

1. Introduction

Based on the information in [1], the coronavirus or Covid-19 affected 214 countries and regions worldwide with 38,404,464 cases, 1,091,569 deaths, and 28,874,023 recovered. This information was taken on October 14, 2020, at 4:17 pm Indonesia time.

Covid-19 pandemic is caused by a coronavirus that first appeared in Wuhan, China, at the end of 2019. Symptoms that often appear are fever, fatigue, and dry cough. This disease attacks all age elements in society worldwide. It is vulnerable to the elderly and people with medical records such as heart disease, diabetes, and respiratory problems [2]. On March 2, 2020, President Joko Widodo announced that the coronavirus exposed two Indonesian citizens. Both are a 64-year-old woman, and her daughter, who is 31 years old, had positive Covid-19 after contact with a Japanese citizen [3]. On August 23, 2020, according to Johns Hopkins University, Indonesia is in the 23rd position in the world confirmed 153,535 positive Covid-19 cases. The data put of 0.66 percent globally, which achieved approximately 23 million. Indonesia is in the 19th position because of the mortality rate, which is 6,680 fatalities [4].

Implementing health protocols, such as washing hands or cleaning it with alcohol, maintaining a minimum distance of two meters from people around them, avoiding crowds, and paying attention to balanced nutrition so that the body is always healthy and fit, can reduce the new cases and death rate [5]. Hence, a prediction model of Covid-19 plays an essential role in making the right decisions to control the outbreak. Many researchers in Indonesia use machine learning-based models to predict the new cases and new deaths in real-time based on the Covid-19 time series dataset. They exploit various methods, such as Gaussian Naive Bayes (GNB), and Support Vector Machine (SVM), and Decision Tree (DT) that provide relatively high accuracies in many cases of predictions. Those methods have advantages and disadvantages. GNB is quite simple to implement, while GNB and DT are relatively complex. SVM works for a continuous dataset, while GNB and DT can be suitable for discrete and continuous datasets. DT is traceable [6], while GNB and SVM are not. They also give different performances in terms of effectiveness (accuracy) and efficiency (cost and time). Since they are statistical data-driven methods, various datasets and parameter tuning may produce different results.

This research investigates three prediction models of GNB, SVM, and DT using the Covid-19 time series dataset in Indonesia. They are implemented using the same tools and run on the same computer to compare their performances in accuracy and processing time.