I}{N} - \epsilon E$$
$$\frac{dI}{dt} = \epsilon E - (\gamma + \alpha) I$$
$$\frac{dR}{dt} = \gamma I$$
$$\frac{dD}{dt} = \alpha I$$

COVID specific models





Works already done

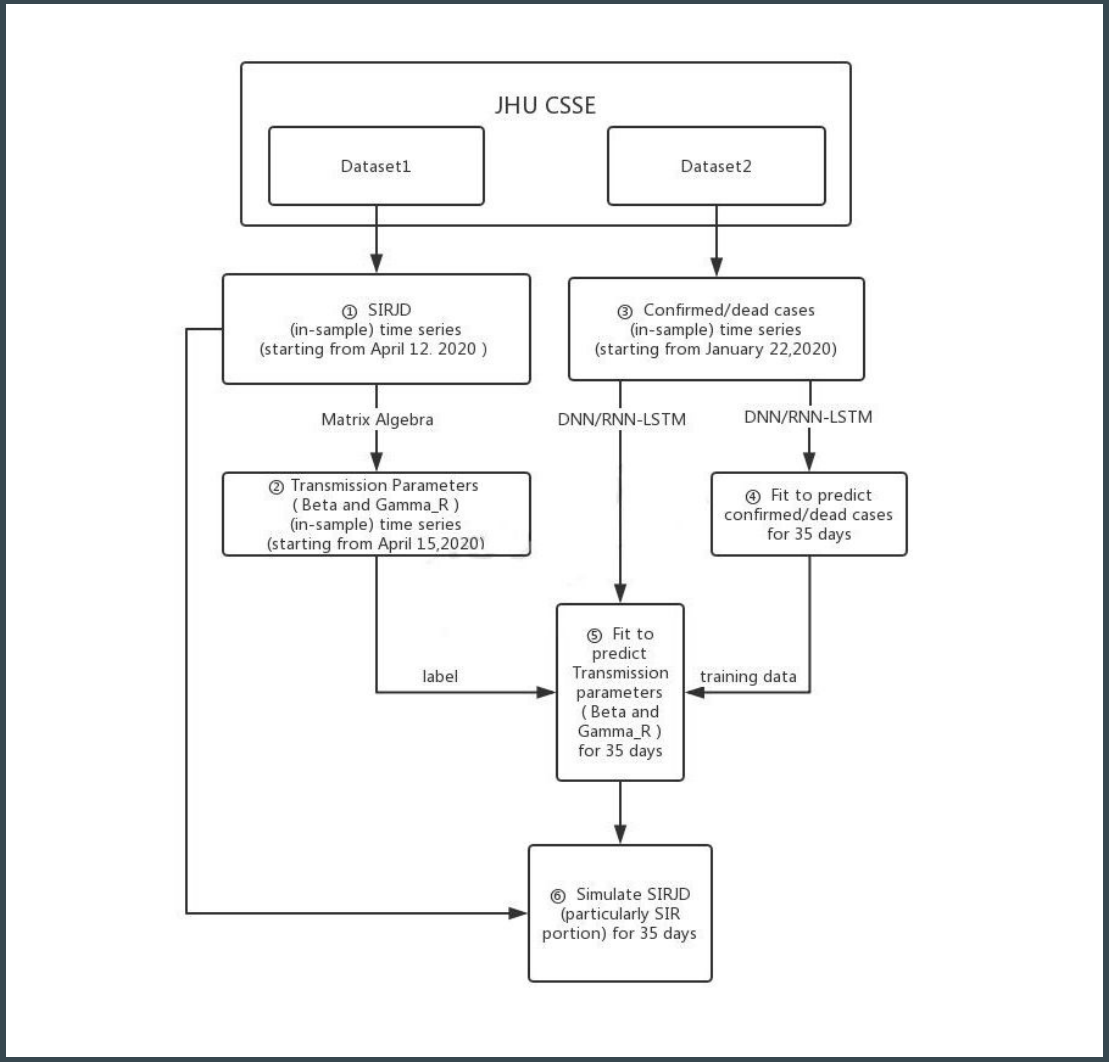
- Worldwide and Regional Forecasting of Coronavirus (Covid-19) Spread using a Deep Learning Model. Cem Direkoglu and Melike Sah
  - Done with deep neural network, which consist of LSTM (Long Short Term Memory) layer, dropout layer, and fully connected layers, to analyze the reported Covid-19 cases and predict the possible future scenarios for the spread in China, Europe, Middle East and worldwide
  - Predictions are done for the next 10 days given the actual time series data of Covid-19
  - Evaluate their approach on the last 3 days of actual data using Root Mean Square Error (RMSE) metric. They present results from the networks that give the minimum RMSE values.
  - As new data arise daily, the network can be re-train in order to adjust the real-time predictions. Further, they present the deep neural network used for forecasting and evaluate their approach with RMSE

Works already done

- Peipei Wang, Xinqi Zheng, Gang Ai, Dongya Liu, Bangren Zhu. Time se-ries prediction for the epidemic trends of COVID-19 using the improved LSTM deep learning method: Case studies in Russia, Peru and Iran,Chaos, Solitons & Fractals,
    - Similar

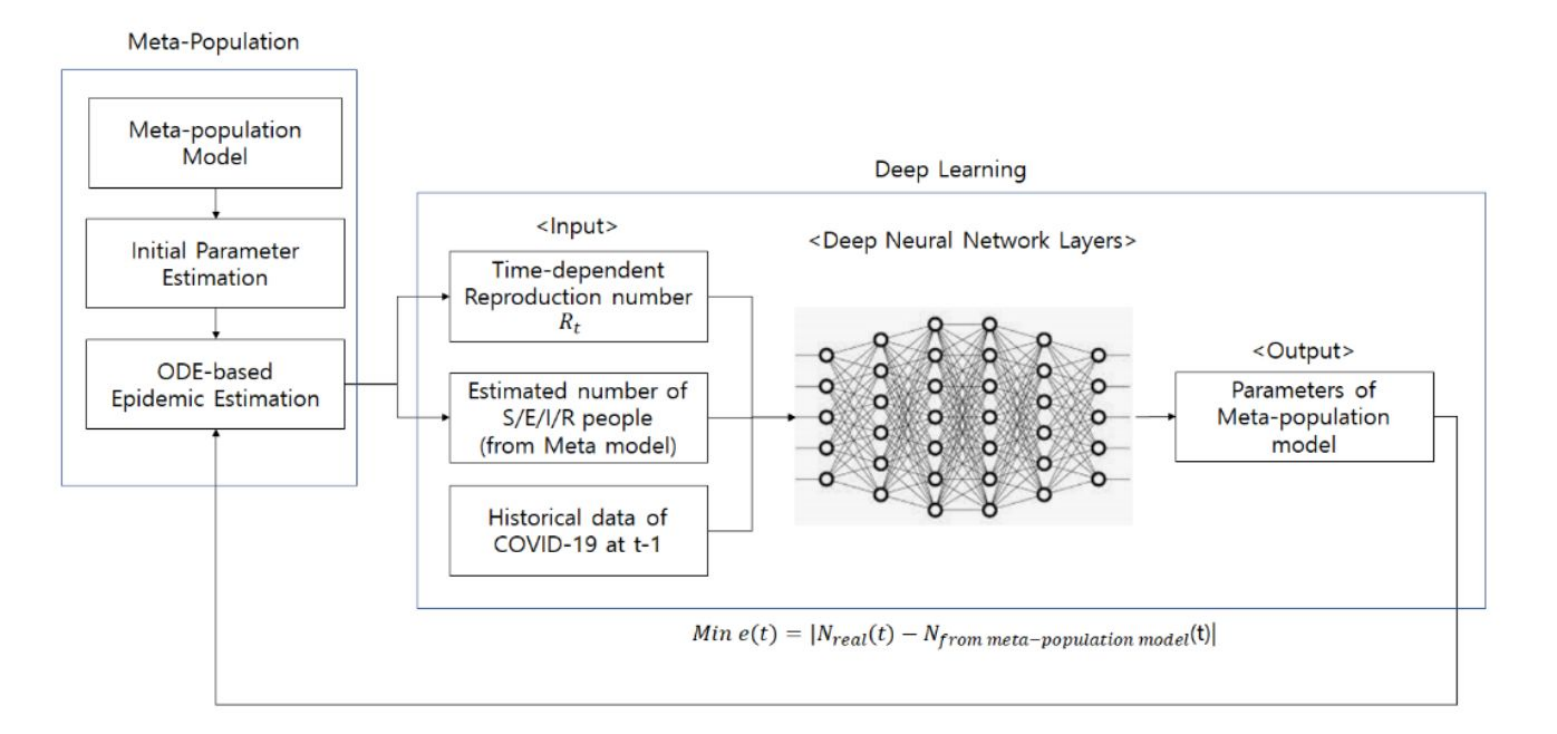
Works already done

- Dynamics and Development of the COVID-19 Epidemic in the United States:A Compartmental Model Enhanced With Deep Learning Techniques, QiDeng.



Works already done

- Hybrid Deep Learning-Based Epidemic Prediction Framework of COVID-19: South Korea Case, Firda Rahmadani and Hyunsoo Lee.



References

- [Kaggle] Shreyas P J, 2021, *Covid forecasting using DL and statistical models.*
- [International Health] Sherry M, Ashok K. P., Md. Arshad & Ubydul H, 2021, *Short-term forecasting of the Covid-19 outbreak in India.*
- [Youtube] Bharani Akella, Great Learning, *Predicting COVID-19 With Machine Learning.*
- [Nature] Yazeed Zoabi, Shira Deri-Rozov & Noam Shomron, 2021, *Machine learning-based prediction of COVID-19 diagnosis based on symptoms.*
- Narayana D, Deepak R, Anwesh R P, Pooja A, H S Nithin, *Forecasting of COVID-19 in India Using ARIMA Model*
- Ritanjali M, Rahul T, Renu PS, Niraj, *Analysis and prediction of COVID-19 trajectory: A machine learning approach*

Thank you