

Triclustering menggunakan Metode Hybrid Delta-TRIMAX Binary Particle Swarm Optimization dan Ontologi Gen pada Data Ekspresi Gen Tiga Dimensi Sel Kanker Pankreas PANC-1 = Triclustering with Hybrid Delta Trimax Binary Particle Swarm Optimization and Gene Ontology on Three-dimensional Gene Expression Data of Pancreatic Cancer Cells PANC-1

Syahputri Riani, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=9999920530610&lokasi=lokal>

Abstrak

Salah satu teknik analisis yang dapat digunakan pada data mining dalam mengelompokkan data adalah Triclustering. Triclustering merupakan metode pengelompokan secara bersamaan pada data tiga dimensi yang terdiri dari observasi, atribut, dan konteks. Triclustering kerap digunakan pada bidang bioinformatika untuk mengelompokkan data ekspresi gen di titik waktu tertentu pada suatu kondisi eksperimen. Triclustering yang diajukan pada penelitian ini menggunakan metode Hybrid – TRIMAX Binary Particle Swarm Optimization. Particle Swarm Optimization (PSO) adalah teknik pengelompokan yang terinspirasi oleh perilaku biologis populasi ikan atau kawanan burung yang bergerak untuk menuju sumber makanan. Setiap individu di dalam populasi disebut sebagai partikel yang didefinisikan sebagai kandidat solusi (tricluster). Istilah “Binary” mengartikan bahwa partikel yang bergerak di ruang pencarian berbentuk vektor biner (bit) yang bernilai 0 atau 1. Tahap inisiasi populasi dilakukan dengan menggunakan algoritma nodes deletion pada – TRIMAX untuk menghasilkan populasi awal yang homogen. Metode – TRIMAX dapat menghasilkan tricluster dengan nilai Mean Residual Square (MSR) lebih kecil dari threshold δ_c sehingga dapat meningkatkan efektifitas komputasi dari metode Hybrid – TRIMAX Binary Particle Swarm Optimization. Algoritma gabungan kemudian diimplementasikan pada data ekspresi gen tiga dimensi sel kanker pankreas PANC-1 yang diberikan obat kemoterapi ATO, JQ1, dan kombinasi keduanya pada 3 titik waktu. Diperoleh tricluster optimum dengan skenario 0,0003; 0,8; 0,2; dan tipe neighbourhood = “Gbest”. Tricluster tersebut memiliki nilai TQI sebesar 1,427E-09 dan volume tricluster sebesar 169.410. Berdasarkan tricluster optimum, diperoleh informasi mengenai kumpulan gen yang tidak merespon baik terhadap pengobatan JQ1 dan JQ1+ATO pada jangka waktu menengah dan panjang. Hasil analisis ontologi gen menunjukkan tiga aspek ontologi yang signifikan dengan p-value < 0,05, yaitu proses biologi, fungsi molekuler, dan komponen seluler. Diperoleh gen yang resisten terhadap pengobatan terlibat dalam proses biologi metabolisme sel dan pengembangan sel yang mempertahankan kehidupan sel. Pada aspek fungsi molekuler, gen berperan dalam proses pengikatan, seperti pengikatan ion, senyawa organik siklik, dan senyawa heterosiklik, serta aktivitas katalitik. Selain itu, juga ditemukan bahwa sebagian besar gen berlokasi pada sitoplasma, organel, dan nukleus dalam komponen seluler. Aspek-aspek dari ontologi gen dapat berkontribusi pada resistensi kumpulan gen dalam sel kanker PANC-1 terhadap pengobatan.

One of the analysis techniques that can be used in data mining to group data is Triclustering. Triclustering is a method of simultaneously grouping three-dimensional data consisting of observations, attributes, and context. Triclustering analysis is often used in the field of

bioinformatics to group gene expression data at certain time points under experimental conditions. The triclustering analysis proposed in this study used the Hybrid – TRIMAX Binary Particle Swarm Optimization method. Particle Swarm Optimization (PSO) is a clustering technique inspired by the biological behavior of fish populations or flocks of birds that move towards food sources. Each individual in the population is referred as particles which are defined as candidate solutions (tricluster). The term "Binary" means that the particles move in the search space in the form of binary vectors (bits) with a value of 0 or 1, the number "1" represents that an individual is present in the particle. The population initialization stage is carried out using the nodes deletion algorithm in $\hat{I}$ -TRIMAX to produce a homogeneous initial population. The $\hat{I}$ -TRIMAX method can generate a tricluster with a Mean Residual Square (MSR) value smaller than the threshold δ_{ϵ} so that it can increase the computational effectiveness of the Hybrid $\hat{I}$ -TRIMAX Binary Particle Swarm Optimization method. The combined algorithm then implemented on three-dimensional gene expression data of PANC-1 pancreatic cancer cells given ATO, JQ1, and a combination of both chemotherapy drugs at three time points. The optimum tricluster was obtained with scenario 0,0003; 0,8; 0,2; and neighborhood type = "Gbest". The tricluster has a TQI value of 1.427E-09 and a tricluster volume of 169,410. Based on the optimum tricluster, information was obtained about the gene pools that did not respond well to JQ1 and JQ1+ATO treatment in the medium and long term. The results of gene ontology analysis showed three significant ontological aspects with p-value <0.05, namely biological processes, molecular functions, and cellular components. It was found that treatment-resistant genes are involved in the biological process of cell metabolism and cell development that maintains cell life. In the aspect of molecular function, genes play a role in binding processes, such as ion binding, cyclic organic compounds, and heterocyclic compounds, as well as catalytic activity. In addition, it was also found that most genes are located in the cytoplasm, organelles, and nucleus in cellular components. These aspects of the gene ontology may contribute to the resistance of the gene pool in PANC-1 cancer cells to treatment.