

## 1 Pendahuluan

### Latar Belakang

Kanker adalah salah satu penyakit yang paling mematikan di dunia. Pada tahun 2012, terdapat 14.1 juta kasus baru kanker dari sekitar 32.6 juta orang yang positif mengidap kanker dan terdapat 8.2 juta kematian yang disebabkan oleh kanker[2]. Menurut Kementerian Kesehatan Republik Indonesia ada beberapa faktor yang dapat menyebabkan risiko Penyakit kanker. Diantaranya adalah : (1) Faktor genetik, (2) Faktor karsinogen, yang diantaranya merupakan zat kimia, radiasi, virus, hormon, dan iritasi kronis, (3) Faktor perilaku atau gaya hidup, diantaranya yaitu merokok, pola makan yang tidak sehat, konsumsi alkohol, dan kurangnya aktifitas fisik [1]. Ada banyak cara yang bisa dilakukan untuk deteksi kanker sejak dini. Salah satu caranya adalah melalui faktor genetik dengan merepresentasikan data kanker pada DNA microarray dimana ribuan untaian molekul dan dinamai dengan fitur. Data dengan jumlah fitur yang sangat banyak memberikan beban komputasi yang sangat besar dan *running time* lama, sehingga metode reduksi dimensi dibutuhkan untuk mengurangi beban komputasi dan waktu *running time*.

Beberapa penelitian mengenai klasifikasi data microarray dengan metode ANN telah dilakukan, salah satunya adalah klasifikasi kanker payudara menggunakan metode *Artificial Neural Network* dengan menggunakan struktur *Backpropagation* dan *Genetics Algorithm* untuk optimalisasi parameter pada data kanker payudara dan menghasilkan rata-rata nilai akurasi sebesar 97%[3]. Melihat hasil akurasi dan kemampuan metode tersebut pada kasus kanker payudara maka metode tersebut dijadikan sebagai acuan untuk penelitian ini dengan merubah beberapa struktur pada penelitian tersebut [3] GA digunakan sebagai metode untuk mencari parameter *learning rate* dan *hidden neuron*. Pada penelitian ini, GA digunakan sebagai metode untuk mereduksi dimensi dengan tetap mempertahankan fitur-fitur yang penting, sedangkan *Artificial Neural Network* (ANN) – *Backpropagation* digunakan sebagai metode classifier.

### Topik dan Batasannya

Permasalahan yang dijadikan di dalam penelitian ini adalah :

1. Bagaimana hasil performansi dari algoritma ANN dan GA untuk diagnosa pada penyakit kanker.
2. Berapa nilai hyperparameter terbaik agar klasifikasi kanker *colon* menjadi optimal.

Batasan masalah dalam penelitian ini adalah :

1. Data kanker *colon* yang diambil dari Kent Ridge Biomedical Dataset Repository.
2. Data dibagi menjadi 80% *training*, 20% *testing*.

### Tujuan

Dari permasalahan yang ada, berikut tujuan yang hendak dicapai dalam penelitian tugas akhir ini, yaitu:

1. Mengetahui dan mencari struktur terbaik dalam *Artificial Neural Network* (ANN) dan *Genetic Algorithm* (GA) dalam mereduksi dimensi pada microarray data.
2. Mengetahui hasil performansi *Artificial Neural Network* (ANN) dan *Genetic Algorithm* (GA) pada data microarray.

### Organisasi Tulisan

Pada bagian 2 berisi mengenai literatur dari penelitian, pada bagian 3 berisi langkah-langkah pengerjaan dalam penelitian, pada bagian 4 berisi mengenai pengerjaan penelitian dan pada bagian 5 berisi kesimpulan dari penelitian.