

Klasifikasi subtype molekular kanker payudara menggunakan naive bayes classifier = Breast cancer molecular subtype classification using naive bayes classifier

Nabilla Ayu Fauziyyah, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20486909&lokasi=lokal>

Abstrak

ABSTRACT

Dewasa ini, sudah banyak rumah sakit modern yang dilengkapi dengan peralatan monitoring yang lengkap, yang menyebabkan makin banyaknya data medis yang tersimpan. Data medis ini memiliki karakteristik khusus, dan biasanya metode statistika biasa tidak dapat diterapkan begitu saja. Dari sinilah kemudian muncul gagasan mengenai Medical Data Mining (MDM) yang sudah terbukti cocok untuk diterapkan dalam analisis data medis. Naive Bayes Classifier (NBC) merupakan salah satu implementasi dari MDM. Kendati terbukti memiliki hasil yang akurat dan memuaskan dalam proses diagnosis medis, metode-metode dalam MDM belum sepenuhnya diterima dalam praktek medis untuk diterapkan. Alasan utama mengapa metode ini belum dapat diterima adalah karena terdapatnya resistansi dari tenaga medis terhadap metode diagnosis yang baru. Tujuan dari penelitian ini adalah untuk menerapkan dan mengevaluasi performa NBC pada data rekam medis pasien kanker payudara di salah satu rumah sakit di Jakarta dalam masalah klasifikasi subtype molekular kanker payudara, serta membandingkan hasil klasifikasi NBC dengan metode MDM lain, yaitu Decision Tree (DT). Hasil analisis menunjukkan bahwa NBC mengungguli DT dengan tingkat akurasi sebesar 92,8%. Selain itu, dapat juga ditunjukkan secara empiris bahwa NBC mampu menangani missing value dengan cukup baik dan tidak membutuhkan data dalam jumlah banyak untuk tetap dapat mengklasifikasikan sebagian besar pasien dengan benar.

ABSTRACT

Nowadays, modern hospitals are well equipped with data monitoring devices, which resulted in an abundant amount of medical data. These medical data possess specific characteristics and usually, statistical methods could not be applied directly. This is what started the notion of Medical Data Mining (MDM), which has proven to be effective in analysing medical data. Naive Bayes Classifier (NBC) is an implementation of MDM. Even though MDM methods produce a sufficiently accurate and satisfying results in diagnosis problems, these methods are still not well accepted in the medical practice. One of the main reasons is because there is a resistance of physicians to a new diagnosis method. The main goal of this study is to apply and evaluate the performance of NBC in classifying breast cancer patients in a private hospital in Indonesia into five classes of molecular subtypes and compare its performance with another popular MDM method, Decision Tree (DT). Results showed that NBC outperformed DT by reaching an accuracy rate of 92.8%. This study could also show empirically that NBC does not need a big dataset to be able to achieve a high accuracy rate and that NBC could handle the problem of missing values just fine.