

Penerapan *Random Forest* untuk Identifikasi Tipe Kepribadian: Studi Kasus Klasifikasi *Introvert* dan *Extrovert*

Fauzan Azim¹, Doni Iskandar Rotama Sihalo¹, Muhammad Rizki¹, Afdanil Hidayat Alghifary¹, Husnul Qadri Habibullah¹, Harya Jaya Kusuma¹, Rahmad Gunawan¹

¹Program Studi Teknik Informatika, Universitas Muhammadiyah Riau, Indonesia

Artikel Info

Kata Kunci:

Klasifikasi Kepribadian;
Machine Learning;
Random Forest.

Keywords:

Personality Classification;
Machine Learning;
Random Forest.

Riwayat Artikel:

Submitted: 4 Agustus 2025
Accepted: 2 November 2025
Published: 10 Desember 2025

Abstrak: Pengelompokan tipe kepribadian menjadi salah satu cara penting untuk memahami bagaimana individu berinteraksi dalam konteks sosial. Penelitian ini mengeksplorasi penerapan algoritma pembelajaran mesin dalam membedakan individu ke dalam dua kategori kepribadian utama: *introvert* dan *ekstrovert*, berdasarkan indikator perilaku sosial yang dapat diukur secara kuantitatif. Dataset yang digunakan mencakup 2.900 observasi dengan variabel-variabel seperti durasi waktu yang diluangkan secara mandiri, frekuensi partisipasi dalam kegiatan sosial, tingkat interaksi di luar lingkungan rumah, serta jumlah teman dalam jaringan sosial. Sebelum analisis, data mengalami proses pra-pemrosesan yang meliputi konversi variabel kategorikal menjadi bentuk numerik dan penskalaan fitur numerik agar seragam. Model klasifikasi yang digunakan adalah *Random Forest Classifier* dengan parameter baku, yang divalidasi menggunakan pembagian data latih dan uji melalui teknik *stratified split* untuk menjaga keseimbangan distribusi kelas. Hasil evaluasi menunjukkan performa model yang baik, dengan tingkat akurasi sebesar 90,86% dan *F1-Score* mencapai 91,03%. Temuan ini menegaskan bahwa pendekatan berbasis pembelajaran mesin, terutama model *ensemble* seperti *Random Forest* memiliki potensi besar dalam mengidentifikasi ciri kepribadian melalui pola perilaku sosial yang terekam.

Abstract: Classifying personality types is an important way to understand how individuals interact in social contexts. This research explores the application of machine learning algorithms in distinguishing individuals into two main personality categories: *introverts* and *extroverts*, based on quantifiable indicators of social behavior. The dataset used includes 2,900 observations with variables such as duration of time spent alone, frequency of participation in social activities, level of interaction outside the home environment, and number of friends in social networks. Prior to analysis, the data underwent pre-processing which included converting categorical variables into numerical form and scaling the numerical features for uniformity. The classification model used was a *Random Forest Classifier* with default parameters, which was validated using a *stratified split* of training and test data to maintain a balanced class distribution. The evaluation results showed good model performance, with an accuracy rate of 90.86% and *F1-Score* reaching 91.03%. This finding confirms that machine learning-based approaches, especially ensemble models such as *Random Forest* have great potential in identifying personality traits through recorded social behavior patterns.

This is an open access article under the [CC BY-SA](#) license



Corresponding Author:

Fauzan Azim

Email: fauzanazim17@umri.ac.id

PENDAHULUAN

Kemajuan teknologi informasi telah membuka peluang baru dalam analisis kepribadian secara efisien dan berbasis data. Kepribadian merupakan aspek psikologis yang memengaruhi cara individu berpikir, berinteraksi, dan bertindak dalam kehidupan sosial. Salah satu pendekatan kategorisasi kepribadian yang banyak diterapkan adalah pembagian menjadi dua tipe utama, yaitu *introvert* dan *ekstrovert*. Pemetaan tipe kepribadian semacam ini menjadi semakin relevan dalam pengembangan sistem berbasis pengguna, seperti sistem rekomendasi, personalisasi layanan digital, hingga penyesuaian strategi komunikasi dalam konteks pendidikan atau dunia kerja.

Seiring meningkatnya akses terhadap data digital, pendekatan berbasis perilaku sosial yang terstruktur mulai dikembangkan untuk mengidentifikasi kepribadian seseorang. Fitur seperti durasi waktu menyendiri, frekuensi menghadiri kegiatan sosial, dan ukuran lingkaran pertemanan dinilai dapat menjadi indikator objektif dalam membedakan karakteristik *introvert* dan *ekstrovert*. Pendekatan ini dinilai lebih efisien dibandingkan metode konvensional berbasis kuesioner yang bersifat subjektif atau metode berbasis analisis teks yang memerlukan pemrosesan linguistik yang kompleks (Azifa et al., 2023).

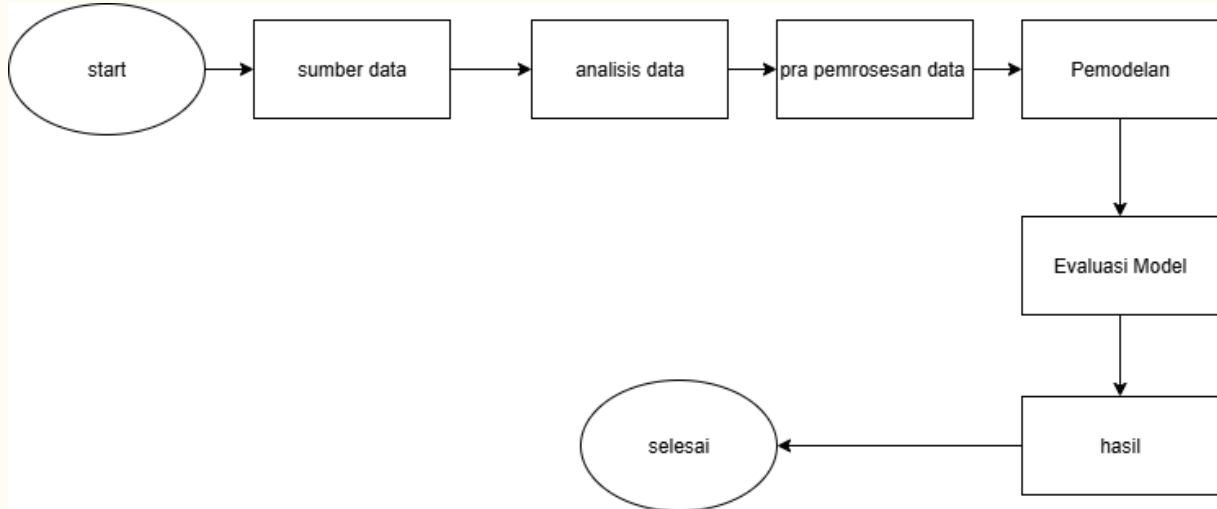
Salah satu algoritma pembelajaran mesin yang terbukti efektif dalam konteks klasifikasi kepribadian adalah *Random Forest*. Metode ini mampu mengelola data *multivariat* secara stabil dan tahan terhadap *overfitting*, sehingga menghasilkan performa klasifikasi yang konsisten (Gumilang & Agustin, 2024). Penerapannya telah diuji dalam berbagai konteks, mulai dari klasifikasi tulisan tangan hingga sistem klasifikasi berbasis interaksi digital melalui platform daring dengan hasil yang menunjukkan akurasi tinggi (Mayasari et al., 2024).

Namun demikian, sebagian besar penelitian terdahulu masih berfokus pada klasifikasi berbasis data teks atau citra, dan belum banyak yang mengeksplorasi penggunaan data perilaku sosial dalam bentuk numerik secara langsung sebagai dasar klasifikasi tipe kepribadian. Padahal, data tersebut merupakan representasi konkret dari interaksi sosial sehari-hari yang berpotensi memberikan gambaran kepribadian secara lebih nyata dan terukur (Azifa et al., 2023; Mayasari et al., 2024).

Berdasarkan latar belakang yang telah diuraikan, penelitian ini bertujuan untuk menyusun model klasifikasi kepribadian yang membagi individu ke dalam dua kelompok utama, yaitu *introvert* dan *ekstrovert*. Model ini dikembangkan dengan memanfaatkan algoritma *Random Forest* yang diterapkan pada data perilaku sosial dalam bentuk numerik. Dataset yang digunakan terdiri dari 2.900 data dengan tujuh atribut kunci, seperti durasi waktu yang dihabiskan secara mandiri, tingkat partisipasi dalam kegiatan sosial, serta frekuensi interaksi sosial. Tahapan pengolahan data meliputi encoding variabel kategorikal, normalisasi fitur numerik, dan pembagian data menggunakan metode stratifikasi untuk memastikan keseimbangan kelas. Kinerja model dievaluasi berdasarkan sejumlah metrik klasifikasi, termasuk akurasi, presisi, *Recall*, dan *F1-score*. Diharapkan penelitian ini dapat memberikan kontribusi dalam penerapan ilmu data untuk analisis kepribadian secara komputasional, serta mendukung pengembangan sistem cerdas yang lebih responsif dan personal terhadap perilaku pengguna.

METODE

Metodologi penelitian ini dirancang secara sistematis mengikuti tahapan standar dalam proyek *data science*, mulai dari persiapan data, prapemrosesan data, pemodelan, hingga evaluasi model, sebagaimana diilustrasikan dalam banyak penelitian serupa (Ayu Firnanda et al., 2025; Ubed, 2025).



Gambar 1. Diagram Alur Penelitian Ilmiah

Pada Gambar 1 tentang diagram alur penelitian dapat kita lihat bahwa alur artikel ilmiah ini dimulai dari mencari sumber data, melakukan pra pemrosesan data agar semua data yang ada pada dataset bersih dan dapat dilanjutkan ke tahap pemodelan, kemudian dari data yang sudah bersih dilakukan evaluasi pada model yang sudah dikerjakan, hingga pada visualisasi hasil.

Sumber Data

Data yang digunakan pada penelitian ini merupakan dataset publik yang diambil dari website Kaggle, berisi berbagai atribut perilaku individu dan sebuah variabel target yang telah diberi label, yaitu tipe kepribadian (*Introvert* atau *Ekstrovert*). *Dataset* serupa telah digunakan dalam penelitian sebelumnya untuk klasifikasi kesehatan mental dan kepribadian karena sifatnya yang terbuka, lengkap, dan relevan untuk tujuan klasifikasi (Arsad et al., 2024)

Analisis data

Dataset yang dianalisis berjumlah 2.900 entri dengan 8 variabel yang digunakan untuk mengkaji hubungan antara kebiasaan sosial seseorang dengan kecenderungan kepribadiannya, yaitu *Ekstrovert* atau *Introvert*. Variabel yang dianalisis mencakup data kategorikal dan numerik:

- *Time_spent_Alone*: Hours spent alone daily (0–11).
- *Stage_fear*: Presence of stage fright (Yes/No).
- *Social_event_attendance*: Frequency of social events (0–10).
- *Going_outside*: Frequency of going outside (0–7).
- *Drained_after_socializing*: Feeling drained after socializing (Yes/No).
- *Friends_circle_size*: Number of close friends (0–15).
- *Post_frequency*: Social media post frequency (0–10).
- *Personality*: Target variable (Extrovert/Introvert).

Penggunaan atribut kategorikal dan numerik ini serupa dengan studi (Zulfikar, n.d.) yang juga mengevaluasi atribut-atribut perilaku sosial dan psikologis untuk klasifikasi kepribadian.

Pra-pemrosesan Data

Tahap pra-pemrosesan sangat krusial untuk memastikan kualitas data dan mengoptimalkan kinerja model. Dua langkah utama yang dilakukan adalah:

- **Encoding Data Kategorikal:** Semua fitur kategorikal seperti jawaban 'Ya' atau 'Tidak' diubah menjadi representasi numerik (misalnya, 1 dan 0) menggunakan teknik label encoding, seperti juga dilakukan dalam penelitian (Hendi Muhammad et al., n.d.).
- **Standardisasi Fitur (Feature Scaling):** Fitur-fitur numerik distandardisasi dengan *StandardScaler* sehingga memiliki nilai rata-rata 0 dan standar deviasi 1, untuk mencegah bias akibat skala yang berbeda-beda, seperti dipraktikkan pula pada studi (Ayu Firnanda et al., 2025) yang menunjukkan pentingnya normalisasi untuk kinerja model.
- **Split Data:** *Dataset* dibagi menjadi dua bagian: 80% sebagai data latih dan 20% sebagai data uji. Pembagian dilakukan secara stratifikasi terhadap variabel target untuk menjaga proporsi kelas *introvert* dan *ekstrovert* tetap konsisten di kedua set data. Hal ini sesuai dengan praktik terbaik pada penelitian klasifikasi dengan *Random Forest* yang juga disarankan oleh (Ayu Firnanda et al., 2025).

Pemodelan

Tahap pemodelan dilakukan dengan melatih dan menguji algoritma *Random Forest*. Algoritma *Random Forest* bekerja dengan membangun banyak pohon keputusan pada subsampel acak dari dataset, dan keputusan akhir diambil dengan mayoritas suara dari pohon-pohon tersebut. Metode ini dikenal mampu meningkatkan akurasi dan mengurangi risiko overfitting dibandingkan dengan satu pohon tunggal (Ayu Firnanda et al., 2025)

Evaluasi Model

Kinerja model yang telah dilatih dievaluasi menggunakan data uji yang belum pernah dilihat sebelumnya. Metrik evaluasi yang digunakan meliputi:

- Akurasi: proporsi total prediksi yang benar.
- Presisi: dari semua prediksi positif, berapa banyak yang benar-benar positif.
- *Recall* (Sensitivitas): dari semua kasus positif aktual, berapa banyak yang berhasil diidentifikasi dengan benar.
- *F1-Score*: Rata-rata harmonik dari presisi dan *recall* memberikan ukuran kinerja yang seimbang.

Selain metrik tersebut, *Confusion Matrix* digunakan untuk memvisualisasikan hasil klasifikasi secara rinci, menampilkan jumlah *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN).

HASIL DAN PEMBAHASAN

Evaluasi Performa Model

Setelah model dilatih menggunakan data latih, performa model diuji menggunakan data uji yang belum pernah dilihat sebelumnya. Tahap evaluasi ini krusial untuk mengetahui seberapa baik model dapat menggeneralisasi pengetahuannya pada data baru. Hasil evaluasi performa model klasifikasi disajikan pada Tabel 1.

Tabel 1. Hasil Metrik Evaluasi

	<i>Introvert</i>	<i>Extrovert</i>
<i>Prevision</i>	0.9000	0.9200
<i>Recall</i>	0.9100	0.9000
<i>F1-Score</i>	0.9100	0.9100
Rata-Rata	0.9100	0.9100

Tabel 1. menunjukkan bahwa model klasifikasi yang dibangun memiliki performa yang sangat baik. Akurasi model tercatat sebesar 0.9086, yang menunjukkan bahwa model mampu memprediksi tipe kepribadian dengan benar pada sekitar 90.86% dari data uji. Tingkat akurasi yang tinggi ini sejalan dengan temuan (Mayasari et al., 2024), yang melaporkan bahwa metode berbasis *Random Forest* dengan data interaksi digital juga dapat mencapai tingkat akurasi tinggi pada klasifikasi kepribadian. Hal ini membuktikan bahwa algoritma *Random Forest* cocok dan relevan digunakan untuk klasifikasi tipe kepribadian berbasis perilaku sosial. Studi oleh Perera & Costa (2023) juga menyebutkan bahwa model *Random Forest* memiliki keunggulan dalam klasifikasi kepribadian karena mampu mengelola banyak variabel masukan dan menghasilkan prediksi yang stabil, terutama saat diterapkan pada data berbasis teks atau perilaku digital (Perera & Costa, 2023).

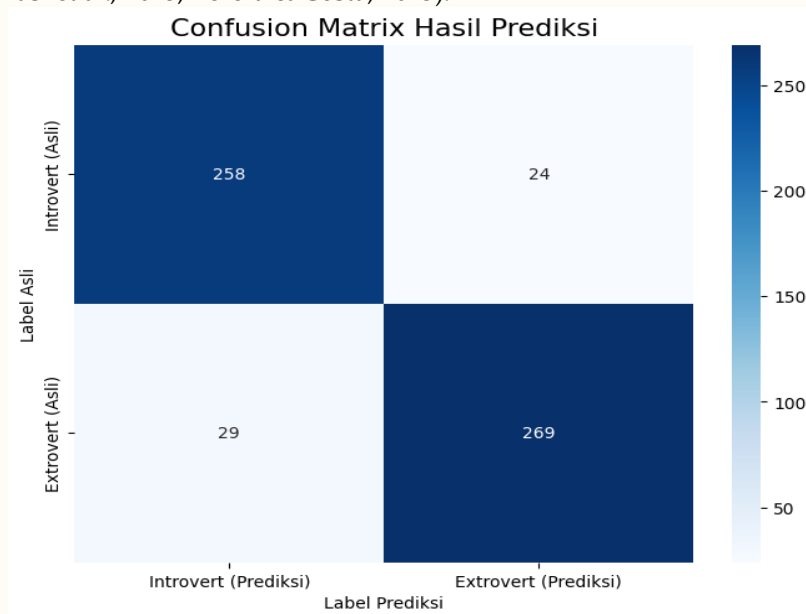
Presisi model mencapai 0.9181, menunjukkan tingkat kepercayaan yang tinggi ketika model memprediksi seorang individu sebagai *Ekstrovert*, prediksi tersebut benar dalam 91.81% kasus. Nilai presisi ini juga diperkuat oleh hasil penelitian (Sher Khan et al., 2020), yang melaporkan bahwa penerapan algoritma berbasis pembelajaran mesin pada data teks daring dapat menghasilkan presisi prediksi yang baik, khususnya dalam mengidentifikasi ciri-ciri kepribadian dari perilaku pengguna. Temuan ini juga sejalan dengan studi oleh Perera & Costa (2023) yang menunjukkan bahwa presisi pada model *Random Forest* tetap tinggi ketika digunakan pada data teks kepribadian, karena pendekatannya yang memanfaatkan kombinasi banyak decision tree yang menghasilkan prediksi yang lebih terpercaya (Perera & Costa, 2023).

Recall model sebesar 0.9027 menunjukkan bahwa model berhasil menemukan sekitar 90.27% dari semua individu yang benar-benar *Ekstrovert* di dalam data uji. Hasil ini selaras dengan temuan (Harahap & Muslim, 2020), yang menyatakan bahwa penggunaan *Random Forest* dengan teknik augmentasi data mampu meningkatkan kemampuan model dalam mendeteksi kelas minoritas secara lebih baik, yang dalam kasus ini adalah kelas *Ekstrovert*. Selain itu, algoritma *Random Forest* telah terbukti mampu mengatasi ketidakseimbangan kelas (imbalanced data) dan tetap mempertahankan tingkat *Recall* yang baik dalam klasifikasi kepribadian berbasis perilaku (Nur Listiawan Dhito Eka Santoso, 2025).

F1-Score merupakan rata-rata harmonis dari Presisi dan *Recall*, diperoleh sebesar 0.9103. *F1-Score* yang tinggi dan seimbang ini menunjukkan bahwa model tidak hanya akurat tetapi juga konsisten dalam menjaga keseimbangan antara kesalahan tipe I (*False Positive*) dan kesalahan tipe II (*False Negative*). Hasil ini konsisten dengan laporan (Li & Chen, 2025), yang menunjukkan bahwa algoritma *ensemble* seperti *Random Forest* mampu memberikan performa stabil dan unggul dalam menangani interaksi variabel yang kompleks dan *non-linear*. Skor F1 yang tinggi juga tercermin dalam studi oleh Perera & Costa (2023), yang menekankan bahwa algoritma *ensemble* seperti *Random Forest* cenderung menjaga keseimbangan antara presisi dan *Recall*, menjadikannya pilihan yang stabil dan kuat dalam klasifikasi kepribadian (Grunenberg et al., 2023; Perera & Costa, 2023).

Hasil evaluasi model pada penelitian ini menunjukkan performa yang baik secara keseluruhan. Hal ini membuktikan bahwa metode *Random Forest* yang digunakan sudah tepat dan relevan untuk tugas klasifikasi kepribadian berbasis atribut perilaku, seperti juga dilaporkan oleh berbagai penelitian terdahulu (Harahap & Muslim, 2020; Li & Chen, 2025; Mayasari et al., 2024; Sher Khan et al., 2020). Hasil ini memperkuat keyakinan bahwa penggunaan *Random Forest* dalam klasifikasi kepribadian secara

umum mendapatkan konsistensi hasil yang baik pada berbagai jenis dataset, baik berbasis teks maupun perilaku (Informasi et al., 2025; Perera & Costa, 2023).



Gambar 2. *Confusion Matrix* Hasil Prediksi Model *Random Forest*

Pada matrix ini, bisa dianalisis 4 *output* utama dari hasil kinerja model:

- *True Negative* (TN): Model sudah bisa mengklasifikasikan ada 258 orang memiliki kepribadian introvert dan sudah benar, tepat seperti label aslinya.
- *True Positive* (TP): Model juga sudah bisa mengklasifikasikan 269 orang sebagai extrovert dan sudah tepat seperti pada label aslinya.
- *False Positive* (FP): Pada bagian ini model membuat kesalahan, model mengklasifikasikan sebanyak 24 orang introvert sebagai extrovert.
- *False Negative* (FN): Pada bagian ini model membuat kesalahan dengan klasifikasi yang dihasilkan model sebanyak 29 orang *introvert* yang seharusnya *extrovert*.

Pembahasan Umum

Validasi konsep

Hasil penelitian membuktikan bahwa algoritma Random Forest sangat efektif (akurasi>90%) untuk mengklasifikasikan tipe kepribadian berdasarkan data perilaku. Hal ini menunjukkan bahwa pola perilaku sosial memiliki daya prediksi yang kuat.

Didukung oleh:

- "Penelitian oleh Maldini et al. menunjukkan bahwa Random Forest memiliki akurasi tinggi dalam mendeteksi gangguan kepribadian berbasis analisis sentimen dengan akurasi mencapai 95,66%" (Arsad et al., 2024; Maldini et al., 2025).
- "Ashawa et al. menggabungkan CNN dan Random Forest dalam model prediksi kepribadian berbasis media sosial, dan menghasilkan MAE rata-rata sebesar 0.1867, menunjukkan performa model yang akurat dan stabil dalam prediksi trait Big Five" (Ashawa et al., 2025).
- "Algoritma Random Forest terbukti efektif dalam klasifikasi tipe kepribadian berdasarkan pola perilaku sosial digital" (Statistician & Applications, 2022).

Implikasi praktis

Keberhasilan model ini membuka potensi penerapan di berbagai bidang, seperti alat bantu asesmen awal dalam psikologi atau analisis kandidat di bidang sumber daya manusia (HR).

Didukung oleh:

- "Prediksi kepribadian berbasis digital membuka aplikasi di berbagai bidang seperti rekrutmen, konseling, dan personalisasi layanan digital" (Ashawa et al., 2025).
- "Penelitian Mayasari et al. menyatakan bahwa klasifikasi kepribadian dapat digunakan untuk sistem rekomendasi dan layanan interaktif berbasis personalisasi pengguna" (Mayasari et al., 2024).

Keterbatasan penelitian

Keterbatasan utama dari studi ini adalah cakupan dataset yang mungkin tidak mewakili seluruh populasi dan fitur yang digunakan terbatas pada aspek perilaku sosial tertentu.

Didukung oleh:

- "Ashawa et al. menggarisbawahi bahwa pendekatan mereka berbasis Instagram hanya berlaku efektif pada populasi urban dan literat, dengan keterbatasan pada generalisasi budaya" (Ashawa et al., 2025).
- "Model kepribadian berbasis perilaku digital sangat dipengaruhi oleh keterbatasan data, seperti noise, bias populasi, dan tidak semua dimensi kepribadian dapat terdeteksi dari aktivitas daring" (Statistician & Applications, 2022).

Saran Pengembangan

Penelitian selanjutnya disarankan untuk menggunakan dataset yang lebih besar dan beragam. Selain itu, perlu dilakukan eksperimen dengan algoritma lain (contoh: *XGBoost*, *SVM*) untuk membandingkan dan mencari performa yang lebih optimal.

KESIMPULAN

Berdasarkan analisis klasifikasi terhadap data perilaku sosial, studi ini menyimpulkan bahwa penerapan metode *machine learning* dengan algoritma *Random Forest* mampu membedakan individu *Introvert* dan *Extrovert* secara akurat. Model yang dikembangkan menunjukkan performa yang sangat baik dan andal, yang ditandai dengan pencapaian akurasi 90.86%. Hasil ini tidak hanya memvalidasi pendekatan yang digunakan, tetapi juga membuka implikasi praktis untuk pengembangan sistem identifikasi kepribadian otomatis di masa depan. Disarankan agar penelitian selanjutnya dapat memperluas cakupan data dan mengeksplorasi algoritma lain untuk penyempurnaan lebih lanjut.

DAFTAR PUSTAKA

- Arsad, Muhammad, A. H., & Hidayat, T. (2024). Classification of Mental Disorders Using Modified Balanced *Random Forest* And Feature Selection. *Jurnal Teknologi Informasi Universitas Lambung Mangkurat (JTIULM)*, 9(2), 45–54. <https://doi.org/10.20527/jtiulm.v9i2.320>
- Arsad, Muhammad, A. H., & Hidayat, T. (2024). Classification of Mental Disorders Using Modified Balanced Random Forest And Feature Selection. *Jurnal Teknologi Informasi Universitas Lambung Mangkurat (JTIULM)*, 9(2), 45–54. <https://doi.org/10.20527/jtiulm.v9i2.320>
- Nur Listiawan Dhito Eka Santoso, R. (2025). *SITEKNIK Sistem Informasi, Teknik dan Teknologi Terapan Optimization of Stress Classification Among Students Using Random Forest Algorithm*. 2(2), 76–87. <https://doi.org/10.5281/zenodo.15130385>
- Ashawa, M., Bryan, J. D., & Owoh, N. (2025). Personality Prediction Model: An Enhanced Machine Learning Approach. *Electronics*, 14(13), 2558. <https://doi.org/10.3390/electronics14132558>

- Ayu Firnanda, P., Shofwatillah, L., Rahma, F., Fauzi, F., Studi Statistika, P., Muhammadiyah Semarang, U., Kedungmundu No, J., Tembalang, K., & Tengah, J. (2025). Analisis Perbandingan Decision Tree dan *Random Forest* dalam Klasifikasi Penjualan Produk pada Supermarket. *Emerging Statistics and Data Science Journal*, 3(1).
- Azifa, A. Z., Maharani, W., & Gani, P. H. (2023). Comparative Analysis of Personality Detection using *Random Forest* and Multinomial Naive Bayes. *Jurnal Media Informatika Budidarma*, 7(1), 554. <https://doi.org/10.30865/mib.v7i1.5592>
- Grunenberg, E., Peters, H., Francis, M. J., Back, M. D., & Matz, S. C. (2023). Machine learning in recruiting: predicting personality from CVs and short text responses. *Frontiers in Social Psychology*, 1. <https://doi.org/10.3389/frsps.2023.1290295>
- Gumilang, A., & Agustin, S. (2024). Deteksi Kepribadian Melalui Margin Pada Tulisan Tangan Menggunakan *Random Forest*. 9(1).
- Harahap, R. N., & Muslim, K. (2020). Peningkatan Akurasi pada Prediksi Kepribadian MbtI Pengguna Twitter Menggunakan Augmentasi Data. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 7(4), 815. <https://doi.org/10.25126/jtiik.2020743622>
- Li, J., & Chen, X. (2025). Modeling student satisfaction in online learning using random forest. *Scientific Reports*, 15(1), 23254. <https://doi.org/10.1038/s41598-025-06686-3>
- Maldini, N., Utomo, D. W., & Tresyani, R. P. (2025). Deteksi Dini Gangguan Kesehatan Mental berdasarkan Sentimen menggunakan Metode Stacking. *Sistemasi: Jurnal Sistem Informasi*, 14(1), 2540–9719. <http://sistemasi.ftik.unisi.ac.id>
- Mayasari, N., Chairul Rizal, & Fitri Wulandari. (2024). Penerapan Metode Klasifikasi Karakteristik Kepribadian Berbasis Website. *Journal Zetroem*, 6(2), 27–30. <https://doi.org/10.36526/ztr.v6i2.3755>
- Perera, H., & Costa, L. (2023). Personality Classification of Text Through Machine Learning and Deep Learning: a Review (2023). *International Journal For Research In Advanced Computer Science And Engineering*, 9(4), 6–12. <https://doi.org/10.53555/cse.v9i4.2266>
- Sher Khan, A., Ahmad, H., Zubair Asghar, M., Khan Saddozai, F., Arif, A., & Ali Khalid, H. (2020). Personality Classification from Online Text using Machine Learning Approach. In *IJACSA International Journal of Advanced Computer Science and Applications* (Vol. 11, Issue 3).
- Statistician, M., & Applications, E. (2022). *Big Five Personality Prediction Using Machine Learning Algorithms*. 71(3), 1128–1133. <http://philstat.org.ph>
- Ubed, I. A. (2025). Analisis Kinerja Algoritma Machine Learning Untuk Klasifikasi Potensi FRAUD Klaim Layanan Kesehatan Rumah Sakit. <https://doi.org/10.47065/bits.v7i1.7349>