

BAB I

PENDAHULUAN

1.1 Latar Belakang

Udara bersih, yang dibutuhkan untuk bernapas, merupakan salah satu komponen vital bagi kehidupan di bumi. Namun, mendapatkan udara bersih kini semakin sulit karena kualitas udara terus menurun. Beberapa faktor yang menyebabkan penurunan ini antara lain meningkatnya jumlah industri yang melepaskan polutan seperti partikel halus (PM_{10}) dan ($PM_{2.5}$), peningkatan emisi gas buang seperti karbon dioksida (CO_2), karbon monoksida (CO), ozon (O_3), dan nitrogen monoksida (NO), serta tingginya penggunaan kendaraan bermotor. Pada bulan juni, Jakarta tercatat sebagai kota dengan kualitas udara terburuk di dunia, dengan indeks kualitas udara (AQI) mencapai angka 166 (Fahmi, et al, 2022). Kota Yogyakarta merupakan salah satu wilayah dengan aktivitas manusia yang tinggi. Pertumbuhan penduduk yang terus meningkat berkontribusi pada bertambahnya jumlah kendaraan sebagai alat mobilitas dalam memenuhi kebutuhan sehari-hari, sehingga memicu tingginya mobilisasi penduduk. Kualitas udara di Yogyakarta sebagian besar dipengaruhi oleh emisi dari kendaraan bermotor, yang menjadi sumber pencemaran utama atau disebut juga sumber bergerak. Hal ini diperkuat oleh Yogyakarta sebagai kota yang identik dengan pariwisata, industri, pusat perkantoran, serta kegiatan Pendidikan.

Menurut Undang-Undang Nomor 32 Tahun 2009, lingkungan hidup mencakup seluruh unsur di sekitar manusia yang dapat memengaruhi cara hidupnya, baik secara langsung maupun tidak langsung. Lingkungan hidup terbentuk dari komponen fisik dan kimia, seperti udara, tanah, dan air. Udara, sebagai salah satu elemen penting dalam lingkungan, memiliki peran vital dalam menunjang kehidupan, sehingga kualitasnya perlu dijaga dan ditingkatkan agar dapat mendukung keberlangsungan hidup makhluk hidup secara optimal menurut Despek (Mayasari, et al, 2020)

Dalam proses pencemaran udara terdapat komponen atau zat lain yang masuk ke dalam udara, komponen tersebut yaitu *Partikulat* (PM_{10} , $PM_{2.5}$), *Karbon Dioksida* (CO_2), *Karbon Monoksida* (CO), *Ozon* (O_3), dan *Nitrogen Monoksida* (NO). Dapat menjadi indikator parameter tingkat baik atau buruk udara pada suatu

daerah termasuk di Yogyakarta. Indikator parameter diatas termasuk dalam aspek pemantauan pada stasiun pemantauan kualitas udara yang tersebar di Yogyakarta pada tahun 2020 dan 2021. Pada data hasil pemantauan di beberapa SPKU Yogyakarta telah dapat memberikan kategori bagus, sedang, dan tidak sehat. Akan tetapi data hasil pemantauan tersebut belum dapat menampilkan informasi berupa prediksi pengaruh terhadap indikator parameter kualitas udara terutama pada kondisi parameter masing-masing bernilai sama.

Pemanfaatan data mining dapat diterapkan untuk mengklasifikasikan kualitas udara di Yogyakarta. Data mining merupakan perpaduan dari berbagai disiplin ilmu yang mengintegrasikan Teknik pembelajaran mesin, pengenalan pola, statistika, serta visualisasi guna menangani permasalahan dalam pengambilan informasi dari basis data berukuran besar (Lumban Gaol, et al, 2021). Algoritma pohon keputusan (*decision tree*) jenis C4.5 merupakan salah satu metode yang digunakan untuk melakukan klasifikasi, segmentasi, atau pengelompokan data, serta memiliki kemampuan *prediktif* (Siahaan et al., 2022). Berdasarkan penelitian (Halim, et al, 2023) yang berjudul “Penerapan Metode *Decision Tree* Untuk Prakiraan Cuaca Kota Bekasi” menghasilkan bahwa Algoritma C4.5 menggunakan situs web dan program php mendapatkan akurasi 82%. Penelitian lain yang dilakukan oleh (Putri, 2021) berjudul “Implementasi Data Mining Efektivitas pada Mesin Injection Menggunakan Algoritma C4.5” menunjukkan bahwa algoritma C4.5 mampu mencapai akurasi sebesar 88,30% dan 98,64%, dengan nilai *recall* 84,91% serta *precision* 93,75%. Kurva ROC/AUC pada data pelatihan mencapai 0,999 yang termasuk dalam kategori klasifikasi sangat baik. Dengan demikian, dapat disimpulkan bahwa penerapan data mining menggunakan algoritma c4.5 untuk memprediksi efektivitas mesin injection menghasilkan klasifikasi yang akurat.

Berdasarkan uraian pada bagian pendahuluan, kedua penelitian ini yang telah dibahas menyimpulkan bahwa algoritma C4.5 tergolong mudah digunakan untuk keperluan klasifikasi, serta memiliki tingkat akurasi yang lebih tinggi dibandingkan dengan algoritma *K-nearest* dan *Naïve Bayes*. Oleh karena itu, penulis memilih untuk melakukan penelitian dengan judul “Implementasi Data Mining untuk Klasifikasi Kualitas Udara di Yogyakarta Menggunakan Algoritma C4.5”.

1.2 Rumusan Masalah

1. Bagaimana keadaan kualitas udara di Yogyakarta jika ditinjau dari parameter partikel seperti PM_{10} , $PM_{2.5}$, SO_2 , CO , O_3 dan NO_2 ?
2. Bagaimana penerapan algoritma C4.5 dalam *Data Mining* untuk mengklasifikasi kualitas udara di Yogyakarta?
3. Seberapa besar nilai *accuracy*, *recall*, *precision*, *f1-score*, yang dihasilkan oleh algoritma C4.5 dalam proses klasifikasi kualitas udara di Yogyakarta?

1.3 Tujuan Penelitian

1. Menganalisis kualitas udara di Yogyakarta berdasarkan parameter seperti PM_{10} , $PM_{2.5}$, SO_2 , CO , O_3 , dan NO_2 .
2. Menerapkan algoritma C4.5 dalam metode data mining untuk klasifikasi kualitas udara.
3. Mengevaluasi performa model klasifikasi yang dibentuk berdasarkan metrik *accuracy*, *recall*, *precision* dan *F1-Score*.

1.4 Manfaat

1.4.1. Manfaat Bagi Penulis

1. Penelitian ini dilakukan sebagian bagian dari persyaratan untuk memperoleh gelar Serjana Komputer pada Fakultas Ilmu Komputer, Universitas Buana Perjuangan Karawang.
2. Sebagai bentuk penilaian atas kompetensi penulis selama menempuh pendidikan di Fakultas Ilmu Komputer, Universitas Buana Perjuangan Karawang.
3. Mempelajari bagian dan pengimplementasian dari Data Mining dan *Knowledge Discovery in Database* (KDD). Mengkaji aspek-aspek serta penerapan Data Mining dan proses *Knowledge Discovery in Database* (KDD).
4. Memahami penerapan algoritma C4.5 dalam melakukan klasifikasi terhadap kualitas udara.

1.4.2. Manfaat Bagi Universitas

1. Mengetahui kemampuan dari penulis dalam menerapkan ilmu-ilmu yang dipelajari selama perkuliahan
2. Menjadi referensi untuk penelitian penelitian selanjutnya.