

ABSTRAK

Kanker paru-paru menjadi permasalahan global dalam menyebabkan kematian akibat kanker. Salah satu penyebab kematian kanker ini yaitu terapi yang seragam untuk semua pasien. Pengobatan presisi perlu dikembangkan untuk mengatasi heterogenitas respons pasien kanker, sehingga efektivitas terapi meningkat dan mencegah efek samping. Integrasi data antara proteomik dengan struktur molekul obat telah menjadi strategi untuk memprediksi respons obat. Pendekatan algoritma *machine learning* diperlukan untuk mengoptimalkan model prediksi. Penelitian ini bertujuan untuk menemukan klasifikasi terbaik dalam prediksi respons obat menggunakan tiga algoritma *machine learning* yaitu Random Forest, XGBoost, dan *Multi-layer Perceptron* (MLP) yang unggul dalam memodelkan data berdimensi tinggi. Algoritma ini dikombinasikan dengan dua representasi molekul yaitu Morgan Fingerprint dan *Graph Neural Networks* (GNN) yang mampu mengubah struktur kimia menjadi vektor numerik. Pendekatan GNN menghasilkan fitur lebih informatif dengan tingkat *sparsity* sangat rendah yaitu mencapai 8,55%, dibandingkan Morgan Fingerprint. Arsitektur model terbaik yaitu kombinasi XGBoost dan GNN dengan nilai evaluasi *Area Under Precision-Recall Curve* (AUPRC) mencapai 95,57%, serta dievaluasi lebih lanjut menggunakan *Confusion Matrix*. Penelitian ini menunjukkan integrasi data proteomik dan representasi molekul berbasis GNN dengan algoritma XGBoost menghasilkan prediksi respons obat yang lebih optimal dan mampu mendukung pengembangan pengobatan presisi.

Kata Kunci: kanker paru-paru, pengobatan presisi, proteomik, respons obat, *machine learning*, representasi molekul, *Graph Neural Networks*, XGBoost.