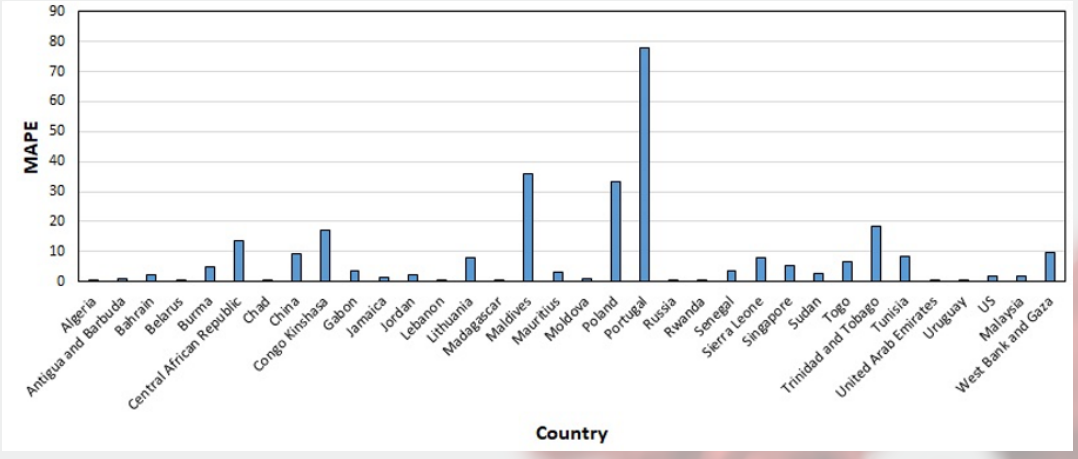
UAEU مكتب النائب المشارك للبحث العلمي Office of the Associate Provost for Research
جامعة الإمارات العربية المتحدة United Arab Emirates University

COVID-19 RESEARCH NEWSLETTER

Issue 27, 14 October 2020

Tailoring Machine Learning Models for Accurate Prediction of COVID-19 Infections - Case Studies of 187 Countries

Computational and Structural Biotechnology Journal 2020, In press
https://doi.org/10.1016/j.csbj.2020.09.015



Is there a second wave of COVID-19 pandemic? Are the current precautionary measures effective? Machine learning methods have been used to predict the number of infections of COVID-19 and to tackle the pandemic situation. For effective pandemic precautionary measures, it is necessary to obtain reliable predictions. Given the significant difference in the geographical characteristics and social behaviors of the countries under study, the use of a single model to predict disease spread and rate of infection becomes questionable, as it is highly likely that the model would fail to generate accurate predictions of the infection rates. Such a failure may lead to greater distress, economic hardship and possibly increased death rate.

In this study, the research team presents a detailed review of the time series models used for the prediction of the coronavirus spread. They conducted 2,805 experiments to develop these time-series models for 187 countries. Then they developed a method to tailor time series machine-learning modeling for accurate forecasting. This method takes into consideration the spatial distribution of infections over time for a country and develops the corresponding accurate machine-learning model. They achieved an accuracy of 98.93%. The results of this study show that the 187 countries exhibit nine different data distribution trends, which were mapped into nine time series models.

The authors made the following conclusions: for countries with *exponentially increasing* data trend, *ARIMA* gives the best accuracy; for data following the *exponential+linear* trend, *Holt's Linear Trend* has the best prediction; the *logistic* data trend maps to *S-curve Trend*; for infections data having *exponential/logarithmic* damping trend, *Damped Trend* has the most accurate forecast; for data following *polynomial* distribution, *Quadratic Trend* has the best prediction; the *stochastic* data trend maps to *LSTM*; for countries with *cyclic-increasing+constant* data trend, *Holt-Winters' Additive* gives the highest accuracy; for data following *cyclic-exponential+constant* trend, *SSM* has the best prediction; the *constant+sharp* increase data trend maps to *Sutte-ARIMA*.

Research Team: **Leila Ismail, Huned Materwala , Taieb Znati, Sherzod Turaev** (College of Information Technology, UAEU), and **Moien Khan** (College of Medicine and Health Sciences, UAEU)

If you are interested in sharing your COVID-19 releated research, please send your contribution to research.office@uaeu.ac.ae

Editors: Prof. Mohsen Sherif & Dr. Ranjit Vijayan