

BAB I PENDAHULUAN

I.1 Latar Belakang

Udara, sebagai salah satu komponen lingkungan yang merupakan kebutuhan yang paling utama untuk mempertahankan kehidupan. Metabolisme dalam tubuh makhluk hidup tidak mungkin dapat berlangsung pada oksigen yang berasal dari udara. Selain oksigen terdapat zat-zat lain yang terkandung di udara, yaitu karbon monoksida, karbon dioksida, formaldehid, jamur, virus, dan sebagainya. Zat-zat tersebut jika masih berada dalam batas-batas tertentu masih dapat dinetralsisir, tetapi jika sudah melampaui ambang batas maka proses netralisir akan terganggu. Peningkatan konsentrasi zat-zat di dalam udara tersebut dapat disebabkan oleh aktivitas manusia (Wulandari, dll., 2008). Udara bersih yang dihirup hewan dan manusia merupakan gas yang tidak nampak, tidak berbau, tidak berwarna maupun berasa. Meskipun demikian, udara yang benar-benar bersih sulit didapatkan terutama di kota besar yang banyak terdapat industri dan lalu lintas yang padat. Udara yang tercemar dapat merusak lingkungan dan kehidupan manusia, yang pada akhirnya akan mengurangi kualitas hidup manusia secara keseluruhan.

Di dalam udara terkandung gas yang terdiri dari 78% nitrogen, 20% oksigen, 0,93% argon, 0,03% karbon dioksida, dan sisanya terdiri dari neon, helium, metana dan hidrogen. Komposisi tersebut dikatakan sebagai udara normal dan dapat mendukung kehidupan manusia. Namun, akibat aktivitas manusia yang tidak ramah lingkungan, udara seing keli menurun kualitasnya. Perubahan ini dapat berupa sifat-sifat fisis maupun kimiawi. Perubahan kimiawi dapat berupa pengurangan maupun penambahan salah satu komponen kimia yang terkandung dalam udara. Kondisi seperti itu lazim disebut dengan pencemaran udara.

Buruknya kualitas udara disebabkan oleh pencemaran udara contohnya terdapat di negara Indonesia terutama pada provinsi DKI Jakarta. Pencemaran udara di DKI Jakarta disebabkan oleh beberapa faktor diantaranya kendaraan bermotor, Industri, Pembakaran batu bara, debu jalanan, dan debu aktivitas konstruksi. Namun yang menjadi penyebab

utama adalah naiknya kebutuhan kendaraan yang menyumbang 31 sampai 40% pencemaran udara pada provinsi DKI Jakarta.

DKI Jakarta, sebagai ibu kota Indonesia, menjadi pusat aktivitas ekonomi, industri, dan transportasi yang berdampak signifikan pada peningkatan emisi polusi udara. Menurut laporan Kementerian Lingkungan Hidup dan Kehutanan (KLHK), Jakarta sering kali menduduki peringkat tertinggi dalam hal polusi udara di Indonesia, terutama akibat tingginya emisi dari kendaraan bermotor dan industri (Sudaryanti, dll., 2020).

Selain itu, Jakarta memiliki kepadatan penduduk yang tinggi, yang menyebabkan volume emisi dari kegiatan rumah tangga dan transportasi meningkat secara signifikan. Menurut data dari World Air Quality Report tahun 2019, Jakarta tercatat sebagai salah satu kota dengan kualitas udara terburuk di dunia (IQAir, 2019). Kondisi ini berdampak serius terhadap kesehatan masyarakat, terutama dalam meningkatkan risiko penyakit pernapasan seperti asma, bronkitis, dan penyakit paru obstruktif kronis (Rosatul, dll., 2021).

Pemilihan Jakarta sebagai objek penelitian didasarkan pada urgensi penanganan kualitas udara yang semakin memburuk, serta kebutuhan akan solusi yang berbasis data. Penggunaan metode klasifikasi seperti algoritma Random Forest memberikan potensi besar dalam memprediksi kualitas udara secara akurat berdasarkan data yang ada. Beberapa penelitian menunjukkan bahwa metode ini memiliki tingkat akurasi yang tinggi dalam memproses data lingkungan yang kompleks (Zhang dll., 2019). Melalui klasifikasi ini, diharapkan penelitian dapat memberikan rekomendasi yang lebih efektif kepada pemerintah daerah dalam rangka menurunkan tingkat polusi dan memperbaiki kualitas udara di DKI Jakarta.

Dari permasalahan yang ada peneliti akan melakukan identifikasi untuk memantau kualitas udara yang terdapat pada data pencemaran kualitas udara Provinsi DKI Jakarta yang akan menjadi acuan kepada pemerintah untuk membuat kebijakan dan pengelolaan yang baik terhadap kualitas udara di Provinsi DKI Jakarta.

Algoritma *data mining* telah diterapkan untuk klasifikasi polusi udara, alasan untuk menggunakan *data mining* adalah untuk menggunakan informasi rahasia di database yang memiliki data besar dengan perhitungan statistika dan model matematika (Meilani & Asadulloh, 2015). Klasifikasi

merupakan salah satu metode dalam pengolahan atau komputasi data untuk mengelompokan data sesuai dengan ciri-ciri yang sama (Aji, dll., 2018).

Hasil klasifikasi yang akurat sangat membantu pemerintah dalam membuat suatu kebijakan. Kebijakan ini bertujuan untuk berusaha mengendalikan pencemaran agar berada pada standar mutu kualitas udara yang dapat bermanfaat sebaik-baiknya terhadap kelangsungan makhluk hidup. Model klasifikasi yang digunakan adalah *random forest*. *Random forest* telah dianggap sebagai yang paling mutakhir. *Random forest* meningkatkan varians model pembelajaran dasar dengan menggabungkankonsep *bagging* dan subruan acak, dengan demikian meningkatkan kinerja model pembelajaran. *Random forest* telah diterapkan secara luas untuk masalah klasifikasi di berbagai bidang penelitian, microarray, segmentasi gambar, dan pemilihan fitur. Manuel Fernandez-Delgado dkk, membandingkan 179 model pembelajaran dari 17 keluarga dengan menggunakan 121 kumpulan data klasifikasi, di antaranya *random forest* mencapai kinerja terbaik (Qiu, dll., 2017).

Dalam penelitian ini menggunakan algoritma random forest karena random forest karena adalah algoritma machine learning yang fleksibel dan mudah digunakan, bahkan tanpa penyetelan hyper-parameter, tapi sering kali mendapat hasil yang bagus. Random forest juga salah satu algoritma yang paling banyak digunakan, karena kesederhanaan dan keragamannya. Ini juga menjadi satu keuntungan besar dari algoritma random forest yaitu dapat digunakan untuk masalah klasifikasi dan regresi, yang merupakan Sebagian besar sistem machine learning saat ini.

Penelitian ini membedakan dirinya dari sistem atau aplikasi yang sudah ada dalam beberapa aspek penting. Banyak aplikasi pemantauan kualitas udara yang telah beredar di pasar, seperti IQAir dan AirVisual, yang umumnya menggunakan data kualitas udara yang bersumber dari stasiun pemantauan resmi dan algoritma prediksi yang relatif sederhana (Kumar dll., 2020). Namun, aplikasi-aplikasi tersebut sering kali tidak mempertimbangkan variabel lokal dan dinamis yang memengaruhi kualitas udara secara real-time.

Sebaliknya, penelitian ini menggunakan algoritma Random Forest yang memungkinkan analisis yang lebih mendalam dan akurat terhadap faktor-

faktor yang mempengaruhi kualitas udara di DKI Jakarta.

I.2 Rumusan Masalah

Berdasarkan latar belakang diatas, maka masalah yang akan diteliti adalah sebagai berikut:

1. Bagaimana mengimplementasikan algoritma Random Forest dalam klasifikasi Indeks Standar Pencemaran Udara DKI Jakarta?
2. Bagaimana tingkat akurasi menggunakan algoritma Random Forest pada klasifikasi kualitas udara data ISPU di Provinsi DKI Jakarta?

I.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang sudah disebutkan pada latar belakang, adapun tujuan dari penelitian ini adalah sebagai berikut :

1. Dapat mengimplementasikan algoritma Random Forest pada analisis klasifikasi kualitas pencemaran udara Provinsi DKI Jakarta.
2. Dapat mengetahui tingkat akurasi algoritma Random Forest dalam melakukan klasifikasi data kualitas udara Provinsi DKI Jakarta.

I.4 Manfaat Penelitian

Dari penelitian ini, diharapkan masyarakat dapat mengetahui informasi mengenai kualitas udara yang ada pada lingkungan sekitarnya dan dapat menjaga kualitas udara tetap berada dalam kondisi yang ideal.

I.5 Batasan Masalah

Adapun Batasan masalah dalam penelitian tugas akhir ini yaitu :

1. Data yang digunakan didapatkan dari website Open Data Jakarta (data.jakarta.go.id).
2. Dalam penelitian data yang akan digunakan dari tahun 2019 hingga tahun 2022.
3. Parameter yang diuji dalam analisis meliputi PM10(Partikulat Matter), SO₂(Sulfur Dioksida), CO(Karbon Monoksida), O₃(Ozon), dan NO₂(Nitrogen Oksida).

I.6 Sistematika Penulisan

Sistematika penulisan dari penelitian ini terdiri dari beberapa bab antara lain:

BAB I PENDAHULUAN

Bab ini membahas latar belakang dari penelitian yang akan dilakukan dilengkapi dengan tujuan dan batasan - batasan yang ditetapkan.

BAB II TINJAUAN PUSTAKA

Bab ini membahas teori - teori relevan yang terkait yang digunakan untuk

mendukung penelitian.

BAB III METODOLOGI PENELITIAN

Bab ini akan menjelaskan metode, model penelitian, prosedur pengerjaan, dan analisis data yang akan digunakan dalam pengerjaan laporan tugas akhir.

BAB IV PERANCANGAN

Bab ini akan menjelaskan skema perancangan dan perhitungan pada proses pengklasifikasian.

BAB V ANALISIS HASIL

Bab ini akan menjelaskan hasil klasifikasi yang terdapat pada BAB IV yang sudah diuji terlebih dahulu.

BAB VI KESIMPULAN DAN SARAN

Bab ini akan menjelaskan tentang kesimpulan dari hasil penelitian dan saran saran yang dapat membangun hasil penelitian lebih baik lagi kedepannya.