

ABSTRAK

Demam berdarah adalah salah satu penyakit endemik yang ditularkan oleh virus *dengue* dan masih sering menyebar di beberapa negara, terutama negara dengan iklim tropis. Hingga saat ini, belum ada obat khusus yang dapat digunakan untuk mengobati demam berdarah. Penelitian pada salah satu obat yang diselidiki secara klinis, yaitu *balapiravir*, dinilai belum efektif dalam menghambat aktivitas replikasi virus *dengue*. Protein *non-structural 3* (NS3) diketahui dapat digunakan sebagai target pengembangan obat yang cocok untuk menemukan komposisi baru yang dapat bekerja lebih baik. Model *quantitative structure activity relationship* (QSAR) dapat digunakan karena model ini dianggap valid untuk memprediksi dan mengklasifikasikan aktivitas biologis terhadap senyawa yang belum diuji. Penelitian ini bertujuan untuk membangun model QSAR dalam memprediksi aktivitas inhibitor NS3. Proses klasifikasi dilakukan menggunakan *feature importance* untuk seleksi fitur serta metode *ensemble* dengan algoritma *random forest*, *adaptive boosting* (AdaBoost) serta *extremely randomized trees* untuk membangun model prediksi. Prosedur *hyperparameter tuning* dilakukan untuk membantu meningkatkan kinerja model. Berdasarkan analisis validasi, model 9 yang berisi tujuh deskriptor molekuler dengan algoritma *extremely randomized trees* menghasilkan akurasi terbaik sebesar 0.73 (73%) dan AUC sebesar 0.82 dibandingkan dengan model lain. Model ini juga bukan karena korelasi kebetulan semata. Hal ini dibuktikan dengan 10 kali percobaan *y-scrambling* menghasilkan nilai koefisien korelasi lebih rendah dibandingkan dengan nilai koefisien korelasi yang asli.

Kata Kunci : protein *non-structural 3* (NS3), *quantitative structure activity relationship* (QSAR), metode *ensemble*, algoritma *random forest*, algoritma *adaptive boosting* (AdaBoost), algoritma *extremely randomized trees*