

Abstrak

Kanker adalah penyakit mematikan yang terkenal di dunia. Berdasarkan *World Health Organization* (WHO), kanker adalah penyebab utama kematian kedua di dunia dan bertanggung jawab atas perkiraan 9,6 juta kematian pada tahun 2018. Salah satu teknik terkenal untuk pendeteksian kanker adalah teknik DNA *microarray*. Teknik DNA *microarray* memberikan kesempatan bagi para peneliti untuk menganalisis ribuan profil *gene expression* dalam waktu yang bersamaan untuk mengetahui apakah seseorang menderita kanker atau tidak. Namun, salah satu masalah yang ada pada data DNA *microarray* adalah jumlah fitur yang sangat banyak sehingga perlu dilakukan *feature selection*. Dalam mengatasi permasalahan tersebut, maka penelitian ini akan menggunakan metode *feature selection* *Support Vector Machine-Recursive Feature Elimination* (SVM-RFE) dan *Chi-Square* dan menggunakan metode klasifikasi *Naïve Bayes*. Hasil akurasi dari penggunaan *feature selection* dengan yang tidak, akan dibandingkan. Hasil akurasi antara penggunaan kedua metode *feature selection* juga akan dibandingkan, untuk mencari metode *feature selection* mana yang lebih baik jika digabungkan dengan metode klasifikasi *Naïve Bayes*. Untuk mendapatkan gambaran keseluruhan tentang perbandingan performa, penelitian ini juga mempertimbangkan *precision*, *recall*, dan *F1-score*. Hasil akurasi terbaik yang didapat untuk data *lung cancer* sebesar 100% dengan SVM-RFE dan *Chi-Square*, *ovarian cancer* sebesar 99.6% dengan SVM-RFE, *breast cancer* sebesar 93.7% dengan SVM-RFE, dan *colon cancer* sebesar 90% dengan SVM-RFE.

Kata kunci : Kanker, *Microarray*, *Feature Selection*, *Support Vector Machine-Recursive Feature Elimination* (SVM-RFE), *Chi-Square*, *Naïve Bayes*.