

# Evaluasi Kinerja Algoritma Berbasis Decision Tree pada Dataset Skala Kecil: Studi Perbandingan Analisis Komparatif

Arie Setya Putra<sup>1</sup>, Ochi Marshella Febriani<sup>2</sup>, Agus Wantoro<sup>3\*</sup>

<sup>1</sup>Teknologi Informasi, Jl. ZA Pagar Alam No.7, Bandar Lampung, Indonesia

<sup>2</sup>Sistem Informasi, Jl. ZA Pagar Alam No.93, Bandar Lampung, Indonesia

<sup>3</sup>Department of Informatics Engineering, Faculty of Technology and Informatics, Aisyah University, Indonesia

## Info Artikel

### Riwayat Artikel:

Received November 05, 2025

Revised January 29, 2026

Accepted January 29, 2026

### \*Corresponding Author:

Agus Wantoro

Email:

[aguswantoro@aisyahuniversity.ac.id](mailto:aguswantoro@aisyahuniversity.ac.id)



This is an open access article under the [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/) license.

**Abstract** – The development of Machine Learning (ML) has made significant contributions in various fields, particularly in the data classification process. One group of widely used algorithms is decision tree-based algorithms, such as J48, Random Forest, and Random Tree. Although these algorithms have been widely researched, most studies have been conducted on large datasets, while the conditions for using small datasets are still rarely explored. In fact, in real practice, data limitations are often a challenge, especially in the fields of health, social, and experimental research. This study aims to run three decision tree algorithms, namely J48, Random Forest, and Random Tree on small datasets. The evaluation was conducted using a student assessment dataset with a total of 100 records. The 5-fold cross-validation method was used to obtain more stable evaluation results. Algorithm performance was compared based on metrics of accuracy, precision, recall, F1-score, and computation time. The results show that Random Forest tends to provide better accuracy and stability than J48 and Random Tree, despite requiring relatively higher computation time. Meanwhile, J48 has advantages in interpretability, while Random Tree offers cruelty with lower complexity and better computation time. This finding emphasizes the importance of selecting the right algorithm according to the dataset size and research objectives.

**Keywords:** Machine learning, J48, Random Forest, Random Tree, Small Dataset

**Abstrak** – Perkembangan Machine Learning (ML) telah memberikan kontribusi signifikan dalam berbagai bidang, khususnya pada proses klasifikasi data. Salah satu kelompok algoritma yang banyak digunakan adalah algoritma berbasis decision tree, seperti J48, Random Forest, dan Random Tree. Meskipun algoritma tersebut telah banyak diteliti, sebagian besar studi dilakukan pada dataset berukuran besar, sementara kondisi penggunaan dataset kecil masih jarang dieksplorasi. Padahal, dalam praktik nyata, keterbatasan data sering menjadi tantangan, terutama pada bidang penelitian kesehatan, sosial, maupun eksperimen laboratorium. Penelitian ini bertujuan untuk mengevaluasi kinerja tiga algoritma decision tree, yaitu J48, Random Forest, dan Random Tree pada dataset kecil. Evaluasi dilakukan dengan menggunakan dataset penilaian mahasiswa dengan jumlah record dari 100. Metode 5-fold cross validation digunakan untuk memperoleh hasil evaluasi yang lebih stabil. Kinerja algoritma dibandingkan berdasarkan metrik akurasi, presisi, recall, F1-score, dan waktu komputasi. Hasil penelitian menunjukkan bahwa Random Forest cenderung memberikan akurasi dan stabilitas yang lebih baik dibandingkan J48 dan Random Tree, meskipun membutuhkan waktu komputasi yang relatif lebih tinggi. Sementara itu, J48 memiliki keunggulan dalam interpretabilitas, sedangkan Random Tree menawarkan fleksibilitas dengan kompleksitas yang lebih rendah dan waktu komputasi yang lebih baik. Temuan ini menegaskan pentingnya pemilihan algoritma yang tepat sesuai dengan ukuran dataset dan tujuan penelitian.

**Kata Kunci:** Machine learning, J48, Random Forest, Random Tree, Dataset Kecil

## I. PENDAHULUAN

Perkembangan teknologi informasi yang semakin pesat dalam satu dekade terakhir telah mendorong lahirnya berbagai metode analisis data yang lebih canggih dan adaptif [1]. Salah satu bidang yang mengalami perkembangan signifikan adalah *Machine Learning* (ML), yang berfungsi untuk membangun model prediksi maupun klasifikasi berdasarkan data historis [2]. Penerapan ML telah menjangkau berbagai disiplin ilmu, mulai dari kesehatan, pendidikan, keuangan, hingga keamanan siber [3]. Hal ini menunjukkan bahwa kemampuan algoritma pembelajaran mesin dalam mengekstraksi pola dari data memiliki nilai yang sangat penting dalam mendukung pengambilan keputusan yang lebih tepat dan efisien [4].

Namun, dalam praktiknya, ketersediaan data sering kali menjadi tantangan utama [5]. Tidak semua domain penelitian memiliki data dalam jumlah besar yang dapat digunakan untuk melatih model ML. Pada kondisi tertentu, peneliti harus bekerja dengan dataset berukuran kecil yang memiliki keterbatasan jumlah sampel maupun atribut. Penggunaan dataset kecil sering menimbulkan permasalahan serius, seperti risiko *overfitting* [6], kurangnya kemampuan generalisasi model, serta keterbatasan dalam menangkap keragaman pola yang sesungguhnya. Hal ini menjadikan evaluasi kinerja algoritma pada dataset skala kecil sebagai aspek penting untuk diperhatikan [7].

Salah satu kelompok algoritma yang banyak digunakan dalam klasifikasi adalah algoritma berbasis Decision Tree. Algoritma ini populer karena mampu menghasilkan model yang mudah dipahami, memiliki *interpretabilitas* tinggi, dan dapat diaplikasikan dalam berbagai konteks penelitian [1]. Beberapa algoritma turunan decision tree yang banyak digunakan adalah J48, Random Forest, dan Random Tree[2]. Masing-masing algoritma memiliki karakteristik yang berbeda. J48 dikenal dengan kesederhanaannya, Random Forest memiliki performa yang stabil karena menggunakan pendekatan *ensemble*, sedangkan Random Tree menawarkan fleksibilitas dengan membangkitkan pohon secara acak [8].

Meskipun telah banyak penelitian yang membandingkan algoritma-algoritma tersebut, sebagian besar fokus dilakukan pada dataset dengan ukuran besar. Penelitian dengan konteks dataset kecil masih relatif terbatas, padahal dalam praktiknya kondisi ini cukup sering dijumpai, terutama di bidang penelitian yang menghadapi keterbatasan data, seperti kesehatan, sosial, atau riset lapangan dengan sumber daya terbatas. Oleh karena itu, penting dilakukan studi evaluasi yang secara khusus menyoroti performa algoritma berbasis Decision Tree pada dataset dengan ukuran terbatas.

Berdasarkan uraian tersebut, penelitian ini difokuskan pada evaluasi kinerja tiga algoritma pohon keputusan, yaitu J48, Random Forest, dan Random Tree, pada dataset skala kecil. Penelitian ini bertujuan untuk membandingkan kinerja masing-masing algoritma berdasarkan metrik evaluasi seperti akurasi, presisi, recall, F1-score, serta waktu komputasi. Dengan adanya penelitian ini, diharapkan dapat diperoleh gambaran yang lebih jelas mengenai kelebihan dan keterbatasan setiap algoritma ketika dihadapkan pada kondisi dataset kecil, sekaligus memberikan rekomendasi bagi peneliti dan praktisi dalam memilih metode klasifikasi yang tepat.

## II. METODE

### A. Jenis Penelitian

Penelitian ini merupakan penelitian kuantitatif dengan pendekatan eksperimen komparatif. Pendekatan ini dipilih untuk membandingkan performa tiga algoritma berbasis Decision Tree seperti (J48, Random Forest, dan Random Tree) pada dataset berukuran kecil. Evaluasi dilakukan dengan mengukur kinerja model berdasarkan sejumlah metrik evaluasi, termasuk akurasi, presisi, recall, F1-score, dan waktu komputasi.

### B. Dataset

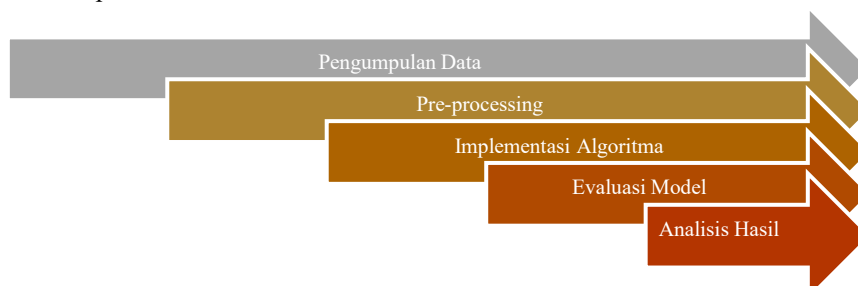
Dataset yang digunakan dalam penelitian ini adalah dataset nilai yang memiliki skala kecil dengan jumlah sampel 100 *record*. Dataset diambil dari Sistem Informasi Akademik (SIKAD) Universitas Aisyah Pringsewu. Pemilihan dataset dilakukan dengan mempertimbangkan keragaman atribut serta jumlah sampel yang sesuai dengan konteks dataset skala kecil. Tabel 1 menampilkan semua fitur

TABLE 1  
FITUR DATASET NILAI MAHASISWA

ID	Fitur	Keterangan
f1	Nama	Atribut
f2	Quiz	Atribut
f3	UTS	Atribut
f4	UAS	Atribut
f5	Huruf Mutu	Label

### C. Tahapan Penelitian

Tahapan penelitian adalah serangkaian langkah terstruktur, sistematis, dan logis yang dilakukan secara berurutan untuk memperoleh pengetahuan dan memecahkan masalah [9]. Tahapan penelitian ini terdiri atas beberapa langkah yang diilustrasikan pada Gambar 1.



Gambar 1. Tahapan penelitian

1. Pengumpulan Data

Data diambil dari website [www.siakadaisyahuniversity.ac.id](http://www.siakadaisyahuniversity.ac.id). Seleksi dataset sesuai kriteria dataset kecil dengan jumlah record=100, memiliki empat atribut, dan satu label.

2. Pre-processing Data

Proses membersihkan data dari nilai yang hilang (*missing values*) atau duplikasi [10]. Normalisasi atau standarisasi atribut jika diperlukan. Membagi dataset menggunakan teknik (*k-fold*) cross validation ( $k=5$ ) agar hasil evaluasi lebih stabil [11]. Hasil percobaan yang telah kami lakukan  $k=5$  menghasilkan kinerja yang lebih optimal

3. Implementasi Algoritma

Algoritma yang digunakan dalam penelitian ini berbasis Decision Tree yaitu J48, Random Forest, dan Random Tree. Setiap algoritma memiliki keunggulan masing-masing [12]. Algoritma Random Forest mampu yang membangun banyak pohon keputusan secara acak untuk meningkatkan akurasi. Random Tree merupakan algoritma pohon keputusan sederhana yang membangkitkan pohon berdasarkan pemilihan atribut acak [7]. Implementasi dilakukan menggunakan perangkat lunak WEKA

4. Evaluasi Model

Setiap model dievaluasi berdasarkan *confusion matrix* pada Tabel 2 berupa akurasi, presisi, dan recall. Akurasi yaitu proporsi prediksi benar terhadap keseluruhan data uji. Presisi yaitu ketepatan model dalam mengklasifikasikan data positif. Recall yaitu kemampuan model dalam mendeteksi data positif secara benar. F1-Score yaitu rata-rata harmonis dari presisi dan recall [13]. Waktu komputasi yaitu lama waktu yang dibutuhkan algoritma untuk membangun model [4].

TABLE 1  
CONFUSION MATRIX

Label	Prediksi Positif	Prediksi Negatif
Positif	True Positives (TP)	False Negatives (FN)
Negatif	False Positives (FP)	True Negatives (TN)

Pengukuran akurasi, presisi, recall adalah metode untuk mengevaluasi dan membandingkan pengklasifikasi [14]. Ketiga nilai tersebut dapat diperoleh dari Confusion Matrix pada Tabel 2 dan dihitung menggunakan persamaan (1) dan (2).

$$Akurasi = \frac{TP + TN}{TP + FP + FN + TN} \quad (1)$$

$$Presisi = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

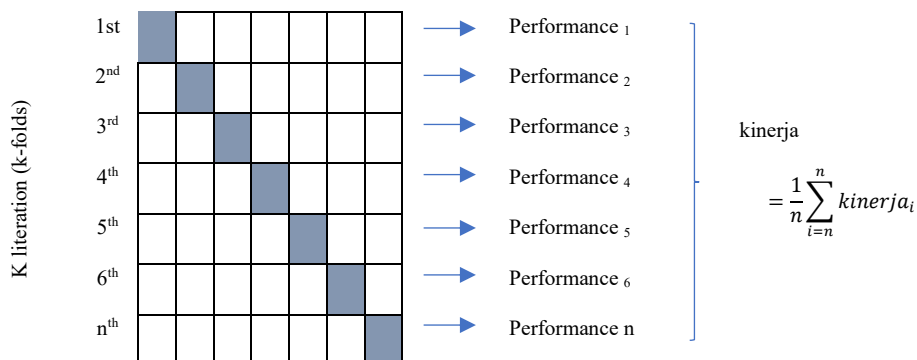
$$F1 - score = \frac{2 * Presisi * Recall}{Presisi + Recall} \quad (4)$$

5. Analisis Hasil

Membandingkan kinerja ketiga algoritma berdasarkan hasil evaluasi. Mengidentifikasi algoritma yang paling efektif pada dataset kecil. Menganalisis kelebihan dan kelemahan tiap algoritma.

6. Rancangan Eksperimen

Eksperimen dilakukan dengan rancangan berupa (a) setiap algoritma dijalankan pada dataset yang sama (b) Validasi dilakukan menggunakan teknik 5-fold *cross-validation*. Ilustrasi proses k-fold ditampilkan pada Gambar 2. (c) Setiap algoritma dibandingkan berdasarkan hasil evaluasi dari metrik kinerja. (d) Hasil diekspresikan dalam bentuk tabel dan grafik untuk memudahkan analisis.



Gambar 2. Proses validasi (k-fold)

### III. HASIL DAN PEMBAHASAN

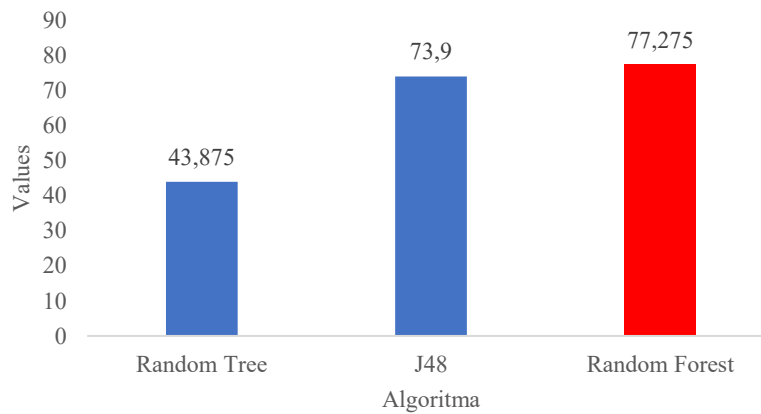
#### A. Perbandingan Kinerja Algoritma

Kami menggunakan tools WEKA karena mampu menyederhanakan konstruksi beberapa data teknik analisis. WEKA memiliki kemampuan mengategorikan, melakukan regresi, klasifikasi, menghilangkan fitur, membuat aturan asosiasi, dan penyesuaian class pada dataset [15]. Kami menggunakan beberapa uji coba pada *k-fold* (5, 7, 10) untuk mendapatkan kinerja optimal dari masing-masing algoritma. Kami menemukan *k-fold* optimal yaitu  $k=5$ . Hasil kinerja algoritma ditampilkan dalam Tabel 2.

TABLE 3  
PERBANDINGAN KINERJA ALGORITMA

No.	Algoritma	Akurasi (%)	Pesisi (%)	Recall (%)	F1-score (%)
1	J48	74	73.9	74	73.7
2	Random Forest	77	78.3	77	76.8
3	Random Tree	37	69.8	37	31.7

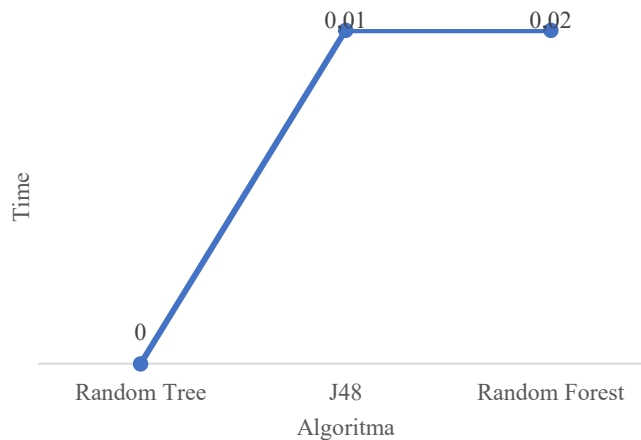
Berdasarkan evaluasi menggunakan confusion matrix, setiap algoritma memiliki kinerja yang berbeda. Algoritma Random Forest berdasarkan nilai akurasi, presisi, recall dan f1-score memiliki kinerja terbaik, disusul J48, dan terburuk Random Tree. Nilai rata-rata algoritma disajikan dalam Gambar 2.



Gambar 3. Nilai rata-rata kinerja algoritma

#### B. Waktu Komputasi

Berdasarkan Gambar 2, algoritma Random Forest memiliki kinerja terbaik. Hal ini dikarenakan algoritma Random Forest memiliki ketahanan terhadap *overfitting*, kemampuan menangani data berdimensi kecil dan besar, fleksibilitas untuk tugas klasifikasi dan regresi, serta evaluasi kepentingan fitur secara otomatis. Selain itu, algoritma ini kuat terhadap data *outlier* dan tidak memerlukan banyak *pre-processing* data, serta bekerja efisien pada dataset besar maupun kecil. Selanjutnya kami melakukan perbandingan waktu komputasi dalam membangun model yang diilustrasikan pada Gambar 4.



Gambar 4. Waktu komputasi algoritma

Meskipun secara akurasi, presisi, dan recall kinerja algoritma Random Tree buruk, namun secara waktu komputasi algoritma ini memiliki waktu yang cepat, disusul algoritma J48, dan Random Forest.

### C. Diskusi

Hasil pengujian yang telah dilakukan pada Gambar 3 menunjukkan bahwa algoritma Random Forest memiliki kinerja terbaik. Random Forest adalah algoritma ML berbasis ensemble (gabungan banyak decision trees) yang unggul dalam akurasi tinggi, ketahanan terhadap overfitting, dan penanganan data hilang (missing values) maupun outlier. Keunggulannya meliputi fleksibilitas untuk klasifikasi/regresi, namun memiliki kelemahan berupa kompleksitas komputasi yang lambat, penggunaan memori besar, dan sulit diinterpretasikan (model black box), oleh karena itu perlu kajian mendalam untuk membuktikan keunggulan algoritma tersebut.

## IV. KESIMPULAN

Berdasarkan hasil evaluasi terhadap algoritma J48, Random Forest, dan Random Tree pada dataset skala kecil, dapat disimpulkan bahwa algoritma Random Forest menunjukkan performa yang lebih baik dari segi akurasi, presisi, recall dan f1-score meskipun membutuhkan waktu komputasi lebih lama dibandingkan algoritma lainnya. Algoritma Random Tree memiliki keunggulan dalam hal efisiensi waktu dan interpretabilitas model, sehingga cocok digunakan pada penelitian yang membutuhkan pemahaman struktur pohon keputusan secara lebih sederhana. Algoritma J48 menawarkan fleksibilitas dengan kompleksitas model yang relatif sedang, namun konsistensi performanya stabil dibandingkan Random Tree. Hasil ini menunjukkan bahwa pemilihan algoritma yang tepat sangat bergantung pada konteks penelitian.

Penelitian ini memiliki beberapa keterbatasan yang perlu diperhatikan seperti dataset yang digunakan terbatas pada dataset dengan ukuran kurang dari 100 record, sehingga generalisasi hasil mungkin tidak sepenuhnya mewakili semua jenis dataset kecil. Parameter algoritma sebagian besar menggunakan nilai default, sehingga belum dilakukan optimasi mendalam yang berpotensi meningkatkan performa model. Penelitian ini hanya berfokus pada tiga algoritma decision tree sehingga belum membandingkan dengan algoritma klasifikasi lain seperti SVM, K-Nearest Neighbor, atau Neural Network.

Berdasarkan keterbatasan yang ada, beberapa saran yang dapat diberikan untuk penelitian selanjutnya adalah menggunakan dataset yang lebih bervariasi, baik dari sisi jumlah atribut, distribusi kelas, maupun jenis domain aplikasi, agar hasil penelitian lebih komprehensif. Melakukan optimasi parameter (hyperparameter tuning) pada masing-masing algoritma untuk memperoleh performa yang lebih maksimal. Membandingkan algoritma decision tree dengan algoritma klasifikasi lain sehingga dapat diketahui posisi relatif keunggulan dan kelemahan setiap metode. Menerapkan pendekatan ensemble atau hybrid model untuk mengatasi kelemahan algoritma tunggal, khususnya pada kondisi dataset berukuran kecil.

## UCAPAN TERIMA KASIH

Penulis menyampaikan penghargaan dan terima kasih yang sebesar-besarnya kepada berbagai pihak yang telah memberikan dukungan dalam penyelesaian penelitian ini. Ucapan terima kasih khusus ditujukan kepada Fakultas Komputer Universitas Mitra Indonesia yang telah menyediakan fasilitas dan sumber daya sehingga penelitian dapat terlaksana dengan baik. Penulis juga berterima kasih kepada BAAK Administrasi sebagai penyedia dataset yang menjadi dasar eksperimen dalam penelitian ini. Tidak lupa, apresiasi diberikan kepada para rekan sejawat dan dosen pembimbing yang telah memberikan masukan, arahan, serta diskusi konstruktif selama proses penelitian berlangsung. Akhirnya, penulis berharap bahwa hasil penelitian ini dapat memberikan kontribusi positif bagi perkembangan ilmu pengetahuan, khususnya di bidang Machine Learning dan analisis data pada dataset skala kecil.

## DAFTAR PUSTAKA

- [1] U. S. Reddy, A. V. Thota, and A. Dharun, "Machine Learning Techniques for Stress Prediction in Working Employees," *2018 IEEE Int. Conf. Comput. Intell. Comput. Res. ICCIC 2018*, no. May, 2018, doi: 10.1109/ICCIC.2018.8782395.
- [2] A. Agliata, D. Giordano, F. Bardozzo, S. Bottiglieri, A. Facchiano, and R. Tagliaferri, "Machine Learning as a Support for the Diagnosis of Type 2 Diabetes," *International Journal of Molecular Sciences*, vol. 24, no. 7, 2023. doi: 10.3390/ijms24076775.
- [3] M. T. H. Sultan, V. R. Kar, S. K. Panda, and K. Jayakrishna, "Advanced Composite Materials and Structures," *Adv. Compos. Mater. Struct.*, no. April, 2022, doi: 10.1201/9781003158813.
- [4] A. Wantoro, A. F. Yuliana, D. Yana, A. Andini, and I. Awaliyani, "Optimizing Type 2 Diabetes Classification with Feature Selection and Class Balancing in Machine Learning," *J. Tek. Inform.*, vol. 6, no. 4, pp. 2625–2637, 2025.
- [5] C. H. B. Liu, B. P. Chamberlain, D. A. Little, and A. Cardoso, "Generalising Random Forest Parameter Optimisation to Include Stability and Cost BT - Machine Learning and Knowledge Discovery in Databases," 2017, pp. 102–113.

- [6] A. Tharwat, "Parameter investigation of support vector machine classifier with kernel functions," *Knowl. Inf. Syst.*, vol. 61, no. 3, pp. 1269–1302, 2019, doi: 10.1007/s10115-019-01335-4.
- [7] B. Mahesh, "Machine Learning Algorithms - A Review," *Int. J. Sci. Res.*, vol. 9, no. 1, pp. 381–386, 2020, doi: 10.21275/art20203995.
- [8] A. B. R. Shatte, D. M. Hutchinson, and S. J. Teague, "Machine learning in mental health: a scoping review of methods and applications," *Psychol. Med.*, vol. 49, no. 9, pp. 1426–1448, 2019, doi: DOI: 10.1017/S0033291719000151.
- [9] E. Susetyarini and A. Fauzi, "Trend of critical thinking skill researches in biology education journals across Indonesia: From research design to data analysis," *Int. J. Instr.*, vol. 13, no. 1, pp. 535–550, 2020, doi: 10.29333/iji.2020.13135a.
- [10] R. Jacob *et al.*, "Evaluating the Role of Breast Ultrasound in Early Detection of Breast Cancer in Low- and Middle-Income Countries: A Comprehensive Narrative Review," *Bioengineering*, vol. 11, no. 3, 2024. doi: 10.3390/bioengineering11030262.
- [11] H. Sulistiani, A. Syarif, K. Muludi, and Warsito, "Performance evaluation of feature selections on some ML approaches for diagnosing the narcissistic personality disorder," *Bull. Electr. Eng. Informatics*, vol. 13, no. 2, pp. 1383–1391, 2024, doi: 10.11591/eei.v13i2.6717.
- [12] N. R. Muntiari and K. H. Hanif, "Klasifikasi Penyakit Kanker Payudara Menggunakan Perbandingan Algoritma Machine Learning," *J. Ilmu Komput. dan Teknol.*, vol. 3, no. 1, pp. 1–6, 2022, doi: 10.35960/ikomti.v3i1.766.
- [13] I. Düntsch and G. Gediga, "Confusion Matrices and Rough Set Data Analysis," *J. Phys. Conf. Ser.*, vol. 1229, no. 1, 2019, doi: 10.1088/1742-6596/1229/1/012055.
- [14] M. Ohsaki, P. Wang, K. Matsuda, S. Katagiri, H. Watanabe, and A. Ralescu, "Confusion-matrix-based kernel logistic regression for imbalanced data classification," *IEEE Trans. Knowl. Data Eng.*, vol. 29, no. 9, pp. 1806–1819, 2017, doi: 10.1109/TKDE.2017.2682249.
- [15] I. Popchev and D. Orozova, "Algorithms for Machine Learning with Orange System," *Int. J. online Biomed. Eng.*, vol. 19, no. 4, pp. 109–123, 2023, doi: 10.3991/ijoe.v19i04.36897.