

ABSTRAK

Osteoporosis merupakan penyakit tulang yang proses penemuan obat konvensional memakan waktu lama dan biaya tinggi. Penelitian ini berfokus pada prediksi bioaktivitas senyawa inhibitor enzim Cathepsin K, yang menjadi target penting untuk terapi osteoporosis. Risiko seumur hidup osteoporosis pada wanita dan pria berkisar antara 40%–50% dan 13%–22%. Namun, proses penemuan obat konvensional seringkali memakan waktu dan mahal. Meskipun penelitian sebelumnya telah menerapkan machine learning, masih terdapat celah untuk peningkatan performa, terutama dalam proses seleksi fitur dari data kimia yang memiliki dimensi sangat tinggi. Optimasi pemilihan fitur menjadi kunci untuk membangun model prediksi yang lebih efisien dan akurat. Solusi yang diusulkan adalah membangun model prediksi dengan mengintegrasikan algoritma Glowworm Swarm Optimization (GSO) untuk seleksi fitur dan XGBoost sebagai model klasifikasi. Dataset inhibitor Cathepsin K diolah dan dibagi menjadi empat kelas bioaktivitas. Algoritma GSO diterapkan untuk memilih sub-set fitur deskriptor molekuler yang paling relevan dan informatif dari ratusan fitur awal. Selanjutnya, model klasifikasi XGBoost dioptimalkan melalui penyesuaian hyperparameter untuk memaksimalkan performa prediksinya berdasarkan fitur-fitur yang telah dipilih oleh GSO. Hasil pengujian menunjukkan model GSO-XGBoost hasil optimasi terbaik mampu mencapai performa puncak dengan akurasi sebesar 0.84 dan F1-Score rata-rata tertimbang 0.84. Penelitian ini memberikan kontribusi berupa kerangka kerja yang efektif dalam mengombinasikan GSO dan XGBoost untuk prediksi bioaktivitas, yang berpotensi meningkatkan efisiensi pada tahap awal proses penemuan obat untuk penyakit tulang.

Kata Kunci: Osteoporosis, Cathepsin K, GSO, XGBoost, Bioaktivitas Inhibitor, Prediksi.