

Klasifikasi kanker paru-paru menggunakan support vector machine dengan pemilihan fitur berbasis support vector machine recursive feature elimination dan artificial bee colony = Classification of lung cancer using support vector machine and feature selection based on support vector machine recursive feature elimination and artificial bee colony

Selly Anastassia Amellia Kharis, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20467570&lokasi=lokal>

Abstrak

Kanker merupakan kelompok penyakit yang ditandai dengan pertumbuhan dan penyebaran sel-sel abnormal yang tidak terkendali. Jika penyebaran sel tersebut tidak terkendali, hal ini dapat menyebabkan kematian. Berdasarkan American Cancer Society, pendeteksian dini terhadap sel kanker dapat meningkatkan angka harapan hidup seorang pasien lebih dari 97%. Banyak penelitian yang telah meneliti mengenai klasifikasi kanker menggunakan microarray data. Microarray data terdiri dari ribuan fitur gen namun hanya memiliki puluhan atau ratusan sampel. Hal tersebut dapat menurunkan akurasi klasifikasi sehingga perlu dilakukannya pemilihan fitur sebelum proses klasifikasi.

Pada penelitian ini dilakukan dua tahap pemilihan fitur. Pertama, support vector machine recursive feature elimination SVM-RFE digunakan untuk prefilter gen. Kedua, hasil pemilihan fitur SVM-RFE diseleksi kembali dengan menggunakan artificial bee colony ABC yang merupakan algoritma optimisasi berdasarkan perilaku lebah madu. Penelitian ini menggunakan dua dataset, yaitu data kanker paru-paru Michigan dan Ontario dari Kent Ridge Biomedical Dataset.

Hasil percobaan dengan menggunakan SVM-RFE dan ABC menunjukkan nilai akurasi klasifikasi yang lebih tinggi daripada tanpa pemilihan fitur, SVM-RFE, dan ABC, yaitu 98 untuk data kanker paru-paru Michigan dengan menggunakan 100 fitur dan 97 untuk data kanker paru-paru Ontario dengan menggunakan 70 fitur.

.....Cancer is a group of diseases characterized by the uncontrolled growth and spread of abnormal cells. If the spread is not controlled, it can result in death. Based on American Cancer Society, early detection of cancerous cells can increase survival rates for patients by more than 97%. Many study showed new aspect of cancer classification based microarray data. Microarray data are composed of many thousands of features genes and from tens to hundreds of instances. It can decrease classification accuracy so feature selection is needed before the classification process

In this paper, we propose two stages feature selection. First, support vector machine recursive feature elimination recursive feature elimination SVM RFE is used to prefilter the genes. Second, the SVM RFE features selection result is selected again using Artificial Bee Colony ABC which is an optimization algorithm based on a particular intelligent behavior of honeybee swarms. This research conducted experiments on Ontario and Michigan Lung Cancer Data from Kent Ridge Biomedical Dataset. Experiment results demonstrate that this approach provides a higher classification accuracy rate than without feature selection, SVM RFE, and ABC, 98 for Michigan lung cancer dataset with using 100 features and 97 for Ontario lung cancer dataset with using 70 features.