

# ANALISIS FAKTOR-FAKTOR YANG MEMPENGARUHI JUMLAH TIMBULAN SAMPAH DI PROVINSI JAWA TIMUR MENGGUNAKAN METODE RANDOM FOREST REGRESSION

*(Analysis of Factors Affecting the Amount of Waste Generation in East Java Province Using  
the Random Forest Regression Method)*

**Meuthia Nur Azizah Alamsyah<sup>1\*</sup>, Sofia Dawani Silaban<sup>2</sup>, Dave Agles Rizky Nor<sup>3</sup>,  
Lalu Muhamad Rifani Fadli<sup>4</sup>, Surya Puspita Sari<sup>5</sup>, Magdalena Effendi<sup>6</sup>**

Institut Teknologi Kalimantan<sup>1,2,3,4,5,6</sup>

E-mail: [16231032@student.itk.ac.id](mailto:16231032@student.itk.ac.id)

## ABSTRAK

Timbulan sampah di Provinsi Jawa Timur terus meningkat seiring pertumbuhan penduduk dan aktivitas ekonomi. Penelitian ini bertujuan menganalisis faktor-faktor yang mempengaruhi timbulan sampah serta membangun model prediksi menggunakan Random Forest Regression. Data yang digunakan merupakan data sekunder tahun 2024 dari Badan Pusat Statistik (BPS) dan Sistem Informasi Pengelolaan Sampah Nasional (SIPSN) dengan unit analisis 38 kabupaten/kota. Variabel penelitian meliputi jumlah penduduk, PDRB, luas wilayah, kepadatan penduduk, dan persentase penanganan sampah. Pemodelan dilakukan setelah pra-pemrosesan data menggunakan standarisasi Z-score dan dievaluasi dengan 5-Fold Cross Validation. Hasil evaluasi menunjukkan nilai  $R^2$  sebesar 0,683, RMSE sebesar 68.766 ton, dan MAE sebesar 34.091 ton. Analisis feature importance berdasarkan Increase in Node Purity menunjukkan bahwa jumlah penduduk merupakan faktor paling dominan yang mempengaruhi timbulan sampah, diikuti oleh PDRB, sementara luas wilayah dan kepadatan penduduk berpengaruh sedang dan persentase penanganan sampah memiliki pengaruh paling rendah. Hasil ini menegaskan bahwa faktor kependudukan dan ekonomi menjadi determinan utama timbulan sampah di Jawa Timur serta mendukung perumusan kebijakan pengelolaan sampah yang lebih berkelanjutan.

**Kata kunci:** timbulan sampah, Jawa Timur, Random Forest Regression, faktor sosial ekonomi, PDRB.

## ABSTRACT

*Waste generation in East Java Province continues to increase along with population growth and economic activity. This study aims to analyze the factors influencing waste generation and to develop a prediction model using Random Forest Regression. The data used are secondary data for 2024 obtained from Statistics Indonesia (BPS) and the National Waste Management Information System (SIPSN), covering 38 districts/cities. The variables include population size, Gross Regional Domestic Product (GRDP), land area, population density, and the percentage of waste handling. After data preprocessing using Z-score standardization, the model was evaluated using 5-Fold Cross Validation. The results show an  $R^2$  value of 0.683, an RMSE of 68,766 tons, and an MAE of 34,091 tons. Feature importance analysis based on Increase in Node Purity indicates that population size is the most dominant factor affecting waste generation, followed by GRDP. Land area and population density have moderate effects, while the percentage of waste handling shows the lowest contribution. These findings confirm that demographic and economic factors are the main determinants of waste generation in East Java and can support the formulation of more sustainable waste management policies.*

**Keywords:** Waste generation, East Java, Random Forest Regression, Socioeconomic factors, GRDP.

## PENDAHULUAN

Permasalahan timbulan sampah merupakan salah satu isu lingkungan yang semakin kompleks di Indonesia. Peningkatan jumlah penduduk, urbanisasi, serta pertumbuhan aktivitas ekonomi berkontribusi terhadap meningkatnya volume sampah yang dihasilkan setiap tahun. Data Kementerian Lingkungan Hidup dan Kehutanan (2024) menunjukkan bahwa timbulan sampah nasional telah mencapai puluhan juta ton per tahun, mencerminkan bahwa sistem pengelolaan sampah masih menghadapi tantangan dari aspek teknis, sosial, dan kelembagaan.

Provinsi Jawa Timur merupakan salah satu wilayah dengan tingkat timbulan sampah yang relatif tinggi karena memiliki jumlah penduduk yang besar dan aktivitas ekonomi yang intensif. Perbedaan karakteristik wilayah antar kabupaten/kota menyebabkan variasi pola timbulan sampah, sehingga diperlukan analisis komprehensif untuk mengidentifikasi faktor-faktor yang mempengaruhinya (Badan Pusat Statistik, 2024).

Sejumlah penelitian terdahulu menunjukkan bahwa faktor sosial ekonomi seperti jumlah penduduk, tingkat pendapatan, dan PDRB berpengaruh terhadap timbulan sampah (Aulia & Hadju, 2024; Putri et al., 2025). Ramdani (2024) juga mengemukakan adanya pola Environmental Kuznets Curve pada hubungan antara pertumbuhan ekonomi dan timbulan sampah di Indonesia. Selain itu, evaluasi sistem pengelolaan sampah nasional menunjukkan bahwa pendekatan pengelolaan masih cenderung reaktif dan belum optimal dalam pengurangan sampah dari sumbernya (Julia Lingga et al., 2024; UNEP, 2018).

Sebagian besar penelitian sebelumnya menggunakan pendekatan regresi linier, yang memiliki keterbatasan dalam menangkap hubungan nonlinier dan interaksi kompleks antarvariabel (Halif et al., 2025; Ilmi & Syaifuddin, 2025). Oleh karena itu, diperlukan pendekatan yang lebih fleksibel, salah satunya melalui metode machine learning. Random Forest Regression merupakan metode ensemble yang mampu menangani hubungan nonlinier, mengurangi risiko overfitting, serta mengidentifikasi variabel yang paling berpengaruh (Breiman, 2001; Fitri & Nugraha, 2024). Penerapan metode ini diharapkan dapat memberikan gambaran yang lebih akurat mengenai faktor-faktor yang mempengaruhi timbulan sampah di Provinsi Jawa Timur.

## METODE

Penelitian ini menggunakan data sekunder yang mencakup wilayah Provinsi Jawa Timur, yang terdiri atas 38 kabupaten dan kota. Wilayah tersebut dipilih karena memiliki ketersediaan data timbulan sampah serta indikator sosial ekonomi yang relatif lengkap pada setiap daerah. Data yang digunakan merupakan data tahun 2024 yang diperoleh dari dua sumber utama, yaitu Sistem Informasi Pengelolaan Sampah Nasional (SIPSN) dan Badan Pusat Statistik (BPS) Provinsi Jawa Timur. Variabel dalam penelitian ini terdiri atas satu variabel dependen dan lima variabel independen. Variabel dependen adalah jumlah timbulan sampah tahunan (Y), sedangkan variabel independen meliputi indikator sosial ekonomi sebagaimana ditunjukkan pada **Tabel 1** berikut.

**Tabel 1.** Variabel Penelitian.

Variabel	Nama Variabel	Satuan	Sumber Data
Y	Timbulan Sampah Tahunan	Ton/tahun	SIPSN
X1	Luas Wilayah	Km <sup>2</sup>	BPS
X2	Jumlah Penduduk	%	BPS
X3	Kepadatan Penduduk	Jiwa/km <sup>2</sup>	BPS
X4	PDRB	Rupiah	BPS
X5	Persentase Penanganan Sampah	%	SIPSN

Sumber: [bps.go.id](https://bps.go.id) dan [sipsn.kemenvh.go.id](https://sipsn.kemenvh.go.id)

### Pre-Processing Data

Metode penelitian ini diawali dengan pra pemrosesan data untuk memastikan konsistensi dan skala yang sebanding antarvariabel. Tahapan ini biasanya mencakup *cleaning* seperti menghapus duplikasi dan menangani nilai hilang, *transforming* seperti normalisasi atau standarisasi, serta *reduction* atau pemilihan fitur agar data lebih konsisten dan berkualitas untuk diproses (Syam, 2024). Data dibersihkan dari nilai hilang (*missing value*) kemudian dilakukan normalisasi menggunakan metode standarisasi Z-Score agar seluruh variabel memiliki skala yang setara, dengan rumus yaitu:

$$X' = \frac{X - \bar{X}}{\sigma} \quad (1)$$

dimana:

$\bar{X}$  = rata-rata

$\sigma$  = simpangan baku

### **Random Forest Regression**

Random Forest Regression (RFR) merupakan metode *ensemble* yang membangun sejumlah pohon keputusan menggunakan data yang di-*bootstrap* dan menghasilkan prediksi akhir melalui proses agregasi untuk meningkatkan akurasi dan stabilitas model (Fitri, 2024). Teknik ini bekerja dengan memilih subset fitur secara acak pada setiap pemisahan node, sehingga mampu mengurangi risiko *overfitting* dan memberikan performa prediksi yang lebih baik pada data dengan pola kompleks dibandingkan model regresi sederhana (Susetianingtias, 2022). Pemodelan dilakukan menggunakan metode *Random Forest Regression* (RFR), yang membangun sejumlah pohon keputusan  $f_b(X)$  melalui teknik *bootstrap sampling*, dengan hasil Prediksi akhir diperoleh dengan menghitung rata-rata dari seluruh hasil prediksi pohon yang terbentuk (Breiman, 2001). Adapun rumus dari prediksi akhir yaitu:

$$\hat{Y} = \frac{1}{B} \sum_{b=1}^B f_b(X) \quad (2)$$

dimana:

$\hat{Y}$  : Nilai prediksi (hasil akhir model).

B : Jumlah pohon.

$f_b(X)$  : Hasil prediksi dari model ke-b terhadap data input X.

$\sum_{b=1}^B$  : Penjumlahan seluruh hasil prediksi dari model ke-1 hingga ke-B.

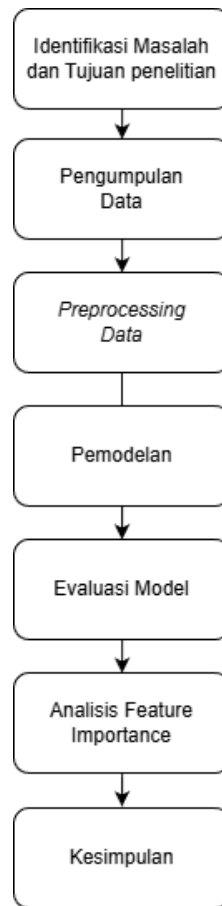
Penentuan parameter model dilakukan secara empiris, dengan jumlah pohon (ntree) ditetapkan sebesar 1000 dan parameter mtry ditentukan melalui proses tuning untuk memperoleh kinerja optimal. Evaluasi model dilakukan menggunakan metode 5-Fold Cross Validation dengan metrik  $R^2$ , RMSE, dan MAE. Pendekatan Random Forest dipilih karena telah terbukti memberikan performa yang baik pada data lingkungan dan sosial ekonomi (Hakim et al., 2023; Indaryono et al., 2024; Sari et al., 2023).

### **Analisis Feature Important**

Analisis feature importance dilakukan untuk mengidentifikasi kontribusi relatif masing-masing variabel independen dalam mempengaruhi timbulan sampah. Dalam penelitian ini, feature importance dievaluasi berdasarkan nilai Increase in Node Purity, yang menunjukkan sejauh mana suatu variabel berkontribusi dalam meningkatkan homogenitas node pada pohon keputusan. Variabel dengan nilai Increase in Node Purity yang lebih tinggi dianggap memiliki pengaruh yang lebih besar dalam pembentukan model Random Forest (Breiman, 2001).

### **Tahap Analisis Data**

Tahapan analisis data dalam penelitian ini disusun melalui rangkaian tahapan yang terstruktur. Berikut diagram alir dari penelitian kami, yang memberikan gambaran umum mengenai alur kerja yang digunakan dalam pelaksanaan penelitian ini.



**Gambar 1.** Diagram Alir Penelitian

Penelitian ini diawali dengan identifikasi permasalahan terkait peningkatan timbulan sampah di Provinsi Jawa Timur serta penetapan tujuan penelitian, yaitu menganalisis faktor-faktor yang memengaruhi timbulan sampah dan memprediksi jumlah timbulan sampah menggunakan metode Random Forest Regression. Selanjutnya dilakukan pengumpulan data sekunder yang bersumber dari Sistem Informasi Pengelolaan Sampah Nasional (SIPSN) dan Badan Pusat Statistik (BPS), dengan timbulan sampah tahunan sebagai variabel dependen dan lima indikator sosial ekonomi sebagai variabel independen. Data yang diperoleh kemudian melalui tahap *preprocessing* yang meliputi pembersihan data dari nilai tidak valid, konversi tipe data, serta standarisasi menggunakan metode Z-Score agar skala antar variabel seragam. Tahap berikutnya adalah pemodelan menggunakan Random Forest Regression dengan membangun sejumlah pohon keputusan melalui *bootstrap sampling* dan pemilihan variabel secara acak, di mana prediksi akhir diperoleh dari rata-rata seluruh pohon. Kinerja model selanjutnya dievaluasi menggunakan metode 5-Fold Cross Validation dengan metrik  $R^2$ , RMSE, dan MAE untuk menilai kemampuan prediksi dan tingkat kesalahan model. Setelah itu, dilakukan analisis *feature importance* berdasarkan nilai *Increase in Node Purity* untuk mengidentifikasi variabel yang paling berpengaruh terhadap timbulan sampah. Tahap akhir penelitian adalah penarikan kesimpulan berdasarkan hasil evaluasi model dan analisis *feature importance* guna mengidentifikasi faktor dominan serta implikasinya terhadap kebijakan pengelolaan sampah.

## HASIL DAN PEMBAHASAN

### *Preprocessing Data*

Pra-pemrosesan data dilakukan untuk memastikan kualitas dan keseragaman data sebelum pemodelan. Pada tahap awal, data diperiksa untuk mengidentifikasi nilai tidak valid, data hilang, serta ketidakkonsistenan format, kemudian dilakukan pembersihan dan konversi seluruh variabel ke dalam bentuk numerik. Sebelum dilakukan standarisasi, setiap variabel memiliki rentang nilai dan satuan yang berbeda-beda, misalnya timbulan sampah dalam satuan ton per tahun dan indikator sosial ekonomi

dalam skala atau ukuran yang beragam, sehingga berpotensi menimbulkan bias dalam proses pembentukan model. Oleh karena itu, data distandarisasi menggunakan metode Z-Score agar setiap variabel memiliki skala yang sebanding dan tidak mendominasi pembentukan model.

### **Pemodelan Menggunakan Random Forest Regression**

Pemodelan timbulan sampah dilakukan menggunakan metode Random Forest Regression dengan bantuan paket *caret* pada perangkat lunak R. Model dibangun menggunakan satu variabel respon, yaitu timbulan sampah tahunan (Y), serta lima variabel prediktor yang meliputi luas wilayah (X1), jumlah penduduk (X2), kepadatan penduduk (X3), Produk Domestik Regional Bruto (PDRB) (X4), dan persentase penanganan sampah (X5).

Data yang telah melalui tahap pra-pemrosesan digunakan secara langsung dalam pemodelan tanpa pemisahan eksplisit antara data latih dan data uji, karena proses evaluasi dilakukan menggunakan metode 5-Fold Cross Validation. Dalam metode ini, data dibagi menjadi lima bagian dengan proporsi yang seimbang, di mana pada setiap iterasi empat bagian digunakan sebagai data pelatihan dan satu bagian sebagai data validasi. Proses ini dilakukan secara berulang hingga seluruh data berperan sebagai data validasi.

Random Forest bekerja dengan membangun sejumlah pohon keputusan (*decision tree*) melalui teknik *bootstrap sampling*, yaitu pengambilan sampel data secara acak dengan pengembalian. Pada setiap pemisahan node, algoritma memilih subset variabel prediktor secara acak untuk menentukan pemisahan terbaik. Mekanisme ini bertujuan untuk mengurangi korelasi antar pohon serta meningkatkan kemampuan generalisasi model.

Penentuan parameter model dilakukan melalui proses *hyperparameter tuning*, khususnya pada parameter *mtry*, yaitu jumlah variabel prediktor yang dipilih secara acak pada setiap pemisahan node. Nilai *mtry* diuji dalam beberapa kombinasi untuk memperoleh kinerja model terbaik. Selain itu, jumlah pohon (*n tree*) ditetapkan sebesar 1000 pohon guna meningkatkan kestabilan hasil prediksi dan mengurangi variabilitas model.

### **Evaluasi Model**

Evaluasi kinerja model dilakukan berdasarkan hasil prediksi dari proses 5-Fold Cross Validation. Nilai prediksi dibandingkan dengan nilai aktual untuk menghitung metrik evaluasi berupa Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), dan koefisien determinasi ( $R^2$ ) seperti yang ditunjukkan pada **Tabel 3.** berikut ini.

**Tabel 3.** Evaluasi Model *Random Forest Regression*.

Metrik	Nilai
RMSE	68766.16
MSE	34090.88
$R^2$	0.683

Hasil evaluasi menunjukkan bahwa model Random Forest Regression menghasilkan nilai RMSE sebesar 68.766, MAE sebesar 34.091, dan  $R^2$  sebesar 0,683. Nilai  $R^2$  tersebut menunjukkan bahwa model mampu menjelaskan sekitar 68,3% variasi timbulan sampah tahunan antar kabupaten/kota di Provinsi Jawa Timur. Nilai RMSE dan MAE mengindikasikan bahwa tingkat kesalahan prediksi model masih berada pada batas yang dapat diterima.

### **Analisis Feature Importance**

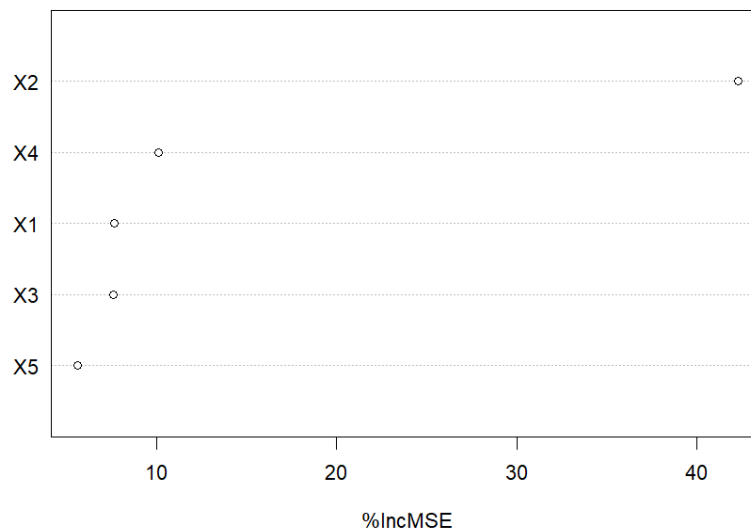
Analisis feature importance dilakukan untuk mengetahui variabel prediktor yang paling berpengaruh dalam pembentukan model Random Forest. Pengukuran feature importance didasarkan pada nilai *Increase in Node Purity*, yang menunjukkan kontribusi masing-masing variabel dalam

meningkatkan kemurnian node pada proses pembentukan pohon keputusan. Hasil analisis tersebut ditampilkan pada **Tabel 4** berikut ini.

**Tabel 4.** Analisis *Feature Importance* Model *Random Forest Regression*.

Variabel	%IncMSE	IncNodePurity
X1 (Luas Wilayah)	7.628956	19972301114
X2 (Jumlah Penduduk)	42.298144	279064552686
X3 (Kepadatan Penduduk)	7.595795	86159263933
X4 (PDRB)	10.103556	55045288850
X5 (Persentase Penanganan Sampah)	5.567129	34421328279

**Feature Importance Random Forest**



**Gambar 2.** Plot Feature Important

Analisis feature importance menunjukkan bahwa jumlah penduduk merupakan variabel paling dominan dalam mempengaruhi timbulan sampah, diikuti oleh PDRB. Temuan ini sejalan dengan penelitian Ilmi dan Syaifuddin (2025) serta Aulia dan Hadju (2024) yang menyatakan bahwa faktor kependudukan dan ekonomi memiliki pengaruh signifikan terhadap peningkatan timbulan sampah. Variabel luas wilayah dan kepadatan penduduk memiliki pengaruh pada tingkat sedang, sedangkan persentase penanganan sampah menunjukkan kontribusi paling rendah.

Rendahnya kontribusi variabel penanganan sampah menunjukkan bahwa peningkatan timbulan sampah masih lebih dipengaruhi oleh faktor hulu seperti konsumsi dan aktivitas ekonomi dibandingkan faktor hilir berupa pengelolaan. Hal ini sejalan dengan evaluasi kebijakan pengelolaan sampah nasional yang menekankan pentingnya pengurangan sampah dari sumbernya (Julia Lingga et al., 2024; UNEP, 2018).

## KESIMPULAN

Penelitian ini menunjukkan bahwa metode Random Forest Regression mampu memodelkan timbulan sampah tahunan pada 38 kabupaten/kota di Provinsi Jawa Timur dengan kinerja yang cukup baik. Nilai koefisien determinasi ( $R^2$ ) yang diperoleh menunjukkan bahwa model dapat menjelaskan sebagian besar variasi timbulan sampah, dengan tingkat kesalahan prediksi yang masih berada pada

batas yang dapat diterima. Hasil ini sejalan dengan temuan Breiman (2001) serta Fitri dan Nugraha (2024) yang menyatakan bahwa Random Forest memiliki kemampuan yang baik dalam menangkap hubungan nonlinier dan kompleks pada data lingkungan dan sosial ekonomi. Dengan demikian, pendekatan berbasis *machine learning* dapat digunakan sebagai alat analisis dan prediksi yang andal dalam kajian pengelolaan sampah daerah.

Hasil analisis *feature importance* mengindikasikan bahwa faktor kependudukan, khususnya jumlah penduduk, merupakan variabel yang paling dominan dalam mempengaruhi timbulan sampah di Provinsi Jawa Timur, diikuti oleh faktor ekonomi yang direpresentasikan oleh Produk Domestik Regional Bruto (PDRB). Temuan ini konsisten dengan penelitian sebelumnya yang menunjukkan bahwa peningkatan populasi dan aktivitas ekonomi berkontribusi signifikan terhadap peningkatan timbulan sampah (Aulia & Hadju, 2024; Ilmi & Syaifuddin, 2025; Putri et al., 2025). Sementara itu, luas wilayah dan kepadatan penduduk memiliki pengaruh pada tingkat sedang, dan persentase penanganan sampah menunjukkan kontribusi yang relatif lebih rendah dalam pembentukan model.

Berdasarkan temuan tersebut, dapat disimpulkan bahwa peningkatan timbulan sampah di Provinsi Jawa Timur lebih banyak dipengaruhi oleh tekanan kependudukan dan aktivitas ekonomi. Oleh karena itu, pemerintah daerah perlu memprioritaskan kebijakan pengelolaan sampah pada wilayah dengan jumlah penduduk dan aktivitas ekonomi yang tinggi, serta mengintegrasikan kebijakan pengelolaan sampah dengan perencanaan pembangunan daerah. Pendekatan ini sejalan dengan rekomendasi penguatan pengelolaan sampah perkotaan dan pengurangan sampah dari sumbernya sebagaimana ditekankan dalam agenda pembangunan berkelanjutan (Bappenas, 2020; United Nations, 2015; UNEP, 2018). Upaya pengurangan sampah dari sumbernya dan penguatan sistem pengelolaan sampah yang berkelanjutan diharapkan dapat menjadi langkah strategis dalam mengendalikan peningkatan timbulan sampah di masa mendatang.

## DAFTAR PUSTAKA

- Aulia, R., & Hadju, S. (2024). Analisis faktor ekonomi terhadap timbulan sampah perkotaan di Indonesia. *Jurnal Pengelolaan Lingkungan*, 12(2), 115–126.
- Bappenas. (2020). *SDGs Indonesia: Pedoman pelaksanaan pencapaian tujuan pembangunan berkelanjutan*. Kementerian PPN/Bappenas.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.
- Diana Tri Susetianingtias, Eka Patriya, & Rodiah. (2022). Model Random Forest Regression Untuk Peramalan Penyebaran Covid-19 di Indonesia. *Decode: Jurnal Pendidikan Teknologi Informasi*, 2(2), 84–95.
- Fitri, E., & Nugraha, S. N. (2024). Optimasi kinerja Linear Regression, Random Forest Regression dan Multilayer Perceptron pada prediksi hasil panen. *INTI Nusa Mandiri*, 18(2), 210–217.
- Hakim, A. R., Atmaja, D. M. U., Basri, A., & Syafii, M. (2023). Implementation of Random Forest Algorithm on Palm Oil Price Data. *Journal of Tech-E*, 6(2). Universitas Budi Dharma.
- Halif, J., Wahiddin, D., Sanjaya, I., & Faisal, S. (2025). Model regresi linear berganda untuk prediksi tingkat pengangguran di Provinsi Jawa Barat. *Jurnal Algoritma*, 22(1), 324–335.
- Ilmi, M. B., & Syaifuddin, A. (2025). Analisis faktor-faktor yang mempengaruhi jumlah timbulan sampah di Provinsi Jawa Timur menggunakan metode multiple linear regression. *Jurnal Ilmiah Teknologi Informasi dan Sains (SUBMIT)*, 5(1), 22–28.
- Indaryono, R., Pratama, B., & Nanda, F. (2024). Penerapan algoritma Random Forest untuk prediksi parameter lingkungan di kawasan perkotaan. *Jurnal Teknologi dan Data*, 8(3), 201–214.
- Julia Lingga, R., Wahyuni, S., & Pratama, D. (2024). Evaluasi sistem pengelolaan sampah nasional dan tantangan implementasinya. *Jurnal Lingkungan Indonesia*, 8(2), 101–115.
- Kementerian Lingkungan Hidup dan Kehutanan. (2024). *Statistik Pengelolaan Sampah Nasional 2024*. KLHK.

- Putri, N., Wijaya, D., & Sihombing, E. (2025). Hubungan antara tingkat pendapatan dan timbulan sampah plastik di wilayah perkotaan. *Jurnal Ekologi dan Pembangunan*, 9(1), 55–68.
- Ramdani, A. (2024). Analisis Environmental Kuznets Curve terhadap timbulan sampah di Indonesia. *Jurnal Ekonomi dan Lingkungan*, 7(2), 98–112.
- Sari, M., Nugraha, T., & Handayani, F. (2023). Perbandingan algoritma machine learning untuk prediksi variabel lingkungan: Kasus kualitas air sungai. *Jurnal Data dan Sains Terapan*, 6(4), 233–244.
- United Nations. (2015). *Sustainable Development Goal 11: Make cities and human settlements inclusive, safe, resilient and sustainable*. United Nations.
- United Nations. (2015). *Sustainable Development Goal 12: Ensure sustainable consumption and production patterns*. United Nations.
- United Nations Environment Programme. (2018). *Single-use plastics: A roadmap for sustainability*. UNEP.