

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pengenalan pola adalah proses mengenali dan mengidentifikasi pola. Pola yang dimaksud merujuk pada suatu urutan, struktur, atau keteraturan yang dapat dikenali dalam data. Tujuannya adalah untuk mengklasifikasikan data berdasarkan pola tertentu. Hal tersebut melibatkan penggunaan algoritma untuk menganalisis kumpulan data yang besar dan mengidentifikasi suatu pola. Contoh dari pengenalan pola adalah pengenalan wajah, pengenalan teks, dan pengenalan suara.

Kumpulan data berdimensi tinggi menghadirkan tantangan baru dalam pemrosesan data. Hampir semua algoritma pembelajaran mesin kinerjanya menurun pada data berdimensi tinggi, karena data berdimensi tinggi berpeluang mengandung *noise*, redundan (berkorelasi antar variabel), dan fitur dengan varian kecil. Hal ini dapat menyebabkan fenomena yang disebut kutukan dimensi (B. Li et al., 2019), sehingga perlu dilakukan reduksi dimensi.

Reduksi dimensi bermanfaat untuk meningkatkan kinerja algoritma pembelajaran mesin dan bertujuan untuk memetakan kumpulan data berdimensi tinggi ke ruang berdimensi lebih rendah tanpa kehilangan informasi yang berarti (He et al., 2021). Analisis Diskriminan Linier (LDA) dan Analisis Komponen Utama (PCA) adalah metode yang banyak digunakan untuk reduksi dimensi. LDA merupakan metode reduksi dimensi yang berbasis kelas (*supervised*), sedangkan PCA merupakan metode reduksi dimensi yang tidak berbasis kelas (*unsupervised*) (He et al., 2021). Pada penelitian ini digunakan metode LDA untuk melakukan reduksi dimensi. LDA bertujuan untuk memaksimalkan pemisahan kelompok data dengan melakukan transformasi data ke ruang proyeksi dalam subruang berdimensi rendah.

Jika dimensi data jauh lebih besar dibandingkan dengan jumlah vektor pelatihan, maka matriks pencar dalam kelas adalah matriks singular, sehingga

metode LDA tidak dapat diterapkan (He et al., 2021). Kelemahan ini dianggap sebagai masalah utama LDA yang umumnya dikenal sebagai masalah singularitas yang disebabkan karena ukuran sampel kecil (Sharma & Paliwal, 2015).

Penelitian ini fokus membahas metode NLDA. Metode ini merupakan variasi dari metode LDA yang bertujuan untuk mengatasi masalah singularitas pada matriks pencar yang disebabkan oleh ukuran sampel kecil (He et al., 2021). Cara bekerjanya yaitu dengan memproyeksikan data ke dalam ruang nul dari matriks pencar dalam kelas dan kemudian memaksimalkan jarak antar kelas dalam ruang nul.

Dalam penelitian ini, penulis membahas tiga variasi metode NLDA dengan menerapkan dekomposisi matriks, yaitu metode NLDA dengan Dekomposisi Nilai Singular (SVD), metode NLDA dengan Dekomposisi Cholesky, dan metode NLDA dengan Dekomposisi QR, yang pada penelitian sebelumnya masing-masing telah diusulkan oleh (Ye & Xiong, 2006), (Lu & Zheng, 2013), dan (Lu & Wang, 2012). Berdasarkan algoritma NLDA, Dekomposisi Nilai Singular (SVD), Dekomposisi Cholesky, dan Dekomposisi QR berperan dalam pencarian ruang nul dari matriks pencar dalam kelas, sehingga matriks hasil transformasi dapat diperoleh.

Untuk mengetahui efektivitas dari masing-masing metode NLDA, perlu dilakukan evaluasi akurasi klasifikasi yang dapat diperoleh dari matriks hasil transformasi. Dalam hal ini, efektivitas yang dimaksud adalah penilaian sejauh mana metode tersebut berhasil meningkatkan pemisahan antar kelas. Akurasi klasifikasi dihitung dengan menerapkan beberapa kumpulan data sebagai bahan pengujian. Dengan demikian, kinerja ketiga metode NLDA dalam menghasilkan representasi variabel yang optimal dapat dibandingkan secara kuantitatif.

1.2 Perumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka permasalahan yang dibahas pada penelitian ini sebagai berikut :

1. Bagaimana langkah-langkah dari tiga metode NLDA menggunakan dekomposisi matriks untuk mengatasi masalah singularitas matriks pencar pada metode LDA?
2. Bagaimana perbandingan kinerja dari metode NLDA menggunakan SVD, Dekomposisi Cholesky, dan Dekomposisi QR?

1.3 Pembatasan Masalah

Penelitian ini dibatasi dengan menjelaskan masalah singularitas pada matriks pencar dari data ukuran sampel kecil yang akan diatasi dengan tiga macam metode NLDA, yaitu metode NLDA menggunakan Dekomposisi Nilai Singular (SVD), Dekomposisi Cholesky, dan Dekomposisi QR.

Data simulasi numerik yang digunakan berupa data sekunder yang tersedia untuk umum. Tiga kumpulan data digunakan untuk menerapkan tiga metode NLDA tersebut.

1.4 Tujuan Penelitian

Tujuan dalam penelitian ini sebagai berikut :

1. Mempelajari langkah-langkah tiga metode NLDA serta mengidentifikasi kontribusi masing-masing metode dekomposisi matriks dalam proyeksi matriks pencar ke ruang nul.
2. Mengevaluasi efektivitas dari tiga metode NLDA dengan membandingkan akurasi klasifikasi yang dihasilkan oleh ketiga metode NLDA yang diterapkan pada kumpulan data yang berbeda untuk menentukan metode yang paling optimal.

1.5 Manfaat Penelitian

Manfaat dari hasil penelitian ini adalah sebagai berikut :

1. Menambah dan memperdalam pengetahuan terkait metode NLDA dan penerapannya pada permasalahan nyata di berbagai bidang.
2. Memberikan referensi baru terkait metode NLDA di berbagai bidang untuk penulisan karya ilmiah yang lebih lanjut dari permasalahan terkait.

3. Memberikan informasi mengenai reduksi dimensi pada data ukuran sampel kecil, sehingga metode NLDA dapat digunakan dalam permasalahan data di berbagai bidang.

1.6 Sistematika Penulisan

Sistematika penulisan dalam penelitian ini terdiri dari tiga bab yaitu Pendahuluan, Tinjauan Pustaka, dan Metodologi Penelitian

BAB I Pendahuluan adalah bab yang terdiri dari latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

BAB II Tinjauan Pustaka adalah bab yang berisi teori penunjang untuk mendukung pembahasan pada penelitian ini, yang meliputi penelitian terdahulu, ruang vektor, ortogonalitas dalam ruang vektor, Dekomposisi Nilai Singular, Dekomposisi Cholesky, Dekomposisi QR, ukuran sampel kecil, reduksi dimensi, analisis diskriminan linier (LDA), LDA klasifikasi, dan metode evaluasi.

Bab III Metodologi Penelitian adalah bab yang menjelaskan langkah-langkah yang digunakan penulis untuk menulis penelitian ini yang meliputi metode penelitian, bahan dan alat penelitian, dan langkah-langkah penelitian.

Bab IV Pembahasan adalah bab yang menjelaskan metode LDA berbasis ruang nul (NLDA), metode NLDA menggunakan Dekomposisi Nilai Singular (SVD), metode NLDA menggunakan Dekomposisi Cholesky, metode NLDA menggunakan Dekomposisi QR, deskripsi data, visualisasi data dari model NLDA, serta evaluasi model NLDA.

Bab V Penutup adalah bab yang berisi kesimpulan dari hasil penelitian dan saran bagi pembaca maupun bagi peneliti selanjutnya.