



Implementasi Algoritma Random Forest untuk Klasifikasi Rentang Harga Ponsel Berdasarkan Spesifikasi Teknis

Yustinus Liguori¹, I Wayan Sudiarsa^{2*}, I Made Jagat Dita³, I Gusti Ngurah Galih Jimbar Baskara⁴, Pande wisnu wijaya putra⁵

¹⁻⁵Rekayasa Sistem Komputer, Institut Bisnis dan Teknologi Indonesia, Indonesia
Email: petrikliguori@gmail.com¹, sudiarsa@instiki.ac.id^{2*}, jagatdita6@gmail.com³,
galihbaskara5588@gmail.com⁴, wijayaputu298@gmail.com⁵

*Penulis Korespondensi: sudiarsa@instiki.ac.id²

Abstract. *The rapid development of smartphone technology today creates challenges for consumers and manufacturers in determining an objective price range based on highly varied technical specifications. This study aims to implement the Random Forest algorithm in classifying smartphone price ranges into four main categories, namely low, mid-range, high, and flagship. The research method was carried out systematically through the stages of loading a dataset of 2,000 entries, exploratory data analysis (EDA) to ensure data integrity, and model training with a training and testing data split of 80:20. The results showed that the Random Forest model achieved a significant overall accuracy rate of 89%. Based on feature importance analysis, it was found that RAM capacity was the most dominant determining factor, contributing 47% to prediction accuracy, followed by battery power and screen resolution as supporting features. These findings have strategic implications for manufacturers to prioritize memory capacity upgrades in determining product pricing in the market, as well as providing guidance for consumers in assessing the fairness of a device's price based on its technical capabilities.*

Keywords: *Classification; Mobile Phone; Pattern Recognition; Price Range; Random Forest.*

Abstract. Pesatnya perkembangan teknologi ponsel cerdas saat ini menciptakan tantangan bagi konsumen dan produsen dalam menentukan rentang harga yang objektif berdasarkan spesifikasi teknis yang sangat bervariasi. Penelitian ini bertujuan untuk mengimplementasikan algoritma Random Forest dalam mengklasifikasikan rentang harga ponsel ke dalam empat kategori utama, yaitu murah, menengah, tinggi, dan flagship. Metode penelitian dilakukan secara sistematis melalui tahap pemuatan dataset sebanyak 2.000 entri, analisis data eksploratif (EDA) untuk memastikan integritas data, serta pelatihan model dengan pembagian data latih dan uji sebesar 80:20. Hasil penelitian menunjukkan bahwa model Random Forest berhasil mencapai tingkat akurasi keseluruhan yang signifikan sebesar 89%. Berdasarkan analisis tingkat kepentingan fitur (feature importance), ditemukan bahwa kapasitas RAM merupakan faktor penentu paling dominan dengan kontribusi sebesar 47% terhadap akurasi prediksi, diikuti oleh daya baterai dan resolusi layar sebagai fitur pendukung. Temuan ini memberikan implikasi strategis bagi produsen untuk memprioritaskan peningkatan kapasitas memori dalam menentukan posisi harga produk di pasar, serta memberikan panduan bagi konsumen dalam menilai kewajaran harga sebuah perangkat berdasarkan kekuatan teknisnya.

Kata Kunci: Klasifikasi; Pengenalan Pola; Ponsel; Random Forest; Rentang Harga.

1. LATAR BELAKANG

Pada era digital saat ini, industri perangkat seluler, khususnya *smartphone*, mengalami perkembangan yang sangat pesat dengan inovasi teknologi yang terus muncul. Di Indonesia, penggunaan *smartphone* terus meningkat dan telah menjangkau hampir dua pertiga populasi (Zein & Ekawati, 2025), di mana perangkat ini tidak hanya berfungsi sebagai alat komunikasi tetapi juga menunjang sektor ekonomi, pendidikan, dan hiburan. Namun, pesatnya variasi merek dan spesifikasi teknis sering kali membuat konsumen mengalami kebingungan dalam menentukan pilihan yang tepat dan harga yang wajar. (Saputra & Sudiatmika, 2024)

Penetapan harga yang akurat menjadi sangat krusial bagi produsen untuk menjaga daya saing serta bagi pedagang dalam merencanakan strategi stok barang di masa depan.

Tantangan utama dalam industri ini adalah menentukan harga jual yang objektif berdasarkan fitur-fitur teknis seperti kapasitas RAM, memori internal, resolusi kamera, hingga jenis prosesor. Untuk mengatasi keterbatasan asumsi manusia, pemanfaatan teknologi *machine learning* melalui pendekatan *supervised learning* menjadi solusi yang efisien (Indrasetianingsih et al., 2025). Berbagai algoritma telah diuji efektivitasnya, seperti Random Forest Regression yang terbukti andal dalam menangani data kompleks dengan akurasi tinggi dan risiko *overfitting* yang minim (Saputra & Sudiatmika, 2024).

Machine learning merupakan salah satu pendekatan yang banyak digunakan untuk menyelesaikan permasalahan klasifikasi (Triansyah et al., 2025). Algoritma Random Forest, sebagai metode ensemble learning, memiliki keunggulan dalam menangani data dengan banyak fitur dan mampu memberikan hasil klasifikasi yang akurat (Didayat et al., 2023). Penelitian ini dilakukan untuk mengisi celah kebutuhan akan sistem klasifikasi harga ponsel yang objektif dan berbasis data, sehingga dapat membantu proses pengambilan keputusan baik bagi konsumen maupun pelaku industri (Warjiyono et al., 2024).

Beberapa penelitian sebelumnya (Aulia & Wahyuni, 2024) menunjukkan bahwa algoritma Random Forest memiliki kinerja yang lebih stabil dibandingkan algoritma klasifikasi lainnya, terutama dalam menghadapi data yang tidak linear dan memiliki banyak fitur (Sinambela et al., 2023). Selain itu, Random Forest juga mampu memberikan informasi mengenai tingkat kepentingan (*feature importance*) dari setiap atribut, sehingga dapat diketahui spesifikasi teknis mana yang paling berpengaruh terhadap penentuan rentang harga ponsel.

2. KAJIAN TEORITIS

Algoritma Random Forest (RF) merupakan metode *ensemble learning* berbasis pohon keputusan (*decision tree*) yang sangat andal untuk tugas klasifikasi dan regresi (Fitri et al., 2025). Dalam konteks klasifikasi rentang harga ponsel berdasarkan spesifikasi teknis, RF bekerja dengan membangun banyak pohon (hutan) untuk menghasilkan prediksi yang lebih akurat dan stabil dibandingkan model tunggal (Rahmada & Susanto, 2025). Algoritma ini sangat efektif karena kemampuannya dalam mengolah dataset yang kompleks dengan banyak fitur teknis (seperti RAM, memori, dan prosesor) serta memberikan interpretasi yang lebih mudah. Mekanisme penentuan pemisahan simpul (*node*) pada setiap pohon biasanya menggunakan Indeks Gini untuk mengukur kemurnian data, di mana nilai yang lebih rendah menunjukkan tingkat kemurnian yang lebih tinggi. Penggunaan RF juga memiliki keunggulan

dalam meminimalkan risiko *overfitting* saat mencapai akurasi maksimal, yang sangat penting untuk memastikan model dapat menggeneralisasi data ponsel baru dengan baik.

Secara teori, Random Forest mampu menangani data berdimensi tinggi dan variabel non-linear karena setiap pohon dibangun dari sampel acak serta fitur terpilih secara acak pula (Turino et al., 2025). Struktur ini membuat Random Forest efektif dalam klasifikasi dan regresi di berbagai domain ilmu, tidak hanya pada bidang teknologi tetapi juga di ilmu kimia dan pola variabel biologis (Aprilianto & Rizal, 2025). Studi teoretis lain menunjukkan bahwa meskipun pemilihan variabel acak meningkatkan kinerja, pentingnya variabel bisa menjadi bias tergantung pada skala variabel jika tanpa penyesuaian tertentu, sehingga pengetahuan tentang karakteristik model tetap penting saat interpretasi hasil dilakukan.

Untuk mencapai performa optimal dalam klasifikasi harga, tahap Prapemrosesan Data menjadi kunci utama. Spesifikasi teknis ponsel yang memiliki rentang nilai beragam perlu diseragamkan melalui teknik Standardisasi (seperti *Z-score normalization* atau *Standard Scaler*) agar tidak ada fitur tertentu yang mendominasi model secara tidak sengaja karena skala nilai yang besar (Oriana et al., 2025). Selain itu, jika terdapat ketidakseimbangan jumlah data pada rentang harga tertentu (misalnya lebih banyak data ponsel harga murah dibanding *flagship*), teknik SMOTE (*Synthetic Minority Oversampling Technique*) dapat diterapkan untuk menciptakan data sintetis pada kelas minoritas sehingga distribusi data menjadi adil dan tidak bias (Oriana et al., 2025).

Kinerja model klasifikasi rentang harga ini kemudian dievaluasi menggunakan Confusion Matrix untuk mengukur efektivitas algoritma (Zildan T.M, 2025). Metrik evaluasi utama meliputi Akurasi untuk melihat proporsi keseluruhan prediksi yang benar, Presisi untuk mengukur keakuratan prediksi pada kategori harga tertentu, dan Recall untuk menilai kemampuan model mendeteksi seluruh data pada kelas harga yang sebenarnya. Selain itu, nilai F1-Score digunakan sebagai rata-rata harmonis untuk menyeimbangkan presisi dan *recall* (Banjarnahor et al., 2025). Dengan integrasi teknik-teknik ini, implementasi Random Forest diharapkan mampu memberikan hasil klasifikasi yang sangat akurat, sebagaimana dibuktikan dalam berbagai studi yang mencapai tingkat akurasi di atas 90% hingga 97% pada berbagai kasus klasifikasi kompleks

3. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen berbasis *machine learning* untuk melakukan klasifikasi rentang harga ponsel. Tahapan awal dimulai dengan pengumpulan data dari dataset "train.csv" yang mencakup 2.000 sampel dengan 21 fitur spesifikasi teknis. Proses prapemrosesan data dilakukan dengan melakukan pemeriksaan terhadap nilai yang hilang (*missing values*) serta analisis statistik deskriptif untuk memastikan kualitas data. Selanjutnya, dataset dibagi menjadi dua bagian utama menggunakan teknik *train-test split*, di mana 80% data digunakan sebagai data latih untuk membangun model, dan 20% sisanya dialokasikan sebagai data uji untuk mengevaluasi kinerja prediksi yang dihasilkan.

4. HASIL DAN PEMBAHASAN

Hasil

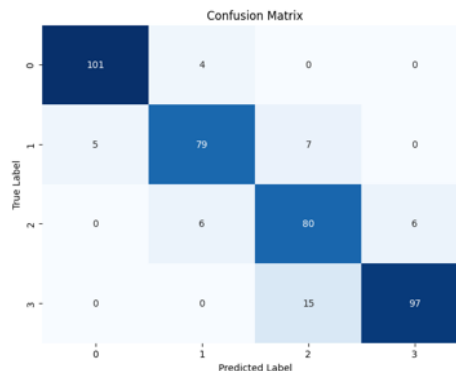
Load Data

index	battery_power	blue	clock_speed	dual_sim	fc	four_g	int_memory	m_dep	mobile_wt	n_cores	pc	px_height	px_width	ram	sc_h	sc_w	talk_time	three_g	touch_screen	wifi
0	842	0	2.2	0	1	0	7	0.6	188	2	2	20	756	2549	9	7	19	0	0	1
1	1021	1	0.5	1	0	1	53	0.7	136	3	6	905	1988	2631	17	3	7	1	1	0
2	563	1	0.5	1	2	1	41	0.9	145	5	6	1263	1716	2603	11	2	9	1	1	0
3	615	1	2.5	0	0	0	10	0.8	131	6	9	1216	1786	2769	16	8	11	1	1	0
4	1821	1	1.2	0	13	1	44	0.6	141	2	14	1208	1212	1411	8	2	15	1	1	0

Gambar 1. Proses load data.

Proses ini dilakukan dengan menggunakan pustaka *pandas* untuk membaca file bernama train.csv ke dalam sebuah *DataFrame*. Tujuan utama dari langkah ini adalah untuk melakukan inspeksi struktur data melalui fungsi *df.head()*, yang menampilkan lima baris pertama dari dataset tersebut. Berdasarkan tampilan tersebut, diketahui bahwa dataset ini memiliki struktur yang terdiri dari 2000 entri dan 21 kolom spesifikasi teknis, seperti daya baterai (*battery power*), kecepatan prosesor (*clock speed*), dan memori internal. Tahap ini sangat krusial sebagai fondasi sebelum melanjutkan ke proses analisis data eksploratif (EDA) dan pelatihan model *Random Forest*.

Confusion Matrix

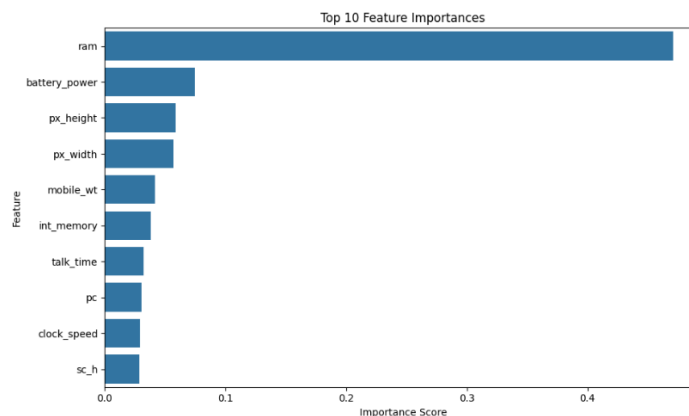


Gambar 2. Confusion Matrix.

Visualisasi *Confusion Matrix* memberikan gambaran detail mengenai performa klasifikasi model Random Forest pada setiap kategori rentang harga (0, 1, 2, dan 3). Secara keseluruhan, model menunjukkan kinerja yang sangat baik, ditandai dengan dominasi angka pada garis diagonal utama, yang merepresentasikan jumlah prediksi yang benar (*True Positives*). Sebagai contoh, pada kategori harga rendah (label 0), model berhasil memprediksi 101 data dengan benar, dan pada kategori harga tertinggi (label 3), model mencatat 97 prediksi akurat. Hal ini mengindikasikan bahwa model memiliki tingkat Akurasi yang tinggi dalam membedakan kelas-kelas harga berdasarkan spesifikasi teknis yang diberikan.

Namun, terdapat sedikit ambiguitas atau kesalahan klasifikasi (*misclassification*) yang terjadi pada kategori yang berdekatan. Misalnya, pada label 3 (harga tertinggi), terdapat 15 data yang salah diprediksi sebagai label 2. Fenomena ini menunjukkan adanya irisan (*overlap*) spesifikasi teknis antara ponsel kelas menengah-atas dengan kelas *flagship*, sehingga model terkadang mengalami kesulitan dalam membedakan kedua kategori tersebut secara absolut. Meski demikian, fakta bahwa tidak ada data yang terlempar jauh dari kategori aslinya (misalnya, label 0 diprediksi sebagai label 3) membuktikan bahwa model Random Forest ini sangat stabil dan memiliki nilai Precision serta Recall yang seimbang di seluruh kelas.

Feature Importance



Gambar 3. Hasil visualisasi *Feature Importance*.

Hasil visualisasi *Feature Importance* menunjukkan kontribusi relatif dari masing-masing spesifikasi teknis terhadap kemampuan model Random Forest dalam mengklasifikasikan rentang harga ponsel. Dalam grafik tersebut, variabel RAM muncul sebagai fitur paling dominan dengan skor kepentingan yang sangat signifikan, yakni mendekati nilai 0,5. Hal ini mengindikasikan bahwa kapasitas memori akses acak merupakan pembeda utama (*key differentiator*) dalam penentuan kasta harga ponsel dibandingkan fitur lainnya. Dominasi RAM ini menunjukkan bahwa algoritma Random Forest secara konsisten

menggunakan fitur ini sebagai titik pemisah (*split point*) utama pada sebagian besar pohon keputusan yang terbentuk dalam model.

Di bawah variabel RAM, terdapat beberapa fitur lain yang memberikan pengaruh moderat, yaitu *battery_power* (kapasitas baterai), diikuti oleh *px_height* dan *px_width* (resolusi layar). Kehadiran fitur-fitur ini menunjukkan bahwa selain performa komputasi (RAM), daya tahan baterai dan kualitas visual layar merupakan faktor teknis penting yang dipertimbangkan model dalam memprediksi kelas harga. Sementara itu, fitur lain seperti *mobile_wt* (berat ponsel), *int_memory* (memori internal), hingga *sc_h* (tinggi layar) memiliki skor kepentingan yang jauh lebih rendah dan relatif seragam. Secara keseluruhan, analisis ini membuktikan bahwa model tidak hanya mengandalkan satu parameter, namun tetap menempatkan spesifikasi perangkat keras inti sebagai parameter paling krusial dalam struktur hirarki pengambilan keputusan klasifikasi harga.

Pembahasan

Berdasarkan hasil pengujian yang telah dilakukan, model *Random Forest* menunjukkan performa yang sangat impresif dalam mengklasifikasikan rentang harga ponsel dengan tingkat akurasi mencapai 89%. Keberhasilan model ini terlihat jelas melalui matriks kebingungan (*confusion matrix*), di mana model mampu mengenali pola spesifikasi teknis pada setiap kategori harga dengan konsisten. Hasil ini menegaskan bahwa algoritma *Random Forest* memiliki ketahanan yang baik terhadap data multivariat yang kompleks, mampu meminimalisir kesalahan klasifikasi secara signifikan dibandingkan dengan model pohon keputusan tunggal.

Analisis lebih mendalam terhadap tiap kelas harga menunjukkan bahwa model memiliki kemampuan prediksi yang hampir sempurna pada segmen harga ekstrem, yaitu Kelas 0 (Murah) dan Kelas 3 (*Flagship*). Hal ini dibuktikan dengan nilai *F1-score* pada Kelas 0 yang mencapai 0.96. Tingginya akurasi pada segmen ini mengindikasikan bahwa ponsel dengan harga terendah dan tertinggi memiliki karakteristik spesifikasi teknis yang sangat kontras dan distingtif, sehingga algoritma dapat dengan mudah membedakan batas-batas nilai fitur pada kedua kategori tersebut.

Namun demikian, terdapat sedikit tantangan pada klasifikasi untuk Kelas 1 (Menengah) dan Kelas 2 (Tinggi), di mana terjadi beberapa misklasifikasi antar kedua kelas tersebut. Sebagai contoh, sejumlah data pada Kelas 2 terprediksi secara salah masuk ke dalam Kelas 1 atau Kelas 3. Fenomena ini menunjukkan adanya persinggungan atau *overlap* pada spesifikasi teknis untuk ponsel di segmen harga menengah-atas. Hal ini selaras dengan kondisi pasar saat ini di mana perbedaan fitur antara ponsel kelas menengah dan kelas atas sering kali sangat tipis,

yang kemudian menjadi tantangan tersendiri bagi model dalam menentukan batas keputusan (*decision boundary*) yang tegas.

Salah satu temuan paling krusial dalam penelitian ini adalah hasil analisis *Feature Importance* yang menunjukkan dominasi mutlak variabel RAM sebagai penentu harga. Kapasitas RAM memberikan kontribusi sebesar 47% terhadap keputusan model, jauh mengungguli fitur lain seperti daya baterai (7,4%) dan resolusi layar (5,8%). Hasil ini memberikan bukti empiris bahwa dalam industri ponsel cerdas, kapasitas memori akses acak merupakan indikator utama yang digunakan produsen untuk memposisikan produk mereka dalam rentang harga tertentu. Hal ini berkaitan erat dengan kebutuhan pengguna akan performa *multitasking* yang semakin meningkat.

Secara keseluruhan, hasil penelitian ini memperkuat teori bahwa pendekatan *machine learning* berbasis *ensemble* sangat efektif dalam memberikan prediksi harga yang objektif. Dibandingkan dengan penelitian terdahulu oleh Saputra & Sudiatmika (2024), penggunaan *Random Forest* dalam studi ini berhasil memetakan fitur-fitur teknis secara lebih granular melalui bobot kepentingan fitur yang jelas. Temuan ini tidak hanya bermanfaat untuk pengembangan model prediksi yang lebih akurat di masa depan, tetapi juga memberikan wawasan strategis bagi konsumen dan pelaku industri dalam memahami struktur harga ponsel cerdas berdasarkan kekuatan spesifikasi teknisnya.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian dan analisis yang telah dilakukan, dapat disimpulkan bahwa implementasi algoritma *Random Forest* terbukti sangat efektif dalam mengklasifikasikan rentang harga ponsel cerdas berdasarkan spesifikasi teknisnya dengan capaian akurasi keseluruhan sebesar 89%. Penggunaan metode *ensemble* ini mampu menangani kompleksitas fitur teknis secara objektif, di mana model menunjukkan performa prediksi yang sangat stabil terutama pada kategori harga ekstrem seperti kelas murah dan *flagship*. Temuan paling signifikan dalam studi ini mengonfirmasi bahwa kapasitas RAM merupakan variabel determinan utama dalam penentuan harga pasar dengan skor kepentingan mencapai 47,09%, yang menunjukkan bahwa performa memori menjadi tolok ukur primer nilai ekonomi sebuah perangkat seluler.

Sebagai langkah pengembangan di masa depan, peneliti menyarankan dilakukannya optimasi model melalui teknik penalaan hiperparameter seperti *GridSearchCV* untuk meningkatkan presisi pada klasifikasi kelas menengah dan tinggi yang masih memiliki sedikit ambiguitas atau tumpang tindih spesifikasi. Selain itu, penambahan variabel fitur modern

seperti jenis *chipset*, dukungan teknologi 5G, serta tipe layar (OLED/AMOLED) sangat dianjurkan agar model tetap relevan dengan tren industri yang dinamis.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Program Studi Rekayasa Sistem Komputer, Universitas Institut Bisnis dan Teknologi Indonesia yang telah memberikan arahan dan bimbingan teknis selama proses pengolahan data. Ucapan terima kasih juga disampaikan kepada tim riset yang telah membantu dalam pengumpulan dataset serta seluruh dosen yang memberikan masukan terhadap metodologi *machine learning* yang digunakan. Semoga hasil penelitian ini dapat memberikan kontribusi nyata bagi perkembangan literatur informatika di masa depan.

DAFTAR REFERENSI

- Aprilianto, D., & Rizal, E. (2025). Klasifikasi penyakit kanker paru menggunakan algoritma random forest berbasis Streamlit. *Metik Jurnal (Akreditasi Sinta 3)*, 9(2), 275-283. <https://doi.org/10.47002/metik.v9i2.1076>
- Aulia, A., & Wahyuni, S. (2024). Analisis data mining dalam pemilihan smartphone dan klasifikasi di berbagai perangkat. 5(4), 333-339. <https://doi.org/10.47065/bit.v5i2.1703>
- Banjarnahor, E., Belferik, R., Cendana, W., & Abraham, Y. A. S. (2025). Analisis implementasi support vector machine dan random forest untuk prediksi kategori indeks kualitas udara Jakarta. *Jurnal INSTEK (Informatika Sains Dan Teknologi)*, 10(1), 175-184. <https://doi.org/10.24252/instek.v10i1.56477>
- Didayat, Sunyoto, A., & Al Fatah, H. (2023). Jantung RF. *VII*(September), 31-40. <https://doi.org/10.47970/siskom-kb.v7i1.464>
- Fitri, R. R., Apriandari, W., Studi, P., Informatika, T., Sukabumi, U. M., Sukabumi, K., & Barat, J. (2025). Penggunaan random forest dalam sistem. 13(3). <https://doi.org/10.23960/jitet.v13i3.6905>
- Indrasetianingsih, A., Hermanto, E. M. P., Ilham, M., Rahmawati, N., & Hariyanti, I. A. (2025). Inti Nusa Mandiri. *LPPM Nusa Mandiri*, 19(2), 146-153. <https://doi.org/10.33480/inti.v19i2.5283>
- Oriana, L., Dwi, A., Wati, S., Ramahdani, A. P., Safira, N. N., & Ismanto, E. (2025). Peningkatan kinerja model random forest untuk deteksi kecurangan kartu kredit menggunakan RandomizedSearchCV. *Jurnal Fasilkom*, 15(2), 229-237. <https://doi.org/10.37859/jf.v15i2.9832>
- Rahmada, A., & Susanto, E. R. (2025). Peningkatan akurasi prediksi penyakit jantung dengan teknik SMOTEENN pada algoritma random forest. *Jurnal Pendidikan Dan Teknologi Indonesia*, 4(12), 795-803. <https://doi.org/10.52436/1.jpti.524>
- Saputra, P., & Sudiatmika, I. (2024). Analisis prediksi harga smartphone tahun 2023 menggunakan model random forest regression berdasarkan fitur-fitur spesifikasi teknis. *Jurnal Komputer Dan Teknologi Sains (KOMTEKS)*, 3(2), 13-17.

<https://www.kaggle.com/datasets/howisusmanali/mo>
<https://doi.org/10.37637/komteks.v3i2.2233>

- Sinambela, D. P., Naparin, H., Zulfadhilah, M., & Hidayah, N. (2023). Implementasi algoritma decision tree dan random forest dalam prediksi perdarahan pascasalin. *Jurnal Informasi Dan Teknologi*, 5(3), 58-64. <https://doi.org/10.60083/jidt.v5i3.393>
- Triansyah, A. A., Mustika, F. T., Meilinda, S., Anjani, S. P., & Dzakkiya, Y. (2025). *Jurnal Inovasi Pendidikan Kreatif*. Perubahan tren digital, teknologi dalam kesehatan jasmani saat pandemi Covid-19, 6(2), 239-250. <https://ijurnal.com/1/index.php/jipk>
- Turino, T., Saputro, R. E., & Karyono, G. (2025). Penerapan model ensemble learning dengan random forest dan multi-layer perceptron untuk prediksi gempa. *Jurnal Pendidikan Dan Teknologi Indonesia*, 5(2), 493-504. <https://doi.org/10.52436/1.jpti.667>
- Warjiyono, W., Rais, A. N., Robi, I. A., Hadi, S. W., & Kurniawan, W. (2024). Prediksi harga rumah menggunakan metode random forest. *Jurnal Sistem Informasi Dan Teknologi Informasi*, 6(2), 416-423. <https://doi.org/10.52005/jursistekni.v6i2.323>
- Zein, A., & Ekawati, F. (2025). Prediksi harga handphone berbasis algoritma supervised learning. *Sainstech: Jurnal Penelitian Dan Pengkajian Sains Dan Teknologi*, 35(2), 27-33. <https://doi.org/10.37277/stch.v35i2.2343>
- Zildan T. M, M. (2025). Penggunaan algoritma random forest untuk deteksi malware berdasarkan pola permission pada aplikasi android komunikasi. *Jurnal Cakrawala Informasi*, 5(1), 30-49. <https://doi.org/10.54066/jci.v5i1.559>