COMPARISON OF DEEP LEARNING METHODS FOR FORECASTING COVID-19 IN INDIA

Cluster Innovation Centre, University of Delhi

Akash Erugurala, Ankit Kumar Bhagat, Vipin Bihari Yadav
akasherugurala5@gmail.com

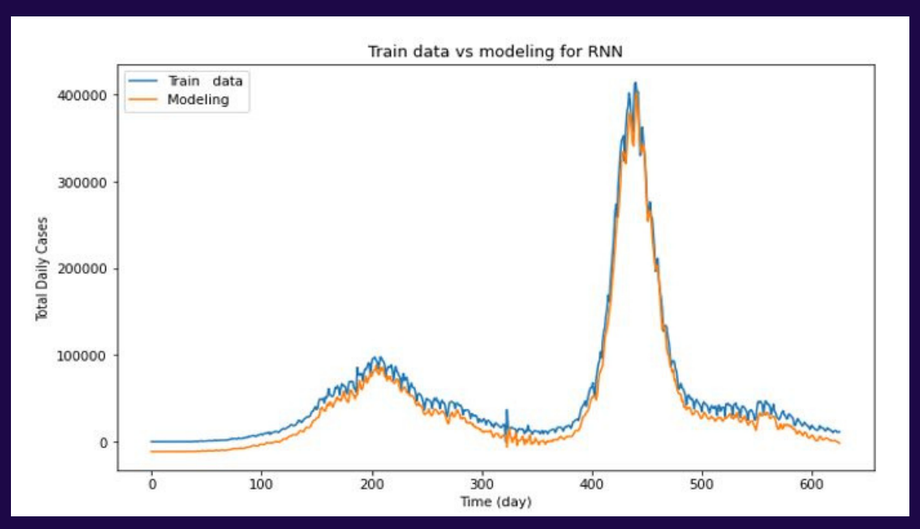
Supervised by: Dr Sachin Kumar

INTRODUCTION

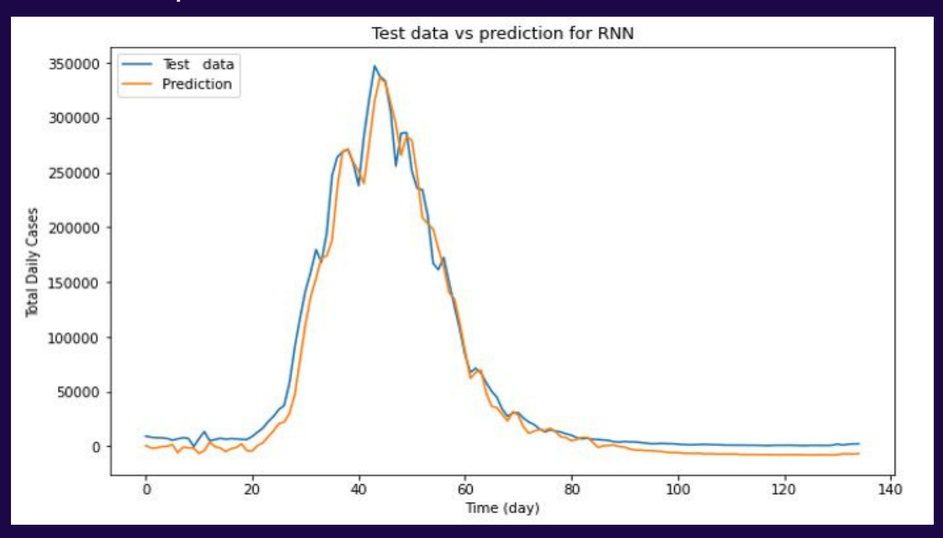
Coronavirus, also known as COVID-19 caused by the SARS-CoV-2 virus. In 2020, World Health Organization (WHO) declared the 2019 novel coronavirus a global pandemic. Covid-19 has caused significant economic, social, and political disruption. In this study, our aim is to model, forecast, and predict the Covid-19 wave using deep learning methods. We have used recurrent neural network (RNN), long short term memory (LSTM), bi-directional LSTM (BiLSTM), and gated recurrent unit (GRU).

EXPERIMENTAL DETAILS

We got the live streaming data from John Hopkins University's GitHub repository. We then modelled our data.



To evaluate the model performance, we plotted test data vs. prediction

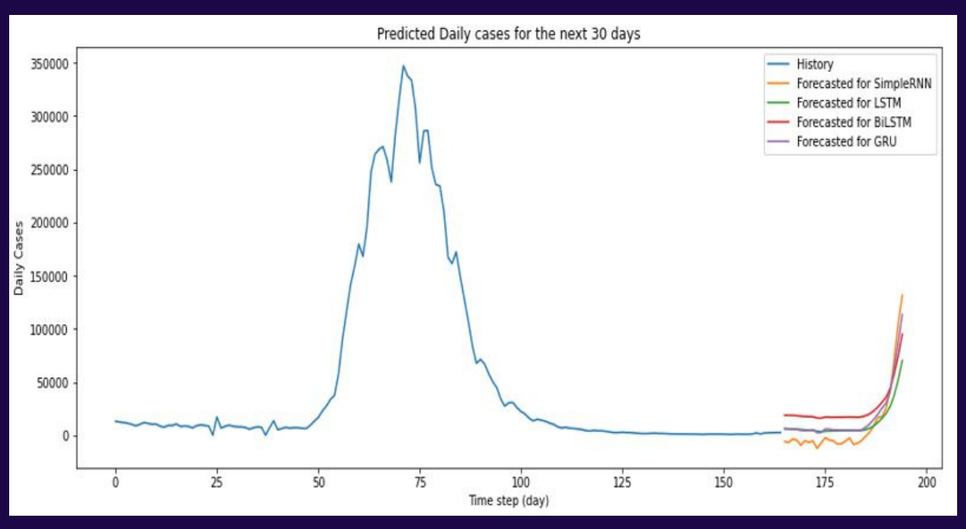


RESULTS AND DISCUSSION

We have calculated the error metrics: MAE, MSE, RMSE, and MDA. We can see that in all the models MDA predicted nearly 60% of data accurately.

	Model	mae	mse	rmse	mda
0	RNN	11291.000768	2.484415e+08	15762.027206	0.696296
1	LSTM	34568.419031	1.463347e+09	38253.720834	0.674074
2	BiLSTM	20093.458740	8.744006e+08	29570.264921	0.525926
3	GRU	11715.162284	4.466697e+08	21134.561326	0.644444

Then we forecasted our data for different interval of days. In the span of 30 days prediction, it is clear that there will be a fourth wave coming soon as all the models showed significant rise in the cases.



CONCLUSION

The choice of error evaluation metric i.e. MAE, MSE, RMSE, and MDA demonstrated that the RNN modelled the data more accurately as there is only one single hidden layer but by resulting graphs, we can see that GRU has achieved better forecasting performance in comparison to all other models. Thus in the future, we will make our dataset multivariate using temperature, population density, and other factors. We will also increase the hidden layers. We will apply ML models like Logistic Regression, Support Vector Machine, and Gradient Boosting algorithm in the future. The plan is to use a combination of other machine learning and deep learning methods to achieve better results.

REFERENCES

Wang, P., Zheng, X., Ai, G., Liu, D., & Zhu, B. (2020). Time series 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, 140, 110214.

Zeroual, A., Harrou, F., Dairi, A., & Sun, Y. (2020). Deep learning methods for forecasting COVID-19 time-Series data: A Comparative Study. Chaos, Solitons & Fractals, 140, 110121.