

ABSTRAK

Tuberkulosis (TB) masih menjadi salah satu penyebab utama kematian di dunia, dan proses diagnosisnya menjadi semakin kompleks ketika terjadi koinfeksi dengan HIV. Dalam beberapa tahun terakhir, profil ekspresi gen dari data microarray telah muncul sebagai pendekatan yang menjanjikan untuk deteksi TB. Namun, tingginya dimensi data dan keterbatasan jumlah sampel menimbulkan tantangan serius bagi model machine learning, yang sering mengakibatkan performa yang tidak stabil. Metode ensemble konvensional seperti Bagging, AdaBoost, dan Uniform Weighted Ensemble telah diterapkan pada domain ini, namun cenderung menghasilkan kinerja yang tidak seimbang—mencapai presisi tinggi namun mengorbankan recall—yang dapat menjadi masalah serius dalam pengambilan keputusan klinis, di mana melewatkan satu kasus positif tidak dapat diterima. Untuk mengatasi tantangan tersebut, penelitian ini mengusulkan kerangka kerja klasifikasi weighted ensemble yang dioptimasi menggunakan Firefly Algorithm (FA). Metode ini mengombinasikan beberapa base classifier dalam kerangka Bagging, dengan bobot adaptif yang ditentukan oleh FA untuk meningkatkan keseimbangan prediksi dan ketahanan model. Dataset yang digunakan terdiri dari profil ekspresi gen pasien TB-HIV koinfeksi, yang terlebih dahulu dinormalisasi dan direduksi dimensinya melalui seleksi fitur sebelum dilakukan pelatihan. Hasil evaluasi menunjukkan bahwa model baseline terbaik, Bagging dengan Decision Tree, mencapai nilai AUC sebesar 0.7755 dan presisi sempurna (1.0000), namun dengan recall rendah (0.4286). Sebaliknya, Decision Tree ensemble yang dioptimasi dengan FA menghasilkan performa yang lebih baik dan seimbang, dengan AUC sebesar 0.8367, akurasi 0.7857, presisi 0.8333, dan recall 0.7143. Temuan ini menunjukkan bahwa penggabungan ensemble learning dengan optimasi metaheuristik dapat secara efektif menangani data biomedis berdimensi tinggi dengan jumlah sampel kecil, sekaligus mempertahankan performa prediksi yang bermakna secara klinis.

Kata kunci: Deteksi tuberkulosis, HIV-positif, data ekspresi gen, weighted ensemble, Firefly Algorithm, bagging, klasifikasi.