

Abstrak

Kanker telah menjadi salah satu penyakit paling mematikan di dunia, yang dimana penyakit ini terutama disebabkan oleh akumulasi mutasi somatik dan turunan. Namun, fenomena ini dapat ditelusuri kembali ke level molekuler, spesifiknya ke ranah protein. Protein adalah molekul yang bertanggung jawab atas berbagai bioproses dalam tubuh manusia melalui interaksinya dengan molekul lain. Abnormalitas dalam interaksi ini dapat menyebabkan hal yang tidak diinginkan, diantaranya termasuk penyakit dan kanker. Peptida memiliki potensi untuk digunakan sebagai molekul dalam interaksi protein untuk mengobati kanker. Namun, identifikasi peptida yang sesuai dengan protein target di laboratorium memakan waktu dan juga biaya. Oleh karena itu, diperlukan metode komputasi dalam membantu identifikasi. TabNet, sebuah metode komputasi berbasis *deep learning* digunakan dalam penelitian ini. Sebagai perbandingan, kami memilih metode dari *ensemble learning*, diantaranya Random Forest dan Extreme Gradient Boosting, dan juga metode dari *deep learning* lainnya seperti Convolutional Neural Network dan Stacked Autoencoder-Deep Neural Network. Prediksi dilakukan pada dataset interaksi peptida-protein multifitur, dengan fiturnya meliputi *position-specific scoring matrices*, *intrinsic disorder*, *amino acid sequence*, dan *physicochemical properties*. Di antara metrik yang digunakan, kami menemukan bahwa TabNet mencapai skor yang lebih baik dalam AUC yaitu 0,7 dan juga *false negative* yang lebih rendah dibandingkan dengan model lain.

Kata Kunci: Kanker, Interaksi Peptida-Protein, TabNet