**KLASIFIKASI INDEKS PENCEMARAN KUALITAS UDARA DI DKI JAKARTA TAHUN 2021 MENGGUNAKAN ALGORITMA RANDOM FOREST**

***AIR QUALITY POLLUTION INDEX CLASSIFICATION IN DKI JAKARTA IN 2021 USING RANDOM FOREST ALGORITHM***

**Rainal Zulian Oktavianto**

Program Studi S1 Teknik Informatika, Fakultas Teknik, Universitas Pelita Bangsa, Cikarang.

Abstrak

Pencemaran udara merupakan suatu masalah yang berdampak buruk bagi kehidupan adalah aktivitas transportasi, industri, jasa, kegiatan lainnya yang meningkat, dan buangan sisa-sisa kegiatan ke udara. Proses metabolisme yang terjadi dalam tubuh makhluk hidup nitrogen oksida, ozon dan zat lain sebagainya. Menurut Air Quality Live Indeks (AQLI), pada bulan April 2021, DKI Jakarta merupakan kota ke-6 (enam) dengan kualitas udara terburuk. Hal ini ditandai dengan nilai indeks AQI Jakarta memiliki nilai 156 dengan kategori tidak sehat. Indeks Standar Pencemaran Udara (ISPU) digunakan sebagai parameter untuk mengukur kualitas udara. Analisis klasifikasi data kualitas udara di DKI Jakarta menggunakan data yang diambil dari Website Jakarta Open Data yang dilakukan oleh Dinas Lingkungan Hidup DKI Jakarta Tahun 2021 untuk memantau kualitas udara yang ada di DKI Jakarta. Berdasarkan penjelasan di atas, pada penelitian ini akan melakukan klasifikasi kualitas udara di DKI Jakarta dengan algoritma Random Forest Penelitian ini diharapkan dapat menghasilkan informasi bagaimana kualitas udara di DKI Jakarta.

Kata Kunci : Pencemaran, ISPU, Udara

1. **Pendahuluan**

Pencemaran udara merupakan suatu masalah yang berdampak buruk bagi kehidupan adalah aktivitas transportasi, industri, jasa, kegiatan lainnya yang meningkat, dan buangan sisa-sisa kegiatan ke udara. Udara yang tercemar dapat memberikan dampak negatif bagi kesehatan bahkan dapat menyebabkan kematian. Udara mengambil peran penting dalam menjaga kehidupan makhluk hidup di bumi. Proses metabolisme yang terjadi dalam tubuh makhluk hidup nitrogen oksida, ozon dan zat lain sebagainya. Beberapa zat - zat tersebut masih dapat ditoleransioleh tubuh dalam konsentrasi yang masih di bawah batas wajar, namun jika melampaui batas wajar dapat menyebabkan masalah serius pada kesehatan manusia seperti penyakit jantung, infeksi pernapasan, stroke, asma hingga penyakit paru. Salah satu faktor penyebab peningkatan konsentrasi zat - zat tersebut yang ada di dalam udara adalah aktivitas manusia.

Menurut Air Quality Live Indeks (AQLI), pada bulan April 2021, DKI Jakarta merupakan kota ke-6 (enam) dengan kualitas udara terburuk. Hal ini ditandai dengan nilai indeks AQI Jakarta memiliki nilai 156 dengan kategori tidak sehat. Polutan utama yang menyebabkan penurunan kualitas udara yaitu PM2.5, di mana polutan ini jumlahnya tidak boleh melebihi 10 mikron saat berada di udara. Di DKI Jakarta, polutan ini tercatat mencapai 57 mikron per kubik yang menandakan bahwa kualitas udara di DKI Jakarta sangat buruk.

Indeks Standar Pencemaran Udara (ISPU) digunakan sebagai parameter untuk mengukur kualitas udara. Pada pengukuran ISPU menetapkan lima parameter pencemaran udara yang digunakan untuk pengamatan, yaitu Karbon Monoksida (CO), Sulfur Dioksida (SO2), Nitrogen Dioksida (NO2), Ozon Permukaan (O3), Partikel Debu (PM10). Berdasarkan pemantauan kualitas udara yang dilakukan oleh Dinas Lingkungan Hidup Provinsi DKI Jakarta laporan dari hasil pemantauan kualitas udara yang memberikan keterangan mengenai kualitas udara yang bersih ataupun tercemar.

Analisis klasifikasi data kualitas udara di DKI Jakarta menggunakan data yang diambil dari Website Jakarta Open Data yang dilakukan oleh Dinas Lingkungan Hidup DKI Jakarta Tahun 2021 untuk memantau kualitas udara yang ada di DKI Jakarta. Dataset ini berisi mengenai Indeks Standar Pencemar Udara (ISPU) yang diukur dari 5 stasiun pemantau kualitas udara (SPKU) yang ada di Provinsi DKI Jakarta Tahun 2021 yang terletak, di Bundaran Hotel Indonesia (HI), Kelapa Gading, Jagakarsa, Lubang Buaya dan Kebon Jeruk.

Berdasarkan penjelasan di atas, pada penelitian ini akan melakukan klasifikasi kualitas udara di DKI Jakarta dengan algoritma Random Forest Penelitian ini diharapkan dapat menghasilkan informasi bagaimana kualitas udara di DKI Jakarta. Selain itu, tujuan penelitian ini adalah diharapkan hasil dari penelitian dapat menjadi acuan untuk Dinas Lingkungan Hidup dapat melakukan klasifikasi data kualitas udara berdasarkan atribut yang ada, agar pihak Dinas Lingkungan Hidup dapat melakukan penangan lebih cepat terhadap pencemaran udara yang terjadi di DKI Jakarta.

1. **Metodologi**
2. **Skema Alur Penelitian**

Penelitian dalam mengklasifikasikan level ISPU yang dilakukan untuk menentukan kualitas udara di DKI Jakarta Tahun 2021 menggunakan skema alur penelitian sebagai berikut :

1. **Objek Penelitian**

Obyek yang diteliti pada penelitian ini adalah data pencemaran udara dari Indeks Pencemar Standar Udara di DKI Jakarta Tahun 2021. Data ini diambil dari Website Jakarta Open Data yang dilakukan oleh Dinas Lingkungan Hidup DKI Jakarta untuk memantau kualitas udara yang ada di DKI Jakarta.

Dari data tersebut diteliti untuk membuat perbandingan antar algoritma klasifikasi untuk mencari algoritma terbaik dalam memprediksi kualitas udara yang baik, sedang, tidak sehat, sangat tidak sehat dan berbahaya.

1. **Studi Literatur**

Tahap awal dalam merumuskan penelitian ini dilakukan dengan cara mengumpulkan informasi yang menunjang penelitian yakni berkaitan dengan kualitas udara, ISPU dan berbagai algoritmeklasifikasi yang diperoleh dari artikel, jurnal, dan dokumen lainnya.

1. **Data Selection**

Dataset Indeks Pencemar Standar Udara DKI Jakarta dikumpulkan dari laman situs terbuka pemerintah DKI Jakarta http://www.data.jakarta.go.id/ berupa CSV yang diambil per SPKU dan per daerah yakni Jakarta Pusat (DKI1), Jakarta Utara (DKI2), Jakarta Selatan (DKI3), Jakarta Timur (DKI4), dan Jakarta Barat (DKI5) dari bulan Januari 2021 hingga bulan Desember 2021.

1. **Data Preparation**

Data yang sudah dikumpulkan dan diseleksi akan memasuki tahap data preparation. Pada tahap ini data akan di reduksi dengan menghapus data yang memiliki nilai kosong, data duplikat, data yang terisi namun tidak memiliki data secara lengkap, dan penghapusan atribut yang tidak digunakan. Kemudian karena klasifikasi masuk dalam supervised learning maka dibutuhkan tahap labeling pada atribut yang akan diklasifikasikan.

1. **Data Transformation**

Pada tahap ini dari dataset yang masih terpisah karena laporan ISPU atau frekuensi penerbitannya selama satu bulan sekali, jadi antar berkas CSV dari bulan januari 2021 hingga bulan desember 2021 digabungkan menjadi satu berkas CSV, sehingga menjadi kesatuan dataset untuk mempermudah dalam proses berikutnya.

1. **Komparasi Model**

Dari evaluasi yang dihasilkan berupa nilai akurasi dari masing – masing model algoritma akan di komparasi dengan menggunakan pengujian hipotesis parametrik T-Test apakah hasil tersebut memiliki perbedaan yang signifikan atau tidak.

1. **Hasil dan Pembahasan**
2. **Persiapan Data**

Data ISPU pada penelitian ini disusun oleh Dinas Lingkungan Hidup Provinsi DKI Jakarta. Secara keseluruhan, dataset ini memuat 450 record data dengan 10 atribut dan 1 kelas. Atribut-atribut yang terdapat dalam dataset ini antara lain tanggal, stasiun, pm10, pm25, so2, co, o3, no2, max, critical, dan categori seperti pada tabel 2 berikut :

1. **Pre-Processing**

Dataset yang digunakan didapatkan dari Website Jakarta Open Data dalam bentuk format CSV (Comma Separated Values). Dataset tersebut berbentuk file terpisahberdasarkan bulan dari bulan Januari hingga bulan Desember 2021. Dataset ini berisi mengenai Indeks Standar Pencemar Udara (ISPU) yang diukur dari 5 stasiun pemantau kualitas udara (SPKU) yang ada di Provinsi DKI Jakarta Tahun 2021. Penjelasan variabel atau parameter yang terdapat pada dataset adalah sebagai berikut :

* tanggal : Tanggal pengukuran kualitas udara
* stasiun : Lokasi pengukuran di stasiun
* pm10 : Partikulat salah satu parameter yang diukur
* pm25 : Partikulat salah satu parameter yang diukur
* so2 : Sulfida (dalam bentuk SO2) salah satu parameter yang diukur
* co : Carbon Monoksida salah satu parameter yand diukur
* o3 : Ozon salah satu parameter yang diukur
* no2 : NItrogen dioksida salah satu parameter yang diukur
* max : Nilai ukur paling tinggi dari seluruh parameter yang diukur dalam waktu yang sama
* critical : Parameter yang hasil pengukurannya paling tinggi
* categori : Kategori hasil perhitungan indeks standar pencemaran udara

1. **Proses Klasifikasi**

proses klasifikasi dengan menggunakan algoritma Decision Tree dan algoritma SVM yang dilakukan menggunakan tools RapidMiner. Data yang diolah sebanyak 1678 baris data, didapatkan setelah melakukan pengumpulan dan pembersihan data menggunakan Microsoft Excel. Selanjutnya untuk proses klasifikasi menggunakan tools RapidMiner. Proses pertama yaitu input dataset yang digunakan kemudan dihubungkan dengan set role untuk penandaan kolom sebagai label. Kemudian split data yang bertujuan untuk membagi menjadi 2 data yaitu data uji dan data latih.

Peneliti menggunakan beberapa rasio pembagian data yang dilakukan pengujian terhadap model. Rasio pembagian data tang digunakan adalah rasio 60:40, 70;30, 80:20, dan 90:10. Rasio tersebut akan diuji satu persatu terhadap model kemudian dibandingkan dengan satu sama lain dengan hasil evalusai terbaik dan dihubungkan ke algoritma yang digunakan yaitu algoritma Decision Tree dan algoritma SVM. Selanjutnya dihubungkan ke Apply Model untuk menguji data testing dan data training. Selanjutnya pada operator performance akan melakukan evaluasi kinerja dan akan memberikan nilai dari klasifikasi.

1. **Kesimpulan**

Berdasarkan analisis data dan pembahasan yang telah kita uji dengan menggunakan metode Algoritma C.45 untuk mengelompokkan atau mengklasifikasi data kualitas udara DKI Jakarta berdasarkan data yang diambil dari Website Jakarta Open Data, data tersebut menggunakan 6 parameter pengukuran kualitas udara yaitu PM10. PM25, SO2, CO, O3,NO2, dan kategori. Hasil dari penerapan algoritma C.45 data mining untuk klasifikasi kualitas udara di DKI Jakarta didapatkan bahwa kualitas udara dikategorikan menjadi tiga macam yaitu sedang, baik, dan tidak sehat dengan rasio yangberbeda-beda setiap harinya.

Setelah dilakukan klasifikasi maka dapat disimpulkan kualitas udara yang dianalisis pada bulan Januari-Desember 2021 masuk kedalam kategori sedang dimana tingkat kualitas udara masih dapat diterima untuk kesehatan manusia, hewan, dan tumbuhan. Berdasarkan kesimpulan tersebut maka diharapkan masyarakat DKI Jakarta untuk meminimalisir pencemaran udara dengan cara menggunakan transportasi publik, melakukan reboisasi, menggunakan bahan bakar ramah lingkungan, dan lain sebagainya. Dengan begitu maka angka harapan hidup dari masyarakat DKI Jakarta akan semakin meningkat.

1. **Daftar Pustaka**