



INTISARI

PREDIKSI STRUKTUR SEKUNDER PROTEIN MENGGUNAKAN CONVOLUTIONAL NEURAL NETWORK

Zaki Indra Sukma
12/330923/PA/14406

Struktur sekunder protein adalah salah satu struktur penting dalam menentukan bentuk tiga dimensi protein yang menentukan fungsi protein. Menentukan struktur sekunder protein dari struktur primernya menggunakan pendekatan konvensional adalah proses yang sukar dan mahal. Pendekatan komputasi dapat digunakan untuk memprediksi struktur sekunder protein, salah satu metodenya yaitu convolutional neural network.

Convolutional neural network adalah salah satu metode dalam *deep learning*, telah digunakan untuk klasifikasi citra dan pengenalan suara dengan hasil yang sangat baik. Convolutional neural network merupakan jaringan syaraf tiruan yang di dalamnya terdapat operasi konvolusi dan pooling, dan memiliki jaringan yang dalam untuk meningkatkan akurasi pembelajaran.

Dalam penelitian ini dibangun model prediksi struktur sekunder protein dengan arsitektur convolutional neural network. Arsitektur yang dibangun memiliki 3 convolutional layer dan diberikan modifikasi dengan teknik *shift and stitch*. Dengan pencarian hyperparameter optimal yang mencakup panjang kernel, dropout, regularisasi L2, dan perbandingan *input profile*, didapatkan hasil berupa akurasi Q8 mencapai 70.36% dan akurasi Q3 mencapai 82.06% pada dataset CullPDB.

Kata kunci: prediksi struktur sekunder protein, convolutional neural network, *deep learning*, akurasi Q8, akurasi Q3, CullPDB



ABSTRACT

PROTEIN SECONDARY STRUCTURE PREDICTION USING CONVOLUTIONAL NEURAL NETWORK

Zaki Indra Sukma
12/330923/PA/14406

The secondary structure of protein is one of the important structure to determine the three-dimensional form of protein in which determine protein's function. Determining secondary structure of protein using conventional approach is difficult, yet expensive. Computational approach can be leveraged to predict the secondary structure of protein, and convolutional neural network is one such method.

Convolutional neural network is one method in deep learning, has been used to image classification, speech recognition with decent results. Convolutional neural network is neural network in which has convolution and pooling, and also has deep network to improve learning accuracies.

In this research, protein secondary structure prediction model is built using convolutional neural network architecture. Architecture which has been build has 3 convolutional layer and modification is given by shift and stitch technique. By optimal hyperparameter search that includes kernel length, dropout, L2 regularization, and input profile comparison, result is achieved. Q8 accuracy of 70.36%B and Q3 accuracy of 82.06% is gained on CullPDB dataset.

Keywords: protein secondary structure prediction, convolutional neural network, deep learning, Q8 accuracy, Q3 accuracy, CullPDB