

Pengujian kinerja Deep Stacking Network (DSN) pada diagnosis tipe kanker payudara menggunakan data microarray = Deep Stacking Network (DSN) performance analysis on breast cancer type diagnosis using microarray data

Nurul Hanifah, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20475515&lokasi=lokal>

Abstrak

Diagnosis kanker payudara umum menggunakan data morfologi dan informasi klinis. Akan tetapi diagnosis kanker yang akurat dibutuhkan untuk menangani dan terapi yang tepat. Deteksi ekspresi gen merupakan cara yang dapat digunakan untuk menganalisis kanker hingga tingkat molekuler. Microarray adalah teknologi yang memungkinkan analisis ribuan ekspresi gen dalam waktu yang bersamaan. Pembelajaran mesin dapat digunakan untuk menganalisis pola pada microarray dataset. Model pembelajaran dengan deep learning, seperti deep stacking network DSN diperlukan untuk data yang kompleks seperti pada microarray. DSN pertama kali diusulkan untuk mengatasi kekurangan dari deep learning pada umumnya. Selain itu dikarenakan tingginya dimensi data microarray, sebelum melalui proses pembelajaran diperlukan reduksi dimensi pada data microarray. Reduksi data menggunakan gene shaving, dimana data akan dikelompokkan menjadi beberapa cluster dan juga dengan PCA. Selanjutnya data yang sudah direduksi melalui proses pembelajaran pada model DSN yang terdiri dari 2-layer, 3-layer dan 4-layer dengan laju pembelajaran 0.01, 0.1 dan 1.0. Parameter yang diukur adalah nilai akurasi, mean square error MSE dan waktu eksekusi, dimana nilai terbaik didapatkan pada DSN 2-layer dengan laju pembelajaran 1.0 pada data hasil reduksi PCA.

.....The diagnosis of breast cancer uses morphological data and clinical information. However, an accurate cancer diagnosis is necessary for the proper treatment. Gene expression detection is a way that can be used to analyze cancer to the molecular level. Microarray is a technology that allows the analysis of thousands of gene expression at the same time. Machine learning can be used to analyze patterns in the microarray dataset. Machine learning with deep architecture deep learning, such as deep stacking learning DSN, is needed for complex data such as microarray. DSN was proposed to overcome the shortage of deep learning. In addition, due to the high dimension of microarray data, dimension reduction in microarray data is required before going through the learning process. Data reduction uses gene shaving, where data will be grouped into clusters and also PCA. Clusters found through the process of gene shaving is a cluster that carries important genetic information in cancer cells so that the gene can represent all data. Furthermore, the data has been reduced going through the learning process on the DSN model consisting of 2 layer, 3 layer and 4 layer with the learning rate of 0.01, 0.1 and 1.0. The parameters measured are the accuracy, mean square error MSE and execution time, which the best value is obtained on 2 layer DSN with learning rate 1.0 on data from the reduction of PCA.