ABSTRACT

Accuracy of Support Vector Machines and K Nearest Neighbor in Predicting the Number of COVID-19 Cases in Indonesia

Chandra [1], Nur Uddin,Ph.D [2], Lathifah Alfat, S.T ., M.T. [3]
[1]*Student of Informatics Department, Universitas Pembangunan Jaya*
[2]*Lecturer of Informatics Department, Universitas Pembangunan Jaya*
[3]*Lecturer of Informatics Department, Universitas Pembangunan Jaya*

Corona virus is a virus that can cause disease in humans and animals. In humans it usually causes respiratory infections, ranging from the common cold to serious illnesses such as MERS syndrome and SARS syndrome. The limited research regarding the prediction of the number of COVID-19 cases in Indonesia makes researchers interested in raising cases regarding the prediction of the number of COVID-19 cases using machine learning methods, namely SVM (Support Vector Machine) regression type which will be compared with KNN (K-Nearest Neighbor) for predicting the Daily Case of COVID-19 in Indonesia. Based on the following, the purpose of this study is to provide predictive modeling and compare the accuracy values using the SVM and KNN algorithms on the number of COVID-19 cases in Indonesia. The method in this research is quantitative. The first stage of data collection is conducting a Literature Study and collecting various historical data on daily COVID-19 cases in Indonesia from March 9, 2020 to February 28, 2022 using Time Series data. The second stage of data collection is carried out during the Processing and Building Machine Learning process using the SVM and KNN algorithms. The third stage of data collection is carried out when predicting the actual data with the predicted data and comparing the accuracy between the SVM algorithm and KNN. This study uses three ratios, namely 90:10, 80:20 and 70:30. Accuracy and prediction results on the Support Vector Regression algorithm have better performance in predicting COVID-19 cases in Indonesia when the ratio is 80:20, which is 96.38% and the RMSE value is 2309, while the K Nearest Neighbor algorithm is at a ratio of 90:10, which is 93.75. % and an RMSE value of 2888, it can be concluded that the SVR and KNN algorithms produce good predictive data.

Keywords : Corona virus, Machine Learning, Support Vector Machine, K-Nearest Neighbor.

Libraries   : 59

Publication Years   : 2017 – 2022