



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: http://it.usu.ac.id

FORM PENGAJUAN JUDUL



Nama : Neha Sabila Nazmira Sitorus

NIM : 211402090

Judul diajukan oleh* : ☐ Dosen
☒ Mahasiswa

Bidang Ilmu (tulis dua bidang) :

1. Data Science and Intelligent System
2. Computer Graphics and Vision

Uji Kelayakan Judul** : ☐ Diterima ☐ Ditolak

Hasil Uji Kelayakan Judul :

Calon Dosen Pembimbing I: Umayya Ramadhani Putri, S.Kom., M.Kom.
(Jika judul dari dosen maka dosen tersebut berhak menjadi pembimbing I)

Calon Dosen Pembimbing II: Fanindia Purnamasari, S.TI., M.IT.

Paraf Calon Dosen Pembimbing I

Medan, 10 Januari 2025

Ka. Laboratorium Penelitian,

* Centang salah satu atau keduanya

** Pilih salah satu

(Fanindia Purnamasari, S.TI., M.IT.)

NIP. 198908172019032023



RINGKASAN JUDUL YANG DIAJUKAN

*Semua kolom di bawah ini diisi oleh mahasiswa yang sudah mendapat judul

Judul / Topik Skripsi	PREDIKSI JUMLAH EMISI KARBON DIOKSIDA (CO2) BERDASARKAN SPESIFIKASI PADA KENDARAAN MENGGUNAKAN METODE EXTREME GRADIENT BOOSTING (XGBOOST)
Latar Belakang dan Penelitian Terdahulu	<p>Latar Belakang</p> <p>Perubahan iklim global telah menjadi masalah mendesak yang berdampak pada ekosistem, kesehatan manusia, dan perekonomian di seluruh dunia. Salah satu penyebab utamanya adalah emisi gas rumah kaca, khususnya karbon dioksida (CO2) yang dihasilkan oleh berbagai aktivitas manusia, termasuk transportasi (Jainal et al., 2022). Dalam Peraturan Menteri Lingkungan Hidup Nomor 12 Tahun 2010 emisi didefinisikan sebagai zat, energi, dan/atau komponen lain yang dihasilkan dari suatu aktivitas yang dapat masuk ke lingkungan dan berpotensi menjadi polutan. Emisi karbon mengacu pada hasil sisa pembakaran bahan bakar yang dilepaskan melalui sistem pembuangan mesin (Nestiti, 2017).</p> <p>Di Indonesia, sektor transportasi menyumbang hampir 30% dari total emisi karbon dioksida (CO2) yang dihasilkan, dengan mayoritas, yaitu sekitar 88%, berasal dari transportasi darat. Transportasi darat ini didominasi oleh penggunaan bahan bakar minyak (BBM), yang mencapai pangsa hingga 99,97%. Data ini menunjukkan bahwa sektor transportasi, khususnya transportasi darat, memiliki kontribusi yang sangat signifikan terhadap produksi emisi karbon di negara ini. Dengan tingginya ketergantungan pada BBM, sektor ini memainkan peran kunci dalam peningkatan emisi gas rumah kaca, yang pada akhirnya berdampak pada perubahan iklim. Oleh karena itu, diperlukan upaya strategis untuk mengurangi emisi dari sektor transportasi, baik melalui pengembangan teknologi ramah lingkungan, transisi ke energi bersih, maupun perbaikan sistem transportasi secara keseluruhan (Anindya & Handayeni, 2024).</p> <p>Setiap jenis kendaraan memiliki spesifikasi unik, seperti jenis transmisi, kelas kendaraan, jumlah silinder, dan tingkat konsumsi bahan bakar. Faktor-faktor ini</p>



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

menyebabkan jumlah emisi karbon yang dihasilkan oleh tiap tipe kendaraan berbeda-beda. Selain itu, jenis bahan bakar yang digunakan juga memberikan pengaruh yang signifikan terhadap besarnya emisi. Sebagai contoh, kendaraan yang menggunakan bahan bakar Pertamina Turbo cenderung menghasilkan emisi karbon yang lebih rendah dibandingkan dengan kendaraan yang menggunakan bahan bakar Pertamina. Perbedaan ini menunjukkan pentingnya pemilihan bahan bakar yang lebih ramah lingkungan untuk mengurangi dampak negatif terhadap lingkungan (Sanusi et al., 2019).

Penelitian oleh (Asgaryansyah & Paniran, 2024) yang bertujuan untuk mengembangkan model prediktif menggunakan algoritma regresi linier untuk mengukur emisi CO₂ kendaraan bermotor berdasarkan jenis bahan bakar (premium, bensin, diesel, etanol) dan jarak tempuh. Dataset mencakup variabel seperti kombinasi konsumsi bahan bakar di jalan raya, emisi CO₂, dan jenis bahan bakar. Hasil penelitian menunjukkan model regresi linier memiliki akurasi tinggi, dengan tipe bahan bakar diesel memberikan performa terbaik dengan nilai MSE mencapai 0,5621 sedangkan untuk R² bernilai 0.9965 menunjukkan bahwa hampir semua variabilitas dalam y dapat dijelaskan oleh variabel x.

Penelitian terkait lainnya oleh (Ji et al., 2024) yang bertujuan untuk memprediksi emisi CO₂ dan kebutuhan energi di sektor transportasi China berdasarkan variabel seperti populasi, kilometer kendaraan, tahun, dan PDB per kapita. Metode yang digunakan meliputi algoritma jaringan saraf tiruan (ANN), dan Support Vector Machines (SVM), serta model matematika regresi linier dan eksponensial. Hasil menunjukkan bahwa algoritma SVM memiliki performa terbaik dalam prediksi emisi CO₂, sementara ANN unggul dalam memprediksi kebutuhan energi. Secara keseluruhan, semua model menunjukkan prediksi yang akurat, dengan rata-rata peningkatan tahunan konsumsi energi transportasi sebesar 3,8% dan emisi CO₂ sebesar 3,66%.

Adapun metode yang diusulkan pada penelitian ini menggunakan metode Extreme Gradient Boosting Regression. XGBoost secara iteratif menggabungkan prediksi dari beberapa pohon lemah untuk membentuk model yang memiliki performa tinggi (Zhou



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

et al., 2020). Metode ini bisa digunakan untuk penyelesaian klasifikasi maupun regresi (Osman et al., 2021). XGBoost Regression adalah pengembangan lanjutan dari algoritma gradient boosting. Metode ini menggunakan model yang lebih terstruktur untuk membangun pohon regresi, sehingga dapat meningkatkan performa sekaligus mengurangi kompleksitas model untuk menghindari overfitting. Prediksi akhir XGBoost diperoleh dengan menjumlahkan hasil prediksi dari setiap pohon regresi. Algoritma ini sangat efektif pada data dengan fitur kategorikal dan tetap menunjukkan kinerja yang baik meskipun menghadapi data dengan kelas yang tidak seimbang (Chairunisa et al., 2024).

Penelitian yang menggunakan metode XGBoost juga telah banyak dilakukan. Pada penelitian yang dilakukan oleh (Chairunisa et al., 2024) untuk memprediksi angka harapan hidup melalui model yang paling akurat dengan menggunakan model decision tree regression, random forest regression, gradient boosting regression, dan XGBoost regression. Model XGBoost regression mengalami peningkatan setelah tuning, dengan RMSE berkurang dari 2.77 menjadi 2.57 dan R^2 meningkat dari 0.85 menjadi 0.87.

Berdasarkan dari latar belakang dan penelitian terdahulu, maka penulis mengajukan penelitian untuk memprediksi jumlah emisi karbon dioksida (CO₂) pada kendaraan dengan judul **“Prediksi Jumlah Emisi Karbon Dioksida (CO₂) Berdasarkan Spesifikasi Pada Kendaraan Menggunakan Metode Extreme Gradient Boosting (XGBoost)”**.

Penelitian Terdahulu

No.	Penulis	Tahun	Metode	Keterangan
1.	Asgaryansyah & Paniran	2024	Linear Regression	Mengembangkan model prediktif menggunakan regresi linier untuk mengukur emisi CO ₂ kendaraan berdasarkan jenis bahan bakar dan konsumsi bahan bakar. Dataset mencakup variabel seperti



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

					konsumsi bahan bakar di jalan raya, jenis bahan bakar, dan emisi CO2. Hasil menunjukkan model regresi linier akurat, dengan bahan bakar diesel memberikan performa terbaik dengan nilai MSE 0,5621.
	2.	Ji et al	2024	ANN, SVM	Memprediksi emisi CO2 dan kebutuhan energi transportasi di China menggunakan variabel seperti populasi, jarak tempuh kendaraan, tahun, dan PDB per kapita. Metode yang digunakan meliputi ANN, SVM, serta regresi linier dan eksponensial. SVM unggul dalam prediksi emisi CO2, sementara ANN lebih baik untuk kebutuhan energi.
	3.	Suwandi et al	2022	SARIMA dan LSTM	Peramalan emisi karbon dunia dengan membandingkan performa metode SARIMA dan LSTM menggunakan data emisi karbon Amerika Utara dari tahun 1949 hingga 2018. Data dianalisis sebagai time series untuk mengidentifikasi pola historis. Hasil evaluasi menunjukkan bahwa model LSTM, dengan nilai MAPE 0,540%, lebih akurat dibandingkan SARIMA yang memiliki MAPE 1,995%. Hal ini



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

					menunjukkan bahwa LSTM lebih optimal untuk memprediksi emisi karbon.
	4.	Chairunisa et al	2024	Decision Tree Regression, Random Forest Regression, Gradient Boosting Regression, dan XGBoost Regression	Penelitian dilakukan untuk memprediksi angka harapan hidup menggunakan berbagai metode seperti decision tree regression, random forest regression, gradient boosting regression, dan XGBoost regression. Setelah tuning, performa XGBoost regression meningkat, dengan RMSE turun dari 2,77 menjadi 2,57 dan R ² naik dari 0,85 menjadi 0,87.
	5.	Prastiyo & Febriandirza	2024	XGBoost dan Random Forest Regression	Memprediksi tingkat kemiskinan di DKI Jakarta menggunakan data dari Badan Pusat Statistik (BPS) untuk periode 2010–2023. Hasil menunjukkan bahwa XGBoost Regression lebih unggul di wilayah Kepulauan Seribu, dengan MAPE sebesar 4,4938 dibandingkan Random Forest yang memiliki MAPE 4,5652. Selain itu, XGBoost Regression memiliki nilai MSE lebih rendah, yaitu 0,446641 dibandingkan 0,489093 pada Random Forest.



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

	6.	Syahreza et al	2024	Random Forest, Support Vector Regression, dan XGBoost	Memprediksi cuaca menggunakan data meteorologi (suhu minimum, suhu maksimum, curah hujan, arah angin, dan kelembaban rata-rata), dengan total 1650 data per variabel. Hasil menunjukkan bahwa XGBoost memberikan performa terbaik dibandingkan dengan model lainnya dengan MAE 0,3744, MSE 0,2278, dan R ² 0,8183.
	<p>Perbedaan penelitian dengan yang sebelumnya yaitu terletak pada data serta metode yang digunakan. Penelitian ini berfokus pada memprediksi jumlah emisi yaitu CO₂ pada kendaraan mobil atau kendaraan beroda empat. Data yang digunakan pada penelitian ini mencakup data spesifikasi kendaraan seperti perusahaan kendaraan, jenis mobil, kelas mobil, jenis transmisi, ukuran mesin, jumlah silinder, tipe bahan bakar, serta konsumsi bahan bakar di kota dan di tol. Metode yang digunakan pada penelitian ini adalah menggunakan extreme gradient boosting yang diketahui belum digunakan pada penelitian sebelumnya dalam memprediksi emisi karbon pada kendaraan.</p>				
Rumusan Masalah	<p>Tingginya emisi Karbon Dioksida (CO₂) yang dihasilkan terutama dari sektor transportasi darat telah menjadi perhatian serius dalam mitigasi perubahan iklim global. Perbedaan spesifikasi kendaraan, seperti jenis bahan bakar, konsumsi bahan bakar, jenis transmisi, serta konfigurasi mesin, menyebabkan variasi signifikan dalam jumlah emisi yang dihasilkan. Untuk mengatasi masalah ini, diperlukan sistem prediksi yang dapat memperkirakan emisi CO₂ yang dihasilkan oleh kendaraan secara akurat. Sistem ini tidak hanya membantu pembuat kebijakan mengelola emisi, tetapi juga membantu pengguna kendaraan memilih opsi yang lebih ramah lingkungan dan efisien.</p>				



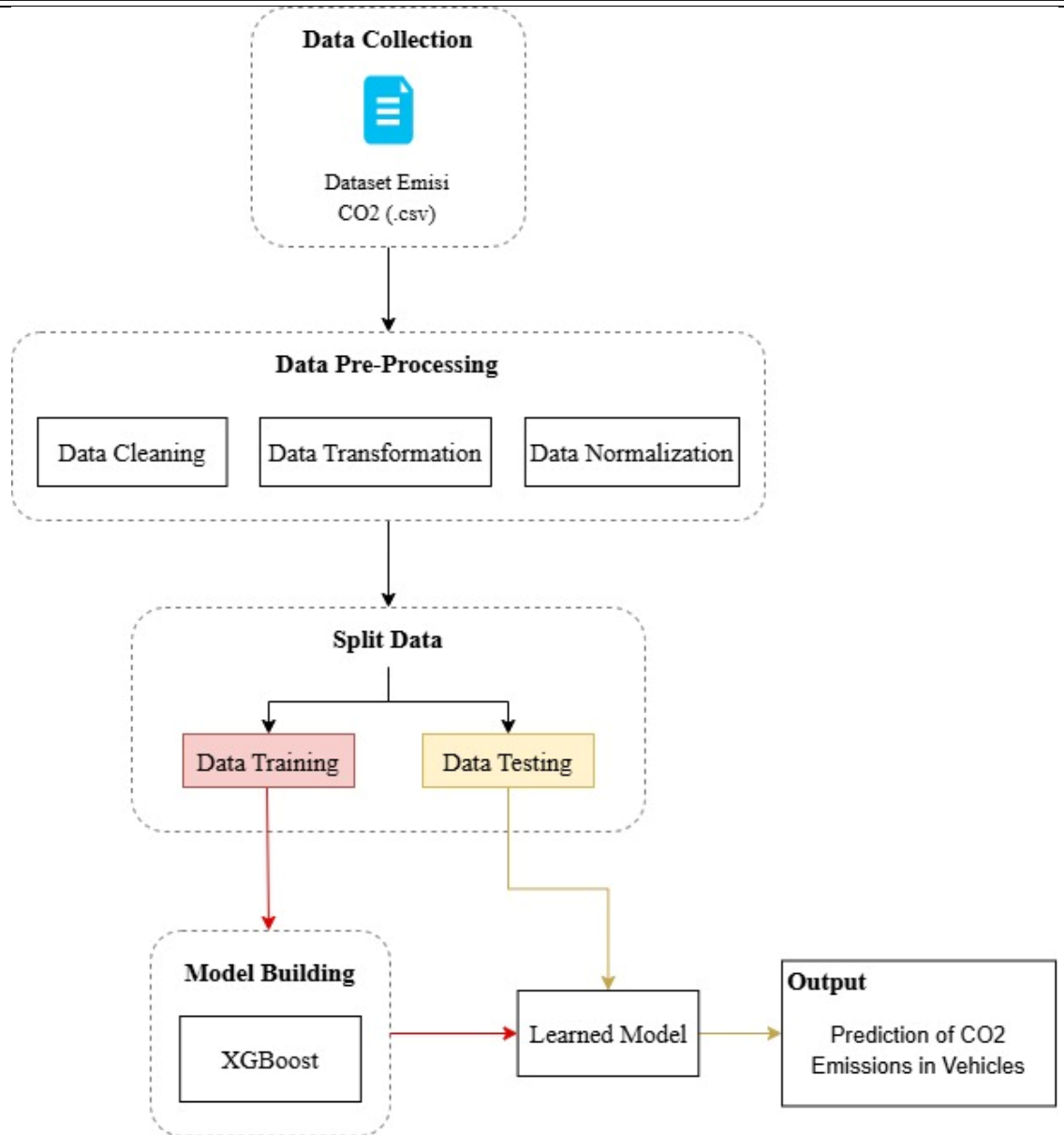
KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

Metodologi



Berdasarkan metode yang digunakan, penelitian ini melewati beberapa tahapan pemrosesan. Tahapan pertama dengan melakukan pengumpulan data mengenai emisi CO2 yang dihasilkan oleh berbagai kendaraan mobil. Data tersebut berisi informasi berupa spesifikasi kendaraan seperti perusahaan kendaraan, jenis mobil, kelas mobil, jenis transmisi, ukuran mesin, jumlah silinder, tipe bahan bakar, konsumsi bahan bakar di kota dan di tol, serta total emisi CO2 yang dihasilkan. Setelah pengumpulan data, dilakukan tahap *pre-processing* untuk membersihkan dan mempersiapkan data sebelum masuk ke dalam proses pelatihan. Data cleaning bertujuan untuk membersihkan data



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

	<p>dengan menghapus informasi yang tidak relevan dan data duplikat, serta menangani missing value dengan mengisi berdasarkan nilai rata-rata atau median dari jumlah data. Data transformation mencakup encoding data kategorikal menjadi bentuk numerik dan data normalisasi membuat data memiliki rentang nilai yang sama sehingga membuat model menganalisis data menjadi cepat. Tahapan selanjutnya dilakukan proses pembagian data menjadi data training dan data testing. Setelah data dibagi, selanjutnya adalah pembuatan model dan melatih model menggunakan data training dengan mempertimbangkan menggunakan hyperparameter tuning. Ketika diberikan data testing, model akan menggunakan informasi dari data training untuk menghasilkan prediksi.</p>
Referensi	<p>Asgaryansyah, N. K., & Paniran, N. P. (2024). Implementasi Algoritma Regresi Linier untuk Mengukur Tingkat Pengeluaran Co2 Pada Kendaraan Bermotor. <i>Deleted Journal</i>, 2(3), 84–93. https://doi.org/10.61132/mars.v2i3.132</p> <p>Anindya, A. I., & Handayani, K. D. M. E. (2024). Pengaruh Perilaku Perjalanan Masyarakat Kawasan Pusat Kota Surabaya terhadap Produksi Emisi Karbon. <i>Jurnal Teknik ITS</i>, 13(2). https://doi.org/10.12962/j23373539.v13i2.130199</p> <p>Chairunisa, G., Najib, M. K., Nurdianti, S., Imni, S. F., Sanjaya, W., Andriani, R. D., Henriyansah, N., Putri, R. S. P., & Ekaputri, D. (2024). Life expectancy prediction using decision tree, random forest, gradient boosting, and XGBooSt regressions. <i>Deleted Journal</i>, 2(2), 71–82. https://doi.org/10.62375/jsintak.v2i2.249</p> <p>Jainal, A., Erwina, A. H., Alwendi. (2022). Pentingnya Kesadaran Untuk Peduli Untuk Menjaga Dan Melestarikan Lingkungan. <i>Jurnal Pengabdian Masyarakat</i>. Vol 1. No. 3.</p> <p>Ji, T., Li, K., Sun, Q., & Duan, Z. (2024). Urban transport emission prediction analysis through machine learning and deep learning techniques. <i>Transportation Research Part D Transport and Environment</i>, 135, 104389. https://doi.org/10.1016/j.trd.2024.104389</p> <p>Nestiti, R. F. (2017). Perubahan emisi karbondioksida dengan pemindahan kendaraan pribadi ke kendaraan umum konvensional di Kota Surabaya. https://repository.its.ac.id/43326/</p> <p>Osman, A. I. A., Ahmed, A. N., Chow, M. F., Huang, Y. F., & El-Shafie, A. (2021).</p>



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

Extreme gradient boosting (Xgboost) model to predict the groundwater levels in Selangor Malaysia. *Ain Shams Engineering Journal*, 12(2), 1545–1556.
<https://doi.org/10.1016/j.asej.2020.11.011>

Prastiyo, I. W., & Febriandirza, A. (2024). Analisis Perbandingan Prediksi Tingkat Kemiskinan Menggunakan Metode XGBoost dan Random Forest Regression. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 8(3), 1694.
<https://doi.org/10.30865/mib.v8i3.7892>

Sanusi, M., Uloli, H., & Arafat, M. Y. (2019, December). PENGARUH VARIASIJENIS BAHAN BAKAR TERHADAP EMISI GAS BUANG PADA SEPEDA MOTOR VIXION 155cc VVA TIPE INJEKSI TAHUN 2018. In *SemanTECH (Seminar Nasional Teknologi, Sains Dan Humaniora)* (Vol. 1, No. 1, pp. 202-209).

Suwandi, S. I. N., Tyasnurita, R., & Muhayat, H. (2022). Peramalan Emisi Karbon Menggunakan Metode SARIMA dan LSTM. *Journal of Computer Science and Informatics Engineering (J-Cosine)*, 6(1), 73-80.

Syahreza, A., Ningrum, N. K., & Syahrazy, M. A. (2024). Perbandingan Kinerja Model Prediksi Cuaca: Random Forest, Support Vector Regression, dan XGBoost. *Edumatic: Jurnal Pendidikan Informatika*, 8(2), 526–534.
<https://doi.org/10.29408/edumatic.v8i2.27640>

Zhou, J., Qiu, Y., Zhu, S., Armaghani, D. J., Khandelwal, M., & Mohamad, E. T. (2020). Estimation of the TBM advance rate under hard rock conditions using XGBoost and Bayesian optimization. *Underground Space*, 6(5), 506–515.
<https://doi.org/10.1016/j.undsp.2020.05.008>

Medan, 10 Januari 2025
Mahasiswa yang mengajukan,

(Neha Sabila Nazmira Siorus)

NIM. 211402090