

BAB I

PENDAHULUAN

Pada bab ini dijelaskan bagian pendahuluan penelitian, yang meliputi latar belakang, rumusan masalah, tujuan, ruang lingkup, metode penelitian, tinjauan Pustaka, dan sistematika penulisan.

1.1. Latar Belakang

Penurunan kualitas udara salah satu penyebabnya adalah tingkat emisi dari kendaraan yang memproduksi partikulat sehingga dapat menurunkan kualitas udara (Kasus, Monginsidi, Kawasan, & Mall, 2020). Salah satu kendaraan bermotor yang sering melintas di jalan adalah angkutan umum, dimana kendaraan angkutan umum merupakan moda transportasi yang digunakan untuk mengangkut orang atau barang dari suatu tempat ke tempat lain. Terdapat suatu cara dalam mendeteksi kualitas dari gas buang yang dihasilkan oleh kendaraan yaitu melakukan uji emisi gas. Parameter yang diuji pada uji emisi gas diantaranya *Hidrokarbon* (HC) dan *Karbon Monoksida* (CO). Dari parameter yang diuji terdapat permasalahan saat pengujian seperti salah satu parameter yang lebih tinggi dari parameter lainnya. Dalam hal ini dapat dilakukan proses pengelompokan agar mengetahui data-data tersebut termasuk kedalam kelompok yang sebenarnya.

Metode pengelompokan seperti *K-Medoids* (Park & Jun, 2009), *Fuzzy MADM* (Kusumadewi Sri, Hartati Sri, Harjoko Agus, 2006) dan *Self Organizing Maps* (Teuvo Kohonen, Erkki Oja, Olli Simula, Ari J.E. Visa, n.d.) diusulkan banyak peneliti untuk melakukan pengelompokan. Algoritma *K-Medoids* memiliki kelebihan handal dibanding dengan algoritma lain ketika ada data *noise* dan pencilan karena k-medoids tidak terlalu dipengaruhi oleh data pencilan atau data ekstrem lainnya. Tetapi algoritma *K-Medoids* memiliki kekurangan yaitu tingkat efisiensi yang kurang efisien. Algoritma *Fuzzy MADM* mempunyai kelebihan kemampuan untuk mengelompokkan data yang besar dan lebih kokoh terhadap data outlier. Tetapi Algoritma *Fuzzy MADM* memiliki kekurangan yaitu *running timenya* $O(NC^2 T)$. Algoritma *Self Organizing Maps* (SOM)

memiliki kelebihan memetakan data berdimensi tinggi kedalam bentuk peta berdimensi rendah. Tetapi Algoritma *Self Organizing Maps* (SOM) juga memiliki kekurangan yaitu hasil akurasi kurang baik dikarenakan data yang digunakan adalah data kategori yang ditransformasikan menjadi data numerik.

Algoritma K-Means dapat memecahkan masalah *K-Medoids*, *Fuzzy MADM* dan *Self Organizing Maps* (SOM), yaitu implementasi yang rumit, lambatnya waktu dalam pemrosesan. Algoritma *K-Means* dapat dilakukan saat pengimplementasian dan dijalankan serta waktu yang di butuhkan untuk melakukan pembelajaran relatif lebih cepat. Tetapi algoritma *K-Means* memiliki kekurangan dalam penentuan titik *centroid* atau pusat kluster dimana pada tahap inisialisasi awal dari Algoritma *K-Means* ditentukan secara acak, pemilihan titik pusat kluster (*intial centroid*) yang berbeda akan menghasilkan kluster yang berbeda pula. Algoritma genetika dapat memecahkan permasalahan dari penentuan titik pusat kluster awal adalah salah satu algoritma untuk optimalisasi algoritma *K-Means* yang memiliki kelemahan dalam menentukan titik pusat kluster awal (Al-shboul & Myaeng, n.d.). Selain memiliki kelemahan dalam menentukan titik kluster awal, terdapat kelemahan lain yaitu menentukan jumlah K terbaik dalam mengkluster suatu data (Putu, Merliana, & Santoso, n.d.). Dalam hal ini digunakan beberapa metode diantaranya *Elbow*, *Silhouette Coefficient* dan *Davies-Bouldin Index* (DBI). Ketiga metode tersebut akan diuji lalu dipilih salah satu metode mana yang hasil dari pencarian jumlah kluster yang optimal. Untuk mengukur jarak data dengan pusat kluster pada Algoritma *K-Means* menggunakan pendekatan metode *Euclidean Distance* karena pada metode tersebut mendapatkan akurasi tertinggi dibandingkan dengan metode *manhattan* dan *minkowski* (Nishom, 2019).

Pada penelitian ini algoritma *K-Means* diterapkan untuk pengelompokan kelaikan angkutan umum yang di optimalisasi Algoritma Genetika dalam mencari titik pusat kluster awal dan metode *Elbow*, *Silhouette Coefficient* dan *Davies-Bouldin Index* (DBI) untuk mencari jumlah kluster yang terbaik sehingga hasil dari pengelompokan lebih optimal.

1.2. Rumusan Masalah

Dalam menjawab permasalahan pengelompokan kendaraan yang tidak laik beroperasi dimana terjadi sebagai akibat dari tidak melakukan peremajaan terhadap angkutan umum yang berdampak menurunnya kualitas udara. Maka Diperlukan sistem untuk mengelompokan kelaikan angkutan umum.

Berdasarkan identifikasi yang telah ditetapkan muncul berbagai masalah diantaranya sebagai berikut :

- a) Bagaimana menentukan jumlah kluster terbaik menggunakan metode *Elbow*, *Silhouette Coefficient* dan *Davies-Bouldin Index*.
- b) Bagaimana Algoritma Genetika dapat menentukan inisialisasi pusat kluster awal atau *initial centroid*.
- c) Bagaimana melakukan pengelompokan menggunakan Algoritma *K-Means*.

1.3. Tujuan

Penelitian ini bertujuan untuk menerapkan algoritma genetika dalam menentukan inisialisasi pusat kluster awal pada *K-Means* serta metode *Elbow*, *Silhouette Coefficient* dan *Davies-Bouldin Index* (DBI) untuk mencari jumlah kluster terbaik dalam pengelompokan kelaikan angkutan umum dengan parameter *Karbon Monoksida* (CO) dan *Hidrokarbon* (HC) yang merupakan hasil uji emisi gas buang sehingga hasil sistem dapat lebih optimal.

1.4. Ruang Lingkup

Batasan-batasan yang membatasi penelitian yang dilakukan adalah sebagai berikut.

- a) Proses pengolahan data dilakukan berbasis *desktop*.
- b) Jenis bahan bakar pada kendaraan umum yang dipilih yaitu bensin.
- c) Angkutan umum yang dipilih trayek Soreang-Banjaran.
- d) Dataset berbentuk CSV.
- e) Parameter uji emisi *Karbon Monoksida* (CO) dan *Hidrokarbon* (HC).

1.5. Sistematika Penulisan

Penulisan laporan penelitian ini terdiri dari lima bab dengan sistematika penulisan sebagai berikut :

BAB I PENDAHULUAN

Bab I berisi latar belakang, rumusan masalah, tujuan, ruang lingkup dan sistematika penulisan laporan.

BAB II LANDASAN TEORI

Bab II dijelaskan mengenai tinjauan pustaka serta teori-teori untuk penunjang dalam pembuatan laporan. Teori sebagai penunjang tersebut ialah seperti Angkutan Umum, Klasterisasi, Normalisasi Data, Algoritma *K-Means*, Algoritma Genetika, *Elbow*, *Silhouette Coefficient* dan *Davies-Bouldin Index*.

BAB III METODE PENELITIAN

Bab III dijelaskan metode yang digunakan dan tahapan-tahapan seperti blok diagram, kebutuhan sistem dan proses pengumpulan data yang digunakan.

BAB IV IMPLEMENTASI DAN PENGUJIAN

Bab IV dijelaskan mengenai implementasi serta pengujian dari sistem yang telah dibuat dari hasil pengujian terhadap fungsionalitas sistem berdasarkan hasil analisa.

BAB V PENUTUP

Bab V ini diuraikan mengenai kesimpulan dari penelitian yang telah dilakukan.