

Deteksi penyakit kanker paru-paru berdasarkan data microarray dengan seleksi fitur Support Vector Machine-Recursive Feature Elimination (SVM-RFE) = Detection of lung cancer based on microarray data using feature selection Support Vector Machine-Recursive Feature Elimination (SVM-RFE)

Muslar Alibasya, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=9999920554845&lokasi=lokal>

Abstrak

Kanker paru-paru merupakan jenis kanker yang dimulai dan tumbuh di dalam paru-paru. Kanker paru-paru terjadi ketika sel-sel yang melapisi bronkus dan bronkiolus tumbuh secara tidak terkendali. Hal ini dapat menyebabkan kematian jika tidak ditangani dengan cepat dan tepat. Pengklasifikasian dini merupakan salah satu solusi yang tepat untuk mengurangi jumlah kematian yang disebabkan oleh kanker paru-paru. Pendekatan machine learning dapat digunakan untuk mengklasifikasi kanker paru-paru. Dalam penelitian ini, pengklasifikasian dilakukan dengan menggunakan data microarray. Data microarray memiliki fitur yang sangat banyak. Oleh karena itu, dibutuhkan seleksi fitur agar proses klasifikasi berlangsung optimal. Pada penelitian ini, penulis mengusulkan metode Support Vector Machine-Recursive Feature Elimination (SVM-RFE) untuk metode seleksi fitur. Data microarray yang digunakan diambil dari National Center for Biotechnology Information (NCBI) yang merupakan sebuah website online database. Pada penelitian ini, penulis menggunakan SVM-RFE sebagai metode seleksi fitur untuk mengeliminasi fitur yang kurang relevan. Setelah itu pendekatan k-fold cross-validation digunakan sebagai pembagian data, dan beberapa machine learning classifier yaitu Support Vector Machine (SVM), Random Forest (RF), Decision Tree (DT), dan Extreme Gradient Boosting (XGBoost) digunakan sebagai metode klasifikasi. Dari hasil simulasi menunjukkan bahwa hasil terbaik berdasarkan nilai akurasi, precision, recall dan running time diperoleh oleh metode klasifikasi SVM dengan nilai akurasi 100%, precision 100%, recall 100% dan running time 5,42 detik.

.....Lung cancer is a type of cancer that begins in the lungs. Lung cancer occurs when the cells that cover the bronchi and bronchioles grow uncontrollably. This can lead to death if not treated quickly and appropriately. Early classification is one of the appropriate solution to reduce the number of deaths caused by lung cancer. Machine learning approach can be used to classify lung cancer. In this research, classification is done using microarray data which has a lot of features. Therefore, feature selection is applied such that the classification process used the optimal number of features. In this study, the researcher proposes the Support Vector Machine-Recursive Feature Elimination (SVM- RFE) method for the feature selection method. The microarray data was taken from the National Center for Biotechnology Information (NCBI), which is an online database website. In this study, the researcher used SVM-RFE as a feature selection method to eliminate irrelevant features. Afterwards, the k-fold cross-validation method and several machine learning classifiers, namely Support Vector Machine (SVM), Random Forest (RF), Decision Tree (DT), and Extreme Gradient Boosting (XGBoost) will be used as classification methods. In the final stage, the researcher will analyze the performance results of the proposed method based on the accuracy and running time of each classifier. The simulation results show that the best results based on the values of accuracy, precision, recall and running time are obtained by the SVM classification method with a value of 100% accuracy, 100%

precision, 100% recall and running time of 5.42 seconds.