

Pemodelan Prediktif Emisi CO2 Kendaraan Kanada: Studi Komparatif *Neural Network* dan *Support Vector Machine*

Rifki Nur Hidayat Putra^{1*}, Nindya Dwi Lestari¹, Dinda Aprillia¹, Sumanto¹, Imam Budiawan¹, Roida Pakpahan¹

¹Program Studi S1 Informatika, Fakultas Teknik Informatika, Universitas Bina Sarana Informatika, Indonesia

*Email: rifkinurh.p@gmail.com

Info Artikel

Kata Kunci :

prediksi emisi CO2;
machine learning; *neural network*;
support vector machine (SVM);
evaluasi model; transportasi data

Keywords :

CO2 emission prediction;
machine learning; neural network;
support vector machine (svm);
model evaluation; data transport

Tanggal Artikel

Dikirim : 6 November 2025
Direvisi : 26 Desember 2025
Diterima : 30 Desember 2025

Abstrak

Sektor transportasi merupakan penyumbang emisi karbon dioksida (CO2) terbesar yang memperparah perubahan iklim. Penelitian ini bertujuan mengembangkan model prediktif yang akurat untuk memperkirakan emisi CO2 kendaraan dengan memanfaatkan pendekatan pembelajaran mesin. Dataset yang digunakan adalah data emisi kendaraan Kanada dari Kaggle. Metode yang diterapkan adalah *Support Vector Machine* (SVM) dan *Neural Network* untuk menganalisis pola kompleks dari berbagai parameter teknis kendaraan, seperti ukuran mesin, jumlah silinder, dan jenis transmisi. Hasil penelitian menunjukkan bahwa *Neural Network* secara konsisten unggul dibandingkan SVM dengan tingkat akurasi prediksi melebihi 90% dan nilai *F1-score* mencapai 0,831 untuk model SVM serta 0,954 untuk model *Neural Network*, yang menunjukkan kinerja klasifikasi yang kuat dan konsisten. *Neural Network* juga terbukti lebih baik dalam menangkap hubungan non-linier antara karakteristik kendaraan dan emisi CO2. Keberhasilan model ini membuka peluang pengembangan model prediktif yang lebih canggih serta dapat menjadi dasar bagi pembuat kebijakan dalam merancang regulasi emisi yang lebih akurat dan berbasis data.

Abstract

The transportation sector is the largest contributor to carbon dioxide (CO2) emissions that exacerbate climate change. This research aims to develop an accurate predictive model to estimate vehicle CO2 emissions by utilizing a machine learning approach. The dataset used is Canadian vehicle emissions data from Kaggle. The methods applied are Support Vector Machine (SVM) and Neural Network to analyze complex patterns of various vehicle technical parameters, such as engine size, number of cylinders, and transmission type. The results showed that the Neural Network consistently excelled over SVM with a prediction accuracy rate exceeding 90% and an F1-score value of 0.831 for the SVM model and 0.954 for the Neural Network model, indicating a strong and consistent classification performance. Neural networks have also been shown to be better at capturing the non-linear relationship between vehicle characteristics and CO2 emissions. The success of this model opens up opportunities for the development of more sophisticated predictive models and can serve as a basis for policymakers to design more accurate and data-driven emissions regulations.

1. PENDAHULUAN

Meningkatnya jumlah kendaraan bermotor berkontribusi pada peningkatan emisi CO₂ dan polusi udara, sebagai salah satu polutan yang dihasilkan oleh kendaraan bermotor [1]. Situasi ini memberikan kesempatan unik untuk menilai dampak nyata dari pengurangan lalu lintas kendaraan terhadap emisi CO₂, terutama di pusat kota yang padat contohnya seperti di Jakarta Pusat [2]. Pusat-pusat perkotaan memiliki masalah ini dengan jelas. Alun-alun kota Batu, yang merupakan landmark dan persimpangan jalan, misalnya, mengalami peningkatan signifikan dalam volume lalu lintas. Akibatnya, konsumsi bahan bakar di wilayah ini meningkat, yang berpotensi menyebabkan peningkatan emisi gas buang kendaraan [3]. Selain itu, gedung-gedung seperti kantor pemerintah pusat, juga terkena dampak masalah ini. Karena banyaknya lalu lintas yang melewati wilayah ini, diperlukan evaluasi menyeluruh terhadap emisi CO₂ yang dihasilkan dan potensi pengurangan emisi melalui vegetasi lokal [4].

Jumlah lalu lintas yang meningkat di berbagai lokasi kegiatan ini secara langsung meningkatkan ancaman peningkatan emisi karbon dioksida di Bumi [5]. Salah satunya yaitu Kendaraan bermotor yang sering digunakan dapat memengaruhi lingkungan secara signifikan. *Traffic jam*, suara, dan gas buang yang dihasilkan oleh kendaraan bermotor yang tidak terkendali [6] Menurut Indonesia Energy Outlook 2017, sektor transportasi mengkonsumsi lebih banyak energi daripada sektor rumah tangga, mencapai 31% dari kebutuhan energi nasional, meningkat sebesar 5,2% per tahun dari tahun 2010 hingga 2015 [7]. Sebagai tanggapan terhadap dampak lingkungan ini, pengembangan jalur hijau dan ruang terbuka hijau (RTH) adalah beberapa contoh konsep infrastruktur hijau yang mulai dimasukkan ke dalam perencanaan kota modern.

RTH ini tidak hanya memberikan tampilan yang indah tetapi juga berfungsi menyerap emisi CO₂ yang dihasilkan oleh kendaraan bermotor [8]. Beberapa studi dalam lima tahun terakhir telah menciptakan model yang dapat memprediksi emisi kendaraan dengan menggunakan *machine learning*. Metode Random Forest dan *Neural Network* telah diterapkan untuk memperkirakan emisi CO₂ dari kendaraan ringan dan dapat mencapai akurasi hingga 94% [9]. Selain itu, model *regresi linear* berganda juga digunakan untuk menilai emisi CO₂ dari kendaraan di Indonesia, meskipun masih terbatas dalam menangkap kompleksitas pola data. Hasil-hasil ini menunjukkan bahwa metode *machine learning* yang lebih fleksibel, seperti *Support Vector Machine* dan *Neural Network*, berpotensi meningkatkan akurasi dalam memprediksi emisi CO₂ [10].

Penelitian ini mengungkapkan tingginya emisi CO dan CO₂ pada persimpangan dengan sinyal menggunakan metode penghitungan standar lingkungan. Untuk pengembangan lebih lanjut, beberapa penelitian lain menerapkan teknik pembelajaran mesin seperti Random Forest dan Jaringan Saraf untuk membangun model dan memprediksi emisi kendaraan, yang dapat mencapai tingkat akurasi prediksi yang tinggi sesuai dengan dataset yang digunakan [11]. Selama kemacetan lalu lintas, pembakaran bahan bakar kendaraan bermotor menjadi tidak efisien dan tidak sempurna, melepaskan senyawa seperti karbon monoksida, nitrogen oksida, sulfur oksida, partikel padat, dan senyawa fosfor-timbal. Akibatnya, senyawa yang dikeluarkan oleh kendaraan bermotor menumpuk, meningkatkan polusi udara [12].

Di sisi lain, karakteristik teknis kendaraan juga memegang peran penting. Angka oktan, yang berkorelasi dengan rasio kompresi mesin sepeda motor, memengaruhi emisi CO₂ dari bahan bakar. Rasio kompresi yang lebih tinggi terkait dengan jumlah oktan yang diperlukan [13]. Kerumitan ini, yang melibatkan berbagai variabel teknis dan kondisi lalu lintas yang dinamis, menunjukkan bahwa metode evaluasi tradisional tidak lagi cukup. Oleh karena itu, diperlukan pendekatan pemodelan prediktif yang lebih canggih, seperti yang ditawarkan oleh *Machine Learning*, untuk mengungkap determinan emisi secara lebih akurat. Berdasarkan hal tersebut, tujuan penelitian ini adalah untuk mengembangkan dan mengevaluasi model prediksi emisi CO₂ berbasis *machine learning* menggunakan data numerik, serta membandingkan kinerja beberapa algoritma guna memperoleh model dengan tingkat akurasi terbaik sebagai dasar pendukung perencanaan lingkungan perkotaan.

Penelitian terkait prediksi emisi karbon dioksida telah mengalami evolusi signifikan, seiring dengan meningkatnya kesadaran global akan perubahan iklim. Studi awal sering kali berfokus pada model dispersi polutan dan pendekatan statistik tradisional untuk memahami penyebaran emisi di lingkungan perkotaan [14] yang mengkaji berbagai model mekanistik dan matematis untuk dispersi polutan kendaraan. Meskipun model-model ini fundamental dalam memahami mekanisme penyebaran, mereka sering kali memiliki keterbatasan dalam mengakomodasi kompleksitas dan non-linearitas data emisi dari berbagai sumber, terutama pada level kendaraan individual. Akibatnya, para peneliti beralih ke pendekatan pembelajaran mesin (*machine learning*) dan pembelajaran mendalam (*deep learning*) yang menawarkan kemampuan superior dalam menangani pola-pola kompleks dan hubungan non-linear dalam data [15].

2. TINJAUAN LITERATUR

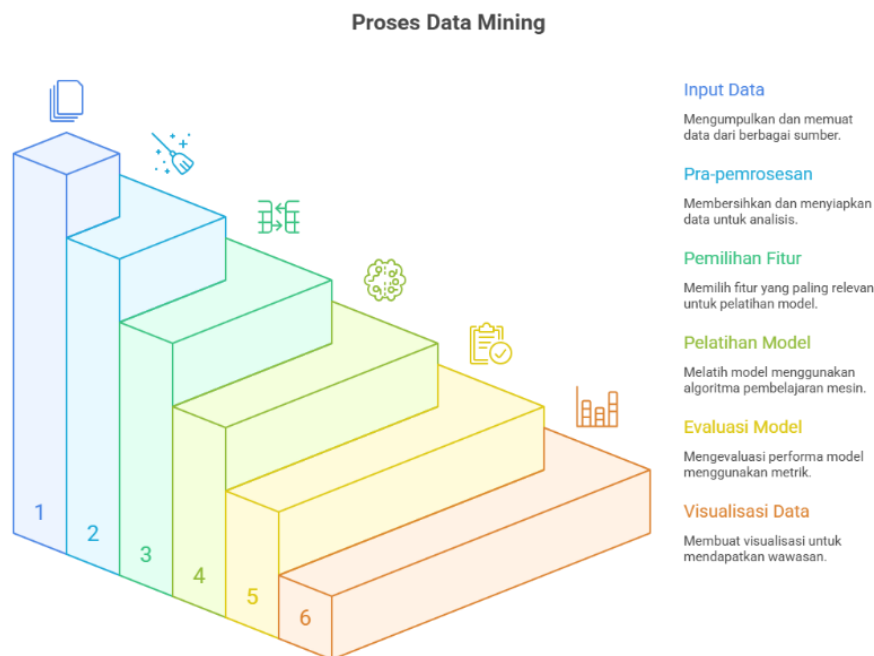
Dalam studinya yang komprehensif membandingkan 14 model berbeda, menemukan bahwa model pembelajaran mesin dan pembelajaran mendalam secara konsisten mengungguli model statistik tradisional seperti ARIMA dalam prediksi emisi CO₂ harian, dengan nilai R^2 yang jauh lebih tinggi. Temuan ini diperkuat oleh penelitian di sektor lain, seperti pertanian, di mana jaringan saraf (*neural network*) berhasil diterapkan untuk memprediksi emisi CO₂ dari peternakan babi dengan

mempertimbangkan variabel-variabel lingkungan yang kompleks [16], menunjukkan fleksibilitas dan kekuatan model ini dalam berbagai konteks.

Di sektor transportasi, kompleksitas emisi kendaraan dipengaruhi oleh multitud faktor, mulai dari parameter teknis hingga kondisi operasional [17] menyoroti pentingnya pengukuran emisi aktual dalam kondisi nyata (*Real Driving Emissions*) dan menemukan bahwa usia kendaraan dan jenis bahan bakar merupakan faktor penentu yang signifikan, yang menunjukkan bahwa model prediktif harus mampu menangani variabilitas ini. Sementara itu, [18] telah mengeksplorasi penggunaan berbagai model pembelajaran mesin, termasuk *Support Vector Machine* (SVM) dan *Convolutional Neural Network* (CNN), untuk memprediksi emisi CO₂ berdasarkan data makro seperti penggunaan energi dan PDB, yang sekali lagi mengonfirmasi potensi besar pendekatan ini.

Meskipun telah banyak bukti keunggulan model pembelajaran mesin, masih terdapat kekosongan dalam penelitian yang secara spesifik dan mendalam membandingkan kinerja SVM dan *Neural Network* khususnya dalam konteks memprediksi emisi CO₂ kendaraan secara eksklusif berdasarkan parameter teknisnya seperti ukuran mesin, jumlah silinder, dan jenis transmisi. Banyak studi yang ada cenderung menggunakan *data agregat* atau kondisi mengemudi dinamis, namun sedikit yang berfokus pada analisis komparatif kedua model ini untuk mengungkap kapabilitasnya dalam menangkap hubungan non-linear intrinsik antara spesifikasi statis kendaraan dan output emisinya.

3. METODE PENELITIAN



Gambar 1. Proses Penelitian

Proses penelitian yang diterapkan dalam studi ini mengikuti langkah-langkah umum dari data mining, yang meliputi pengumpulan pra-pemrosesan data, pemodelan, evaluasi, dan visualisasi hasil [18]. Pada langkah pemodelan, algoritma *Support Vector Machine* (SVM) digunakan karena kemampuannya dalam mengatasi hubungan *nonlinier* dalam data *numerik* serta menciptakan solusi terbaik dengan membentuk *hyperplane* untuk masalah prediksi dan regresi [19]. Selain itu, metode *Neural Network* dimanfaatkan untuk memodelkan hubungan kompleks antara variabel *input* dan *output*, yang terbukti efektif dalam model regresi terawasi (*supervised regression*) pada berbagai penelitian yang berbasis *machine learning* [20].

3.1 Perangkat Lunak yang Digunakan

Penelitian ini menggunakan Orange Data Mining versi 3.36.0 untuk membangun workflow pembelajaran mesin. Orange dipilih karena menyediakan antarmuka *visual* yang memudahkan dalam pra-pemrosesan data, pelatihan model, dan evaluasi tanpa pemrograman.

3.2 Pembelajaran Mesin dan Regresi

Permasalahan estimasi emisi karbon dioksida (CO_2) yang bersifat kontinu dan dipengaruhi oleh berbagai variabel teknis serta lingkungan, menjadikan pendekatan pembelajaran mesin (machine learning) sebagai pilihan yang tepat. Sebagai cabang kecerdasan buatan, pembelajaran mesin memungkinkan penemuan pola kompleks dari data tanpa bergantung pada pemrograman aturan yang eksplisit. Dalam penelitian ini diterapkan metode supervised learning dengan teknik regresi. Regresi dipilih karena secara inheren dirancang untuk memodelkan hubungan antara satu atau lebih variabel masukan, dalam hal ini parameter teknis kendaraan dan kondisi lalu lintas dengan sebuah variabel keluaran target, yaitu nilai emisi CO_2 . Melalui pelatihan pada data berlabel, model regresi ini diharapkan mampu memberikan perkiraan yang akurat mengenai besaran emisi, memberikan wawasan kuantitatif yang berharga.

Pada penelitian ini, algoritma *Support Vector Machine* diterapkan dalam bentuk regresi, yang dikenal sebagai *Support Vector Regression* (SVR). Model SVR dimanfaatkan untuk menggambarkan hubungan nonlinier antara variabel input dan angka emisi karbon dioksida (CO_2). Jenis fungsi *kernel* yang digunakan adalah *Radial Basis Function* (RBF) karena kemampuannya dalam mengolah data *numerik* yang memiliki pola rumit. Parameter utama yang digunakan dalam model ini mencakup parameter regularisasi (C), parameter kernel (γ), serta parameter epsilon (ϵ) yang berfungsi untuk mengatur toleransi kesalahan dalam proses regresi.

3.3 Support Vector Machine (SVM)

Untuk tugas regresi dan klasifikasi, *Support Vector Machine* (SVM) adalah algoritma pembelajaran terawasi. Tujuan utama adalah menemukan *hyperplane* yang ideal yang dapat membagi data ke dalam berbagai kelas. *Hyperplane* yang paling efektif adalah yang memiliki margin terbesar, atau jarak terjauh antara *hyperplane* dan titik data terdekat dari setiap kelas. Vektor pendukung merujuk pada titik data ini.

Keunggulan SVM terletak pada kemampuannya untuk mengatasi data yang tidak dapat dipisahkan secara linier dengan menggunakan metode kernel. Fungsi kernel memungkinkan data input dipetakan ke dalam ruang berdimensi lebih tinggi, di mana data dapat dipisahkan secara linier. Dengan kemampuan ini, SVM efektif dalam menangani hubungan kompleks antar fitur, seperti yang mungkin terjadi pada data performa mesin.

3.4 Neural Network dan Deep Learning

Mekanisme kerja jaringan saraf otak manusia, di mana neuron-neuron terhubung satu sama lain dan tersusun dalam berbagai lapisan untuk memudahkan pemrosesan data, adalah inspirasi dari model komputasi yang dikenal sebagai *neural network*. Struktur dasar dari model ini terdiri dari lapisan input, yang menerima data fitur awal untuk digunakan sebagai dasar untuk analisis berikutnya. Selanjutnya, ada lapisan tersembunyi yang melakukan pemrosesan mendalam dan mengekstrak fitur non-linear dari data tersebut. Jaringan yang terdiri dari banyak lapisan tersembunyi ini disebut *Neural Network*. Dalam, dan ini merupakan fokus utama dalam pembelajaran rumit. Terakhir, lapisan output menghasilkan prediksi akhir, termasuk probabilitas untuk setiap kategori dampak yang relevan. Ini memungkinkan model untuk memberikan gambaran yang akurat tentang pola yang dipelajari.

Setiap koneksi antar neuron memiliki bobot yang mempengaruhi sinyal. Selama proses pelatihan, jaringan menyesuaikan bobot ini dengan algoritma seperti *backpropagation*, untuk mengurangi kesalahan antara prediksi dan label yang sebenarnya. Kemampuan *Neural Network* dalam memahami representasi data yang kompleks dan abstrak menjadikannya sangat sesuai untuk tugas klasifikasi seperti menemukan pola halus dalam data performa mesin yang menunjukkan perubahan dari kondisi Aman, Berisiko, atau Merusak.

3.5 Evaluasi Performa Model

Untuk menilai seberapa baik suatu model klasifikasi bekerja, kita memerlukan berbagai parameter evaluasi yang memberikan perspektif berbeda mengenai performa model tersebut. Parameter-parameter ini sangat krusial dalam memilih model yang paling sesuai, karena masing-masing menekankan aspek tertentu dari hasil prediksi yang dihasilkan. Salah satu parameter evaluasi yang mendasar adalah *confusion matrix*, yaitu matriks yang memperlihatkan secara rinci jumlah prediksi yang tepat dan tidak tepat untuk setiap kategori, yang menjadi dasar perhitungan parameter lainnya. Melalui *confusion matrix*, kita dapat menghitung akurasi, yang merupakan perbandingan sederhana antara jumlah prediksi yang benar dan total prediksi.

keseluruhan. Meskipun akurasi memberikan informasi umum tentang kinerja model, metrik ini bisa menyesatkan jika data yang digunakan tidak seimbang, di mana satu kategori lebih mendominasi.

Evaluasi kinerja model mengandalkan metrik presisi dan recall. Presisi mengukur proporsi prediksi positif yang benar, menjawab seberapa andal model saat memberikan prediksi. Sementara itu, recall menilai kemampuan model dalam mengidentifikasi seluruh kasus positif yang sebenarnya ada. Untuk menjamin estimasi kinerja yang tidak biasa dan *robust*, penelitian ini menerapkan teknik validasi silang, di mana dataset dibagi menjadi beberapa bagian yang bergantian berfungsi sebagai data uji dan data latih, sehingga menghasilkan perkiraan yang lebih akurat untuk data baru.

4. HASIL DAN PEMBAHASAN

4.1 Input Data

Proyek ini berfokus pada pembuatan model prediksi emisi CO₂ kendaraan dengan menggunakan dataset Kaggle [21] yang berisi 6.951 catatan spesifikasi kendaraan untuk mendukung upaya mengurangi dampak perubahan iklim. Alur kerja dimulai dengan memuat data awal melalui *node File*. Kemudian, *node DataTable* langsung memeriksa struktur data untuk memastikan integritas dan volumenya. Untuk mendapatkan pemahaman yang lebih baik tentang karakteristiknya, data diperiksa secara visual dengan modul *Distribusi* dan *Plot Scatter*. Ini menunjukkan pola penting seperti korelasi antara ukuran mesin dan tingkat emisi. Data yang telah siap digunakan untuk melatih dua model pembelajaran mesin, *Neural Network* dan SVM, setelah tahap prapemrosesan di *node Preprocess* dan pemilihan fitur di Kolom Pilihan.

4.2 Pra-Pemrosesan

Mengubah data mentah menjadi bentuk yang siap digunakan oleh model prediksi adalah tujuan dari tahap pra-pemrosesan ini, yang merupakan tahap penting dalam alur kerja pembelajaran mesin yang dirancang. Proses ini dimulai dengan *node File* memuat dataset mentah, yang kemudian diperiksa untuk keamanan dan struktur di *node DataTable*. Selanjutnya, data masuk ke inti pra-pemrosesan di *node Preprocess*, di mana langkah-langkah penting seperti penanganan nilai yang hilang, normalisasi, dan konversi variabel kategorikal dilakukan untuk membersihkan dan menstandarkan data.

Setelah dibersihkan, *node Select Columns* bertugas menyaring fitur-fitur yang paling penting. Ini memastikan bahwa informasi yang paling penting akan digunakan dalam pelatihan. Hasil dari rangkaian ini adalah dataset yang telah dimurnikan yang siap dikirim secara bersamaan untuk melatih dua model prediksi, *Neural Network* dan SVM. Proses ini mencapai puncaknya pada tahap evaluasi di *node Tes dan Skor*, di mana kinerja kedua model diukur dan dibandingkan untuk menentukan metode yang paling efisien.

4.3 Pemilihan Fitur (Feature Selection)

Dataset sering kali memiliki banyak kolom atau variabel (fitur), tetapi tidak semua fitur tersebut memiliki relevansi atau memberikan kontribusi yang berarti dalam hasil prediksi. Proses pemilihan fitur adalah langkah untuk mengidentifikasi serta memilih kelompok fitur yang paling berpengaruh terhadap variabel target. Dengan menghilangkan fitur yang tidak relevan atau berlebihan, model menjadi lebih sederhana, lebih efisien selama pelatihan, dan dapat mengurangi kemungkinan terjadinya *overfitting*.

Dalam penelitian ini, pemilihan fitur dilakukan dengan menggunakan pendekatan pemilihan fitur berbasis filter yang menilai hubungan antara setiap fitur input dengan variabel target emisi CO₂. Fitur yang menunjukkan relevansi rendah atau memiliki korelasi lemah terhadap variabel target akan dihapus, sementara fitur yang memiliki pengaruh signifikan akan dipertahankan untuk proses pemodelan. Metode ini dipilih karena merupakan pendekatan yang sederhana, efisien dalam hal komputasi, dan tidak tergantung pada algoritma prediksi tertentu.

4.4 Pelatihan Model

Dalam analisis berbasis pembelajaran mesin, ini adalah tahap penting di mana algoritma diterapkan pada data yang telah diproses. Pada awalnya, model *K-Nearest Neighbors* (KNN) dipertimbangkan sebagai alternatif. Namun, saat melakukan prediksi pada dataset berukuran 6951 baris, KNN dianggap tidak cocok karena sifatnya yang *malas belajar* dan kemungkinan beban komputasi yang tinggi. Oleh karena itu, perhatian difokuskan pada dua model yang lebih kuat *Support Vector Machines* (SVM) dan *Neural Networks*.

Metode SVM ini dipilih karena kemampuan mereka untuk menangani data yang sangat besar dan menemukan batas keputusan yang optimal, sementara *Neural Network* dipilih karena kemampuan untuk memodelkan hubungan yang sangat kompleks dan *non-linear* yang mungkin ada antara fitur kendaraan dan emisi CO2. Pada langkah ini, tujuan adalah untuk melatih kedua model sehingga mereka dapat mempelajari pola, hubungan, dan tren penting dari data. Pada akhirnya, ini akan menghasilkan model terlatih yang dapat membuat prediksi emisi CO2 yang akurat untuk data baru.

4.5 Evaluasi Model

Hasil evaluasi dari kedua model disajikan dalam tabel dan kurva *ROC* berikut.

Tabel 1. Hasil Evaluasi Kinerja Model SVM dan Neural Network

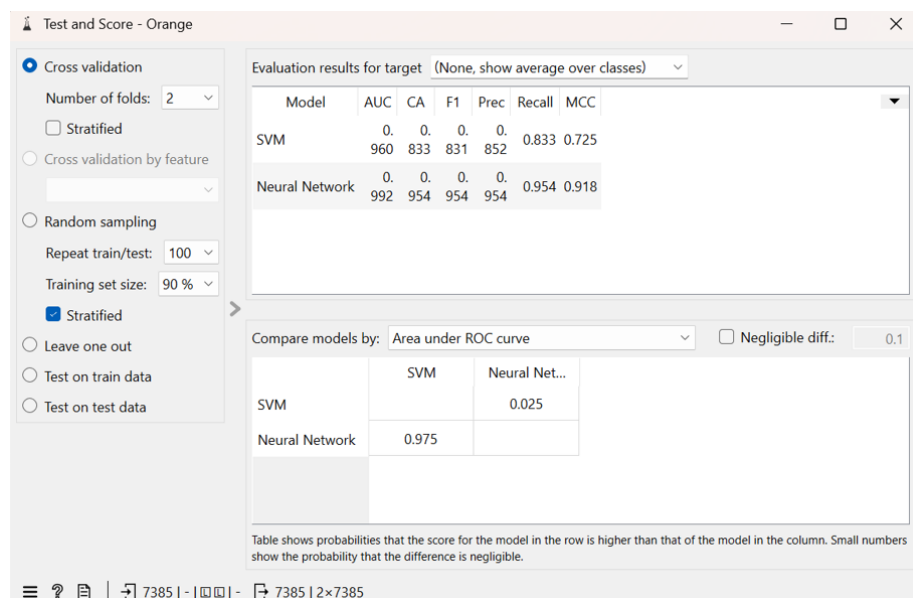
Model	AUC	CA(AKURASI)	F1-SCORE	PRECISION	RECALL
Support Vector Machine (SVM)	0.960	0.833	0.831	0.852	0.833
Neural Network	0.992	0.954	0.954	0.954	0.954

Berdasarkan Tabel 1, dapat dilihat bahwa model *Neural Network* secara mencolok mengungguli model *Support Vector Machine* (SVM) di hampir semua ukuran evaluasi. *Neural Network* membuktikan diri sebagai model unggulan dalam penelitian ini, khususnya akurasi, F1-score, precision, dan recall dengan meraih akurasi mencapai 95,4%, sebuah angka yang melampaui target dan jauh menonjol dibandingkan dengan SVM yang hanya tercatat sebesar 83,3%.

Namun, jika dilihat dari AUC, SVM justru memperoleh nilai yang lebih tinggi, yaitu 0,960, jika dibandingkan dengan *Neural Network* yang hanya mencapai 0,992. Ini menunjukkan bahwa SVM memiliki kemampuan yang sangat baik dalam memisahkan kelas. Walaupun begitu, keunggulan *Neural Network* dalam metrik evaluasi lainnya mengindikasikan bahwa model ini lebih efektif dalam memberikan prediksi yang tepat dan seimbang pada data secara keseluruhan, sehingga model ini bisa dianggap sebagai yang terbaik untuk diterapkan dalam penelitian ini.

4.6 Visualisasi Data

Akhir dari analisis ini adalah mengubah hasil yang kompleks dari model terlatih menjadi cerita yang mudah dilihat. Kami memberikan gambaran jelas tentang karakteristik umum populasi kendaraan yang diteliti dengan menggambarkan sebaran data emisi CO2 dan fitur penting seperti ukuran mesin melalui visualisasi distribusi. Sementara itu, hubungan kausal yang kuat, seperti korelasi positif antara ukuran mesin dan volume emisi CO2, dapat diungkapkan dengan menggunakan Scatter Plot. Pola ini menunjukkan pola yang telah dipelajari oleh model *SVM* dan *Neural Network*. Pemangku kepentingan teknis dan non-teknis dapat dengan cepat memahami temuan penting dan membuat keputusan yang berbasis bukti berkat cerita data yang kuat yang disajikan oleh kombinasi visualisasi kedua ini.



Gambar 2. Visualisasi Perbandingan Kinerja Model Klasifikasi Menggunakan Test and Score

Gambar 2 menampilkan hasil dari penilaian kinerja model *Support Vector Machine* (SVM) dan *Neural Network* melalui fitur *Test and Score* dalam Orange Data Mining dengan menggunakan metode *cross-validation* dua lipatan. Penilaian ini dilakukan untuk mengevaluasi sejauh mana kedua model dapat mengklasifikasikan data dengan tepat dan konsisten.

Dari hasil pengujian, terlihat bahwa *Neural Network* menunjukkan kinerja yang lebih *superior* dibandingkan SVM di hampir semua metrik yang digunakan untuk evaluasi. *Neural Network* mencatat nilai AUC sebesar 0,992, sementara akurasi, *F1-score*, *precision*, dan *recall* masing-masing adalah 0,954, dan nilai MCC sebanyak 0,918, yang menunjukkan kemampuan klasifikasi yang sangat baik dan stabil. Di sisi lain, SVM berhasil mencapai nilai AUC sebesar 0,960, dengan akurasi 0,833, *F1-score* 0,831, *recall* 0,803, dan MCC 0,725, sehingga kinerjanya masih di bawah *Neural Network*.

Perbandingan yang dilakukan berdasarkan AUC juga menunjukkan adanya *probabilitas* sebesar 0,975 bahwa *Neural Network* memiliki performa yang lebih baik dibandingkan SVM. Secara keseluruhan, temuan ini menggarisbawahi bahwa *Neural Network* lebih akurat dan dapat diandalkan untuk digunakan pada dataset dalam penelitian ini.

4.7 Pembahasan

Hasil evaluasi ini memberikan pemahaman mendalam yang menampilkan bahwa model *Neural Network* merupakan pilihan yang jauh lebih unggul untuk studi ini. Selain mampu mencapai target akurasi di atas 90%, *Neural Network* juga menunjukkan tingkat konsistensi tinggi di semua metrik lainnya, dengan nilai AUC (0.992) dan Akurasi (0.918) yang sebanding, yang menunjukkan model yang kuat dan terpercaya. Di sisi lain, temuan yang paling menarik justru berasal dari model SVM, yang menunjukkan perbedaan signifikan antara nilai AUC yang sangat tinggi (0.960) dan akurasi yang cukup rendah (0.725).

Fenomena ini menunjukkan bahwa meskipun SVM secara alami dapat memahami struktur data dan membedakan kelas dengan baik (AUC tinggi), ia kurang berhasil dalam aplikasi praktis ketika mengubah skor probabilitas menjadi prediksi kelas akhir, mungkin disebabkan oleh ketidakseimbangan kelas, ambang batas keputusan yang tidak ideal, atau kelemahan dari penggunaan 2-fold yang tidak terstratifikasi. Sebagai kesimpulan, penelitian ini berhasil menunjukkan bahwa arsitektur *Neural Network* sangat efisien dan sudah siap untuk diterapkan, sedangkan untuk penelitian berikutnya sangat dianjurkan untuk menggunakan metode evaluasi yang lebih kuat seperti *10-fold cross-validation* dengan stratifikasi dan melakukan optimasi ambang batas pada SVM untuk mendapatkan estimasi performa yang lebih tepat dan terpercaya.

Melalui Orange, *workflow* dikembangkan untuk melaksanakan setiap fase penelitian. Data dimasukkan, di pra-proses, dan model dilatih dengan menggunakan widget yang ada. Proses prediksi dilakukan dengan memasukkan data baru ke dalam model yang sudah dilatih melalui widget Prediksi, yang kemudian memberikan estimasi emisi CO₂.

5. KESIMPULAN

Penelitian ini berhasil mencapai tujuan utamanya, yaitu menciptakan model klasifikasi untuk memprediksi emisi karbon dioksida (CO₂) dari kendaraan berdasarkan *parameter teknis* dengan menggunakan pendekatan *machine learning*. Dua model, yaitu *Neural Network* dan *Support Vector Machine* (SVM), telah berhasil dikembangkan dan dievaluasi dengan teliti. Hasil dari pengujian menunjukkan bahwa *Neural Network* memiliki performa yang lebih baik dibandingkan dengan SVM. *Neural Network* mencapai akurasi sebesar 95,4%, dengan *F1-score*, *presisi*, dan *recall* masing-masing sebesar 0,954, sedangkan model SVM mendapatkan akurasi 83,3%, dengan *F1-score* 0,831, *presisi* 0,852, dan *recall* 0,833.

Perbedaan dalam kinerja ini menunjukkan bahwa *Neural Network* lebih handal dalam memahami hubungan *non-linear* antara karakteristik kendaraan dan tingkat emisi CO₂, sehingga mampu memberikan prediksi yang lebih tepat dan konsisten. Berdasarkan hasil ini, *Neural Network* disarankan sebagai model yang paling efektif untuk meramalkan emisi kendaraan. Penemuan dari penelitian ini diharapkan dapat menjadi dasar bagi pengembangan model prediksi emisi yang lebih canggih serta mendukung pembuatan kebijakan pengendalian emisi kendaraan yang berbasis data.

DAFTAR PUSTAKA

- [1] Sudarti, Yushardi, and N. Kasanah, "Analisis Potensi Emisi CO₂ Oleh Berbagai Jenis Kendaraan Bermotor di Jalan Raya Kemantren Kabupaten Sidoarjo," vol. 9, 2022, doi: <http://dx.doi.org/10.21776/ub.jsal.2022.009.02.4>.
- [2] F. Permatasari, S. Hadisusanto, and E. Haryono, "Emisi CO₂ Kendaraan Bermotor Periode Kebijakan Pembatasan Sosial Berskala Besar (Studi Kasus: Ruas Jalan di Jakarta Pusat)," vol. 15, 2021, doi: 10.20886/jklh.2021.15.1.31-44.

- [3] N. Ambarsari, K. Sari, and C. Maulidi, "Emisi CO₂ Kendaraan Bermotor Kawasan Alun-Alun Kota Batu," vol. 8, 2019, Accessed: Nov. 08, 2025. [Online]. Available: <https://purejournal.ub.ac.id/index.php/pure/article/view/385>
- [4] P. Sarasidehe, D. Jati, and Jumiaty, "Analisis Kemampuan Vegetasi pada Ruang Terbuka Hijau dalam Menyerap Emisi CO₂ Kendaraan Bermotor di Area Kantor Gubernur Kalimantan Barat," vol. 6, 2022, doi: <https://doi.org/10.26760/jrh.v6i3.219-228>.
- [5] C. Handayani, A. Sukmono, and H. Firdaus, "Analisis Ketersediaan Ruang Terbuka Hijau Terhadap Emisi CO₂ Oleh Gas Buang Kendaraan Bermotor Di Kelurahan Tembalang Dan Sumurboto," vol. 9, 2020, doi: <https://doi.org/10.14710/pilars.%25v.%25i.%25Y.%25p>.
- [6] Z. Raffinet, "Analisis Emisi Karbondioksida Ditinjau Dari Penggunaan Kendaraan Berbasis Aplikasi Di Kecamatan Sukolilo Surabaya." Accessed: Nov. 08, 2025. [Online]. Available: <https://repository.upnjatim.ac.id/5841/>
- [7] A. Rismana, R. Budiarto, and A. Harto, "Analisis Energi dan Emisi CO₂ Rencana Bus Listrik di YOGYAKARTA Studi Kasus Trans Jogja," vol. 11, 2019, doi: 10.5614/joki.2019.11.1.1.
- [8] F. Mutia, A. Soegianto, and Nurina Fitriani, "Studi Penyerapan Emisi Karbondioksida (Co₂) Kendaraan Bermotor Di Jalur Hijau Jalan Dr. Soetomo Kota Surabaya." Accessed: Nov. 08, 2025. [Online]. Available: <https://repository.unair.ac.id/102847/>
- [9] K. Asgaryansyah and P. Paniran, "Implementasi Algoritma Regresi Linier untuk Mengukur Tingkat Pengeluaran Co₂ Pada Kendaraan Bermotor," vol. 2, 2024, Accessed: Dec. 24, 2025. [Online]. Available: <https://journal.artei.or.id/index.php/Mars/article/view/132>
- [10] R. D. Purnomoasri, T. Yuono, S. Sumina, and F. D. L. Utama, "Analisis Emisi CO dan CO₂ pada Simpang Bersinyal Bejen Kabupaten Karanganyar," vol. 25, 2023, doi: <https://doi.org/10.20961/enviro.v25i1.78524>.
- [11] R. Adhianti, K. Sari, and C. Meidiana, "Peningkatan Biokapasitas RTH Publik Dalam Upaya Pengurangan Emisi CO₂ Ruas Jalan Ranugrati Kota Malang," vol. 9, 2020, Accessed: Nov. 08, 2025. [Online]. Available: <https://purejournal.ub.ac.id/index.php/pure/article/view/116>
- [12] Efa and Ayu Handrini, "Studi Emisi Gas Buang CO₂ Terhadap Bahan Bakar Pada Sepeda Motor Empat Langkah dan Dua Langkah." Accessed: Nov. 08, 2025. [Online]. Available: <http://scholar.unand.ac.id/52015/>
- [13] M. Liang, Y. Chao, Y. Tu, and T. Xu, "Vehicle Pollutant Dispersion in the Urban Atmospheric Environment: A Review of Mechanism, Modeling, and Application," 2023, Accessed: Nov. 08, 2025. [Online]. Available: file:///C:/Users/Dinda%20Aprillia/AppData/Local/Temp/cbe3322c-dc15-4fee-839b-d52e2f8eff31_referensi%20jurnal.zip.f31/referensi%20jurnal/atmosphere-14-00279-v3.pdf
- [14] A. A. Ajala, O. L. Adeoye, O. M. Salami, and A. Y. Jimoh, "An examination of daily CO₂ emissions prediction through a comparative analysis of machine learning, deep learning, and statistical models," 2025, Accessed: Nov. 08, 2025. [Online]. Available: file:///C:/Users/Dinda%20Aprillia/AppData/Local/Temp/e24591af-a175-435f-86d8-545c8ea87982_referensi%20jurnal.zip.982/referensi%20jurnal/s11356-024-35764-8.pdf
- [15] M. R. Rodriguez, R. Besteiro, J. A. Ortega, M. D. Fernandez, and T. Arango, "Evolution and Neural Network Prediction of CO₂ Emissions in Weaned Piglet Farms," 2022, Accessed: Nov. 08, 2025. [Online]. Available: file:///C:/Users/Dinda%20Aprillia/AppData/Local/Temp/081f2df5-f752-4f03-a09f-24121b3aaf44_referensi%20jurnal.zip.f44/referensi%20jurnal/sensors-22-02910.pdf
- [16] P. Pryciński, P. Pielecha, J. Korzeb, J. Pielecha, M. Kostrzewski, and A. Eliwa, "Air Pollutant Emissions of Passenger Cars in Poland in Terms of Their Environmental Impact and Type of Energy Consumption," 2024, Accessed: Nov. 08, 2025. [Online]. Available: file:///C:/Users/Dinda%20Aprillia/AppData/Local/Temp/02a34d55-9d9b-45b5-a3fc-b2f3356e768c_referensi%20jurnal.zip.68c/referensi%20jurnal/energies-17-05357.pdf
- [17] K. BUSSABAN1, K. KULARBPHEHTONG2, and C. BOONSENG3, "Prediction Of CO₂ Emissions Using Machine Learning," 2023, Accessed: Nov. 08, 2025. [Online]. Available: file:///C:/Users/Dinda%20Aprillia/AppData/Local/Temp/969f35b5-7e18-4811-9b8b-4cb75da2a3c7_referensi%20jurnal.zip.3c7/referensi%20jurnal/VASSI-abstracts_129.pdf
- [18] J. Han, M. Kamber, and J. Pei, "Data Mining: Concepts and Techniques," 2011, Accessed: Dec. 25, 2025. [Online]. Available: https://ia600603.us.archive.org/2/items/datamining_201811/DS-book%20u5.pdf
- [19] S. Abe, "Minimal Complexity Support Vector Machines for Pattern Classification," 2020, Accessed: Dec. 25, 2025. [Online]. Available: <https://www.mdpi.com/2073-431X/9/4/88>

- [20] Y. G. Tejada and H. A. Mayer, "Deep Learning with Convolutional Neural Networks: A Compact Holistic Tutorial with Focus on Supervised Regression," 2024, Accessed: Dec. 25, 2025. [Online]. Available: <https://www.mdpi.com/2504-4990/6/4/132>
- [21] D. Podder, "CO2 Emission by Vehicles." Accessed: Nov. 08, 2025. [Online]. Available: <https://www.kaggle.com/datasets/debajyotipodder/co2-emission-by-vehicles>