Proceeding Series of the Brazilian Society of Computational and Applied Mathematics, v. 11, n. 1, 2025.

Trabalho apresentado no XLIII CNMAC, Centro de Convenções do Armação Resort - Porto de Galinhas - PE, 2024

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

Mobility Networks and GNNs in Forecasting COVID-19 Cases in Brazil and their Relationship with Socioeconomic Factors

Fernando H. O. Duarte[1], Antonio Pedro[2], Gladston J. P. Moreira[3], Eduardo J. S. Luz[4], Leonardo B. L. Santos[5], Vander L. S. Freitas[6]

[1,2,3,4,6]Department of Computing, Federal University of Ouro Preto, Ouro Preto, Brazil

[5]National Center for Monitoring and Early Warning of Natural Disasters, São José dos Campos, Brazil

This summary explores predicting COVID-19 case time series in Brazil using GCN (Graph Convolutional Network) based models, a type of Graph Neural Networks (GNN), along with mobility networks [1, 2]. Individual city predictions are made by incorporating city-specific time series data and leveraging subgraphs derived from the connections in the mobility network to evaluate the temporal COVID-19 data. Additionally, the study employs two other models dedicated solely to time series prediction: Prophet [5] and Long Short-Term Memory (LSTM) [3]. The Root Mean Square Error (RMSE) values of COVID-19 forecast models applied in the Brazilian context are summarized in Table 1. Upon analyzing these values, the models can be ranked from the most to the least robust as follows: GCLSTM exhibits the smallest standard deviation (452.59), indicating consistent and reliable performance. Following GCLSTM is GCRN, which has a standard deviation value of 500.39 and achieves the lowest maximum RMSE value of $3,699.74$. Prophet shows competitive performance, with a mean RMSE value of 480.74, similar to LSTM's mean RMSE of 396.71, but with lower maximum and standard deviation RMSE. Lastly, LSTM demonstrates the lowest mean RMSE (396.71) but is characterized by significant variability in errors due to its extremely high Max RMSE ($250,275.07$) and for having the highest standard deviation RMSE ($4,574.69$).

Table 1: RMSE values of the forecast models of COVID-19 cases in Brazil.

Model	**RMSE (Cases)**			
	Mean	Max	Min	Stand.
GCRN	3,059.50	**3,699.74**	2,108.77	500.39
GCLSTM	3,583.88	4,569.97	2,847.56	**452.59**
LSTM	**396.71**	250,275.07	**0.001**	4,574.69
Prophet	480.74	5,1597.08	1.32	1,703.10

We also delve into the correlations between epidemiological time series predictions and the influence regions of Brazilian cities (REGIC) [4], including variables such as: POPMUN (population size), VAR03 (Gross Domestic Product), VAR19 (Territory Management Centrality Score), VAR56 (General Attraction Score), and VAR79 (Quantity of Commercial Categories). Results reveal significant correlations (p-values < 0.05) between Root Mean Square Error (RMSE) and various

[1]fernando.hod@aluno.ufop.edu.br
[2]antonio.pedro@aluno.ufop.edu.br
[3]gladston@ufop.edu.br
[4]eduluz@ufop.edu.br
[5]santoslbl@gmail.com
[6]vander.freitas@ufop.edu.br

variables across the entire Brazilian territory, including POPMUN (population size), VAR03 (Gross Domestic Product), VAR19 (Territory Management Centrality Score), VAR56 (General Attraction Score), and VAR79 (Quantity of Commercial Categories), some obtained from the 2022 Brazilian census [4].

The identified correlations highlight the influence of various factors on the accuracy of COVID-19 prediction models in Brazil. Centrally located municipalities with higher connectivity and larger populations tend to exhibit less precise predictions. Additionally, socioeconomic variables such as GDP (VAR03), municipal attractiveness (VAR56), and governance centrality (VAR19) show significant correlations with prediction model accuracy, suggesting that the complexity of municipalities impacts prediction precision. These findings provide valuable insights into the challenges of modeling and forecasting the spread of COVID-19 in Brazil.

Acknowledgements

The authors thank the CNPq, grants 441016/2020-0, 307151/2022-0, 308400/2022-4, FAPEMIG, grants APQ-01518-21, APQ-01647-22, CAPES, grant 88887.506931/2020-00, and Universidade Federal de Ouro Preto (UFOP).

References

[1] Fernando Henrique Oliveira Duarte, Gladston J. P. Moreira, Eduardo J. S. Luz, Leonardo B. L. Santos, and Vander L. S. Freitas. "Time Series Forecasting of COVID-19 Cases in Brazil with GNN and Mobility Networks". In: **Intelligent Systems**. Ed. by Murilo C. Naldi and Reinaldo A. C. Bianchi. Cham: Springer Nature Switzerland, 2023, pp. 361–375. ISBN: 978-3-031-45392-2.

[2] Fernando Henrique Oliveira Duarte, Gladston J.P. Moreira, Eduardo J.S. Luz, Leonardo B.L. Santos, and Vander L.S. Freitas. "Correlations between epidemiological time series forecasting and influence regions of Brazilian cities". In: Cited by: 0. 2023, pp. 363–368. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85181098306&partnerID=40&md5=f60d73eee623d21c9b7119bbc4407319.

[3] Sepp Hochreiter and Jürgen Schmidhuber. "Long short-term memory". In: **Neural computation** 9.8 (1997), pp. 1735–1780.

[4] IBGE. **Regiões de influência das cidades : 2018**. IBGE, Coordenação de Geografia Rio de Janeiro, Brazil, 2020.

[5] Sean J Taylor and Benjamin Letham. "Forecasting at scale". In: **The American Statistician** 72.1 (2018), pp. 37–45.