

Analisis Triclustering Menggunakan Metode Gabungan Fuzzy Cuckoo Search dengan I-Trimax pada Data Ekspresi Gen Tiga Dimensi = Triclustering Analysis Using Hybrid Fuzzy Cuckoo Search with I-Trimax Method on Three Dimensional Gene Expression Data

Airlangga Muhammad Putrapradana, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20529261&lokasi=lokal>

Abstrak

Analisis triclustering merupakan salah satu metode data mining yang bertujuan mengelompokkan data berbentuk tiga dimensi. Triclustering kerap digunakan pada bidang bioinformatika untuk menganalisis kesamaan ekspresi gen suatu eksperimen pada titik waktu tertentu. Analisis triclustering yang dilakukan pada penelitian ini menggunakan metode gabungan Fuzzy Cuckoo Search dengan I-Trimax. Metode ini merupakan penggabungan algoritma nodes deletion pada I-Trimax dengan algoritma optimasi Fuzzy Cuckoo Search. Cuckoo Search merupakan metode optimasi yang sudah baik dalam menghasilkan himpunan tricluster yang menggunakan konsep parasitisme spesies cuckoo. Fuzzy Cuckoo Search menggunakan fungsi objektif fuzzy c-means untuk mengatasi ketidakjelasan (indiscernibility) yang biasa terjadi dalam data ekspresi gen sehingga masalah kesulitan membedakan objek karena kurangnya pengetahuan dari informasi yang tersedia dapat diatasi. Algoritma nodes deletion pada I-Trimax digunakan pada fase pembentukan populasi awal dari metode gabungan Fuzzy Cuckoo Search dengan I-Trimax. Hal ini dilakukan demi mendapatkan populasi awal yang sudah baik yaitu memiliki MSR yang minimum karena konsep dari algoritma nodes deletion yaitu dapat menghasilkan himpunan tricluster dengan Mean Square Residue (MSR) kecil yaitu di bawah threshold. Berdasarkan itu proses komputasi algoritma Fuzzy Cuckoo Search yang dilakukan pada fase optimasi dapat berjalan dengan efektif sehingga menghasilkan himpunan tricluster yang berkualitas baik secara efisien. Analisis triclustering menggunakan metode gabungan Fuzzy Cuckoo Search dengan I-Trimax digunakan pada data ekspresi gen tiga dimensi sel kanker paru-paru fase stabil (A549) yang berkaitan dengan pemberian obat kemoterapi Motexafin Gadolinium (MGd), di mana ekspresi gen diamati pada 6 kondisi dan 3 titik waktu. Pada penelitian ini, himpunan tricluster yang memiliki kualitas terbaik berdasarkan Triclustering Quality Index (TQI) adalah himpunan tricluster yang dihasilkan dengan nilai α dan β . Berdasarkan himpunan tricluster tersebut, didapatkan informasi penting mengenai kumpulan gen yang memiliki respon baik terhadap pemberian MGd tapi tidak bertahan setiap titik waktu. Hal ini dapat dijadikan acuan penelitian terkait terapi kanker menggunakan obat kemoterapi MGd yang perlu dilakukan pengembangan agar dapat tetap efektif pada seluruh titik waktu. Terdapat juga kumpulan gen yang memiliki respon cepat dan bertahan hingga jangka panjang dengan pemberian MGd dan mannitol. Gen-gen tersebut merupakan gen yang menunjukkan respon baik pemberian obat kemoterapi MGd tetapi efektivitasnya tidak terlalu maksimal karena responnya beririsan dengan subjek yang hanya diberikan mannitol. Hal ini dapat dijadikan bahan untuk penelitian lebih lanjut dalam pengembangan obat MGd supaya dapat lebih efektif.

.....Triclustering analysis is a data mining method that aims to group data in three dimensions. Triclustering is often used in the field of bioinformatics to analyze the similarity of gene expression under experimental conditions at a certain point in time. The triclustering analysis carried out in this study used the combined Fuzzy Cuckoo Search method with I-Trimax. This method is a combination of node deletion algorithm on -

Trimax with Fuzzy Cuckoo Search optimization algorithm. Cuckoo Search is a good optimization method in generating tricluster sets that use the concept of parasitism of cuckoo species. Fuzzy Cuckoo Search uses the fuzzy c-means objective function to overcome the indiscernibility that usually occurs in gene expression data so that the problem of difficulty distinguishing objects due to lack of knowledge from available information can be overcome. The nodes deletion algorithm on I-Trimax is used in the initial population formation phase from the combined Fuzzy Cuckoo Search method with I-Trimax. This is done in order to get a good initial population, which has a minimum MSR because the concept of the nodes deletion algorithm is that it can produce a tricluster set with a small Mean Square Residue (MSR), which is below the threshold. Based on that, the computational process of the Fuzzy Cuckoo Search algorithm which is carried out in the optimization phase can run effectively so as to produce a good quality tricluster set efficiently. Triclustering analysis using the combined Fuzzy Cuckoo Search method with I-Trimax was used on three-dimensional gene expression data of stable phase lung cancer cells (A549) associated with the administration of the chemotherapy drug Motexafin Gadolinium (MGd), where gene expression was observed in 6 conditions and 3 time points. In this study, the tricluster set that has the best quality based on the Triclustering Quality Index (TQI) is the resulting tricluster set with values. Based on these tricluster sets, important information was obtained regarding gene pools that responded well to MGd administration but did not persist at any point in time. This can be used as a reference for research related to cancer therapy using MGd chemotherapy drugs that need to be developed in order to remain effective at all time points. There is also a gene pool that responds quickly and persists in the long term with MGd and mannitol administration. These genes are genes that show a good response to MGd chemotherapy drugs but their effectiveness is not maximal because their responses coincide with subjects who are only given mannitol. This can be used as material for further research in the development of MGd drugs so that they can be more effective.