

ABSTRAK

Sebuah sistem yang mampu menganalisis dan mengidentifikasi seseorang terkena suatu penyakit akan sangat membantu pada dunia medis, hal ini dikarenakan tingkat kematian setiap harinya selalu bertambah. Faktor kematian salah satunya adalah kurangnya penanganan dini pada pasien yang telah terkena suatu penyakit. Hal ini dapat terjadi karena pasien tidak mengetahui bahwa dirinya mengidap penyakit yang mematikan. Adapun lima penyakit yang digolongkan mematikan adalah Kanker, Jantung, Diabetes, AIDS, dan TBC.

Oleh karena itu, pada tugas akhir ini dibangun sistem yang mampu memprediksi seorang pasien apakah terjangkit penyakit atau tidak. Sistem ini sangat membantu dalam penanganan dini pasien yang terkena penyakit. Data penyakit Kanker, Jantung, Diabetes, AIDS, dan TBC bersumber dari website *Kent Ridge Bio-medical Data Set Repository* akan digunakan untuk membangun sistem ini, yang mana data tersebut adalah data berdimensi tinggi untuk setiap penyakit. Dimana data tersebut memiliki ribuan atribut yang akan dibagi menjadi dua data, yaitu data *training* dan data *testing* selanjutnya dilakukan reduksi dengan *Genetic Algorithm* (GA) dan klasifikasi dengan *Naïve Bayes Classifier*.

Dengan prediksi menggunakan model tersebut, didapatkan hasil yang akan menunjukkan seorang pasien yang terkena penyakit atau tidak. Selanjutnya dilakukan uji akurasi menggunakan data *testing* untuk mendapatkan hasil akurasi yang valid. Sehingga hasil akhir yang didapat menunjukkan metode *crossvalidation* lebih baik dengan nilai akurasi dari data *colon tumor* 88.89% dan *leukemia* 100% dibandingkan metode *percentage split* dengan akurasi dari data *colon tumor* 78.95% dan *leukemia* 77.27%.

Kata kunci: *Genetic Algorithm, Naïve Bayes, Evolutionary Datamining.*