

Implementasi spectral clustering-self organizing map pada data microarray ekspresi gen karsinoma = Implementation of spectral clustering self organizing map on microarray data of carcinoma genes expression / M. Abdul Rivai

M. Abdul Rivai, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20455942&lokasi=lokal>

Abstrak

ABSTRAK

Clustering adalah metode pembagian data ke dalam kelompok homogen yang disebut cluster. Spectral clustering merupakan salah satu algoritma clustering modern yang memiliki kelebihan dapat mereduksi dimensi data. Pada penelitian ini metode partisi yang diterapkan pada spectral clustering yaitu self-organizing map SOM. SOM memiliki keunggulan tahan terhadap data noise dan outlier, serta SOM dapat mengatasi dataset yang besar. Penelitian ini bertujuan untuk mengimplementasikan spectral clustering-self organizing map pada data microarray ekspresi gen karsinoma yang terdiri dari 7457 gen dari 18 sampel normal dan 18 sampel penderita kanker karsinoma. Sebelum dilakukan spectral clustering-SOM, data microarray ekspresi gen karsinoma dinormalisasi menggunakan normalisasi min-max. Spectral clustering-SOM dilakukan dengan tahapan-tahapan berikut: menghitung matriks similaritas W , menghitung matriks laplacian ternormalisasi L_{sym} , menghitung eigenvalue dari L_{sym} , membentuk matriks U yang terdiri dari k eigenvector terkecil, membentuk vektor unit U_{norm} dari vektor baris pada matriks U sehingga vektor unit memiliki norm 1, mengelompokkan gen pada matriks U_{norm} menggunakan SOM dan menghitung nilai indeks Davies-Bouldin IDB_k . Penentuan jumlah cluster terbaik berdasarkan nilai indeks Davies-Bouldin yang paling minimum. Dengan menggunakan perangkat lunak R, hasil penelitian ini menunjukkan bahwa data microarray ekspresi gen karsinoma terbagi menjadi dua cluster dengan nilai indeks Davies-Bouldin yaitu 0,5843429. Berdasarkan indeks Davies-Bouldin, hasil clustering menggunakan metode spectral clustering-SOM lebih baik daripada hasil clustering yang menggunakan metode SOM tanpa spectral clustering.

ABSTRACT

Clustering is a method the dividing data into a homogeneous group called a cluster. Spectral clustering is one of the modern clustering algorithms that has the advantage of reducing dimensions of data. In this study the partitioning method applied to spectral clustering is self organizing map. SOM has the advantage of robust to noise and outlier, and SOM can handle large datasets. This study aims to implement spectral clustering self organizing map on microarray data of carcinoma gene expression consisting of 7457 genes from 18 normal samples and 18 samples of carcinoma cancer patients. Before spectral clustering SOM, the microarray data of carcinoma genes expression was normalized using min max normalization. The Spectral clustering SOM is done by the following steps calculate similarity matrix W , calculate the normalized Laplacian matrix L_{sym} , calculate the eigenvalue of L_{sym} , forming a vector unit U_{norm} of the row vector of the matrix U so that the vector unit has norm 1, grouping the genes in the matrix U_{norm} and calculate the Davies Bouldin index values IDB_k . Determination of the best number of clusters based on the minimum value of the Davies Bouldin index. By using software R, the result of this research is microarray data of

carcinoma gene expression is divided into two clusters with Davies Bouldin index value is 0.5843429. Based on the Davies Bouldin index values, clustering using spectral clustering SOM is better than clustering using only SOM method without spectral clustering.