PENERAPAN METODE *SUPPORT VECTOR MACHINE* (SVM) DALAM MENGKLASIFIKASI DATA BIOINFORMATIKA

Oleh: Ilham Safitrah

Program Studi Statistika Fakultas MIPA

Universitas Islam Indonesia

INTISARI

Human Immunodeficiency Virus atau HIV adalah virus yang menyerang sel darah putih di dalam tubuh yang mengakibatkan turunnya kekebalan tubuh manusia. Orang yang dalam darahnya terdapat virus HIV dapat tampak sehat dan belum tentu membutuhkan pengobatan. Epidemiologi demensia (penurunan kemampuan fungsi otak) terkait infeksi yang ter-dokumentasikan dengan baik saat ini adalah demensia terkait HIV-1 (HIV-1 *associated dementia*/HAD). Prevalensi HAD di Asia Pasifik sebesar 12%, di Afrika Selatan dan Uganda masing-masing sebesar 25,4% dan 31%, dan di negara maju sebesar 10%. Terkait permasalahan HIV-1 data diambil dari website NCBI, GSE28160, yang terdiri dari 35 pasien dengan 9 orang tidak terjangkit dan 26 orang terjangkit HIV-1 dengan banyaknya gen 54675. Sebelum analisis dilakukan *pre-processing, filtering, dan feature selection*. Support Vector Machine (SVM) menggunakan kernel linear memberikan akurasi sebesar 85,7%.

Kata Kunci :*HIV, HIV-1, Klasifikasi, Suppor Vector Machine*.

SUPPORT VECTOR MACHINE (SVM) APLICATION FOR BIOINFORMATICS DATA

By: Ilham Safitrah

Departement of Statistics, Faculty of Mathematics and Natural Sciences

Islamic University of Indonesia

ABSTRACT

Human Immunodeficiency Virus or HIV is a virus that attacks white blood cells in the body resulting in decreased human immunity. People who have HIV in their blood can look healthy and do not necessarily need treatment. The epidemiology of dementia (decreased ability of brain function) related to infection which is well documented at the moment is dementia related to HIV-1 (HIV-1 associated dementia / HAD). The prevalence of HAD in the Asia Pacific is 12%, in South Africa and Uganda respectively at 25.4% and 31%, and in developed countries by 10%. Regarding HIV-1 issues, data was taken from the NCBI website, GSE28160, which consisted of 35 patients with 9 people not infected and 26 people infected with HIV-1 with a large number of genes 54675. Before the analysis was carried out pre-processing, filtering, and feature selection. Support Vector Machine (SVM) using a linear kernel gives an accuracy of 85.7%.

Keywords :*HIV, HIV-1, Classification, Suppor Vector Machine*.