

Prediksi Interaksi antara Protein HIV-1 dan Manusia Menggunakan Algoritma POLS dan Machine Learning = Prediction of Protein Interactions between HIV-1 Proteins and Humans Using POLS Algorithms and Machine Learning

Tesdiq Prigel Kaloka, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20493636&lokasi=lokal>

Abstrak

ABSTRAK

Protein merupakan bagian penting dari organisme dan memiliki fungsi yang berbeda. Fungsi dan sifat interaksi protein dapat diketahui dengan mengelompokkan protein-protein yang saling berinteraksi. Objek penelitian ini adalah interaksi antara protein HIV-1 dan manusia. Biclustering merupakan metode yang dapat digunakan untuk menyelesaikan permasalahan interaksi protein. Interaksi dibagi menjadi dua, yaitu interaksi positif dan negatif, selanjutnya diubah menjadi graf bipartit dengan simpul merupakan protein HIV-1 dan protein manusia, sedangkan busur merupakan jenis interaksi yang terjadi. Algoritma POLS merupakan algoritma biclustering yang menggunakan pendekatan teori graf. Hasil bicluster dianalisis menggunakan Gene Ontology (GO) untuk memperoleh fungsi protein pada satu bicluster. Proses terakhir adalah prediksi interaksi protein berdasarkan analisis fungsi-fungsi protein. Metode yang digunakan adalah Support Vector Machine (SVM) karena SVM merupakan metode prediksi machine learning yang robust. Berdasarkan hasil penelitian, dataset interaksi positif terdapat 297 bicluster dengan bicluster terbanyak berukuran 2 X 2 dan bicluster terbesar berukuran 7 X 7. Dataset interaksi negatif diperoleh 203 bicluster dengan 110 bicluster berukuran 2 X 2 dan satu bicluster berukuran 7 X 7. Berdasarkan hasil analisis GO, terdapat protein dalam satu bicluster yang belum diketahui fungsinya. Akurasi model prediksi untuk interaksi positif = 92% dan interaksi negatif = 88%.

ABSTRACT

Protein is an important part of the organism. The function of protein can be known by grouping the interact proteins. This research discusses the interaction between HIV-1 and human protein. Biclustering is a method to solve protein interaction problem. The interaction is divided into two types, called positive and negative interactions. The interaction is transformed into a bipartite graph with vertices are HIV-1 and human protein, while the edges are the interaction. POLS algorithm is a biclustering method based on graph theory. The result of a bicluster is analyzed using Gene Ontology (GO). The last process is the prediction of protein interactions based on analysis of GO. We used Support Vector Machine (SVM) because SVM is a robust machine learning method for prediction. Based on the results, we get 297 biclusters, with 171 biclusters sized 2 X 2 and the largest bicluster sized 7 X 7 for the positive interactions. For the negative interaction, we get 203 biclusters, with 110 biclusters sized 2 X 2 and the largest bicluster sized 7 X 7. Based on GO analysis there were an unknown function in a bicluster. Accuracy of prediction models for positive and negative interaction are 92% and 88% respectively.