

BAB I

PENDAHULUAN

1.1. Latar Belakang

Analisis diskriminan merupakan teknik menganalisis data, dimana variabel dependen merupakan data kategori (nominal dan ordinal) sedangkan variabel independen berupa data interval atau rasio. Analisis diskriminan bertujuan untuk mengklasifikasikan suatu individu atau observasi ke dalam kelompok yang saling bebas (*mutually exclusive*) atau dikenal juga dengan metode dependensi yaitu variabel-variabelnya tidak saling bergantung satu dengan yang lain. Berbeda dengan metode interdependensi yaitu metode yang antarvariabelnya ada saling ketergantungan. Sistem tersebut juga dapat menganalisa dan mengidentifikasi bentuk, tekstur, atau bahkan warna dari suatu objek.

Dalam industri telekomunikasi, penerapan analisis diskriminan pun sangat diperlukan. Berbagai macam data yang penting mengenai infrastruktur pembangunan jaringan telekomunikasi maupun data yang berkaitan dengan pelanggan akan dikumpulkan dalam suatu database yang dikumpulkan oleh manusia atau pekerja industri telekomunikasi tersebut. Hal ini kerap kali menimbulkan kesalahan atau *human error* secara tidak sengaja. Meskipun jumlah error sedikit, tetapi cukup berdampak pada infrastruktur telekomunikasi lain yang masih berkaitan. Maka analisis perlu dilakukan agar dapat diketahui letak data yang error atau bersifat sebagai *outlier*.

Sebagai suatu metode yang alternative dibandingkan metode-metode yang sudah ada digunakan metode yang memanfaatkan hubungan antara fitur-fitur pada suatu kelas untuk mengklasifikasikan suatu sampel pada kelas tertentu dengan data yang berjumlah besar yang tidak dapat dengan mudah diselesaikan oleh metode-metode klasifikasi yang lain. Maka akan dilakukan suatu analisis yang bertujuan untuk merealisasikan penggunaan metode Linear Discriminant Analysis dalam keakurasian klasifikasi data termasuk menentukan munculnya *data error* dalam bentuk outlier dengan menggunakan perangkat lunak MATLAB. Dari keunggulan metode Linear Discriminant Analysis ini kemudian akan diaplikasikan Cross Validation. Untuk membuktikan kemampuan klasifikasi, akan dilakukan pengujian pada metode ini dengan menggunakan data set *Wine*. Penggunaan dataset tersebut dikarenakan

inter-relation fitur pada data set tersebut lebih kuat daripada contoh data set lain. Hal ini semakin menunjang untuk menampilkan keunggulan penggunaan metode klasifikasi.

Pembangunan klasifikasi dataset dengan menggunakan metode Linear Discriminant Analysis dan *Principal Component Analysis* dengan *Cross Validation* ini dibagi menjadi dua bagian, yaitu tahap pelatihan dan tahap pengujian. Tahap pelatihan bertujuan untuk mendapatkan fitur-fitur penting yang akan menjadi masukan untuk MATLAB, dimana pada tahap pelatihan proses klasifikasi ini bertujuan untuk mendapatkan *classifier* yang dapat memisahkan kelas-kelas yang ada pada dataset. Kemudian tahap pengujian bertujuan untuk melihat bagaimana sistem berjalan dari awal sampai akhir dengan beberapa parameter pengujian. Diharapkan dengan adanya penelitian ini dapat menghasilkan sistem klasifikasi data dengan akurasi yang baik dan penggunaan metode yang tepat.

1.2. Tujuan

Adapun tujuan dari penelitian dan analisis dari Tugas Akhir ini adalah:

1. Melakukan pengklasifikasian data menggunakan Linear Discriminant Analysis yang bersifat distribusi normal dan homoskedastik yang dapat memaksimalkan sistem pengenalan pola dalam mengklasifikasikan data.
2. Mengklasifikasikan data dari dataset ke dalam bentuk yang ternormalisasi, baik pada Linear Discriminant Analysis maupun pada Principal Component Analysis.
3. Mengaplikasikan teknik klasifikasi yang telah ternormalisasi dalam metode *Cross Validation*.

1.3. Rumusan Masalah

Permasalahan yang dijadikan obyek penelitian dan pengembangan pada Tugas Akhir ini adalah:

1. Bagaimana mengklasifikasikan data dengan menggunakan metode Linear Discriminant Analysis dan Principal Component Analysis?
2. Bagaimana membentuk sebuah dataset ke dalam bentuk data yang telah ternormalisasi?

3. Bagaimana merancang algoritma metode klasifikasi pada perangkat lunak MATLAB?

1.4. Batasan Masalah

Untuk menghindari meluasnya materi dalam pembahasan Tugas Akhir ini, maka penulis membatasi permasalahan dalam Tugas Akhir ini dengan hanya mencakup hal-hal berikut:

- a. Data yang digunakan dalam penelitian berasal dari UCI Repository.
- b. Penelitian ini menggunakan metode Linear Discriminant Analysis.
- c. Klasifikasi yang telah dihasilkan akan dinormalisasi dan diiterasi dengan teknik Cross Validation.

1.5. Metode Penelitian

Metode penelitian yang digunakan untuk menyelesaikan tugas akhir ini adalah metode ekperimental, sedangkan prosedur penelitian untuk menyelesaikan Tugas Akhir ini adalah:

- a. Studi Literatur
Prosedur ini dilakukan untuk mempelajari konsep dan teori-teori yang mendukung proses perancangan sistem
- b. Survey lapangan
Analisis permasalahan dan kebutuhan
- c. Perancangan dan realisasi
Merancang suatu analisis dengan menggunakan algoritma pada *software* MATLAB
- d. Analisis dan pengujian
Berdasarkan teori yang sudah diperoleh sebelumnya, kemudian melakukan pengujian terhadap rancangan analisis pada *software*.
- e. Konsultasi
Dilakukan dengan dosen pembimbing menyangkut petunjuk dan pertimbangan-pertimbangan praktis mengenai perancangan analisis tersebut.

1.6. Sistematika Penulisan

Penulisan proposal ini akan dibagi menjadi beberapa bagian sebagai berikut:

BAB I Pendahuluan

Berisi latar belakang, tujuan, rumusan masalah, batasan masalah, metode penelitian, dan sistematika penulisan

BAB II Dasar Teori

Mengenai dasar-dasar teori yang diperlukan serta literature-literatur yang mendukung dalam penelitian dan analisis penggunaan metode Linear Discriminan Analysis pada sistem *error rate classifier*.

BAB III Perancangan dan Implementasi

Mengenai pembahasan perancangan awal system yang kemudian diimplementasikan dengan menggunakan algoritma pada perangkat lunak MATLAB.

BAB IV Pengujian dan Analisis Hasil Implementasi

Membahas mengenai pengujian dan analisis hasil implementasi yang telah dilakukan sebelumnya. Hal ini bertujuan untuk mengetahui kecocokan antara analisis yang dilakukan secara teori dengan menggunakan perangkat lunak MATLAB.

BAB V Kesimpulan dan Saran

Berisi kesimpulan dari penelitian Tugas Akhir ini serta saran yang diperlukan untuk pengembangan sistem lebih lanjut.