

Implementasi Metode Ensemble pada Klasifikasi Inhibitor CDK2 sebagai Agen Anti Kanker

Mela Mai Anggraini, Isman Kurniawan, Erwin Budi Setiawan

Fakultas Informatika, Universitas Telkom, Bandung
melamaianggraini@students.telkomuniversity.ac.id, ismankrn@telkomuniversity.ac.id,
erwinbudisetiawan@telkomuniversity.co.id

Abstrak

Pengobatan kemoterapi secara jangka panjang memiliki efek samping dan resistensi sel terhadap obat tertentu. Oleh karena itu, dibutuhkan obat baru yang dapat mengurangi efek samping dan memberikan efek pengobatan yang lebih baik. Pada umumnya, obat-obat anti kanker dikembangkan dengan mempertimbangkan *Cyclin Dependent Kinase 2* (CDK2) sebagai targetnya. Desain obat secara konvensional diketahui tidak efektif dan efisien karena untuk mendapatkan senyawa baru dengan aktivitas biologis tertentu perlu dilakukan sintesis terlebih dahulu untuk mengetahui aktivitasnya. Untuk mendesain obat baru, dibutuhkan suatu model yang dapat memprediksi aktivitas kandidat obat. Proses prediksi tersebut dapat dilakukan dengan menggunakan suatu model matematika yang dapat menentukan korelasi struktur dan aktivitas atau dikenal dengan *Quantitative Structure-Activity Relationships* (QSAR). Penelitian ini bertujuan membangun model prediksi untuk klasifikasi dengan metode *ensemble* yaitu *XGBoost*, *Random Forest* dan *Adaboost*. Penelitian dilakukan dengan mengimplementasikan metode ansambel tersebut pada setiap *fingerprint* (*Estate*, *Extended*, *Maccs* dan *Pubchem*). Berdasarkan hasil yang didapatkan, *Random Forest* dengan *fingerprint Pubchem* merupakan model prediksi terbaik dibandingkan dengan metode lainnya. Hal ini ditunjukkan dengan nilai *Matthews Correlation Coefficient* (MCC) dan *Area Under the ROC Curve* (AUC) yang paling besar yaitu berturut-turut, MCC=0.979 dan AUC=0.999.

Kata Kunci: QSAR, CDK2, *XGBoost*, *Random Forest*, *Adaboost*