

COVID-19 RESEARCH NEWSLETTER

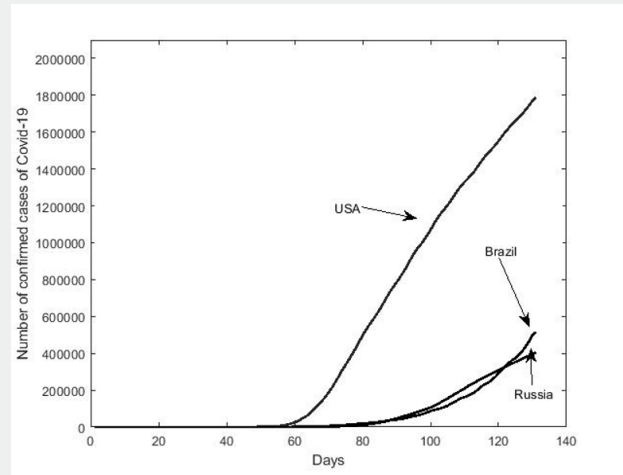
Issue 23, 6 September 2020

Predicting the Number of Confirmed Cases of COVID-19 by using Machine Learning: Methods and Challenges

Archives of Computational Methods in Engineering, 2020 (In press)

<https://doi.org/10.1007/s11831-020-09472-8>

COVID-19 is one of the biggest health challenges that the world has ever faced. Public health policy makers need reliable predictions of the future Covid-19 cases to plan medical facilities. Machine learning methods learn from the historical data and make predictions about the events and hence have been used to predict the number of confirmed cases of Covid-19. The authors present a detailed review of related research papers. They present a taxonomy that groups the papers into four categories. The taxonomy identifies four research themes: *traditional machine learning regression, deep learning regression, network analysis, and social media and search queries data-based methods.*



The authors made the following conclusions:

1. Time series framework is very popular in prediction of the confirmed cases. In many countries the first case of COVID-19 came in January 2020 or after. Therefore, till 30th May 2020, the number of data points is less than 150. It is difficult to train accurate machine learning models with such small datasets. Deep learning methods are successful because of large training data which is not available for Covid-19 confirmed cases prediction task.
2. Pandemics are rare and the characteristics of COVID-19 are different than those of other coronaviruses such as SARS and MERS. Therefore, that data cannot be used for COVID-19.
3. About 80% of the people with COVID-19 have mild or asymptomatic cases. Many of them do not get tested. Therefore, the numbers of confirmed cases in countries are not accurate. However, these people contribute to the new cases.
4. Social stigma attached to COVID-19 in many countries force suspected COVID-19 patients to stay away from medical facilities. Therefore, many confirmed cases are not recorded. However, they contribute to new confirmed cases.
5. Models in one country cannot be applied to other countries easily as they have different factors.
6. Machine learning experts should work with epidemiologists to understand the data well and to select the parameters of machine learning methods.
7. Models trained on different types of data such as times series data, social media data, etc. may be combined to predict the number of confirmed cases.

Research Team: Amir Ahmad (College of Information Technology, United Arab Emirates University, UAE), Sunita Garhwal (Department of Computer Science and Engineering, Thapar University, India), Santosh Kumar Ray (Department of Information Technology, Khawarizmi International College, UAE), Gagan Kumar (Department of Physics, Indian Institute of Technology Guwahati, India), Sharaf Jameel Malebary (Faculty of Computing and Information Technology, King Abdulaziz University, Saudi Arabia), and Omar Mohammed Barukab (Faculty of Computing and Information Technology, King Abdulaziz University, Saudi Arabia).

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