

Abstract

The main cause of Tuberculosis (TB), a specific infectious disease that affects people worldwide, is Mycobacterium Tuberculosis (MTB). An estimated 30% of the population worldwide has a TB infection, which causes over 20 million deaths annually. Also, 37.7 million people are afflicted with HIV and TB together. Detecting TB in HIV patients is crucial due to the high risk associated with TB. To identify HIV-positive patients, RNA-based methods are used to find host gene expression signatures associated with different aspects of the disease. Nevertheless, no group in this method describes gene signatures that can be used to identify patients who are co-infected with TB and HIV. Therefore, a method is needed to identify TB in HIV patients. This study aims to classify high-dimensional microarray data using Grey Wolf Optimization (GWO) with Support Vector Machines (SVM). To improve the performance of the model, hyperparameter tuning was carried out. Based on the results, we obtained the optimal SVM model using a linear kernel that outperforms other kernels in terms of accuracy, with F1-score values of 0.78 and 0.80, respectively.

Keywords: Tuberculosis, HIV, Microarray, Data Ekspresi Gen, *Grey Wolf Optimization* (GWO), *Support Vector Machine* (SVM)
