Perbandingan Kinerja Algoritma K-Nearest Neighbor dan Random Forest untuk Klasifikasi Citra Sampah Organik dan Anorganik Berbasis Web

Baitul Anam[1]*, Adi Susanto, M.[2], A. Hamdani, M. Kom

[1]Teknologi Informasi, Universitas Ibrahimy, Situbondo, Jawa Timur

*Korespondensi Penulis. E-mail: baitulanam190603@gmail.com , Telp: +6281553686203

Abstrak

Permasalahan pengelolaan sampah organik dan anorganik masih menjadi tantangan di berbagai daerah akibat rendahnya tingkat pemilahan sampah oleh masyarakat. Pemanfaatan teknologi kecerdasan buatan melalui machine learning dapat membantu proses klasifikasi sampah secara otomatis dan lebih efisien. Penelitian ini bertujuan untuk membandingkan kinerja algoritma K-Nearest Neighbor (KNN) dan Random Forest dalam klasifikasi sampah organik dan anorganik berbasis web. Penelitian ini menggunakan desain penelitian eksperimen dengan pendekatan kuantitatif. Dataset terdiri dari 2.513 citra sampah organik dan anorganik yang dibagi menjadi data latih sebanyak 2.010 data dan data uji sebanyak 503 data menggunakan rasio split 80:20. Teknik pengumpulan data dilakukan melalui dataset citra digital sampah. Variabel independen dalam penelitian ini adalah algoritma KNN dan Random Forest, sedangkan variabel dependen adalah hasil klasifikasi sampah. Analisis data dilakukan menggunakan confusion matrix dengan parameter accuracy, precision, recall, dan F1-score. Hasil penelitian menunjukkan bahwa algoritma Random Forest memiliki performa lebih baik dibandingkan algoritma KNN. Algoritma KNN memperoleh nilai accuracy sebesar 83.90%, precision 89.64%, recall 80.36%, dan F1-score 84.75%. Sementara itu, algoritma Random Forest memperoleh accuracy sebesar 86.48%, precision 85.10%, recall 91.79%, dan F1-score 88.32%. Hasil confusion matrix menunjukkan bahwa Random Forest mampu mengidentifikasi sampah organik dan anorganik dengan tingkat klasifikasi yang lebih stabil. Algoritma Random Forest memiliki performa klasifikasi yang lebih baik dibandingkan KNN dalam klasifikasi sampah organik dan anorganik berbasis web. Implementasi sistem berbasis web dapat membantu proses identifikasi sampah secara otomatis sehingga mendukung pengelolaan sampah yang lebih efektif.

Kata kunci: Klasifikasi Sampah, K-Nearest Neighbor, Random Forest, Machine Learning

Abstract

The problem of organic and inorganic waste management remains a challenge in various regions due to the low level of waste sorting by the community. The use of artificial intelligence technology through machine learning can help the waste classification process automatically and more efficiently. This study aims to compare the performance of the K-Nearest Neighbor (KNN) and Random Forest algorithms in web-based organic and inorganic waste classification. This study uses an experimental research design with a quantitative approach. The dataset consists of 2,513 images of organic and inorganic waste divided into 2,010 training data and 503 test data using a split ratio of 80:20. The data collection technique is carried out through a digital waste image dataset. The independent variables in this study are the KNN and Random Forest algorithms, while the dependent variable is the waste classification results. Data analysis was carried out using a confusion matrix with parameters of accuracy, precision, recall, and F1-score. The results show that the Random Forest algorithm has better performance than the KNN algorithm. The KNN algorithm achieved an accuracy of 83.90%, precision of 89.64%, recall of 80.36%, and F1-score of 84.75%. Meanwhile, the Random Forest algorithm achieved an accuracy of 86.48%, precision of 85.10%, recall of 91.79%, and F1-score of 88.32%. The confusion matrix results show that Random Forest is able to identify organic and inorganic waste with a more stable classification level. The Random Forest algorithm has better classification performance than KNN in web-based organic and inorganic

waste classification. The implementation of a web-based system can help the process of automatic waste identification, thereby supporting more effective waste management.

Keyword: Garbage Classification, K-Nearest Neighbor, Random Forest, Machine Learning

PENDAHULUAN

Permasalahan sampah menjadi salah satu isu lingkungan yang terus meningkat setiap tahunnya, terutama di wilayah perkotaan dan daerah dengan tingkat konsumsi masyarakat yang tinggi. Pertumbuhan jumlah penduduk, perkembangan industri, serta pola konsumsi masyarakat menyebabkan volume sampah organik dan anorganik mengalami peningkatan secara signifikan(Fatmawati, Dicky Firmansyah 2025:5). Sampah yang tidak dikelola dengan baik dapat menimbulkan berbagai dampak negatif terhadap lingkungan, seperti pencemaran tanah, pencemaran air, bau tidak sedap, hingga gangguan kesehatan masyarakat. Salah satu faktor utama penyebab permasalahan tersebut adalah rendahnya tingkat pemilahan sampah sejak awal proses pembuangan. Banyak masyarakat masih mencampurkan sampah organik dan anorganik sehingga proses pengolahan dan daur ulang menjadi kurang efektif(Utami et al. 2023:1108).

Pengelolaan sampah yang dilakukan secara manual memiliki berbagai keterbatasan, terutama dalam hal efisiensi waktu, tenaga, dan tingkat akurasi pemilahan. Proses pemilahan manual membutuhkan tenaga manusia dalam jumlah besar dan sering kali menghasilkan kesalahan klasifikasi akibat kurangnya konsistensi dalam identifikasi jenis sampah. Kondisi tersebut menyebabkan proses pengolahan sampah menjadi lebih lambat dan meningkatkan beban tempat pembuangan akhir. Oleh karena itu, diperlukan suatu sistem otomatis yang mampu membantu proses klasifikasi sampah secara cepat dan akurat untuk mendukung pengelolaan sampah yang lebih efektif dan berkelanjutan.

Perkembangan teknologi kecerdasan buatan atau Artificial Intelligence (AI), khususnya pada bidang machine learning, memberikan peluang besar dalam pengembangan sistem klasifikasi otomatis berbasis citra digital. Machine learning memungkinkan komputer untuk mempelajari pola tertentu dari data sehingga mampu melakukan proses klasifikasi tanpa harus diprogram secara eksplisit(Zaenuddiin dan Bani 2024:129). Teknologi ini telah banyak diterapkan dalam berbagai bidang seperti kesehatan, pendidikan, keamanan, dan pengolahan citra digital. Dalam konteks pengelolaan sampah, machine learning dapat dimanfaatkan untuk mengidentifikasi jenis sampah berdasarkan karakteristik visual seperti warna, bentuk, dan tekstur objek sehingga proses klasifikasi dapat dilakukan secara otomatis dan lebih efisien.

Salah satu algoritma machine learning yang sering digunakan dalam proses klasifikasi adalah K-Nearest Neighbor (KNN). Algoritma KNN bekerja dengan mengklasifikasikan data berdasarkan kedekatan jarak antara data baru dengan data latih yang telah ada sebelumnya. Algoritma ini dikenal sederhana, mudah diterapkan, dan memiliki kemampuan klasifikasi yang cukup baik pada berbagai jenis data. Menurut penelitian sebelumnya, KNN mampu memberikan hasil klasifikasi yang cukup efektif dalam berbagai kasus pengolahan data dan citra digital. Namun demikian, performa KNN sangat dipengaruhi oleh jumlah data, pemilihan nilai parameter K, serta kompleksitas dataset yang digunakan. Pada dataset dengan ukuran besar dan variasi citra yang tinggi, algoritma KNN cenderung membutuhkan waktu komputasi yang lebih besar(Santi et al. 2024:689).

Selain algoritma KNN, Random Forest juga menjadi salah satu algoritma klasifikasi yang banyak digunakan dalam penelitian machine learning karena memiliki tingkat akurasi dan stabilitas yang baik. Random Forest merupakan algoritma berbasis ensemble learning yang bekerja dengan membangun banyak decision tree kemudian menghasilkan prediksi berdasarkan hasil voting mayoritas dari seluruh pohon keputusan. Algoritma ini memiliki kemampuan dalam menangani data berdimensi tinggi serta mengurangi risiko overfitting(Riansah et al. 2025:4243).

Penelitian yang dilakukan oleh Melania Velisita Lau dkk. dengan judul "Implementasi Metode Random Forest untuk Klasifikasi Sampah Organik dan Anorganik Berbasis Citra Digital" membahas permasalahan pengelolaan sampah yang masih dilakukan secara manual sehingga membutuhkan waktu lama, biaya tinggi, dan bergantung pada tenaga manusia. Penelitian ini bertujuan untuk mengembangkan sistem klasifikasi sampah otomatis menggunakan algoritma Random Forest berbasis citra digital. Dataset yang digunakan berasal dari Kaggle sebanyak 1.000 citra yang terdiri dari 500 citra sampah organik dan 500 citra sampah anorganik dengan ukuran 64x64 piksel. Proses penelitian dilakukan melalui tahapan praproses citra, normalisasi, ekstraksi fitur menggunakan metode Histogram of Oriented Gradients (HOG), pembagian data 80% data latih dan 20% data uji, kemudian dilakukan pelatihan model Random Forest dengan 150 estimator. Evaluasi dilakukan menggunakan confusion matrix, accuracy, precision, recall, dan F1-score. Hasil penelitian menunjukkan bahwa algoritma Random Forest mampu mencapai tingkat akurasi sebesar 96% dengan nilai precision, recall, dan F1-score yang tinggi pada kedua kelas. Kesimpulan penelitian ini menyatakan bahwa Random Forest memiliki kemampuan generalisasi yang baik dalam mengklasifikasikan sampah organik dan anorganik berbasis citra digital sehingga berpotensi diterapkan dalam sistem pengelolaan sampah otomatis yang lebih efektif dan efisien(Lau et al. 2025:204).

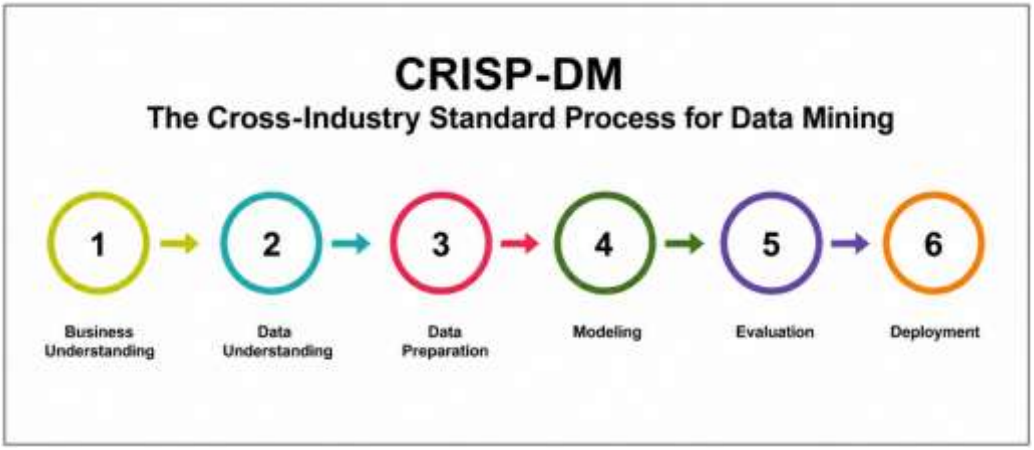
Penelitian berikutnya dilakukan oleh Elza Sahelvi dkk. dengan judul "Perbandingan Algoritma K-Nearest Neighbors dan Random Forest untuk Rekomendasi Gaya Hidup Sehat dalam Mencegah Penyakit Jantung". Penelitian ini dilatarbelakangi oleh tingginya angka penyakit jantung yang menjadi salah satu penyebab utama kematian di dunia akibat pola hidup tidak sehat. Penelitian bertujuan membandingkan performa algoritma K-Nearest Neighbors (KNN) dan Random Forest (RF) dalam memberikan rekomendasi gaya hidup sehat berbasis data kesehatan. Dataset yang digunakan terdiri dari 1.025 data dengan 14 fitur kesehatan seperti usia, tekanan darah, kadar kolesterol, dan jenis nyeri dada. Tahapan penelitian meliputi exploratory data analysis, preprocessing data berupa normalisasi dan feature selection, pembagian data dengan rasio 80:20 dan 70:30, kemudian dilakukan pemodelan menggunakan algoritma KNN dan Random Forest. Evaluasi dilakukan menggunakan accuracy, precision, recall, F1-score, serta confusion matrix. Hasil penelitian menunjukkan bahwa Random Forest memiliki performa yang lebih baik dibandingkan KNN dengan tingkat akurasi sebesar 99% pada rasio 80:20 dan 98% pada rasio 70:30, sedangkan KNN memperoleh akurasi sebesar 83% dan 86%. Kesimpulan penelitian menyatakan bahwa algoritma Random Forest lebih efektif, stabil, dan akurat dalam memberikan rekomendasi gaya hidup sehat untuk mencegah penyakit jantung dibandingkan algoritma KNN(Sahelvi et al. 2025:838).

Penelitian yang dilakukan oleh Salmon dkk. dengan judul "Perbandingan Kinerja Algoritma K-Nearest Neighbor dan Algoritma Random Forest Untuk Klasifikasi Data Mining Pada Penyakit Gagal Ginjal" membahas pentingnya diagnosis dini penyakit gagal ginjal yang menjadi salah satu penyakit kronis dengan tingkat komplikasi tinggi. Penelitian ini memanfaatkan teknologi data mining untuk membantu proses klasifikasi data medis menggunakan algoritma KNN dan Random Forest. Penelitian dilakukan dengan pendekatan klasifikasi data mining melalui beberapa tahapan seperti pengumpulan data, preprocessing, pembagian data training dan testing, serta pengujian model menggunakan confusion matrix. Algoritma KNN digunakan dengan konsep kedekatan jarak Euclidean Distance, sedangkan Random Forest menggunakan metode ensemble learning dengan banyak decision tree untuk menghasilkan klasifikasi yang lebih stabil. Evaluasi model dilakukan menggunakan accuracy, precision, recall, dan F1-score. Hasil penelitian menunjukkan bahwa kedua algoritma memiliki tingkat akurasi di atas 90%, namun Random Forest memberikan performa terbaik dengan tingkat akurasi mencapai 99,75%. Kesimpulan penelitian menyatakan bahwa Random Forest memiliki kinerja yang lebih baik dibandingkan KNN dalam klasifikasi penyakit gagal ginjal sehingga model yang dihasilkan dapat digunakan untuk membantu proses diagnosis penyakit secara lebih efektif dan akurat(Salmon, Azahari, dan Ekawati 2024: 1952).

Namun demikian, penelitian yang secara khusus membandingkan performa kedua algoritma tersebut pada klasifikasi sampah organik dan anorganik berbasis web masih relatif terbatas bahkan belum ada. Berdasarkan permasalahan tersebut, penelitian ini dilakukan bertujuan untuk membandingkan kinerja algoritma K-Nearest Neighbor dan Random Forest dalam klasifikasi sampah organik dan anorganik berbasis web. Sistem dikembangkan menggunakan dataset citra sampah digital dan diimplementasikan dalam bentuk aplikasi berbasis web sehingga dapat mempermudah proses analisis dan pengujian model klasifikasi. Evaluasi performa algoritma dilakukan menggunakan parameter accuracy, precision, recall, dan F1-score untuk mengetahui algoritma yang memiliki performa terbaik dalam proses klasifikasi sampah. Penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan teknologi pengelolaan sampah berbasis kecerdasan buatan yang lebih efektif, efisien, dan mudah diterapkan di lingkungan masyarakat maupun instansi terkait.

METODE

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen dan model pengembangan CRISP-DM (Cross Industry Standard Process for Data Mining). Metode CRISP-DM digunakan karena memiliki tahapan sistematis dalam proses pengolahan data dan pengembangan model machine learning. Tahapan penelitian terdiri dari Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, dan Deployment(Singgalen 2023:238).



Gambar 1. Metode CRISP-DM

Business Understanding

Tahap ini dilakukan untuk mengidentifikasi permasalahan penelitian terkait rendahnya efektivitas pemilahan sampah organik dan anorganik secara manual. Penelitian bertujuan untuk membandingkan performa algoritma K-Nearest Neighbor (KNN) dan Random Forest dalam klasifikasi sampah berbasis web menggunakan parameter Accuracy, Precision, Recall, dan F1-Score.

1. Data Understanding

Dataset yang digunakan dalam penelitian ini berasal dari platform Kaggle berupa dataset citra sampah organik dan anorganik. Dataset terdiri dari 2.513 citra digital sampah organik dan sampah anorganik yang dibagi menjadi data latih sebanyak 2.010 data dan data uji sebanyak 503 data menggunakan rasio 80:20. Pada tahap ini dilakukan identifikasi jumlah data, distribusi kelas, format citra, dan pemeriksaan kualitas dataset. Sampling data dapat dilihat pada gambar dibawah ini

2. Data Preparation

Tahap persiapan data dilakukan sebelum proses pelatihan model. Proses ini meliputi resize citra, normalisasi nilai piksel, pembersihan data rusak, serta ekstraksi fitur menggunakan metode

Histogram of Oriented Gradients (HOG). Dataset kemudian dibagi menjadi data latih sebesar 80% dan data uji sebesar 20%.

3. Modeling

Tahap modeling dilakukan menggunakan algoritma K-Nearest Neighbor (KNN) dan Random Forest. Proses pelatihan model dilakukan menggunakan library Scikit-learn pada bahasa pemrograman Python. Pada algoritma KNN dilakukan pengujian nilai parameter K, sedangkan pada Random Forest dilakukan pengaturan jumlah decision tree (n_estimators) untuk memperoleh performa model terbaik.

4. Evaluation

Evaluasi model dilakukan menggunakan confusion matrix dengan parameter Accuracy, Precision, Recall, dan F1-Score. Hasil evaluasi digunakan untuk membandingkan performa kedua algoritma dalam mengklasifikasikan sampah organik dan anorganik.

5. Deployment

Tahap deployment dilakukan dengan mengimplementasikan model klasifikasi ke dalam sistem berbasis web menggunakan framework Flask dan Streamlit. Sistem memungkinkan pengguna melakukan upload citra sampah dan memperoleh hasil klasifikasi secara otomatis beserta visualisasi hasil evaluasi model.

HASIL DAN PEMBAHASAN

Contoh Citra Dataset per Kelas

Pada tahap awal penelitian, dilakukan observasi terhadap dataset citra sampah organik dan anorganik untuk memastikan kualitas data yang digunakan dalam proses pelatihan model. Dataset yang digunakan berasal dari Kaggle dan terdiri dari dua kelas utama, yaitu sampah organik dan sampah anorganik. Setiap citra memiliki karakteristik visual yang berbeda berdasarkan jenis sampahnya. Citra sampah organik umumnya menampilkan objek seperti sisa makanan, kulit buah, sisa sayuran, dan bahan alami lainnya yang memiliki tekstur tidak beraturan serta warna yang cenderung beragam. Sedangkan citra sampah anorganik terdiri dari objek seperti plastik, kaleng, sedotan, tisu, dan bahan non-organik lainnya yang memiliki bentuk lebih tetap dan tekstur lebih sederhana.

Tahap observasi citra dilakukan untuk memastikan bahwa dataset memiliki representasi visual yang baik dan seimbang pada setiap kelas. Selain itu, proses ini juga membantu sistem dalam mengenali pola visual sebelum dilakukan ekstraksi fitur menggunakan metode Histogram of Oriented Gradients (HOG). Variasi bentuk, warna, dan tekstur pada citra menjadi faktor penting dalam meningkatkan kemampuan model klasifikasi.

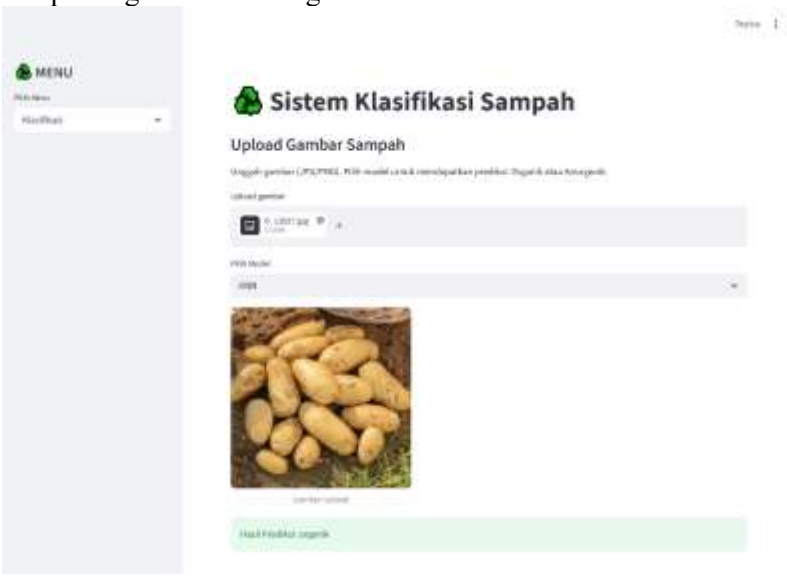
Gambar 2. Contoh Citra Sampah Organik

Gambar 3. Contoh Citra Sampah Anorganik

Implementasi Sistem Klasifikasi Sampah Berbasis Web

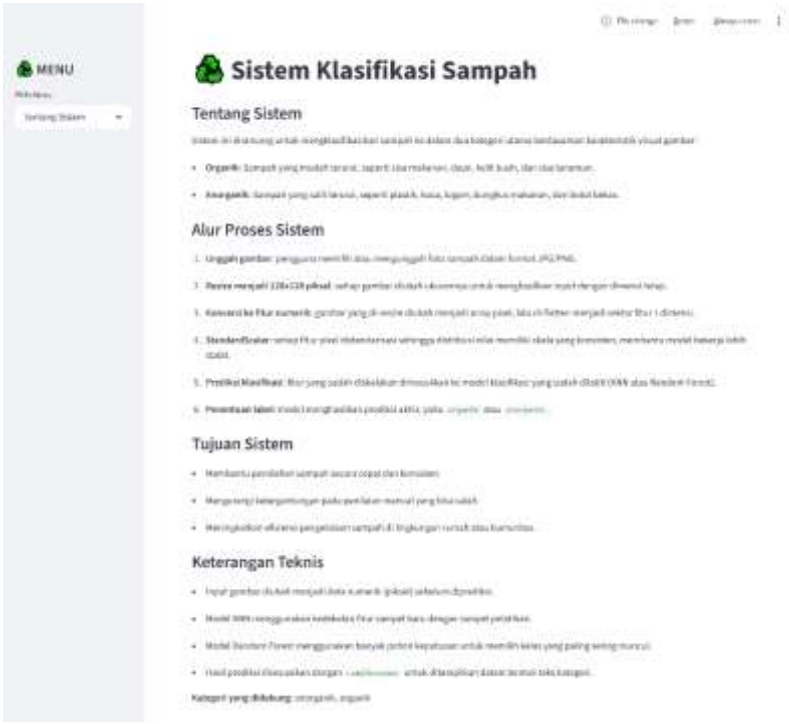
Sistem klasifikasi sampah dikembangkan berbasis web menggunakan framework Streamlit. Sistem dirancang untuk memudahkan pengguna dalam melakukan klasifikasi sampah secara otomatis berdasarkan citra yang diunggah.

Pada halaman klasifikasi, pengguna dapat mengunggah gambar sampah dalam format JPG atau PNG, kemudian memilih model klasifikasi yang akan digunakan, yaitu K-Nearest Neighbor (KNN) atau Random Forest. Setelah gambar diproses, sistem akan menampilkan hasil prediksi berupa kategori sampah organik atau anorganik.



Gambar 4. Halaman Klasifikasi Sistem

Selain halaman klasifikasi, sistem juga menyediakan halaman penjelasan sistem yang berisi informasi mengenai kategori sampah, alur proses klasifikasi, tujuan sistem, serta penjelasan teknis terkait penggunaan algoritma machine learning. Halaman ini bertujuan untuk membantu pengguna memahami proses kerja sistem secara keseluruhan.



Gambar 5 Halaman Penjelasan Sistem

Proses Pengolahan dan Pelatihan Model

Sebelum dilakukan proses klasifikasi, seluruh citra melewati tahap preprocessing yang meliputi resize gambar menjadi ukuran 128×128 piksel, normalisasi data, dan konversi citra menjadi data numerik. Selanjutnya dilakukan ekstraksi fitur menggunakan metode Histogram of Oriented Gradients (HOG) untuk menangkap pola bentuk, tekstur, dan tepi objek pada citra.

Dataset yang telah diproses kemudian dibagi menjadi data latih sebesar 80% dan data uji sebesar 20%. Total dataset yang digunakan sebanyak 2.513 citra, terdiri dari 2.010 data latih dan 503 data uji. Tahap pelatihan model dilakukan menggunakan algoritma K-Nearest Neighbor (KNN) dan Random Forest dengan menggunakan framework streamlit pada bahasa pemrograman Python. Pada algoritma Random Forest digunakan beberapa decision tree untuk meningkatkan stabilitas prediksi, sedangkan KNN melakukan klasifikasi berdasarkan kedekatan jarak antar data.

Evaluasi Model Klasifikasi

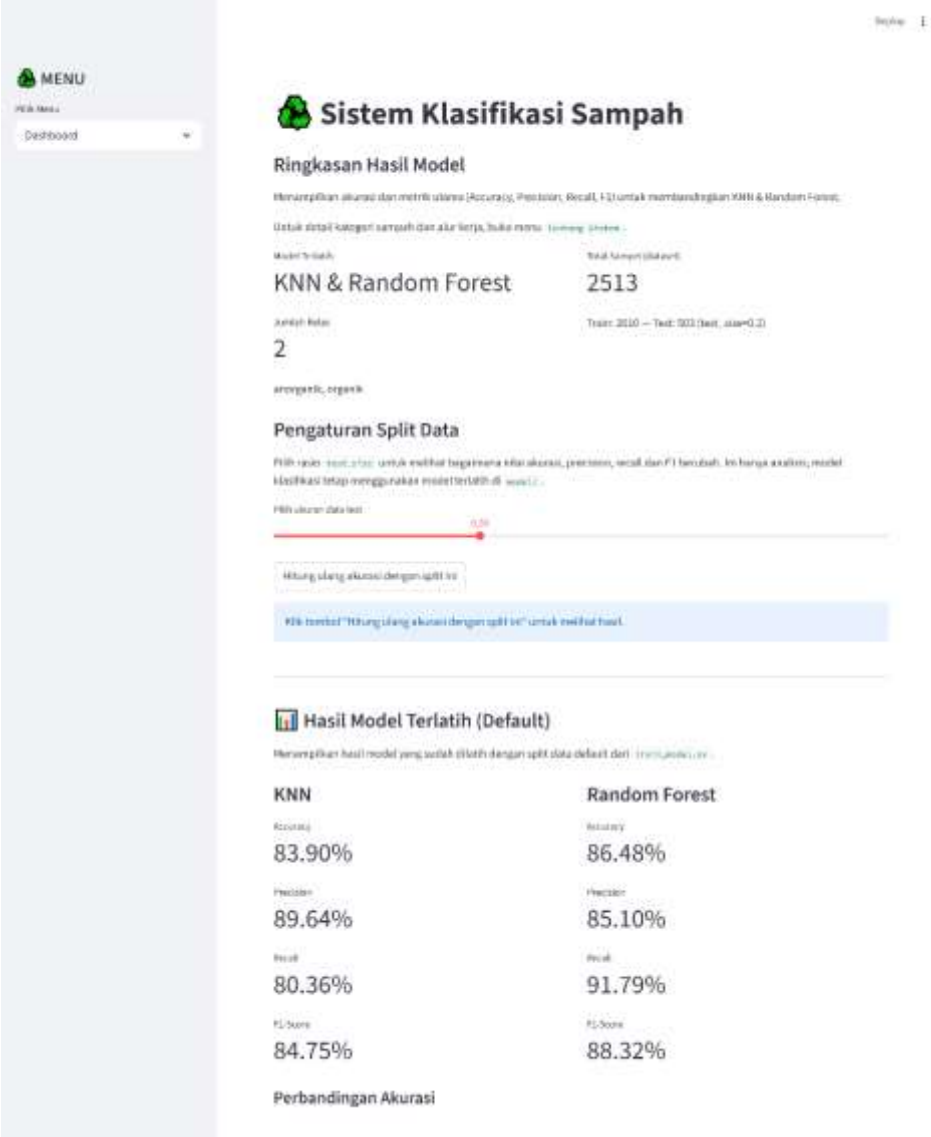
Tahap evaluasi dilakukan untuk mengetahui performa model dalam mengklasifikasikan sampah organik dan anorganik. Evaluasi menggunakan parameter Accuracy, Precision, Recall, dan F1-Score berdasarkan confusion matrix.

Hasil pengujian menunjukkan bahwa algoritma Random Forest memiliki performa yang lebih baik dibandingkan KNN. Algoritma KNN memperoleh nilai accuracy sebesar 83.90%, precision sebesar 89.64%, recall sebesar 80.36%, dan F1-score sebesar 84.75%. Sedangkan algoritma Random Forest memperoleh nilai accuracy sebesar 86.48%, precision sebesar 85.10%, recall sebesar 91.79%, dan F1-score sebesar 88.32%.

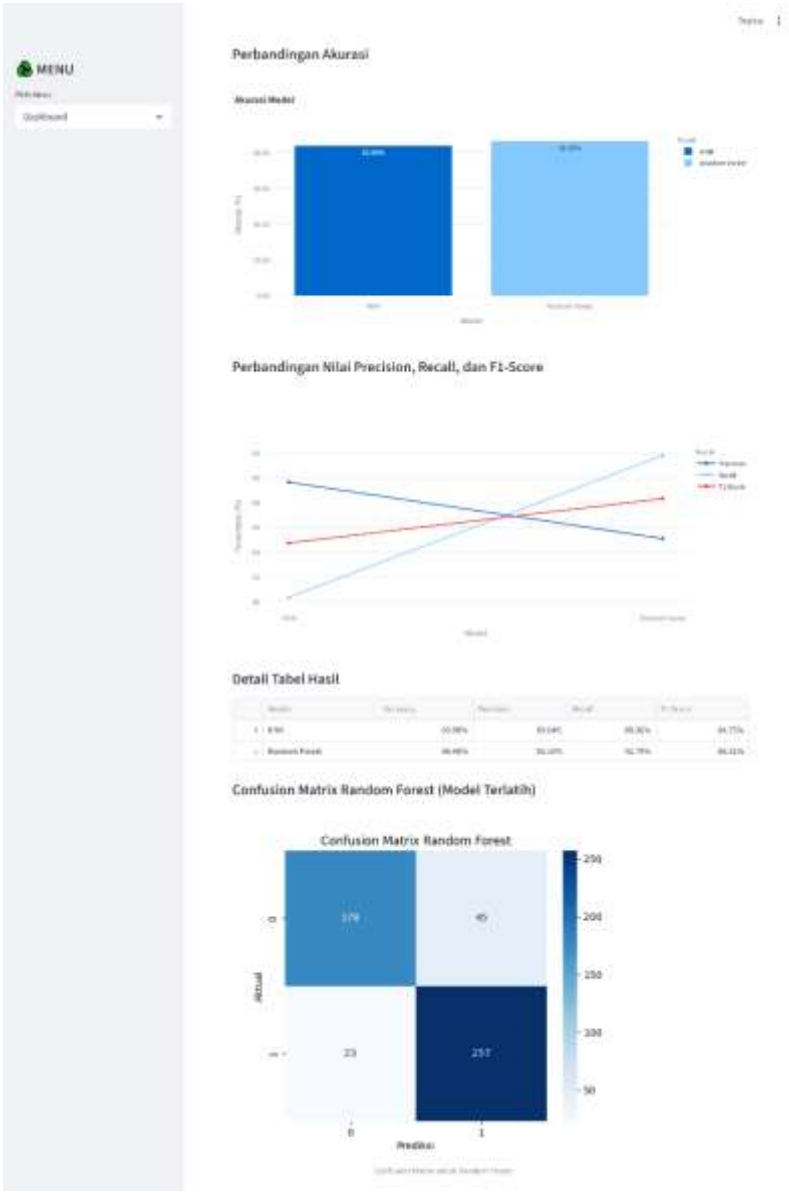
Table 1 Hasil Evaluasi Model Klasifikasi

Model	Accuracy	Precision	Recall	F1-Score
KNN	83.90%	89.64%	80.36%	84.75%
Random Forest	86.48%	85.10%	91.79%	88.32%

Hasil evaluasi divisualisasikan dalam bentuk grafik perbandingan accuracy, precision, recall, dan F1-score untuk mempermudah analisis performa masing-masing algoritma. Selain itu, confusion matrix digunakan untuk mengetahui jumlah prediksi benar dan salah pada setiap kelas.



Gambar 6. Dashboard Hasil Evaluasi Model



Gambar 7. Grafik Perbandingan dan Confusion Matrix

Berdasarkan confusion matrix Random Forest, model berhasil memprediksi 178 data kelas negatif dan 257 data kelas positif secara benar. Sedangkan kesalahan klasifikasi terjadi pada 45 data false positive dan 23 data false negative. Hasil tersebut menunjukkan bahwa model memiliki kemampuan klasifikasi yang cukup baik dan stabil terhadap dataset yang digunakan.

Analisis Kesalahan Klasifikasi

Meskipun Random Forest menunjukkan performa yang lebih baik dibandingkan KNN, masih terdapat beberapa kesalahan klasifikasi pada proses pengujian. Kesalahan prediksi umumnya terjadi pada citra yang memiliki kemiripan warna, bentuk, dan tekstur antara sampah organik dan anorganik.

Selain itu, faktor pencahayaan, kualitas gambar, latar belakang yang tidak seragam, dan ukuran objek yang terlalu kecil juga memengaruhi hasil klasifikasi. Pada beberapa kasus, objek sampah tidak

terfokus dengan baik sehingga sistem mengalami kesulitan dalam mengenali pola visual utama dari citra.

Namun secara keseluruhan, model Random Forest mampu menunjukkan performa klasifikasi yang stabil dan tidak mengalami overfitting secara signifikan. Hal ini terlihat dari hasil evaluasi yang cukup konsisten antara data latih dan data uji.

Keunggulan Random Forest disebabkan karena algoritma ini menggunakan banyak decision tree yang bekerja secara ensemble sehingga mampu meningkatkan stabilitas prediksi dan mengurangi risiko overfitting. Selain itu, kombinasi metode Histogram of Oriented Gradients (HOG) dan Random Forest mampu menangkap pola visual penting seperti bentuk, tekstur, dan tepi objek pada citra sampah.

Sementara itu, algoritma KNN menunjukkan nilai precision yang lebih tinggi dibandingkan Random Forest. Hal ini menunjukkan bahwa KNN cukup baik dalam meminimalkan kesalahan prediksi positif. Namun, nilai recall yang lebih rendah menunjukkan bahwa model masih mengalami kesulitan dalam mengenali beberapa data sampah secara optimal.

SIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa algoritma Random Forest memiliki performa yang lebih baik dibandingkan algoritma K-Nearest Neighbor (KNN) dalam klasifikasi sampah organik dan anorganik berbasis web. Hal ini dibuktikan dengan nilai accuracy Random Forest sebesar 86.48% yang lebih tinggi dibandingkan KNN sebesar 83.90%, serta nilai recall dan F1-score yang menunjukkan kemampuan klasifikasi yang lebih stabil dan seimbang pada kedua kelas sampah. Kombinasi metode Histogram of Oriented Gradients (HOG) dan Random Forest juga mampu menangkap karakteristik visual citra seperti bentuk, tekstur, dan tepi objek dengan baik sehingga meningkatkan performa model dalam proses klasifikasi. Selain itu, implementasi sistem berbasis web mampu membantu proses klasifikasi sampah secara otomatis dan interaktif sehingga mempermudah pengguna dalam melakukan identifikasi sampah organik dan anorganik. Berdasarkan hasil tersebut, algoritma Random Forest dapat dijadikan alternatif yang efektif dalam pengembangan sistem klasifikasi sampah berbasis machine learning. Untuk penelitian selanjutnya, performa sistem dapat ditingkatkan dengan menambah jumlah dataset, melakukan augmentasi citra, serta menerapkan metode deep learning agar hasil klasifikasi menjadi lebih optimal dan akurat.

DAFTAR PUSTAKA

Fatmawati, Dicky Firmansyah, Ade Kamaludin dkk. 2025. *BUKU SAKU OPTIMALISASI SISTEM PENELOLAAN SAMPAH PERKOTAAN*. indonesia: kms.kemkes.go.id.

Lau, Melania Velisita, Aryani Dwi Handayani, Miyosagusti Yudo Widodo, Rani Irma Handayani, dan Risca Lusiana Pratiwi. 2025. "Implementasi Metode Random Forest untuk Klasifikasi Sampah Organik dan Anorganik Berbasis Citra Digital." *RIGGS: Journal of Artificial Intelligence and Digital Business* 4(4):402. doi: 10.31004/riggs.v4i4.3410.

Riansah, Adam, Odi Nurdiawan, Ruli Herdiana, Sistem Informasi, Manajemen Informatika, Teknik Informatika, Kota Cirebon, Random Forest, Decision Tree, dan Prediksi Penjualan. 2025. "PENERAPAN ALGORITMA RANDOM FOREST DAN DECISION TREE UNTUK MENINGKATKAN AKURASI KLASIFIKASI PENJUALAN PADA TOKO BANGUNAN." *JATI (Jurnal Mahasiswa Teknik Informatika)* 9(3):4243.

Sahelvi, Elza, Putri Cikita, Riska Mela Sapitri, Rahmaddeni Rahmaddeni, dan Lusiana Efrizoni. 2025. "Perbandingan Algoritma K-Nearest Neighbors dan Random Forest untuk Rekomendasi Gaya Hidup Sehat dalam Mencegah Penyakit Jantung." *MALCOM: Indonesian Journal of Machine Learning and Computer Science* 5(3):838. doi: 10.57152/malcom.v5i3.1972.

Salmon, Salmon, Azahari Azahari, dan Hanifah Ekawati. 2024. "Perbandingan Kinerja Algoritma

K-Nearest Neighbor dan Algoritma Random Forest Untuk Klasifikasi Data Mining Pada Penyakit Gagal Ginjal." *Building of Informatics, Technology and Science (BITS)* 6(3):1952. doi: 10.47065/bits.v6i3.6476.

Santi, Santi, Cucut Susanto, Muhardi Muhardi, Madyana Patasik, dan Nurlina Nurlina. 2024. "Penerapan Algoritma K-Nearest Neighbor (KNN) Dalam Pengklasifikasian Tingkat Kematangan Buah Nangka Berdasarkan Citra Warna Kulit." *Digital Transformation Technology* 4(1):689. doi: 10.47709/digitech.v4i1.4550.

Singgalen, Yerik Afrianto. 2023. "Penerapan CRISP-DM dalam Klasifikasi Sentimen dan Analisis Perilaku Pembelian Layanan Akomodasi Hotel Berbasis Algoritma Decision Tree (DT)." *Jurnal Sistem Komputer dan Informatika (JSON)* 5(2):238. doi: 10.30865/json.v5i2.7081.

Utami, Ajeng Putri, Universitas Islam, Negeri Sumatera, Nafisah Nur, Addini Pane, Universitas Islam, Negeri Sumatera, Abdurrozzaq Hasibuan, Universitas Islam, dan Sumatera Utara. 2023. "ANALISIS DAMPAK LIMBAH / SAMPAH RUMAH TANGGA." *IAI SAMBAS* 6(2):1108.

Zaenuddiin, Imam, dan Riyan Bani. 2024. "Perkembangan Kecerdasan Buatan (AI) Dan Dampaknya Pada Dunia Teknologi." *JITU: Jurnal Informatika Utama* 2(2):129.