World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:17, No:02, 2023

Improved Computational Efficiency of Machine Learning Algorithm Based on Evaluation Metrics to Control the Spread of Coronavirus in the UK

Authors : Swathi Ganesan, Nalinda Somasiri, Rebecca Jeyavadhanam, Gayathri Karthick

Abstract : The COVID-19 crisis presents a substantial and critical hazard to worldwide health. Since the occurrence of the disease in late January 2020 in the UK, the number of infected people confirmed to acquire the illness has increased tremendously across the country, and the number of individuals affected is undoubtedly considerably high. The purpose of this research is to figure out a predictive machine learning archetypal that could forecast COVID-19 cases within the UK. This study concentrates on the statistical data collected from 31st January 2020 to 31st March 2021 in the United Kingdom. Information on total COVID cases registered, new cases encountered on a daily basis, total death registered, and patients' death per day due to Coronavirus is collected from World Health Organisation (WHO). Data preprocessing is carried out to identify any missing values, outliers, or anomalies in the dataset. The data is split into 8:2 ratio for training and testing purposes to forecast future new COVID cases. Support Vector Machines (SVM), Random Forests, and linear regression algorithms are chosen to study the model performance in the prediction of new COVID-19 cases. From the evaluation metrics such as r-squared value and mean squared error, the statistical performance of the model in predicting the new COVID cases is evaluated. Random Forest outperformed the other two Machine Learning algorithms with a training accuracy of 99.47% and testing accuracy of 98.26% when n=30. The mean square error obtained for Random Forest is 4.05e11, which is lesser compared to the other predictive models used for this study. From the experimental analysis Random Forest algorithm can perform more effectively and efficiently in predicting the new COVID cases, which could help the health sector to take relevant control measures for the spread of the virus.

Keywords : COVID-19, machine learning, supervised learning, unsupervised learning, linear regression, support vector machine, random forest
Conference Title : ICDSTA 2023 : International Conference on Data Science Techniques and Applications
Conference Location : London, United Kingdom
Conference Dates : February 16-17, 2023