

ABSTRAK

Data ekspresi gen atau DNA *microarray* merupakan data berdimensi tinggi yang menjelaskan proses mengenai penguraian DNA sehingga menghasilkan protein. Ekspresi gen atau DNA *microarray* digunakan untuk memprediksi bagaimana ekspresi dari setiap gen tersebut. Ekspresi gen sangat banyak manfaatnya salah satunya adalah untuk mengetahui jenis perkembangan penyakit dalam tubuh manusia. Penelitian data ekspresi gen sebelumnya sudah pernah dilakukan pada salah satu jurnal yang telah di publikasi yaitu “*Review on Feature Selection of Gene Expression Data for Autism Classification*”[8]. Penelitian tersebut menjelaskan mengenai data ekspresi gen dengan menggunakan *Genetic Algorithm*.

Pada tugas akhir ini permasalahan yang diselesaikan adalah mengenai data ekspresi gen dengan menggunakan algoritma *K-Means Clustering* yang digunakan sebagai klasifikasi dalam menentukan kelas dari suatu sample data dan metode WPCA digunakan sebagai seleksi fitur untuk mereduksi dimensi dari masing-masing atribut sehingga jumlah atribut akan semakin lebih sederhana untuk di proses ke tahap selanjutnya.

Adapun seleksi fitur yang diaplikasikan dengan menggunakan fungsi bobot pada metode WPCA yaitu mampu memberikan pengaruh yang cukup baik terhadap masing-masing atribut yang telah direduksi dimana hasil performansi yang diperoleh yaitu 78,94% untuk 37 atribut dengan proporsi data 80% training,20% testing dan 71.05% untuk 16 atribut dengan proporsi data 80% training, 20% testing.

Kata Kunci : Ekspresi Gen, WPCA, K-means Clustering