



Implementasi *Random Forest* Untuk Identifikasi Jenis Sampah Organik Dan Non-Organik

Helmina Yeni Da Silva

Institut Keguruan dan Teknologi Larantuka

Alfian Nara Weking

Institut Keguruan dan Teknologi Larantuka

Dominikus Boli Watomakin

Institut Keguruan dan Teknologi Larantuka

Alamat: Jl. Ki Hajar Dewantara Kec. Larantuka - Kab. Flores Timur - Prov. Nusa Tenggara Timur

Korespondensi penulis: helmindadasilva@gmail.com

Abstract. Waste is a growing environmental problem, especially if it is not managed properly starting from the sorting process. One effort to improve the effectiveness of waste management is through automatic identification of waste types. This study aims to implement the Random Forest algorithm in the process of classifying waste into two categories: organic and non-organic waste. The data used are waste images that have gone through the preprocessing stage and the extraction of color and texture features. The Random Forest model was chosen because it has advantages in handling diverse data and providing stable classification results. Test results show that this model is capable of classifying with a fairly good level of accuracy, with the highest accuracy of 87% on the test data. In addition, this model is also integrated into a mobile application to facilitate users in identifying waste types in real-time. This implementation is expected to help the community sort waste more efficiently and contribute to sustainable environmental management.

Keywords: Organic and Non-Organic Waste, Waste Classification, Random Forest.

Abstrak. Sampah merupakan permasalahan lingkungan yang semakin meningkat, terutama jika tidak dikelola dengan baik sejak dari proses pemilahan. Salah satu upaya untuk meningkatkan efektivitas pengelolaan sampah adalah melalui identifikasi jenis sampah secara otomatis. Penelitian ini bertujuan untuk mengimplementasikan algoritma Random Forest dalam proses klasifikasi sampah menjadi dua kategori, yaitu sampah organik dan non-organik. Data yang digunakan berupa gambar sampah yang telah melalui tahap praproses dan ekstraksi fitur warna serta tekstur. Model Random Forest dipilih karena memiliki keunggulan dalam menangani data beragam dan memberikan hasil klasifikasi yang stabil. Hasil pengujian menunjukkan bahwa model ini mampu melakukan klasifikasi dengan tingkat akurasi yang cukup baik, dengan akurasi tertinggi sebesar 87% pada data uji. Selain itu, model ini juga diintegrasikan ke dalam aplikasi mobile untuk memudahkan pengguna dalam mengidentifikasi jenis sampah secara real-time. Implementasi ini diharapkan dapat membantu masyarakat dalam memilah sampah

Received May 13, 2025; Revised May 17, 2025; Accepted June 28, 2025

* Helmina Yeni Da Silva, helmindadasilva@gmail.com

secara lebih efisien dan berkontribusi terhadap pengelolaan lingkungan yang berkelanjutan.

Kata kunci: Sampah Organik dan Non-Organik, klasifikasi sampah, *Random Forest*.

LATAR BELAKANG

Sampah merupakan masalah utama bagi lingkungan hidup yang semakin mendesak untuk diatasi seiring dengan pertumbuhan populasi dan peningkatan konsumsi. Kurangnya kepedulian manusia terhadap pentingnya proses pemilahan sampah sebelum dibuang, sehingga sampah yang seharusnya dapat didaur ulang justru terkontaminasi dan tidak bisa lagi digunakan. Di Larantuka, lebih dari 70% sampah rumah tangga adalah sampah organik, namun masih tercampur dengan non-organik karena kurangnya edukasi dan alat bantu pemilahan otomatis. Pengelolaan sampah yang baik, bukan untuk kesehatan saja, tetapi juga untuk keindahan lingkungan. Pengolahan sampah yang sering dilakukan yaitu pengumpulan, pengangkutan yang sebagian besar berakhir di tempat pembuangan akhir (TPA), sampai dengan pemusnahan yang menyebabkan dampak pada lingkungan.(Ibnul Rasidi et al. 2022)

Pengenalan jenis sampah organik dan non-organik secara manual membutuhkan waktu dan tenaga, sementara keakuratan bisa bervariasi. Dengan menggunakan teknologi untuk mendukung identifikasi ini, kita dapat meningkatkan efisiensi dan akurasi dalam proses pengelolaan sampah. Algoritma seperti Random Forest telah terbukti efektif dalam klasifikasi berbasis gambar. Dengan menggunakan teknik ini, kita dapat mengembangkan model yang mampu mengenali pola-pola kompleks yang terkait dengan sampah organik dan non-organik dari gambar yang diambil menggunakan kamera ponsel.(Siburian and Mulyana 2018)

Penggunaan aplikasi mobile menawarkan aksesibilitas yang luas dan kemudahan penggunaan bagi masyarakat umum. Dengan menggunakan teknologi ini, kita dapat menghadirkan solusi yang praktis dan mudah digunakan untuk memfasilitasi pengenalan jenis sampah secara langsung dari perangkat pengguna. Implementasi aplikasi ini tidak hanya memberikan manfaat langsung bagi

pengguna dalam mengidentifikasi sampah, tetapi juga berpotensi untuk meningkatkan kesadaran lingkungan dan kepedulian terhadap pengelolaan sampah yang berkelanjutan.(Alden and Sari 2023)

KAJIAN TEORITIS

a. Sampah Organik dan Non-organik

Sampah organik adalah sampah yang mudah terurai dan membusuk yaitu dari makhluk hidup, baik manusia, hewan dan tumbuhan. Sampah organik terbagi menjadi dua yaitu sampah organik basah dan sampah organik kering. Sampah yang mengandung air yang cukup tinggi seperti kulit buah dan sisa sayuran termasuk dalam sampah basah. Sampah kering merupakan sampah yang kandungan airnya sedikit seperti kayu, ranting pohon, dan daun kering.

Sampah anorganik merupakan sampah yang sulit untuk membusuk dan sulit terurai. Sampah anorganik berasal dari bahan yang terbuat dari plastik dan logam. Dan sampah anorganik juga memiliki nilai ekonomis bagi perusahaan-perusahaan yang mendaur ulang sampah plastik.(Imbarraga dan Reinaldi 2019).

b. Pengumpulan Data

Pada penelitian ini data yang digunakan yaitu data sekunder. Data sekunder merupakan sumber data yang didapatkan peneliti dengan media perantara atau tidak secara langsung. Data sekunder pada penelitian ini menggunakan *roboflow universe dataset*.(Maulidah et al. 2024)

c. Pengolahan Citra Digital

Pengolahan Citra Digital adalah serangkaian proses menganalisis data primer maupun sekunder untuk memperbaiki kualitas gambar dan melakukan pemilihan citra ciri yang optimal untuk tujuan analisis dan memodifikasi citra sesuai yang diinginkan. Dalam konteks klasifikasi jenis sampah organik dan non-organik, pengolahan citra dilakukan untuk memperoleh data yang siap dianalisis meliputi tahap; *resizing, segmentasi, normalisasi*.(Apel et al. 2019)

d. Ekstraksi Fitur

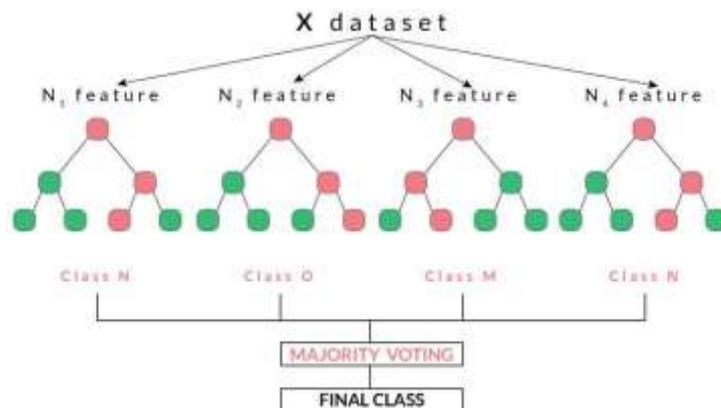
Tujuan utama ekstraksi fitur adalah untuk memperoleh informasi yang paling relevan dari data asli dan merepresentasikan informasi tersebut dalam

ruang dimensi yang lebih rendah. Fitur yang sering diekstrak antara lain fitur warna, fitur ukuran, fitur bentuk, dan fitur tekstur. Dimana setiap fitur yang diekstrak mempunyai perbedaan parameter untuk mendapatkan fitur tersebut.(Aziz 2021)

e. Model Random Forest

Random Forest adalah *algorithm machine learning* berbasis ensemble learning yang menggunakan banyak *decision tree* untuk melakukan klasifikasi atau *regresi*. *Random Forest* dibentuk dari kumpulan *decision tree* yang bekerja secara independen, lalu hasil klasifikasinya digabungkan melalui proses *voting*(untuk klasifikasi) atau rata-rata(untuk *regresi*). (Andi Muhammad Akbar, Muhammad Basri, and Wahyuddin Wahyuddin 2025)

Model *Random Forest* dilatih menggunakan data pelatihan yang telah disiapkan. *Algoritma Random Forest* digunakan karena kemampuan dalam menangani dataset yang kompleks dan memberikan hasil yang baik dalam klasifikasi. Pohon keputusan secara *ensemble* digunakan dalam model ini, di mana setiap pohon memberikan suara untuk keputusan akhir. (Dewi, Mulyadi, and Syafitri 2012)



Gambar 1 *Algoritma Random Forest*

f. Evaluasi Model

Setelah pelatihan, model yang dihasilkan dievaluasi menggunakan data pengujian yang terpisah. Metrik evaluasi seperti akurasi, *presisi*, *recall*, dan skor F1 digunakan untuk mengukur performan model dalam mengklasifikasi sampah organik dan non-organik. Selain itu, matriks konfusi juga digunakan untuk

memvisualisasikan hasil prediksi dan klasifikasi yang dilakukan oleh model.(Posangi, Yahya, and Wungguli 2023)

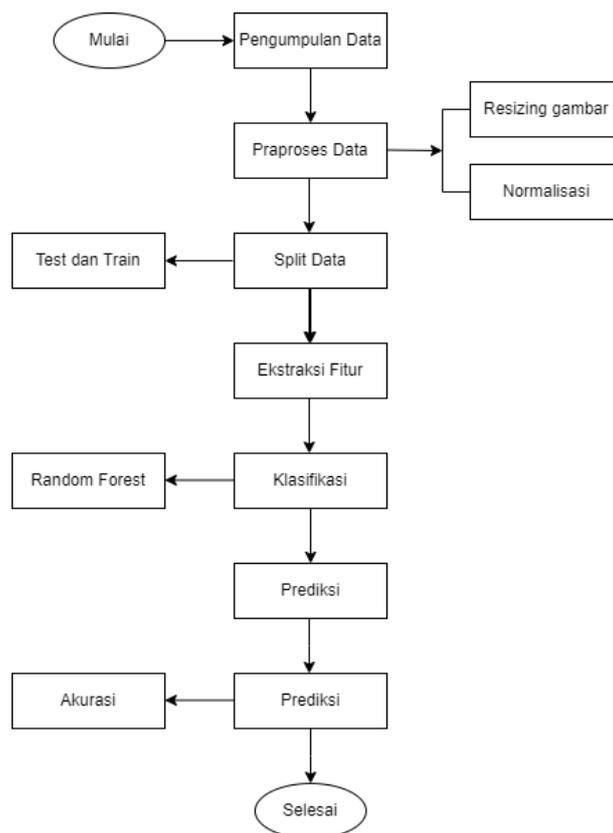
g. Penelitian terdahulu

1. (Fadli and Saputra 2023): Menggunakan model *Random Forest* untuk klasifikasi dan evaluasi untuk prediksi stroke, dengan akurasi tertinggi 93,6%.
2. (Posangi et al. 2023): Implementasi algoritma Random Forest dengan *Forward Selection* untuk klasifikasi indeks pembangunan manusia, dengan akurasi tertinggi 80%.
3. (Sahibu and Taufik 2024): Menggunakan algoritma CNN untuk mengklasifikasikan jenis sampah, dengan akurasi tertinggi 96%.
4. (Ashfania et al. 2022): Penggunaan algoritma *Random Forest* untuk klasifikasi berbasis kinerja efisiensi energi pada sistem pembangkit tenaga, dengan akurasi tertinggi 92,8%.
5. (Dewi et al. 2012): Penerapan Metode Random Forest dalam Driver analysis, dengan nilai rataannya sebesar 34,5%.

Berdasarkan teori dan penelitian terdahulu, penelitian ini dilakukan untuk menunjukkan bahwa algoritma Random Forest memiliki performa yang baik untuk klasifikasi gambar, penelitian ini dilakukan untuk mengimplementasikan model tersebut dalam mengklasifikasi citra sampah menjadi organik atau non-organik.

METODE PENELITIAN

Penelitian ini menggunakan metode Random Forest untuk mengklasifikasi jenis sampah dengan membangun sejumlah pohon keputusan (*decision tree*) selama proses pelatihan dan menghasilkan prediksi berdasarkan hasil voting atau rata-rata dari seluruh gambar. Berikut diagram alur penelitian:



Gambar 2. Metode Penelitian

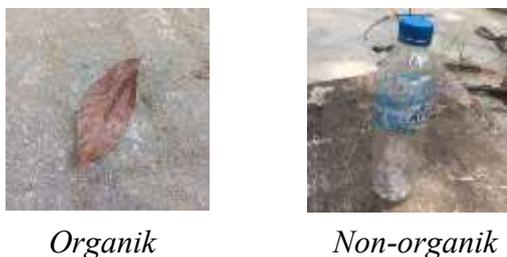
Penelitian ini menggunakan pendekatan kuantitatif eksperimen atau terapan dengan tujuan utama yaitu: (1) Mengumpulkan data berupa gambar sampah organik dan non-organik. Data ini bisa berasal dari dataset terbuka seperti *TrashNet*, *Kaggle*, atau hasil foto kamera peneliti. Pada penelitian ini peneliti menggunakan dataset dari *roboflow universe dataset*, (2) Praproses citra menggunakan tiga langkah yaitu; resizing, segmentasi dan normalisasi, (3) pembagian data menjadi data latih untuk melatih model dan data uji untuk mengevaluasi performa model, (4) ekstraksi fitur warna dan bentuk, (5) klasifikasi dengan model Random Forest, (6) setelah model dilatih, digunakan untuk

memprediksi jenis sampah pada data uji, (7) evaluasi hasil prediksi untuk mengetahui seberapa baik model dalam mengklasifikasi gambar sampah.(Kurniawan et al. 2023)

HASIL DAN PEMBAHASAN

1. Dataset Gambar Sampah Organik dan Non-organik

Penelitian ini dilaksanakan pada 14 Mei sampai 22 Juni 2025 di tempat pembuangan terakhir sampah (TPA) Riangkoli, Kabupaten Flores Timur. Karena sampah yang di TPA Riangkoli merupakan sampah yang sudah tercampur, sehingga peneliti menggunakan media perantara atau tidak secara langsung, dataset sampah yang diperoleh dari *roboflow universe dataset* untuk diproses selanjutnya. Dari data yang diperoleh jumlah total citra 1.310 gambar sampah, dengan rincian 281 gambar sampah organik, dan 1029 gambar sampah non-organik. Dari dua kategori citra sampah di beri label 0 dan 1, dimana sampah organik diberi label 0 dan sampah non-organik diberi label 1.



Gambar 3. Citra sampah

2. Praproses Citra

Pada tahap praproses citra ini bertujuan untuk mempersiapkan citra agar siap digunakan pada proses ekstraksi fitur dan klasifikasi model. Praproses citra yang dilakukan meliputi penyesuaian ukuran(*resizing*) menjadi dimensi seragam 100x100 piksel, sehingga data citra akan memiliki resolusi yang sama. Selain itu, dilakukan proses segmentasi citra didasarkan pada perbedaan derajat keabuan citra, untuk mengubah citra berwarna yang mempunyai nilai masing-masing r,g, dan b menjadi citra *grayscale*. Proses normalisasi juga dilakukan dengan memuat sebagai skala abu-abu untuk pemerataan *histogram* dan mengubah nilai intensitas piksel ke dalam rentang 0-1. Langkah ini penting untuk meminimalkan pengaruh

variasi pencahayaan pada data citra yang digunakan. Tahapan praproses citra dapat dilihat pada gambar berikut ini:



Gambar 4. Praproses citra

3. Ekstraksi Fitur

Setelah tahap praproses, dilakukan ekstraksi fitur warna dan tekstur sebagai klasifikasi jenis sampah. Fitur tekstur diekstraksi untuk semua gambar, lalu impor *library pandas*, buat *dataframe* dengan menggabungkan fitur warna dan tekstur ke dalam satu set data. Kombinasi fitur-fitur RGB dan *Gray-Level Co-occurrence Matrix (GLCM) texture*, dengan menghitung rata-rata RGB, mengukur perbedaan intensitas antara piksel dan tetangganya (*Contrast*), mengukur korelasi linear antar piksel dalam tekstur, nilai diantara -1 dan 1 (*Correlation*), keseragaman dalam tekstur (*Energy*), mengukur kesamaan piksel terhadap tetangganya (*Homogeneity*). Hasil ekstraksi fitur warna dan tekstur ditampilkan beberapa mewakili data yang lain pada tabel ini:

Tabel 1. Ekstraksi fitur warna dan tekstur

Filename	Label	Mean_R	Mean_G	Mean_B	Contrast	Correlation	Energy	Homogeneity
organik_daun_01.jpg	Organik	112.4	85.3	70.1	23.51	0.76	0.021	0.28
organik_kayu_02.jpg	Organik	128.6	97.5	78.9	19.32	0.80	0.027	0.31
nonorganik_kaca_03.jpg	Non-Organik	172.1	178.3	190.5	12.22	0.90	0.033	0.43
organik_kardus_04.jpg	Organik	140.3	110.1	90.4	17.87	0.79	0.024	0.35
nonorganik_kaleng_05.jpg	Non-Organik	165.9	160.3	155.0	13.67	0.85	0.030	0.41
organik_daun_06.jpg	Organik	104.2	78.4	63.2	26.45	0.72	0.019	0.26
nonorganik_botol_07.jpg	Non-Organik	150.7	170.6	185.3	11.94	0.91	0.034	0.45
organik_kayu_08.jpg	Organik	135.3	100.0	76.4	20.18	0.81	0.026	0.33
nonorganik_bungkus_09.jpg	Non-Organik	180.1	170.0	160.0	10.33	0.89	0.038	0.46
.....
organik_kardus_1310.jpg	Organik	138.0	107.5	85.6	18.12	0.78	0.025	0.32

4. Split Data

Pada penelitian ini, digunakan sebanyak 1.310 data citra sampah yang telah dikategorikan menjadi sampah organik dan non-organik. Berikut ini hasil dari split data untuk dibagi menjadi data latih 90% dan data uji 10% sebagai berikut:

```

from sklearn.model_selection import train_test_split
import numpy as np

# Misalnya X adalah data fitur (array atau list), dan y adalah
label-nya
# FIX: Define X and y with example data. Replace with your actual
data.
X = np.random.rand(1310, 10) # Example feature data
y = np.random.randint(0, 2, 1310) # Example label data

X_train, X_test, y_train, y_test = train_test_split(
    X, y,
    test_size=0.1,          # 1/10 = 10% data untuk testing
    random_state=42,

```

```

        stratify=y                # menjaga proporsi organik/non-organik
        tetap seimbang
    )

# Cek jumlah data hasil split
print(f"Jumlah data latih: {len(X_train)}") # Harusnya 1179
print(f"Jumlah data uji:   {len(X_test)}")  # Harusnya 131

Jumlah data latih: 1179
Jumlah data uji:   131

```

5. Implementasi Model *Random Forest*

Model *Random Forest* di latih menggunakan *RandomForestClassifier* karena peneliti ingin melakukan klasifikasi terhadap jenis sampah, langkah pertama dengan mengimpor modul *from sklearn.ensemble import RandomForestClassifier* untuk klasifikasi, selanjutnya buat bentuk objek *RandomForestClassifier* ini dengan parameter, parameter pertama (*n_estimators*) diberi nilai 100, *decision tree* yang akan di training berjumlah 100 sampel. Agar eksperimen ini dilakukan dengan hasil yang konsisten, maka dibutuhkan *random_state* atau *random_state number*, untuk klasifikasi ini *random_state=42*, objek yang dibentuk ini kita tampung ke dalam variabel model, untuk selanjutnya model ini akan di *training* dengan *model.fit(x_train, y_train)*, proses model akan cepat tergantung jumlah datasetnya. Hasil *RandomForestClassifier* (<http://localhost:8501/>) dapat dilihat sebagai berikut;



Gambar 5. Hasil prediksi

6. Evaluasi Model

Model Random Forest menggunakan data uji sebanyak 131 data, dengan metrik akurasi, presisi, recall, dan F1-score. Model mencapai akurasi sebesar 87%, yang berarti sekitar 114 dari 131 data uji berhasil diklasifikasikan dengan benar. Ini menunjukkan kemampuan model yang cukup kuat dalam membedakan kedua jenis sampah berdasarkan fitur warna dan tekstur.. Hasil laporan klasifikasi sebagai berikut:

=== Evaluasi Model Random Forest ===

Akurasi : 0.8702
 Presisi : 0.8824
 Recall : 0.8571
 F1-score : 0.8696

Laporan Klasifikasi:

	precision	Recall	f1-score	Support
Organic	0.86	0.88	0.87	64
Non-organik	0.88	0.86	0.87	67
Accuracy			0.87	131
macro avg	0.87	0.87	0.87	131
weighted avg	0.87	0.87	0.87	131

KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian yang telah dilakukan, maka peneliti menyimpulkan bahwa algoritma Random Forest mampu digunakan secara efektif untuk melakukan klasifikasi jenis sampah ke dalam dua kategori, yaitu organik dan non-organik, berdasarkan data citra sampah yang telah melalui tahapan preprocessing dan ekstraksi fitur. Proses praproses data seperti resizing gambar, normalisasi dan ekstraksi fitur berperan penting dalam meningkatkan kualitas data untuk pelatihan model klasifikasi. Model Random Forest yang dibangun mampu melakukan klasifikasi jenis sampah organik dan non-organik secara akurat dan seimbang. Dengan akurasi di atas 87%, model ini layak digunakan dalam aplikasi mobile untuk membantu sistem pemilahan sampah berbasis

citra, dengan potensi dikembangkan lebih lanjut agar dapat mengenali lebih banyak jenis sampah atau diterapkan dalam kondisi nyata.

DAFTAR REFERENSI

- Alden, Sita, and Betha Nurina Sari. 2023. "Implementasi Algoritma CNN Untuk Pemilahan Jenis Sampah Berbasis Android Dengan Metode CRISP-DM." *Jurnal Informatika* 10(1):62–71. doi: 10.31294/inf.v10i1.14985.
- Andi Muhammad Akbar, Muhammad Basri, and Wahyuddin Wahyuddin. 2025. "Implementasi Machine Learning Menggunakan Algoritma Klasifikasi Untuk Mendeteksi Jenis Sampah." *Jurnal Publikasi Manajemen Informatika* 3(3):152–61. doi: 10.55606/jupumi.v3i3.3751.
- Apel, Daun, Departemen Matematika, Fakultas Matematika, and Sains Data. 2019. "Penerapan Random Forest Untuk Mengukur Tingkat Keparahan Penyakit." 8(2).
- Ashfania, Ghiffari Awliya Muhammad, Toni Prahasto, Achmad Widodo, and Tarwaji Warsokusumo. 2022. "Penggunaan Algoritma Random Forest Untuk Klasifikasi Berbasis Kinerja Efisiensi Energi Pada Sistem Pembangkit Daya." *Rotasi* 24(3):14–21.
- Aziz, W. A. 2021. *Implementasi Metode Random Forest Pada Klasifikasi Data Ulasan Konsumen Perusahaan (Studi Kasus: Aplikasi Kai Access)*.
- Dewi, Nariswari Karina, Soni Yadi Mulyadi, and Utami Dyah Syafitri. 2012. "Penerapan Metode Random Forest Dalam Driver Analysis." *Forum Statistika Dan Komputasi* 16(1):35–43.
- Fadli, Muhamad, and Rizal Adi Saputra. 2023. "Klasifikasi Dan Evaluasi Performa Model Random Forest Untuk Prediksi Stroke." *JT: Jurnal Teknik* 12(2):72–80.
- Ibnul Rasidi, Abdurrahman, Yolanda Al Hidayah Pasaribu, Afzal Ziqri, and Faisal Dharma Adhinata. 2022. "Klasifikasi Sampah Organik Dan Non-Organik Menggunakan Convolutional Neural Network." *Jurnal Teknik Informatika Dan Sistem Informasi* 8(1):142–49. doi: 10.28932/jutisi.v8i1.4314.
- Imbarraga dan Reinaldi, Radhya Prakasa Imbarraga dan Harry. 2019. "Universitas Pasundan." *Kebudayaan* (022):1–47.

- Kurniawan, Ilham, Duwi Cahya Putri Buani, Abdussomad Abdussomad, Widya Apriliah, and Rizal Amegia Saputra. 2023. "Implementasi Algoritma Random Forest Untuk Menentukan Penerima Bantuan Raskin." *Jurnal Teknologi Informasi Dan Ilmu Komputer* 10(2):421–28. doi: 10.25126/jtiik.20231026225.
- Maulidah, Nurlaelatul, Mawadatul Maulidah, Riki Supriyadi, Hiya Nalatissifa, Sri Diantika, and Ahmad Fauzi. 2024. "Prediksi Kualitas Air Menggunakan Metode Random Forest, Decision Tree, Dan Gradient Boosting." *Jurnal Khatulistiwa Informatika* 12(1):1–6. doi: 10.31294/jki.v12i1.16004.
- Posangi, Tiara, Lailany Yahya, and Djihad Wungguli. 2023. "Implementasi Algoritma Random Forest Dengan Forward Selection Untuk Klasifikasi Indeks Pembangunan Manusia." *Jambura Journal of Probability and Statistics* 4(2):85–91. doi: 10.37905/jjps.v4i2.18460.
- Sahibu, Supriadi, and Imran Taufik. 2024. "Implementation of the Convolutional Neural Network Algorithm for Classifying Types of Organic and Non-Organic Waste Implementasi Algoritma Convolutional Neural Network Untuk Klasifikasi Jenis Sampah Organik Dan Non Organik." 4(July):840–52.
- Siburian, Vanissa Wanika, and Ika Elvina Mulyana. 2018. "Prediksi Harga Ponsel Menggunakan Metode Random Forest." *Annual Research Seminar (ARS) 2018* 4(1):144–47.