

Metode THD-Tricluster dengan New Residue Score pada Data Ekspresi Gen Human Immunodeficiency Virus-1 = THD-Tricluster Method with New Residue Score at Human Immunodeficiency Virus-1 Gene Expression Data

Elke Annisa Octaria, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20509469&lokasi=lokal>

Abstrak

Triclustering merupakan teknik analisis pada data 3D yang bertujuan untuk mengelompokkan data secara bersamaan pada baris dan kolom di sepanjang waktu/kondisi yang berbeda. Hasil dari teknik ini disebut dengan tricluster. Tricluster merupakan subruang berupa subset dari baris, kolom, dan waktu/kondisi. Triclustering biasanya digunakan untuk menganalisis data ekspresi gen. Studi dan analisis data ekspresi gen selama perkembangan suatu penyakit merupakan masalah penelitian yang penting dalam bioinformatika dan aspek klinis. Oleh karena itu, penelitian ini mengimplementasikan metode THD-Tricluster dengan new residue score pada data ekspresi gen perkembangan penyakit HIV-1 yang terdiri dari 22283 probe id, 40 observasi, dan 4 kondisi. Pada tahap pertama dilakukan pencarian bicluster dengan lift algorithm berdasarkan nilai new residue score dengan threshold $I = 0.08$. Pada tahap kedua dilakukan pencarian tricluster dengan menentukan minimum probe dan minimum observasi $\tilde{\alpha} = 0.8$ sehingga memperoleh 33 tricluster. Hasil evaluasi tricluster menggunakan Inter-temporal Homogeneity dengan threshold $\tilde{\alpha} = 0.8$ diperoleh 32 tricluster yang menunjukkan 3 gen yang terkait dengan HIV-1 yaitu HLA-C, ELF-1, dan JUN.

Triclustering is an analysis technique on 3D data that aims to group data simultaneously on rows and columns across different times/conditions. The result of this technique is called a tricluster. Triclusters are a subspace consisting of a subset of rows, columns, and time/conditions. Triclustering is commonly used to analyze gene expression data. The study and analysis of gene expression data during disease progression is an important issue in the research of bioinformatics and clinical aspects. Therefore, this study implements the THD-Tricluster method with a new residue score on the gene expression data for HIV-1 disease progression consisting of 22283 probe id, 40 observations, and 4 conditions. In the first stage, a bicluster search was carried out with a lift algorithm based on the new residue score with a threshold of $I = 0.08$. In the second stage, the tricluster search was carried out by determining the minimum probe = 5 and the minimum observation = 2 to obtain 33 triclusters. The results of the tricluster evaluations using Inter-temporal Homogeneity with a threshold of $\tilde{\alpha} = 0.8$ obtained 32 triclusters which shows 3 genes related to HIV-1, namely HLA-C, ELF-1, and JUN.