

1. Pendahuluan

Bagian ini akan menjelaskan latar belakang dipilihnya permasalahan penelitian, perumusan masalah, batasan masalah, tujuan dan metodologi penyelesaian masalah dan sistematika penulisan.

Latar Belakang

Kanker adalah penyakit yang disebabkan oleh pertumbuhan sel-sel abnormal yang tidak terkendali, yang menyebabkan jaringan tubuh normal rusak dan penyebab kematian tertinggi di dunia. Pada tahun 2018, terdapat 18,1 juta jiwa dengan banyaknya kematian akibat kanker sebesar 9,6 juta jiwa [1]. Oleh karena itu, diperlukan teknologi untuk mendeteksi penyakit kanker sejak dini agar mendapat penanganan lebih awal dengan hasil analisis yang akurat. Namun, mendeteksi kanker bukan suatu hal yang mudah, sehingga untuk menganalisis penyakit kanker akan dilakukan dengan teknik microarray. Microarray adalah teknologi modern yang digunakan untuk membaca ekspresi gen pada tubuh makhluk hidup khususnya manusia. Teknologi microarray digunakan untuk mengklasifikasikan ekspresi gen kaitannya dengan penyakit tertentu, termasuk penyakit kanker. Klasifikasi ini dilakukan dengan cara menganalisis dan mengelompokkan sampel terhadap kelas kanker atau tidak kanker.

K-Nearest Neighbor (KNN) adalah algoritma yang berfungsi untuk melakukan klasifikasi suatu data berdasarkan data latih (*training datasets*), yang diambil dari K tetangga terdekatnya (*nearest neighbors*). Dalam penentuan tetangga terdekat melibatkan perhitungan korelasi antar atribut dan nilai fitness sebagai langkah awal dan pertimbangan pembobotan. KNN merupakan metode yang banyak digunakan karena memiliki kelebihan diantaranya untuk data yang berdimensi besar dan MKNN memperbaiki kekurangan KNN dalam memilih atribut yang relevan terhadap observasi dengan melibatkan bobot pada hasil fitness. *Genetic Algorithm* (GA) adalah teknik pencarian dalam bidang komputasi untuk menemukan solusi dengan melibatkan pendekatan optimasi dalam pencariannya. Teknik dalam GA didasarkan pada biologi evolusioner seperti pewarisan, mutasi, seleksi dan *crossover*. Keunggulan menggunakan GA adalah kemampuannya untuk memilih atribut-atribut yang relevan pada data berdimensi tinggi berdasarkan proses mutasi pada komposisi individu yang merepresentasikan atribut pada data. GA pernah digunakan untuk meningkatkan kinerja KNN dan Naïve Bayes dalam deteksi diabetes untuk pemilihan fitur oleh R. N. Patil dan S. C. Tamane (2018) [2]. Untuk deteksi tumor otak oleh M. Poornima, C. Kuyin, M. Revathy (2018) [3]. Kemudian, GA pernah digunakan untuk *minimizing the cost of two-tier cellular network with queuing handoff calls in microcell using genetic algorithm* oleh P. Goel dan D. K. Lobiyal (2018) [4].

Topik dan Batasannya

Berdasarkan masalah yang telah disampaikan maka rumusan masalah pada penelitian ini adalah untuk mengetahui bagaimana mengimplementasikan proses seleksi atribut dengan *Genetic Algorithm* (GA) pada data microarray dan bagaimana hasil akurasi klasifikasi microarray dengan *K-Nearest Neighbor*.

Adapun batasan masalah dalam pengerjaan Tugas Akhir yaitu Data yang digunakan dalam penelitian ini adalah antara lain data *colon tumor*, *lung cancer* dan *leukimia* yang didapat dari Kent-Ridge Bio-medical Data Set Repository (<http://leo.ugr.es/elvira/DBCRepository/>).

Tujuan

Tujuan dari penelitian ini adalah untuk mengetahui bagaimana mengimplementasikan proses seleksi atribut dengan *Genetic Algorithm* (GA) pada data microarray dan bagaimana hasil akurasi klasifikasi microarray dengan *K-Nearest Neighbor*.

Tabel 1. Keterkaitan antara tujuan, pengujian dan kesimpulan

No	Tujuan	Pengujian	Kesimpulan
1	Bagaimana mengimplementasikan proses seleksi atribut dengan GA pada data microarray	Pengujian menggunakan matlab	Tujuan tercapai
2	Bagaimana hasil akurasi klasifikasi microarray dengan KNN	Pengujian menggunakan matlab	Tujuan tercapai

Organisasi Tulisan

Laporan tugas akhir ini disusun dengan sistematika penulisan sebagai berikut:

1. Pendahuluan

- Bab ini berisi penjelasan latar belakang dipilihnya permasalahan penelitian, topik dan batasannya, tujuan dan organisasi tulisan.
2. Studi Terkait
Bab ini berisi landasan teori yang mendukung penelitian pada tugas akhir ini. Landasan teori ini meliputi penjelasan mengenai microarray, reduksi dimensi, *genetic algorithm*, *K-Nearest Neighbor*.
 3. Sistem yang Dibangun
Bab ini berisi dataset spesifikasi, gambaran umum sistem, evaluasi dan implementasi sistem.
 4. Evaluasi
Bab ini menguraikan hasil pengujian terhadap sistem yang dibuat beserta hasil analisis dari pengujian yang telah dilakukan.
 5. Kesimpulan
Bab ini berisi kesimpulan dari hasil pengujian dan skenario yang telah dilakukan.
 6. Lampiran
Bab ini berisi lampiran *screenshot* tampilan sistem