

ABSTRAKSI

Kanker adalah penyebab utama kematian di seluruh dunia. Akan tetapi, sebagian besar dapat disembuhkan jika kanker terdeteksi dini. Dalam beberapa dekade terakhir, teknologi bernama microarray berperan penting dalam diagnosis kanker. Dengan microarray, ribuan gen dapat dianalisis secara bersamaan. Dengan menggunakan teknik data mining, klasifikasi data dapat dilakukan untuk meningkatkan akurasi diagnosis kanker dibandingkan teknik tradisional. Karakteristik data microarray adalah jumlah sampel yang kecil dengan dimensi yang sangat besar. Karena itu, ada tantangan bagi peneliti untuk memberikan solusi bagi klasifikasi data microarray yang memiliki kinerja tinggi baik dalam akurasi maupun waktu berjalan. Beberapa penelitian menunjukkan bahwa pengurangan dimensi dapat menurunkan waktu berjalan secara signifikan sementara juga berpotensi mengatasi kebisingan dalam data sehingga dapat meningkatkan akurasi. Tesis ini mengusulkan penggunaan Principal Component Analysis (PCA) sebagai metode reduksi dimensi bersamaan dengan Support Vector Method (SVM) yang dioptimalkan oleh fungsi kernel sebagai klasifikasi untuk klasifikasi data microarray. Skema yang diusulkan diterapkan pada tujuh data set dengan menggunakan *5-fold cross validation*. Hasil penelitian menunjukkan bahwa skema yang diusulkan dapat mencapai akurasi 100% untuk data kanker ovarium dan paru-paru saat fungsi kernel *Linear* dan *Cubic* digunakan. Dari segi *running time*, PCA mampu menurunkan *running time* untuk setiap data set.

Kata Kunci: Fungsi Kernel, *k-fold cross validation*, klasifikasi data microarray, *Principal Component Analysis*, *Support Vector Machine*.