

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kepadatan penduduk, aktivitas industri, praktik pertanian, pembangkit listrik tenaga panas, sektor energi, industri otomotif dan transportasi semuanya mempengaruhi polusi udara secara berbeda (Ravinda, 2019). Polusi udara menjadi perhatian utama karena efek berbahaya yang disebabkan oleh polusi udara sangat signifikan dalam mempengaruhi kehidupan sehari-hari. Polusi udara juga mempunyai efek buruk pada Kesehatan manusia termasuk kematian dini, penyakit kulit, infeksi paru-paru, infeksi saluran pernapasan, pneumonia, kanker paru-paru dan gagal jantung (Manisalidis et al., 2020). Pencemaran udara pun dapat memberikan efek buruk pada perkembangan tumbuhan, meningkatnya *global warming*, biosfer mengalami perubahan, meningkatnya air laut dan lapisan *ozone* mengalami penipisan. Tentunya efek buruk ini sangat berdampak pada setiap lapisan kehidupan di bumi.

Satu dari banyaknya penyebab polusi udara yaitu hutan yang terbakar, lahan yang terbakar, hal ini disebut sebagai karhutla. karhutla merupakan suatu kejadian yang menyebabkan kawasan hutan ataupun lahan dalam skala kecil maupun luas terbakar. Terkhusus untuk daerah Provinsi Sumatera Selatan, karhutla setiap tahunnya akan selalu terjadi seperti “tradisi” tahunan namun dengan tingkat polusi yang berbeda setiap tahunnya. Pada saat ini tanggal 2 Oktober 2023 berdasarkan data dari website IQAir yang di akses pada jam 10.00 PM bahwa Kota Palembang Provinsi Sumatera Selatan menempati posisi pertama dari 10 kota di Indonesia yang paling buruk tingkat polusi udaranya dengan nilai AQI US atau rata-rata unit 259 berdasarkan parameter ISPU. Sejak Januari hingga Agustus 2023, seluas 4.082,8 hektare lahan di Sumatera Selatan terbakar, hal itu terjadi dikarenakan berada di lahan mineral dan gambut.

Dalam menentukan seberapa parah tingkat polusi pada suatu daerah, pemerintah melalui Kementerian Lingkungan Hidup dan Kehutanan (KLHK) mengeluarkan Peraturan Menteri Lingkungan Hidup dan Kehutanan nomor 14 tahun 2020 tentang Indeks Standar Pencemaran Udara yang merupakan pengganti dari Keputusan Menteri Lingkungan Hidup No. 45 Tahun 1997 tentang Perhitungan dan Pelaporan serta Informasi Indeks Standar Pencemaran Udara. Pada regulasi yang baru dijelaskan untuk kalkulasi ISPU menggunakan parameter-parameter sebagai tolak ukur untuk menentukan tingkat keparahan polusi udara yaitu Partikulat, Karbon Monoksida, Nitrogen Dioksida, Ozon, Sulfur Dioksida dan Hidrokarbon. ISPU dihitung menggunakan metode dari data hasil stasiun pemantau kualitas udara sekitar lingkungan yang dibandingkan dengan alat pemantau kualitas udara yang otomatis atau disebut dengan *Air Quality Monitoring System*.

Berdasarkan Indeks Standar Pencemaran Udara, masyarakat mendapatkan informasi tentang kualitas udara yang benar dan tepat sebagai bentuk usaha untuk mengendalikan pencemaran udara. Hasil kalkulasi variable PM_{2.5} dilaporkan setiap 24 jam dan hasil untuk variable lain dilaporkan minimal dua kali dalam 24 jam.

Angka AQI yang tinggi menunjukkan bahwa lingkungan yang paling berbahaya bagi manusia dan kehidupan dalam bahaya. Oleh karena itu, pemantauan dan prakiraan AQI telah menjadi alat yang sangat penting untuk Pembangunan berkelanjutan internasional (Rybarczyk and Zalakeviciute, 2021). Dalam melakukan prediksi AQI menggunakan basis dari statistical, deterministik, fisika, machine learning dan deep learning oleh sejumlah peneliti. Penggunaan *machine learning* cukup presisi dan konsisten untuk memprediksi kualitas udara dalam segala kondisi lingkungan dikarenakan pertumbuhan jumlah historical data yang dapat diakses untuk penelitian. (Ravindiran et al., 2023).

Untuk melakukan prediksi kualitas udara, algoritma Xtreme Gradient Boost salah satu algoritma *machine learning* yang banyak digunakan oleh

peneliti untuk melakukan prediksi terutama prediksi kualitas udara. XGBoost menggunakan pendekatan *boosting* secara *gradient descent* yang secara efektif dapat menangani kasus *machine learning* dalam skala besar sehingga dapat melakukan prediksi terhadap kualitas udara.

Dalam melakukan prediksi pada objek kualitas udara, banyak dari peneliti terdahulu yang telah melakukan penelitian dengan berbagai metode algoritma *machine learning* ataupun *data mining*. Penelitian yang dilakukan oleh Ahmad Toha, Purwono dan Windu Gata pada tahun 2022 melakukan prediksi kualitas udara dengan algoritma Support Vector Machines dengan Optimasi Hyperparameter GridSearch CV. Berdasarkan hasil penelitiannya bahwa akurasi algoritma Support Vector Machines sebelum dilakukan optimasi sebesar 73,31%, dan hasil setelah dilakukan optimasi mendapatkan nilai 94,8%. Dalam hal ini menunjukkan bahwa ada peningkatan akurasi yang signifikan sebesar 21,5% setelah menerapkan metode *Grid Search CV* pada model SVM di klasifikasi kualitas udara. (Toha et al., 2022)

Pada penelitian yang dilakukan Adinda Amalia, Ati Zaidiah dan Ika Nurlaili Isnainiyah pada tahun 2022 dalam melakukan prediksi kualitas udara menggunakan algoritma *K-Nearest Neighbor* mendapatkan hasil penelitian bahwa dari nilai pengujian $K = 3$ sampai $K = 9$, didapatkan bahwa nilai $K = 7$ memiliki performa yang paling baik dengan akurasi tertinggi sebanyak 96%, presisi 92%, recall 95%, dan *f-measure* 93%.

Pada penelitian yang dilakukan oleh Ali Khumaidi, Ridwan Raafi'udin dan Indra Permana Solihin pada tahun 2020 dalam melakukan pengujian algoritma *Long Short Term Memory* untuk prediksi kualitas udara dan suhu kota Bandung. Dalam hasil penelitian tersebut menunjukkan bahwa pemodelan menghasilkan akurat prediksi yang cukup baik untuk tiga parameter (suhu, kelembapan, dan ISPU). Hal ini ditunjukkan dengan nilai RSME prediksi yang lebih kecil dari nilai standar deviasi uji dataset. Berdasarkan parameter pengujian yaitu kelembapan, suhu, ISPU dan PM 10,

hasil prediksi yang paling baik adalah prediksi kelembapan dan hasil yang prediksi yang kurang baik yaitu PM10.(Khumaidi et al., 2020)

Pada penelitian yang dilakukan oleh Agung Hot Iman, Fransisco Ready Permana, Gito Putro Wardana, Raihan Kemmy Rachmansyah dan Mayanda Mega Santoni pada tahun 2022 dalam melakukan perbandingan algoritma klasifikasi random forest dan extreme gradient boosting pada dataset cuaca provinsi DKI Jakarta Tahun 2018. Berdasarkan hasil penelitiannya menyimpulkan bahwa algoritma *Random Forest* lebih baik daripada *Extreme Gradient Boosting* jika digunakan pada dataset tersebut dikarenakan waktu pemrosesan yang lebih cepat dan akurasi yang dihasilkan cukup tinggi yaitu 68% dengan perbedaan 1-2% dengan algoritma *Extreme Gradient Boosting*.

Berdasarkan permasalahan dan penelitian yang sudah dilakukan sebelumnya, ini menjadi dasar untuk melakukan penelitian lanjutan skripsi mengenai **Prediksi Kualitas Udara di Sumatera Selatan berdasarkan Indeks Standar Pencemar Udara menggunakan algoritma extreme gradient boost.**

1.2 Rumusan Masalah

Pada uraian sebelumnya, penelitian ini merumuskan masalah yaitu menghasilkan model prediksi kualitas udara menggunakan algoritma *Extreme Gradient Boosting* berdasarkan indeks standar pencemar udara.

1.3 Batasan Masalah

Dalam permasalahan yang dihadapi di laporan skripsi bertujuan untuk tidak menambahkan variable baru selain variable yang telah ditentukan yaitu:

1. Data yang digunakan bersumber dari Kementrian Lingkungan Hidup dan Kehutanan dengan menggunakan data berlokasi di Sumatera Selatan khususnya Kota Palembang pada stasiun pemantau kualitas udara Palembang Bukit Kecil dengan rentang waktu 2020-2023

2. Pada penelitian ini akan berfokus pada menghasikan model prediksi algoritma *Extreme Gradient Boosting Classifier* dan visualisasi hasil prediksi dengan format *time series forecasting*

1.4 Tujuan dan Manfaat Penelitian

1.4.1 Tujuan Penelitian

Berikut adalah tujuan yang ingin dicapai adalah:

1. Melakukan prediksi pada kualitas udara dengan variabel ISPU di Sumatera Selatan khususnya Kota Palembang untuk lima tahun kedepan (2024-2029)
2. Mengetahui performa algoritma XGBoost dalam melakukan prediksi kualitas udara berdasarkan ISPU di Sumatera Selatan khususnya Kota Palembang

1.4.2 Manfaat Penelitian

Berikut adalah manfaat dari penelitian yang sedang diteliti bagi,

A. Akademisi

1. Memberikan informasi mengenai performa dari algoritma XGBoost untuk melakukan prediksi kualitas udara berdasarkan ISPU di Sumatera Selatan khususnya Kota Palembang
2. Dapat digunakan sebagai referensi ilmiah kepada peneliti lain yang akan menggunakan algoritma XGBoost dalam melakukan prediksi kualitas udara

B. Praktisi

1. Menjadi informasi yang berguna untuk menginformasikan perkembangan kualitas udara berdasarkan ISPU di Sumatera Selatan khususnya Kota Palembang

1.5 Sistematika Penulisan

Sistematika penulisan laporan penelitian skripsi merupakan pembahasan secara garis besar, dalam isi laporan penelitian skripsi ini yang terdiri dari lima bab diantaranya:

BAB I: PENDAHULUAN

Pada bab ini berisi tentang latar belakang, tujuan dan manfaat penelitian, Batasan masalah, waktu dan tempat pelaksanaan, metode penelitian yang digunakan, metode pengambilan data, metode analisis data, metode pengujian data dan sistematika penulisan.

BAB II: TINJAUAN PUSTAKA

Pada bab ini akan membahas landasan teori umum dan khusus yang berkaitan dengan topik penelitian guna membantu penulisan laporan penelitian skripsi.

BAB III: METODE PENELITIAN

Pada bab ini berisi penjelasan mengenai metode penelitian yang digunakan dalam melakukan penelitian ini. Bab ini meliputi model penelitian yang digunakan, tahapan dalam menjalankan penelitian serta hipotesa dalam penelitian ini.

BAB IV: HASIL DAN PEMBAHASAN

Pada bagian ini akan membahas, menjelaskan dan memberikan hasil dari proses analisis data yang telah teliti dengan teori pembahasan

BAB V: KESIMPULAN DAN SARAN

Pada bagian ini merupakan hasil kesimpulan dari penelitian yang telah dilakukan.

DAFTAR PUSTAKA

Pada bagian ini mencantumkan sumber referensi yang digunakan dalam penelitian.

LAMPIRAN

Pada lampiran akan berisikan mengenai catatan aktivitas kegiatan selama pelaksanaan penelitian.