Edisi XLIII, Vol 9, No 1, Bulan Maret,Tahun 2024
p-ISSN : 1978-6832
e-ISSN : 2622-5948

PENGUKURAN INDEKS STANDAR PENCEMARAN UDARA MENGGUNAKAN *SUPPORT VECTOR MACHINE*

Dimas Dzaky Daniswara, Aviolla Terza Damaliana, I Gede Susrama Mas Diyasa
UPN "Veteran" Jawa Timur Jl.Raya Rungkut Madya No. 1, Gunung Anyar, Surabaya
E-mail correspondence : igsusrama.if@upnjatim.ac.id

Abstrak

Kualitas udara merupakan salah satu faktor penting yang mempengaruhi kesehatan dan kesejahteraan manusia. Udara yang tercemar dapat menyebabkan berbagai penyakit pernapasan, kardiovaskular, dan kanker. Sebagai ibukota negara Indonesia, Jakarta adalah provinsi di Indonesia yang memiliki tingkat pencemaran udara tertinggi, hal ini berdampak pada penurunan kualitas udara di Jakarta. Indeks Standar Pencemar Udara (ISPU) adalah cara untuk memberikan informasi kepada masyarakat tentang kualitas udara. ISPU adalah angka tanpa satuan yang menunjukkan kondisi kualitas udara ambien di lokasi tertentu. ISPU dihitung dengan menghitung konsentrasi beberapa zat pencemar udara, seperti partikulat mater (PM10), sulfur dioksida (SO2), nitrogen dioksida (NO2), dan karbon monoksida (CO). Untuk mengklasifikasikan ISPU, diperlukan suatu metode yang dapat mempelajari pola dari data sensor yang mengukur konsentrasi polutan udara. Oleh karena itu, dengan menggunakan 1837 data dari indeks standar pencemaran udara (ISPU) DKI Jakarta tahun 2021, penulis ingin melakukan pengukuran indeks standar pencemaran udara (ISPU) menggunakan algoritma *Support Vector Machine*. Data yang tersedia dibagi menjadi data *train* dan data *test*, di mana 80% digunakan sebagai data *train* dan 20% sebagai data *test*. Setelah dilakukan pembagian data, dilakukan pelatihan model menggunakan data *train* dan pengujian model menggunakan data *test*. Pada uji coba dengan menggunakan algoritma *Support Vector Machine* mendapatkan akurasi model dengan presentase 94%.
Kata Kunci : Pencemaran Udara, ISPU, Klasifikasi, *Machine Learning*, *Support Vector Machine*..

Abstract

Air quality is one of the most important factors affecting human health and well-being. Polluted air can cause various respiratory, cardiovascular and cancer diseases. As the capital city of Indonesia, Jakarta is the province in Indonesia that has the highest level of air pollution, this has an impact on the decline in air quality in Jakarta. The Air Pollutant Standard Index (ISPU) is a way to provide information to the public about air quality. ISPU is a unitless number that indicates ambient air quality conditions at a particular location. ISPU is calculated by calculating the concentration of several air pollutants, such as particulate matter (PM10), sulfur dioxide (SO2), nitrogen dioxide (NO2), and carbon monoxide (CO). To classify ISPU, a method is needed that can learn patterns from sensor data that measure the concentration of air pollutants. Therefore, using 1837 data from the air pollution standard index (ISPU) of DKI Jakarta in 2021, the author wants to measure the air pollution standard index (ISPU) using the Support Vector Machine algorithm. The available data is divided into train data and test data, where 80% is used as train data and 20% as test data. After dividing the data, model training is carried out using train data and model testing using test data. In the trial using the Support Vector Machine algorithm, the model accuracy was obtained with a percentage of 94%.
Keywords*: Air Pollution, ISPU, Classification, Machine Learning, Support Vector Machine.*

PENDAHULUAN

Salah satu faktor utama yang mempengaruhi kesehatan dan kesejahteraan manusia adalah kualitas udara. Penyakit pernapasan, penyakit kardiovaskular, dan kanker adalah contoh penyakit akibat dari udara yang tercemar (Purba). Sebagai ibukota negara Indonesia, Jakarta adalah provinsi di Indonesia yang memiliki tingkat pencemaran udara tertinggi (Agus et al.). Kualitas udara di Jakarta terpengaruh oleh tingkat polusi yang diukur dari konsentrasi PM10 dan PM25. Faktor-faktor penyebabnya meliputi emisi dari kendaraan bermotor dan kegiatan industri. Dengan menjadi pusat pemerintahan, perdagangan, dan industri serta tempat berbagai aktivitas masyarakat dengan populasi lebih dari sepuluh juta orang dalam wilayah yang terbatas hanya 699,5 kilometer persegi, ini menyebabkan penurunan kualitas udara di Jakarta. 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.).

Indeks Standar Pencemar Udara (ISPU) merupakan suatu angka tanpa satuan yang mencerminkan kondisi kualitas udara di suatu lokasi (Subagiyo et al.), ISPU dihitung dengan mempertimbangkan dampaknya terhadap kesehatan manusia, nilai estetika, dan keberlangsungan makhluk hidup lainnya. Perhitungan ISPU melibatkan evaluasi konsentrasi beberapa parameter pencemar udara, seperti partikulat mater (PM10), sulfur dioksida (SO2), nitrogen dioksida (NO2), karbonmonoksida (CO), dan ozon (O3) (Insani and Darlianti). Jika melebihi batas aman, senyawa-senyawa tersebut dapat membahayakan kesehatan hingga menyebabkan kematian terutama pada sistem pernapasan (Fath). ISPU dibagi menjadi lima tingkatan, yakni kondisi baik, moderat, kurang sehat, sangat tidak sehat, dan berbahaya (Aljuaid and Alwabel). 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

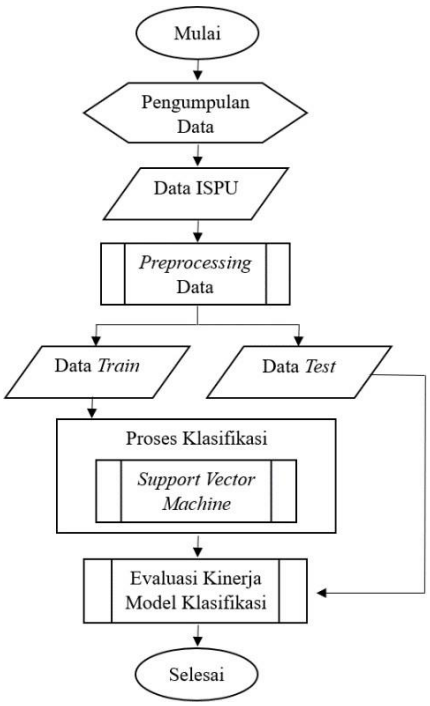
satu indikator penting yang perlu dipantau secara berkala dan akurat.

Pengklasifikasian ISPU memerlukan metode yang dapat mempelajari pola data dari sensor yang mengukur konsentrasi polutan udara. Metode yang umum digunakan dalam bidang ini adalah pendekatan pembelajaran mesin (*machine learning*). Pendekatan ini merupakan suatu cabang dari ilmu kecerdasan buatan yang memungkinkan komputer untuk memperoleh pemahaman dari data tanpa perlu diprogram secara eksplisit (Rahman Sya'ban, Hamzah, and Susanti). Metode pembelajaran mesin non-parametrik yang diterapkan dalam penelitian ini melibatkan penggunaan *Support Vector Machine* (SVM). SVM merupakan suatu pendekatan yang mencari *hyperplane* (bidang pemisah) optimal untuk memisahkan data menjadi dua kelas atau lebih, dengan tujuan mencapai margin terbesar.

Penulis ingin melakukan pengukuran indeks standar pencemaran udara (ISPU) menggunakan algoritma *Support Vector Machine*. Data yang digunakan pada penelitian ini adalah data ISPU DKI Jakarta tahun 2021. Pengukuran ISPU menggunakan algoritma *Support Vector Machine* mempertimbangkan faktor-faktor seperti nilai *accuracy, precision, recall, f1-score,* dan *confusion matrix*. Sehingga dapat diketahui apakah model tersebut dapat mengklasifikasikan ISPU dengan baik. Penelitian mengenai Pengukuran ISPU menggunakan algoritma *Support Vector Machine* dilakukan untuk mengetahui performa algoritma tersebut pada klasifikasi ISPU. Oleh karena itu penulis mengusulkan penelitian berjudul "Pengukuran Indeks Standar Pencemaran Udara Menggunakan *Support Vector Machine*".

METODE

Beberapa tahap penelitian akan dilaksanakan sesuai dengan urutan yang tergambar dalam diagram alir yang ditunjukkan dalam Gambar 1.



Gambar 1. Diagram Alir Penelitian

Pengumpulan Data

Pada langkah ini, dilakukan pengumpulan data yang akan digunakan dalam penelitian. Data yang digunakan merupakan Data Indeks Standar Pencemaran Udara (ISPU) DKI Jakarta tahun 2021 yang dapat diakses melalui situs *web* Jakarta Open Data, dengan URL https://data.jakarta.go.id/. *Dataset* ini mencakup nilai ISPU yang tercatat dari lima Stasiun Pemantauan Kualitas Udara (SPKU) di Provinsi DKI Jakarta selama tahun 2021.

Preprocessing Data

Dalam proses *preprocessing data*, terdapat beberapa langkah yang dilakukan untuk mengolah data. Pengolahan data dilaksanakan dengan tujuan agar data dapat

digunakan secara optimal dalam melakukan klasifikasi. Langkah-langkah *preprocessing* data tersebut yaitu penggabungan data, seleksi data, *data cleaning*, *label encoding* dan normalisasi data.

Split Data Train* dan *Test

Setelah proses *preprocessing data*, data dipisahkan menjadi dua bagian: data *training* dan data *testing*. Klasifikasi dilakukan dengan memanfaatkan data *training* menggunakan algoritma *Support Vector Machine*. Kemudian, pada data *testing* dimanfaatkan untuk mengevaluasi performa dan hasil akhir dari model. Dalam penelitian ini, data dibagi dengan proporsi 80% untuk data *train* dan 20% untuk data *test*.

Support Vector Machine

Dalam melakukan klasifikasi kategori ISPU dengan metode *Support Vector Machine*, prinsip dasar SVM adalah mencari fungsi pemisah (*hyperplane*) terbaik di antara banyak fungsi yang tersedia. SVM juga terlibat dalam beberapa *hyperparameter* seperti nilai c atau *gamma*. Secara keseluruhan, nilai yang dipilih untuk *hyperparameter* memiliki dampak terhadap kinerja prediksi model. Dalam penelitian ini, model klasifikasi *Support Vector Machine* dibuat menggunakan modul *library* sklearn yang disebut SVC atau *Support Vector Classifier*. Model dibuat dengan parameter *default*. Selanjutnya, model dilatih menggunakan data *train* (X_*train* dan y_*train*). Setelah pelatihan, model digunakan untuk membuat prediksi pada data *test* (X_*test*), dan hasilnya disimpan dalam variabel pred_svc. Setelah itu, untuk mengevaluasi kinerja model, dilakukan perhitungan terhadap nilai *accuracy*, *precision*, *recall*, dan *f1-score*.

Evaluasi Kinerja Model Klasifikasi

Pada tahap ini, dilakukan evaluasi kinerja model dari algoritma *Support Vector Machine* pada pengukuran ISPU. Analisis kinerja model *Support Vector Machine* dilakukan dengan melihat hasil model *Support Vector Machine* untuk *accuracy*, *precision*, *recall*, dan *f1-score*. Proses ini sesuai dengan *confusion matrix* yang

dihasilkan dari metode yang digunakan. Setelah memperoleh nilai-nilai tersebut, kesimpulan ditarik dari hasil analisis performa model *Support Vector Machine*.

HASIL DAN PEMBAHASAN

Hasil pembahasan ini akan membahas tentang kinerja algoritma *Support Vector Machine* dalam pengukuran indeks standar pencemaran udara (ISPU) melalui analisis nilai *accuracy, precision, recall,* dan *f1-score*. Nilai-nilai tersebut didapatkan dari hasil perhitungan *confusion matrix* dari metode yang diterapkan.

Support Vector Machine

	precision	recall	f1-score	support
0	0.82	0.82	0.82	34
1	0.96	0.97	0.96	242
2	0.95	0.91	0.93	43
accuracy			0.94	319
macro avg	0.91	0.90	0.91	319
weighted avg	0.94	0.94	0.94	319

Gambar 2. Hasil pengujian *Support Vector Machine*

Pada gambar 2, merupakan hasil pengujian dari model klasifikasi *Support Vector Machine* (SVM). Dari pengujian tersebut didapatkan hasil evaluasi yang baik pada tiga kelas yang berbeda. Untuk kelas 0, *precision* sebesar 82% dan *recall* sebesar 82%, dengan *F1-Score* mencapai 82%, menandakan keseimbangan yang baik antara *precision* dan *recall*. Kelas 1 menunjukkan hasil yang sangat tinggi dengan *precision* dan *recall* masing-masing sebesar 96% dan 97%, serta *F1-Score* sebesar 96%. Kelas 2 juga menunjukkan kinerja yang baik dengan *precision* sebesar 95%, *recall* sebesar 91%, dan *F1-Score* sebesar 93%. Dengan akurasi keseluruhan sebesar 94%, menunjukkan bahwa model klasifikasi *Support Vector Machine* dapat melakukan klasifikasi dengan baik.

Nilai *accuracy, precision, recall* dan *F1-score* tersebut didapatkan melalui perhitungan *confusion matrix* yang ada. Pada Tabel 1 berikut merupakan nilai dari *confusion matrix* dari model klasifikasi *Support Vector Machine*.

Tabel 1. *Confusion Matrix Support Vector Machine*

		Prediksi		
		0	1	2
Aktual	0	28	6	0
	1	6	234	2
	2	0	4	39

Evaluasi Kinerja Model Klasifikasi

Tabel 2. Kinerja Model Klasifikasi Support Vector Machine

Model	Kelas	Accuracy	Precision	Recall	F1_score
Support Vector Machine	0	94%	82%	82%	82%
	1		96%	97%	96%
	2		95%	91%	93%

Dalam evaluasi kinerja dari model algoritma *Support Vector Machine* pada pengukuran ISPU, terhadap tiga kelas yang telah ditetapkan (0, 1, dan 2), dapat ditemukan nilai performa *accuracy*, *precision*, *recall* dan *F1-score* model *Support Vector Machine*. Berdasarkan tabel 2 diatas dapat dilihat bahwa model SVM memiliki akurasi 94%, dengan perbedaan yang mencolok dalam metrik *precision*, *recall*, dan *F1-score* antara kelas-kelasnya. Kelas 0 pada SVM memiliki *precision* dan *recall* sebesar 82%, sementara Kelas 1 menunjukkan performa yang lebih tinggi dengan *precision* 96% dan *recall* 97%. Kelas 2 menunjukkan hasil yang baik dengan *precision* 95%, meskipun *recall*-nya sedikit lebih rendah pada 91%. Secara keseluruhan, SVM memberikan hasil yang baik, tetapi terdapat variasi yang cukup besar dalam kinerja antar kelas.

PENUTUP

Kesimpulan

Berdasarkan evaluasi kinerja model dari algoritma *Support Vector Machine* pada pengukuran ISPU, terhadap tiga kelas yang telah ditetapkan (0, 1, dan 2), dapat disimpulkan bahwa *Support Vector Machine* memberikan hasil yang baik dengan akurasi 94%, namun terdapat variasi yang mencolok dalam metrik *precision*, *recall*, dan *F1-score* antara kelas-kelasnya. *Support Vector Machine* menunjukkan variasi yang besar dalam kinerja antar kelas. Oleh karena itu, dalam konteks klasifikasi untuk tiga kelas yang ditetapkan, *Support Vector Machine* dapat dianggap sebagai pilihan yang baik dalam mencapai tingkat keakuratan dan konsistensi yang tinggi untuk setiap kelas. Kesimpulan ini dapat menjadi pedoman dalam memilih model untuk melakukan pengukuran indeks standar pencemaran udara yang sesuai dengan kebutuhan dan karakteristik data yang dihadapi.

DAFTAR PUSTAKA

Agus, A et al. "Analisis Dampak Diterapkannya Kebijakan Working From Home Saat Pandemi Covid-19 Terhadap Kondisi Kualitas Udara Di Jakarta." *Jurnal Meteorologi …* 6.3 (2019): 6–14. Web.

Aljuaid, Hanan, and Norah Alwabel. "Air Pollution Prediction Using Machine Learning Algorithms." *International Journal of Engineering and Advanced Technology* 8.6s3 (2019): 160–164. Web.

Dian, Febryana et al. "Klasifikasi Tingkat Pencemaran Udara Kota Jakarta Tahun 2021 Menggunakan Algoritma Decision Tree." *Seminarnasionalstatistikaaktuariaii(2023)* 2 (2023): 127–131. Web.

Fath, Muhammad Arsya. "Pengaruh Kualitas Udara Dan Kondisi Iklim Terhadap Perekonomian Masyarakat (Literature Review)." *Media Gizi Kesmas* 10.2 (2021): 329. Web.

Insani, F, and S I Darlianti. "Pembentukan Model Regresi Linier Menggunakan Algoritma Genetika Untuk Prediksi Parameter Indeks Standar Pencemar Udara (ISPU)." *Jurnal CoreIT: Jurnal Hasil Penelitian …* 5.2 (2019): 110–117. Web.

Purba, Annisa Aprilia. "Urgensi Pengetatan Baku Mutu Udara Ambien Indonesia (Studi Kasus Gugatan Pemulihan Udara DKI Jakarta)." *Padjadjaran Law Review* 8 (2020): 100. Web.

Rahman Sya'ban, Dino, Amir Hamzah, and Erma Susanti. "Klasifikasi Buah Segar Dan Busuk Menggunakan Algoritma Convolutional Neural Network Dengan Tflite Sebagai Media Penerapan Model Machine Learning." *Prosiding Snast* November (2022): F7-16. Web.

Subagiyo, Heri et al. "Rancang Bangun Sensor Node Untuk Pemantauan Kualitas Udara." *Jurnal Sains, Teknologi dan Industri* 18.1 (2021): 72. Web.