

ABSTRAK

COVID-19 merupakan penyakit yang pertama kali ditemukan di Wuhan, China dan mengakibatkan pandemi *coronavirus* 2019-2020. Virus ini dapat menyebabkan infeksi saluran pernafasan seperti flu ketika menyerang manusia. Negara Indonesia dinyatakan memiliki kasus konfirmasi tertinggi di ASEAN. Beberapa penelitian telah dilakukan untuk menangani kasus-kasus COVID-19 dengan algoritma *Decision Tree* ataupun Regresi Linear. Pada penelitian ini, peneliti mencoba melakukan analisis terhadap kasus-kasus COVID-19 di Indonesia dan melakukan sebuah uji coba memprediksi kasus-kasus baru dengan algoritma *Decision Tree* beserta Regresi Linear. Kemudian peneliti akan membandingkan nilai dari kedua algoritma ini dengan menggunakan R^2 Score. Hasil analisis ini menyatakan bahwa provinsi DKI Jakarta memiliki jumlah kasus positif, kesembuhan dan kematian tertinggi di Indonesia. Nilai hasil perbandingan dari R^2 Score yang didapatkan pada algoritma *Decision Tree* mencapai 95.69% (*training*) dan 92.15% (*testing*), sedangkan algoritma Regresi Linear mencapai 79.93% (*training*) dan 77.25% (*testing*).

Kata Kunci: *COVID-19, Data Mining, Decision Tree, Regresi Linear*

ABSTRACT

COVID-19 is a disease that was first discovered in Wuhan, China and caused the 2019-2020 *coronavirus* pandemic. This virus can cause respiratory tract infections such as flu when infecting humans. Indonesia is declared to have the highest confirmed cases in ASEAN. Several studies have been carried out to handle COVID-19 cases with the *Decision Tree* or Linear Regression algorithm. In this study, we tried to analyze COVID-19 cases in Indonesia and conducted a trial predicting new cases with the *Decision Tree* and Linear Regression algorithms. Then we will compare the values of these two algorithms using R^2 Score. The results of this analysis state that DKI Jakarta province has the highest number of positive cases, cures and deaths in Indonesia. The value of the comparison results from the R^2 Score obtained in the *Decision Tree* algorithm reached 95.69% (*training*) and 92.15% (*testing*) while the Linear Regression algorithm reached 79.93% (*training*) and 77.25% (*testing*).

Keywords: *COVID-19, Data Mining, Decision Tree, Linear Regression*