ABSTRAK

COVID-19, first identified in Wuhan, China in December 2019, quickly spread worldwide and was declared a pandemic by WHO in March 2020. Indonesia reported its first case on March 2, 2020, and the pandemic has had a significant impact on the country's economic, social, and health sectors. This study aims to predict the death rate due to COVID-19 in Jakarta using multiple linear regression method. The dataset collected from Andra Farm - Go Green website includes COVID-19 cases recorded in all sub-districts in Jakarta on November 1, 2023. Pre-processing was performed to improve the quality and accuracy of the model. The method used was multiple linear regression. The analysis results show that variables such as total travel and discarded trip have a significant influence in predicting the number of positive cases. The study found that lowering the correlation threshold for selecting independent variables reduced the mean squared error (MSE) and improved model performance, highlighting the importance of variable selection in developing accurate predictive models. These findings provide important insights for the government in making informed decisions regarding post-pandemic healthcare. This research underscores the value of robust data processing and variable selection techniques in enhancing predictive accuracy for public health planning.

Kata Kunci: Covid-19; Correlation matrix; Multiple linear regression; Jakarta; Positive;