

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Pengenalan teknologi DNA *microarray* membuat perolehan data *microarray* semakin mudah, hal ini memicu persoalan tentang bagaimana cara terbaik dalam melakukan ekstraksi dan seleksi fitur dari data yang berdimensi sangat besar tersebut. Berbagai macam metode telah diusulkan dalam melakukan ekstraksi dan seleksi fitur dari data *microarray*. Akan tetapi, metode-metode terdahulu mengabaikan adanya hubungan antar gen (*interrelation*) sehingga memungkinkan hilangnya informasi penting yang tersimpan pada suatu gen pada saat ekstraksi fitur. Hal ini mengakibatkan metode tersebut masih belum dapat membantu para ilmuwan biologi untuk menemukan informasi penting pada gen, terutama mengenai masalah klasifikasi kanker [1].

Meskipun berbagai macam metode telah digunakan untuk mengekstrak dan memilih fitur dari data *microarray*, pengembangan metode ekstraksi dan seleksi fitur dari data *microarray* yang lebih *powerful* dan efisien untuk meningkatkan performa klasifikasi kanker masih menjadi persoalan yang perlu diselesaikan [1].

Oleh karena itu dalam tugas akhir ini diusulkan sebuah implementasi metode ekstraksi fitur yang memperlakukan informasi pasangan gen, yaitu pasangan gen yang memiliki kolerasi tinggi pada satu jenis sampel jaringan (*tissue sample*) dan memiliki perbedaan yang signifikan pada tipe *tissue sample* lain, sebagai suatu kesatuan yang digunakan untuk ekstraksi fitur dari model klasifikasi berbasis informasi pasangan gen.

1.2 Rumusan Masalah

Permasalahan yang diangkat dalam pembuatan tugas akhir ini antara lain:

1. Bagaimana cara membuat model klasifikasi berbasis informasi pasangan gen.
2. Bagaimana cara melakukan ekstraksi fitur dari model klasifikasi yang berbasis informasi pasangan gen.
3. Bagaimana mengimplementasikan metode WSCA (*Within Sample Classification Accuracy*) untuk mengevaluasi model klasifikasi.

1.3 Batasan Masalah

Adapun batasan ruang lingkup permasalahan dari pengerjaan tugas akhir ini adalah sebagai berikut:

1. Input berupa dataset *microarray*.
2. Dataset *microarray* yang digunakan sebagai percobaan terhadap metode ini adalah data *colon cancer* dan *diffuse large B cell lymphoma* (DLBCL).
3. Hanya terdapat dua kelas yang terlibat dalam proses klasifikasi.

1.4 Tujuan

Tujuan dari pembuatan tugas akhir ini adalah untuk mengimplementasikan metode ekstraksi fitur dari data *microarray* berbasis informasi pasangan gen yang mampu membantu dalam proses klasifikasi sehingga diharapkan dapat memberikan tingkat akurasi klasifikasi yang tinggi dengan menggunakan fitur yang lebih sedikit.

1.5 Metodologi

Pembuatan tugas akhir ini dilakukan dengan menggunakan metodologi sebagai berikut :

A. Studi literatur

Pada tahap ini dilakukan pencarian dan pemahaman literatur yang diperlukan untuk pengumpulan data dan perancangan sistem yang akan dibuat. Literatur yang didapat berasal dari buku, jurnal, materi kuliah, dan materi-materi lain yang berhubungan dengan ekstraksi fitur data *microarray*, dan model klasifikasi berbasis informasi pasangan gen.

B. Perancangan perangkat lunak

Pada tahap ini diawali dengan melakukan analisis awal terhadap permasalahan utama yang muncul pada topik tugas akhir ini kemudian dilakukan perancangan perangkat lunak yang meliputi penentuan data yang akan digunakan dan proses-proses yang akan dilaksanakan.

C. Implementasi dan pembuatan sistem

Pada tahap ini dilakukan implementasi metode yang diusulkan dari rancangan yang telah dibuat pada tahap sebelumnya dengan menggunakan MATLAB.

D. Uji coba dan evaluasi

Pada tahap ini dilakukan uji coba dengan menggunakan bermacam data masukan yang bervariasi untuk mencoba jalannya aplikasi telah sesuai dengan rancangan dan desain implementasi yang dibuat juga untuk mencari kesalahan program yang mungkin terjadi untuk selanjutnya dapat dilakukan penyempurnaan.

E. Penyusunan laporan tugas akhir

Pada tahap ini dilakukan penyusunan laporan yang berisi dasar teori, dokumentasi dari perangkat lunak, dan hasil-hasil yang diperoleh selama pengerjaan tugas akhir.

1.6 Sistematika Penulisan

Sistematika penulisan buku tugas akhir dibagi menjadi beberapa bab sebagai berikut:

- A. Bab I, Pendahuluan, berisi latar belakang, rumusan masalah, batasan masalah, tujuan, metodologi, dan sistematika penulisan.
- B. Bab II, Landasan Teori, yang berisi dasar penyelesaian ekstraksi fitur dari data *microarray*, penjelasan mengenai model klasifikasi berbasis informasi pasangan gen, evaluasi model klasifikasi.
- C. Bab III, Metodologi, akan dibahas deskripsi umum sistem, perancangan data, serta pemodelan proses yang digunakan pada tugas akhir ini dengan menggunakan diagram alir.
- D. Bab IV, Implementasi, akan dilakukan pembuatan aplikasi yang dibangun dengan MATLAB sesuai dengan permasalahan dan batasannya yang telah dijabarkan pada bab pertama.
- E. Bab V, Uji coba dan evaluasi hasil, akan dilakukan uji coba dari aplikasi yang dibuat dengan melihat data keluaran yang dihasilkan oleh aplikasi. Dari data keluaran tersebut dapat dilakukan analisis hasil untuk mengetahui kemampuan dari aplikasi yang dibuat serta membandingkan hasil penyelesaian permasalahan dari skenario uji coba yang dilakukan.
- F. Bab VI, Kesimpulan dan saran, berisi kesimpulan yang dapat diambil dari tugas akhir ini beserta saran untuk pengembangan selanjutnya.