

Klasifikasi data kanker menggunakan fuzzy c-means dengan pemilihan fitur menggunakan fisher's ratio = Classification of cancer data using fuzzy c means with feature selection using fisher's ratio

Arvan Aulia Rachman, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20430776&lokasi=lokal>

Abstrak

Klasifikasi data kanker dilakukan untuk menemukan terapi yang tepat yaitu memaksimalkan efektivitas dan meminimalkan toksisitas. Pada umumnya, data kanker terdiri dari banyak fitur. Namun, tidak semua fitur tersebut informatif. Oleh karena itu, fitur-fitur tersebut akan diseleksi menggunakan metode Fisher's Ratio untuk memilih fitur-fitur yang paling informatif. Fitur-fitur terbaik akan dibentuk data baru. Data, sebelum dan setelah dilakukan pemilihan fitur, diklasifikasi menggunakan metode Fuzzy C-Means. Akurasi dari proses klasifikasinya akan dibandingkan. Hasilnya, tanpa melakukan pemilihan fitur, diperoleh rata-rata akurasi sebesar 82.92%. Setelah dilakukan pemilihan fitur, diperoleh akurasi terbaik dengan menggunakan 150 fitur dengan rata-rata akurasi sebesar 89.68%.

.....

Classification of cancer data is done to find the right therapy that maximize efficacy and minimize toxicity. In general, cancer data consists of many features. However, not all of these features are informative. Therefore, these features will be selected using Fisher's Ratio to choose features that are most informative. The best features to be formed new data. Data, before and after feature selection, are classified using Fuzzy C-Means. The accuracy of the classification process will be compared. As a result, without doing feature selection, the accuracy is 82.92%. After doing feature selection, the best accuracy is obtained by using 150 features with the accuracy is 89.68%.