

Model integrasi data gene expression dan DNA copy number pada cluster gen kanker menggunakan intelligent kernel k means = Integration data model of gene expression and dna copy number on cluster genes of cancer using intelligent kernel k means

Teny Handhayani, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20350269&lokasi=lokal>

Abstrak

Integrasi data gene expression dan DNA copy number berbasis kernel digunakan untuk menganalisis pola gen pada penyakit kanker payudara cell line. Clustering pada data integrasi dilakukan tanpa adanya informasi jumlah k cluster, teknik ini disebut fully unsupervised clustering. Pada penelitian ini, intelligent kernel K-Means dikembangkan dengan menggabungkan teknik intelligent K-Means dan kernel K-Means. Berdasarkan hasil eksperimen, nilai pada kernel RBF mempengaruhi jumlah cluster yang ditemukan. Hasil clustering dievaluasi menggunakan nilai R, global silhouette, indeks Davies-Bouldien, akurasi LS-SVM dan visualisasi. Hasil eksperimen terbaik yaitu 3 cluster yang memperoleh akurasi LS-SVM sebesar 97.3% dengan standar deviasi 0.2%.

<hr>

In this thesis, kernel based data integration of gene expression and DNA copy number would be utilized to analyze pattern of genes in breast cancer cell line. The cluster analysis on the integrated data will be conducted with has no prior information with regards the number of k clusters which is called fully unsupervised clustering technique. In this work, intelligent kernel K-Means is proposed by combining intelligent K-Means and kernel K-Means. From the experiments, the value of of Radial Basis Function (RBF) has important role for finding the optimal of number of cluster. The clusters those to be found will be evaluated based on global silhouette, Davies-Bouldien Index, LS-SVM accuracy and visualization. The experiment result show that three clusters are successfully to be found. Those clusters produce average accuracy of LS-SVM around 97.3 % with standard deviation 0.2 %.