

BAB 1

PENDAHULUAN

1.1. Latar Belakang

Kualitas udara merupakan salah satu faktor penting yang mempengaruhi kesehatan dan kesejahteraan manusia. Udara yang tercemar dapat menyebabkan berbagai penyakit pernapasan, kardiovaskular, dan kanker (Purba, 2020). Sebagai ibukota negara Indonesia, Jakarta adalah provinsi di Indonesia yang memiliki tingkat pencemaran udara tertinggi (Agus et al., 2019). Tingkat polusi ini diukur berdasarkan konsentrasi PM10 dan PM25, dan ini disebabkan oleh beberapa faktor penyebab, termasuk emisi dari kendaraan bermotor dan aktivitas industri. Sebagai pusat pemerintahan, perdagangan, dan industri, serta pusat berbagai aktivitas masyarakat dengan populasi lebih dari sepuluh juta orang dalam wilayah yang luasnya hanya 699,5 kilometer persegi, hal ini berdampak pada penurunan kualitas udara di Jakarta. Untuk mengatasi masalah pencemaran udara ini, diperlukan tindakan seperti mengurangi emisi kendaraan, meningkatkan pengawasan di sektor industri, dan memperbaiki manajemen sampah. Oleh karena itu, Pemerintah Provinsi DKI Jakarta secara rutin memantau kualitas udara di lingkungan sekitar dan telah menetapkan Indeks Standar Pencemar Udara (ISPU) sesuai dengan peraturan dalam PP 41 Tahun 1999 (Dian et al., 2023).

Indeks Standar Pencemar Udara (ISPU) adalah angka tanpa satuan yang menggambarkan kondisi kualitas udara ambien di lokasi tertentu (Subagiyo et al., 2021), yang didasarkan pada dampak terhadap kesehatan manusia, nilai estetik, dan makhluk hidup lainnya. ISPU dihitung berdasarkan konsentrasi beberapa parameter pencemar udara, seperti partikulat mater (PM10), sulfur dioksida (SO₂), nitrogen dioksida (NO₂), karbon monoksida (CO), dan ozon (O₃) (Insani and Darlianti, 2019). Jika melebihi batas aman, senyawa-senyawa tersebut dapat membahayakan kesehatan hingga menyebabkan kematian terutama pada sistem pernapasan (Fath, 2021). ISPU memiliki lima kategori, yaitu baik, sedang, tidak sehat, sangat tidak sehat, dan berbahaya (Aljuaid and Alwabel, 2019). ISPU dapat digunakan sebagai acuan untuk mengambil keputusan terkait dengan kesehatan dan lingkungan, seperti memberikan peringatan kepada masyarakat, mengatur aktivitas industri, dan melakukan mitigasi pencemaran udara. Oleh karena itu, ISPU merupakan salah satu indikator penting yang perlu dipantau secara berkala dan akurat.

Pengklasifikasian ISPU memerlukan metode yang dapat mempelajari pola data dari sensor yang mengukur konsentrasi polutan udara. Metode yang sering digunakan dalam bidang ini adalah metode pembelajaran mesin (*machine learning*), yaitu suatu cabang ilmu kecerdasan buatan yang memungkinkan komputer untuk belajar dari data tanpa perlu diprogram secara eksplisit (Rahman Sya'ban et al., 2022). Metode pembelajaran mesin dapat dibagi menjadi dua jenis utama, yaitu metode parametrik dan non-parametrik (Salamah, 2021). Metode parametrik adalah metode yang mengasumsikan adanya suatu model matematika tertentu yang menggambarkan hubungan antara variabel-variabel dalam data, sedangkan metode non-parametrik adalah metode yang tidak mengasumsikan adanya model matematika tertentu dan lebih fleksibel dalam menyesuaikan diri dengan karakteristik data.

Jenis metode pembelajaran mesin non-parametrik yang digunakan pada penelitian ini adalah *Random Forest* (RF) dan *Support Vector Machine* (SVM). SVM adalah metode yang mencari suatu *hyperplane* (bidang pemisah) optimal yang dapat memisahkan data menjadi dua kelas atau lebih dengan margin terbesar. RF adalah metode yang menggabungkan banyak pohon keputusan (*decision tree*) yang dibangun secara acak dari sebagian data dan variabel, lalu melakukan *voting* untuk menentukan kelas dari data baru. Kedua metode tersebut memiliki kelebihan dan kekurangan masing-masing, sehingga perlu dilakukan perbandingan untuk mengetahui mana yang lebih baik untuk klasifikasi ISPU.

Pada tahun 2021 (Sang et al., 2021) menerapkan algoritma *Decision Tree* dan *Support Vector Machine* untuk klasifikasi data kualitas udara DKI Jakarta. Data yang digunakan yaitu data kualitas udara DKI Jakarta tahun 2020. Pada algoritma *Decision Tree* didapatkan nilai *Precision* sebesar 99,02%, *Recall* sebesar 99,73%, *F1-Measure* sebesar 99,37%, dan *Accuracy* sebesar 99,40%. Pada algoritma SVM didapatkan nilai *Precision* sebesar 95,82%, *Recall* sebesar 88,89%, *F1-Measure* sebesar 92,22% dan *Accuracy* sebesar 94,93%.

Selanjutnya penelitian tentang klasifikasi tingkat kualitas udara DKI Jakarta berdasarkan *open government* data menggunakan algoritma *random forest* pernah dilakukan oleh (Nugroho, 2022). Data yang digunakan yaitu data kualitas udara DKI Jakarta dari bulan Februari 2021 hingga Oktober 2021. Pada penelitian tersebut didapatkan hasil akurasi tertinggi yaitu 90%. Pada kategori sedang memiliki nilai *precision* 90,9%, *recall* 89,28%, *f1-score* 90,09%, dan kategori tidak sehat memiliki nilai *precision* 89,09%, *recall* 90,74%, dan *f1-score* 89,9%.

Berdasarkan referensi penelitian-penelitian sebelumnya, menunjukkan bahwa algoritma *Random Forest* dan *Support Vector Machine* dapat melakukan klasifikasi dengan baik, khususnya pada klasifikasi ISPU. Selain itu *Support Vector Machine* maupun *Random Forest*

merupakan metode klasifikasi *supervised*, yang berarti kedua algoritma tersebut memerlukan data berlabel untuk melatih model dan melakukan prediksi.

Berdasarkan pada beberapa penelitian terdahulu tersebut, penulis ingin melakukan perbandingan algoritma *Random Forest* dan *Support Vector Machine* pada klasifikasi indeks standar pencemaran udara (ISPU). Data yang digunakan pada penelitian ini adalah data ISPU DKI Jakarta tahun 2021. Perbandingan tersebut mempertimbangkan faktor-faktor seperti nilai *accuracy*, *precision*, *recall*, *f1-score*, dan *confusion matrix*. Berikutnya pada tahap akhir akan dilakukan *deployment* berupa aplikasi komputer berbasis *web* menggunakan algoritma terbaik dari hasil perbandingan kedua model tersebut, sehingga dapat diketahui apakah model tersebut dapat mengklasifikasikan ISPU dengan baik. Penelitian mengenai perbandingan *Random Forest* dan *Support Vector Machine* algoritma tersebut dilakukan karena belum ada penelitian sebelumnya yang membandingkan secara langsung kedua algoritma tersebut pada klasifikasi ISPU. Oleh karena itu penulis mengusulkan penelitian berjudul “Klasifikasi Indeks Standar Pencemaran Udara (ISPU) di DKI Jakarta Menggunakan Algoritma *Random Forest*”

Penelitian ini diharapkan dapat memberikan wawasan yang berharga kepada masyarakat dan penulis dalam pengetahuan masalah kualitas udara serta pemantauan dan pengklasifikasian ISPU. Serta membantu dalam pengambilan keputusan yang lebih baik dalam menghadapi masalah pencemaran udara yang semakin serius di Jakarta.

1.2. Rumusan Masalah

Dari uraian latar belakang yang telah disajikan, dapat diidentifikasi beberapa permasalahan yang akan menjadi fokus pembahasan, yaitu:

- 1) Bagaimana proses dan hasil analisis perbandingan algoritma *Random Forest* (RF) dan *Support Vector Machine* (SVM) dalam mengklasifikasi indeks standar pencemaran udara (ISPU)?
- 2) Bagaimana proses penerapan model klasifikasi terbaik dari hasil perbandingan dalam *deployment* aplikasi klasifikasi indeks standar pencemaran udara (ISPU)?

1.3. Batasan Masalah

Dalam penelitian ini, penting untuk menentukan batasan-batasan agar penelitian tidak menjadi terlalu umum dan dapat difokuskan pada isu-isu yang sedang diteliti. Oleh karena itu, berikut adalah batasan-batasan yang akan diterapkan dalam penelitian ini:

- 1) Data yang digunakan pada penelitian ini adalah data sekunder (data yang tidak diambil secara langsung). Menggunakan data indeks standar pencemaran udara (ISPU) di DKI Jakarta tahun 2021.
- 2) Peneliti menggunakan data parameter pencemaran udara yaitu partikulat mater (PM10), partikulat mater (PM25), sulfur dioksida (SO₂), nitrogen dioksida (NO₂), karbon monoksida (CO), dan ozon (O₃).
- 3) Peneliti membandingkan algoritma *Random Forest* (RF) dan *Support Vector Machine* (SVM) dalam mengklasifikasi indeks standar pencemaran udara (ISPU) pada penelitian ini.

1.4. Tujuan Penelitian

Dengan merujuk pada permasalahan yang telah diuraikan sebelumnya, tujuan utama dari penelitian ini adalah sebagai berikut:

- 1) Melakukan proses dan hasil analisis perbandingan algoritma *Random Forest* (RF) dan *Support Vector Machine* (SVM) untuk mengklasifikasi indeks standar pencemaran udara (ISPU)
- 2) Melakukan proses penerapan model klasifikasi terbaik dari hasil perbandingan untuk *deployment* aplikasi klasifikasi indeks standar pencemaran udara (ISPU)

1.5. Manfaat Penelitian

Dari tujuan penelitian yang telah dijelaskan sebelumnya, berikut adalah gambaran mengenai manfaat yang dapat diperoleh dari penelitian ini:

- 1) Manfaat teoritis
 - a. Mengetahui penerapan algoritma *Random Forest* (RF) dan *Support Vector Machine* (SVM) untuk mengklasifikasi indeks standar pencemaran udara (ISPU).
 - b. Memberikan hasil analisis perbandingan performa algoritma *Random Forest* (RF) dan *Support Vector Machine* (SVM) dalam mengklasifikasi indeks standar pencemaran udara (ISPU).
- 2) Manfaat praktis
 - a. Dapat digunakan sebagai bahan evaluasi maupun referensi pada penelitian selanjutnya, khususnya penelitian yang berkaitan dengan klasifikasi indeks standar pencemaran udara (ISPU).

- b. Dapat memberikan wawasan yang berharga kepada pemerintah dan peneliti khususnya dalam pemantauan dan pengklasifikasian ISPU, serta membantu dalam pengambilan keputusan yang lebih baik dalam menghadapi masalah pencemaran udara yang semakin serius di Jakarta.