

# Prediksi interaksi protein berdasarkan barisan asam amino pada human immunodeficiency virus menggunakan global encoding dikombinasikan dengan rotation forest ensemble classifier = Sequence based prediction of protein protein interactions in human immunodeficiency virus using global encoding combined with rotation forest ensemble classifier

Mohamad Irvan Septiar Musti, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20467562&lokasi=lokal>

---

## Abstrak

### <b>ABSTRAK</b><br>

HIV Human Immunodeficiency Virus adalah sebuah jenis retrovirus obligat intraseluler yang menyerang sistem kekebalan tubuh manusia. Virus ini menyerang dengan cara melakukan interaksi antara protein virus dengan protein manusia. Penelitian ini menggunakan data berupa barisan asam amino dari protein yang akan diubah fiturnya menggunakan metode global encoding. Hasil ekstraksi fitur tersebut kemudian akan digunakan sebagai masukan untuk metode rotation forest guna memprediksi interaksi protein HIV dengan manusia. Selain itu pula, penelitian ini juga membandingkan performa metode rotation forest yang menggunakan Principal Component Analysis RF PCA dengan rotation forest yang menggunakan Independent Principal Component Analysis RF IPCA sebagai metode transformasi peubah bebas dalam metode tersebut. Hasil dari penelitian ini menunjukkan bahwa RF PCA memperoleh hasil performa tertinggi dalam memprediksi interaksi protein HIV dengan protein manusia, yaitu dengan nilai akurasi sebesar 79,50 , sensitivitas 79,91 , spesifisitas 79,07 dan presisi sebesar 79,77 . Sementara itu, metode RF IPCA memperoleh hasil performa tertinggi yaitu dengan nilai akurasi sebesar 77,20 , sensitivitas 76,65 , spesifisitas 77,81 , dan presisi sebesar 79,40 . Selain itu pula, dalam penelitian ini ditemukan sebanyak 2.619 protein manusia yang terprediksi berinteraksi dengan protein HIV melalui model terbaik RF PCA , dan juga ditemukan sebanyak 3.071 protein manusia yang terprediksi berinteraksi dengan protein HIV melalui model terbaik RF IPCA dari total sebanyak 7.678 protein manusia yang diteliti.

<hr />

### <b>ABSTRACT</b><br>

HIV Human Immunodeficiency Virus is a type of retrovirus obligate intracellular that attacks the human body immune system. This virus attacks by doing interaction between virus and human proteins. This research uses data of amino acids sequence from protein that the feature will be extracted using Global Encoding. The result of feature extraction then would be used as an input for Rotation Forest in order to predict interaction between HIV and human proteins. In addition, this research also compares the performance of Rotation Forest that using Principal Component Analysis RF PCA with Independent Principal Component Analysis RF IPCA as a method of transformation in that method. The result shows that RF PCA produced highest performance in classifying protein interactions between HIV and human, with accuracy value of 79,50 , 79,91 sensitivity, 79,07 specificity and 79,77 precision. While the RF IPCA produced highest performance with 77,20 accuracy, 76,65 sensitivity, 77,81 specificity, and 79,40 precision. In addition, there are 2.619 human protein which is predicted has an interaction with HIV protein through RF PCA best model, and there are 3.071 human protein which is predicted has an interaction with HIV protein through RF IPCA best model from the total of 7.678 human protein. All of that can be found in this

research.