

Prediksi Bioaktivitas Inhibitor Cathepsin K untuk Pengobatan Penyakit Tulang Menggunakan Long Short-Term Memory yang Dioptimalkan dengan Simulated Annealing

Alfiansyah Hafidz Arbi Putra¹, Isman Kurniawan²

Fakultas Informatika, Universitas Telkom, Bandung

hanifaditia@student.telkomuniversity.ac.id,

ismankrn@telkomuniversity.ac.id

Abstrak

Osteoporosis, suatu penyakit tulang yang mempengaruhi lebih dari 200 juta orang di seluruh dunia, merupakan tantangan terapeutik yang signifikan, dengan Cathepsin K (CatK) sebagai target utama untuk pengembangan inhibitor karena perannya dalam resorpsi tulang. Sementara metode penemuan obat konvensional sering kali lambat dan mahal, pembelajaran mesin menawarkan alternatif yang menjanjikan. Penelitian ini menjawab kebutuhan akan model prediktif yang lebih akurat dengan mengembangkan kerangka kerja yang kokoh untuk menilai bioaktivitas inhibitor CatK. Jaringan Long Short-Term Memory (LSTM), yang dipilih karena kemampuannya dalam menangani data berurutan yang kompleks khas struktur molekul, dioptimalkan menggunakan metaheuristik Simulated Annealing (SA). Model ini dilatih menggunakan dataset yang terdiri dari 1568 molekul dari database ChEMBL, dengan bioaktivitas diklasifikasikan berdasarkan nilai pIC_{50} ke dalam empat kategori: Poten ($pIC_{50} \geq 9$), Aktif ($9 > pIC_{50} \geq 7$), Menengah ($7 > pIC_{50} \geq 6$), dan Tidak Aktif ($pIC_{50} < 6$). Model LSTM yang dioptimalkan dengan SA secara signifikan mengungguli tiga model LSTM dasar, yang mencapai akurasi rata-rata puncak sebesar 0,77. Konfigurasi SA yang optimal (skema col_rate95) mencapai akurasi rata-rata dan skor F1 sebesar 0,81. Yang menarik, model ini menunjukkan kinerja yang luar biasa dalam mengidentifikasi inhibitor Poten, dengan skor F1 sebesar 0,92. Namun, keterbatasan utama adalah kesulitan dalam membedakan antara kelas Aktif dan Menengah, di mana kesalahan klasifikasi lebih sering terjadi. Penelitian ini menyoroti efektivitas pendekatan SA-LSTM dalam mempercepat penemuan senyawa bioaktivitas tinggi untuk pengobatan osteoporosis. Pekerjaan di masa depan dapat berfokus pada peningkatan ketangguhan model dengan mengintegrasikan deskriptor molekul tambahan atau mengeksplorasi arsitektur pembelajaran mendalam alternatif untuk meningkatkan akurasi klasifikasi.

Kata Kunci: osteoporosis, cathepsin, long short-term memory, simulated annealing
