

## INTISARI

### **PREDIKSI STRUKTUR SEKUNDER PROTEIN MENGGUNAKAN *SUPPORT VECTOR MACHINES* (SVM)**

Oleh

DEFFI ALFIANI MULYANINGRUM

12/334811/PA/15020

Protein merupakan elemen yang sangat penting di dalam tubuh makhluk hidup. Fungsi protein dapat terlihat apabila telah terjadi pelipatan sekuens protein menjadi struktur tiga dimensi. Sehingga, dengan memahami bagaimana pelipatan tersebut terjadi, akan sangat membantu untuk mengetahui fungsi protein, yang selanjutnya dapat dimanfaatkan lebih lanjut, seperti dalam pembuatan obat untuk penyakit tertentu. Pada penelitian ini dilakukan prediksi struktur sekunder protein menggunakan SVM.

*Preprocessing* yang dilakukan menggunakan metode *sliding window* dan dua metode pengkodean, yaitu *orthogonal input profile* dan *BLOSUM62 matrix input profile*. Fungsi kernel digunakan dalam penelitian ini untuk memetakan fitur ke ruang dimensi tinggi. Terdapat tiga fungsi yang digunakan, yaitu *linear kernel*, *polynomial kernel*, dan *RBF kernel*. Adapun untuk menghitung akurasi, digunakan  $Q_3$  yang dapat menghitung akurasi sistem prediksi struktur sekunder protein untuk klasifikasi dengan tiga kelas.

Hasil akurasi  $Q_3$  terbaik adalah sebesar 63,6638% yang diperoleh dengan *preprocessing* menggunakan *sliding window* berukuran tiga belas asam amino dan pengkodean *BLOSUM62 matrix input profile*. Fungsi kernel yang digunakan yaitu *RBF kernel* dengan parameter  $c=32$  dan  $\gamma=0,0078125$ .

**Kata kunci:** *SVM, klasifikasi, prediksi struktur sekunder protein*

## ABSTRACT

### PROTEIN SECONDARY STRUCTURE PREDICTION USING SUPPORT VECTOR MACHINES (SVM)

By

DEFFI ALFIANI MULYANINGRUM

12/334811/PA/15020

Protein is a very important element in the body of living things. The function of protein can be seen if there has been a folding of the protein sequence into a three-dimensional structure. So, by understanding how folding occurs, it is help us to know the function of proteins, which can be used, such as in making drugs for certain diseases. In this study, the protein secondary structure prediction using SVM method.

Preprocessing is done using the sliding window method and two encoding methods, orthogonal input profile and BLOSUM62 matrix input profile. The kernel function is used in this study to map features to high-dimensional space. There are three functions that are used, linear kernel, polynomial kernel, and RBF kernel. As for calculating its accuracy, Q3 is used which can calculate the accuracy of protein secondary structure prediction for classification with three classes.

The best Q3 accuracy results were 63.6638% obtained by preprocessing using thirteen amino acid sliding windows and BLOSUM62 matrix input profile encoding. The kernel function used is the RBF kernel with parameters  $c = 32$  and  $\gamma = 0.0078125$ .

**Keyword:** *SVM, classification, protein secondary structure prediction.*