

Klasifikasi kanker paru-paru menggunakan support vector machine dengan pemilihan fitur berbasis fungsi kernel = Classification of lung cancer using support vector machine with feature selection based on kernel function

Melati Vidi Jannati, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20444878&lokasi=lokal>

Abstrak

Klasifikasi data kanker menggunakan microarray data menjadi salah satu cara untuk mendapatkan pengobatan yang lebih tepat. Kendala yang terdapat adalah karakteristik dari microarray yang memiliki fitur yang sangat banyak. Seringkali fitur tersebut tidak begitu informatif bagi pengklasifikasian sehingga perlu adanya suatu cara untuk memilih fitur-fitur yang mengandung informasi yang penting. Salah satu cara tersebut adalah dengan pemilihan fitur. Pada penelitian ini, metode pemilihan fitur yang digunakan berdasarkan clustering dengan fungsi kernel. Fitur-fitur yang sudah terpilih kemudian diklasifikasikan menggunakan metode Support Vector Machine.

Evaluasi dari klasifikasi pada penelitian ini melibatkan K-Fold Cross Validation, metode tersebut akan membagi data secara acak, tetapi merata sehingga akurasi yang didapat juga merata. Hasil akurasi tersebut dilakukan dengan berbagai uji terhadap parameter yang berkaitan seperti K partisi, nilai dan fitur-fitur yang digunakan. Pada proses klasifikasi tanpa pemilihan fitur tingkat akurasinya mencapai 89.68 dengan k partisi sebanyak 6 sementara dengan 5 fitur akurasinya menjadi 95.87 pada partisi sebanyak 10.

.....Classification of cancer using microarray data is one way to get a more precise treatment. The obstacle on classification data is the characteristics of microarray data that is having many features. These features are often not so informative for classification, so it needs a way to select the features that contain important information. One way is by selection feature. In this research, the method of selection features that are used based on clustering with kernel function. Features that are already selected then classified using Support Vector Machine.

Evaluation of classification in this research involves a K Fold Cross Validation, that methods split data randomly but uniformly so that it can reach all of accuracy. The results of accuracy data was done with different test against related parameters such as K partition, the value of and the features that are used. On the classification process without selection features rate of accuracy reached on 89.68 with k partition number 6 while with the 5 features obtained 95.87 on partition number 10.