

1. Pendahuluan

Kanker adalah salah satu penyakit berbahaya yang dapat menyebabkan kematian. Kanker muncul dari perubahan sel normal menjadi sel tumor yang berkembang menjadi tumor ganas. Menurut WHO, kanker merupakan penyakit mematikan kedua di dunia [1]. Sekitar 70% kematian akibat kanker terjadi di negara berpenghasilan rendah dan menengah [1]. Seperti halnya angka kejadian penyakit kanker di Indonesia yang berada di urutan 8 di Asia Tenggara [2]. Namun, peningkatan kasus kanker dapat dikurangi dengan deteksi kanker sejak dini dan menghindari faktor resiko utamanya seperti mengatur pola makan dengan baik dan menghindari konsumsi alkohol [1]. Dengan demikian, teknologi deteksi kanker sangat dibutuhkan agar upaya pencegahan maupun pengobatan kanker dapat dilakukan sesegera mungkin.

Terdapat suatu teknologi dalam bidang bioinformatika yang digunakan untuk mendeteksi kanker dalam beberapa tahun terakhir ini. DNA microarray adalah teknologi yang menggunakan data microarray untuk mengetahui pola ekspresi gen di berbagai kondisi yang berbeda [3]. Kanker dapat dideteksi dengan memahami pola ekspresi gen karena ekspresi gen dapat mendeskripsikan sel kanker yang mengalami abnormalitas yang bisa menyebabkan kanker tumbuh dan berkembang. Deteksi kanker melalui ekspresi gen jauh lebih terpercaya dan akurat dibandingkan deteksi kanker secara tradisional yang berdasarkan analisis kemunculan tumor [4]. Sehingga, teknologi ini dapat membantu para medis memahami pola ekspresi gen seseorang yang terkena kanker maupun tidak [5].

Dalam mendeteksi kanker menggunakan data microarray, terdapat masalah yang harus diatasi yaitu dimensi data yang sangat besar dimana jumlah gen lebih besar daripada jumlah sampel [6]. Solusinya yaitu dilakukan reduksi dimensi data karena dapat membantu dalam memaksimalkan data dengan mengetahui gen yang sangat berpengaruh untuk mendeteksi kanker [7]. Pada penelitian yang berjudul "Feature Selection Algorithm Based on Mutual Information and Lasso for Microarray Data" [8] telah dilakukan penelitian tentang klasifikasi data microarray menggunakan Mutual Information (MI), Least Absolute Shrinkage and Selection Operator (LASSO) dan Improved LASSO untuk metode reduksi dimensinya dan menggunakan empat metode klasifikasi yaitu Support Vector Machine (SVM), K- Nearest Neighbors KNN, C4.5 Decision Tree and Random Forest. Data yang digunakan yaitu Colon, Prostate, Lymphoma, Leukimia, dan Lung Cancer. Hasil dari penelitian tersebut diperoleh rata-rata akurasi yang menggunakan LASSO sebagai metode seleksi fiturnya adalah 93,562% dan yang menggunakan MI diperoleh rata-rata akurasi sebesar 92,97%. LASSO telah banyak digunakan dalam berbagai penelitian untuk pemilihan fitur karena kinerjanya yang efisien [8, 9]

Pada penelitian yang berjudul "Implementation of Minimum Redundancy Maximum Relevance (MRMR) and Genetic Algorithm (GA) for Microarray Data Classification with C4.5 Decision Tree." [10] telah dilakukan klasifikasi data microarray menggunakan Minimum Redundancy Maximum Relevance (MRMR) dan Genetic Algorithm (GA) dengan metode klasifikasi C4.5 Decision Tree. Pada percobaan pertama menggunakan MRMR yang dikombinasikan dengan GA sebagai metode optimasi dan klasifikasi C4.5 menghasilkan rata-rata akurasi sebesar 79%. Pada percobaan kedua menggunakan GA sebagai metode seleksi fitur dan klasifikasi C4.5 menghasilkan rata-rata akurasi sebesar 78%. Pada penelitian yang berjudul "Implementation of Mutual Information and Bayes Theorem for Classification Microarray Data." [11] telah dilakukan klasifikasi data microarray menggunakan Mutual Information (MI) dan Teorema Bayes. Hasil yang didapatkan yaitu F1-score tertinggi sebesar 91.06% menggunakan Bayesian Network dan 88.85% menggunakan Naive Bayes.

Pada penelitian yang berjudul "Deteksi Kanker berdasarkan Klasifikasi Data Microarray menggunakan Functional Link Neural Network dengan Seleksi Fitur Genetic Algorithm." [12] telah dilakukan klasifikasi data microarray menggunakan Genetic Algorithm (GA) sebagai metode reduksi dimensi dan Functional Link Neural Network (FLNN) sebagai metode klasifikasinya. Pada penelitian tersebut dijelaskan bahwa FLNN merupakan salah satu model Artificial Neural Network (ANN) yang berhasil dalam melakukan klasifikasi karena memiliki arsitektur layer tunggal yang mampu mengoptimasi proses. Berbeda dengan metode ANN yang menggunakan multi layer, waktu komputasi FLNN dapat lebih cepat karena memiliki arsitektur layer tunggal. Pada klasifikasi FLNN, terdapat fungsi basis untuk mengekspansi vektor masukan. Kumar, et.al [13] melakukan perbandingan empat jenis fungsi basis pada FLNN dan mereka menyimpulkan bahwa fungsi basis Legendre Polynomial adalah fungsi basis terbaik dibandingkan dengan tiga basis lainnya.

Dalam penelitian ini, FLNN dengan fungsi basis Legendre Polynomial digunakan pada klasifikasi dan LASSO pada reduksi dimensinya. Kemudian, performansi hasil klasifikasi dianalisis berdasarkan pengaruh dari metode seleksi fitur dan metode klasifikasi.

Pada bab pendahuluan ini penulis menjelaskan latar belakang, rumusan masalah, dan tujuan. Kemudian pada bab studi terkait, penulis membahas penelitian terkait metode yang digunakan. Selanjutnya pada bab sistem yang dibangun, penulis membahas alur system. Setelah itu pada bab evaluasi, penulis membahas hasil penelitian. Pada bab terakhir, penulis menjelaskan kesimpulan dari penelitian ini.