

Prediction of Tuberculosis on HIV Patients Based on Gene Expression Data Using Grey Wolf Optimization-Support Vector Machine

Hana Amani Fatihah¹, Hasmawati², Isman Kurniawan³

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

¹hanaamanii@students.telkomuniversity.ac.id, ²hasmawati@telkomuniversity.ac.id,

³ismankrn@telkomuniversity.ac.id

Abstrak

Tuberculosis merupakan penyakit yang disebabkan oleh *Mycobacterium tuberculosis* (MTB) yang mana penyakit infeksi tersebut menjadi penyebab yang utama di dunia. Diperkirakan 30% dari populasi dunia terinfeksi TB dengan jumlah kematian sekitar 20 juta orang dan 37,7 juta orang yang terinfeksi HIV dan TB. Pentingnya dilakukan deteksi TB pada pasien HIV karena tingginya risiko pada TB. Metode yang digunakan untuk mendeteksi pasien yang terjangkit HIV adalah menggunakan metode berbasis asam ribonukleat (RNA) yang berfungsi untuk mendeteksi spesifik host gene expression signature yang berkorelasi dengan berbagai aspek penyakit, namun pada metode ini belum adanya kelompok yang menggambarkan gene signature yang mampu mengidentifikasi pasien co-infeksi HIV dan TB. Oleh karena itu, dibutuhkan metode yang mampu mengidentifikasi TB pada penderita HIV. Pada penelitian ini, penggunaan Grey Wolf Optimization (GWO) yang digabungkan dengan Support Vector Machines (SVM) bertujuan untuk mengklasifikasikan data microarray yang berdimensi tinggi. Hyperparameter tuning digunakan untuk meningkatkan kinerja model. Dalam hasil eksperimen, kami menemukan bahwa kinerja SVM dengan kernel linear lebih baik daripada kernel lain dengan nilai akurasi dan F1-score masing-masing 0,78 dan 0,80.

Kata kunci: Tuberculosis, HIV, Microarray, Gene Expression Data, Grey Wolf Optimization, Support Vector Machine
