

Evaluasi Kinerja Metode Random Forest dan Decision Tree dalam Klasifikasi Menu Fast Food di Restoran Indonesia

Siti Arnila Pebrianti¹, Putri AnandaSiti Arnila Pebrianti², Nani Sulistianingsih^{3*}

^{1,2,3}Sistem dan Teknologi Informasi, Universitas Muhammadiyah Mataram, Indonesia

¹arnilapebrianti@gmail.com , ²putrianandaa175@gmail.com, ³nani.sulistianingsih@ummat.ac.id

ABSTRACT

Keywords:

Fast Food
Classification
Random Forest
Decision Tree

Abstract: *Fast food menu classification is an important aspect of the restaurant industry that helps improve operational efficiency, marketing strategies, and the understanding of customer consumption patterns. This research evaluates the performance of two machine learning algorithms, namely Decision Tree and Random Forest, in classifying fast food menus based on attributes such as calories, fat, protein, and sodium content. The dataset used consists of 17 features and 516 menu items, which have undergone cleaning and pre-processing. The research result indicate that the Random Tree algorithm outperforms the Decision Tree in classifying fast food menus. Based on the evaluation metrics of accuracy, precision, recall, and F1-score, Random Forest achieved 93% accuracy, while Decision Tree achieved 64% accuracy. Correlation between features were also analyzed to understand the relationship between nutritional components in fast food. The conclusion of this research highlights that Random Forest is superior in handling data complexity and can be reliably implemented in restaurant menu recommendation systems.*

Kata Kunci:

Makanan Cepat Saji
Klasifikasi
Random Forest
Decision Tree

Abstrak: Klasifikasi menu makanan cepat saji merupakan aspek penting dalam industri restoran untuk meningkatkan efisiensi operasional, strategi pemasaran, dan pemahaman pola konsumsi pelanggan. Penelitian ini mengevaluasi kinerja dua algoritma pembelajaran mesin, yaitu *Decision Tree* dan *Random Forest*, dalam mengklasifikasikan menu makanan cepat saji berdasarkan atribut seperti kandungan kalori, lemak, protein, dan natrium. Dataset yang digunakan terdiri dari 17 fitur dan 516 data menu yang telah melalui tahapan pembersihan dan pra pemrosesan. Hasil penelitian menunjukkan bahwa algoritma Random Forest memiliki performa lebih baik dibandingkan dengan *Decision Tree* dalam klasifikasi menu makanan cepat saji. Berdasarkan metrik evaluasi akurasi, presisi, recall, dan F1-Score, *Random Forest* mencapai akurasi 93%, sedangkan *Decision Tree* memperoleh akurasi 64%. Korelasi antar fitur juga dianalisis untuk memahami hubungan antara kandungan nutrisi dalam makanan cepat saji. Kesimpulan penelitian ini menunjukkan bahwa Random Forest lebih unggul dalam menangani kompleksitas data, sehingga dapat diandalkan untuk implementasi dalam sistem rekomendasi menu restoran.

Article History:

Received : 25-02-2025

Accepted : 27-03-2025



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

A. LATAR BELAKANG

Klasifikasi data merupakan salah satu teknik utama dalam *data mining* yang memungkinkan pengelompokan objek berdasarkan karakteristik tertentu. Pada industri makanan cepat saji, klasifikasi menu menjadi elemen penting dalam analisis data restoran untuk mendukung pengambilan keputusan yang lebih efektif. Penerapan metode klasifikasi, restoran dapat mengoptimalkan strategi pemasaran, meningkatkan efisiensi operasional, serta mengidentifikasi pola konsumsi pelanggan. Teknik pembelajaran mesin ataupun *Data Mining* seperti *Decision Tree* dan *Random Forest* telah banyak digunakan dalam proses klasifikasi karena kemampuannya dalam menangani data yang kompleks dan menghasilkan keputusan berbasis data yang akurat (Zebua et al., 2023).

Klasifikasi menu makanan cepat saji memiliki peran krusial dalam analisis data restoran, karena memungkinkan pengambilan keputusan yang lebih efektif terkait pengoptimalan menu, strategi penjualan, serta pemasaran. Pengelompokan item menu restoran dapat mengidentifikasi hidangan yang perlu dipromosikan, dimodifikasi, atau dihapus, sehingga meningkatkan profitabilitas dan kepuasan pelanggan. Salah satu metode yang umum digunakan dalam klasifikasi menu adalah pendekatan rekayasa menu, yang membagi item menjadi beberapa kategori, seperti "bintang (*star*)", "bajak (*plow horse*)", "teka-teki (*puzzle*)" dan "anjing (*dog*)", yang membantu manajemen dalam menentukan strategi yang tepat. Selain itu, penggunaan algoritma *clustering* seperti *K-means* dapat memberikan wawasan yang lebih dalam tentang pola penjualan dan kinerja menu, memungkinkan restoran untuk memahami produk mana yang paling populer dan mana yang kurang diminati (Syahputra, 2022). Pendekatan berbasis data ini mendukung manajemen inventaris dan alokasi sumber daya secara lebih optimal.

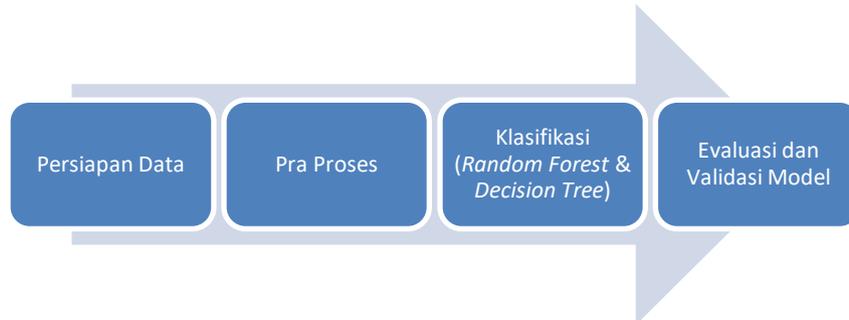
Tren konsumsi makanan cepat saji di Indonesia mengalami peningkatan signifikan akibat urbanisasi, pertumbuhan ekonomi, serta perubahan gaya hidup. Studi menunjukkan bahwa 95% pria dan 98% wanita mengonsumsi makanan cepat saji secara teratur, dengan faktor demografis seperti status perkawinan, tingkat pendidikan, dan pendapatan yang turut mempengaruhi pola konsumsi (Andriyani et al., 2024). Implikasi dari tren ini berkontribusi terhadap meningkatnya prevalensi obesitas, terutama di kalangan remaja, dengan 81,8% dari mereka yang sering mengonsumsi makanan cepat saji dikategorikan sebagai obesitas (Sumiyati et al., 2022). Pada konteks ini, analisis berbasis pembelajaran mesin menjadi relevan untuk memahami pola diet dan faktor risiko obesitas. Teknik seperti *Extreme Gradient Boosting (XGBoost)* telah digunakan untuk mengolah dataset besar dan mengidentifikasi faktor-faktor utama yang mempengaruhi obesitas, sehingga meningkatkan akurasi prediksi risiko kesehatan berbasis demografi dan gaya hidup (Rombe et al., 2022). Mengintegrasikan pembelajaran mesin maupun *data mining*, dengan pendekatan ini dapat dimanfaatkan untuk mendukung kebijakan kesehatan dan intervensi yang lebih efektif.

Analisis menu yang efektif tidak terbatas pada mendukung pengelolaan operasional tetapi juga berfungsi sebagai alat pemasaran yang dapat meningkatkan profitabilitas restoran. Melalui penerapan strategi berbasis data, restoran dapat menyesuaikan harga, memperkenalkan menu baru, serta menghapus hidangan yang kurang menguntungkan. Salah satu pendekatan yang digunakan dalam optimalisasi menu adalah analisis ABC, yang memungkinkan restoran mengidentifikasi hidangan "pemimpin" dan "orang luar" sebagai dasar dalam merancang strategi pemasaran yang lebih terarah. Namun, beberapa pihak berpendapat bahwa ketergantungan berlebihan pada klasifikasi berbasis data dapat menghambat inovasi dalam pengembangan menu. Oleh karena itu, keseimbangan antara keputusan berbasis data dan kreativitas kuliner sangat penting dalam mencapai kesuksesan jangka panjang dalam industri makanan cepat saji (Ivanenko et al., 2022).

Meskipun beberapa penelitian sebelumnya menyoroti pentingnya klasifikasi menu makanan cepat saji serta dampaknya terhadap strategi pemasaran dan kesehatan, masih terdapat celah dalam eksplorasi metode klasifikasi yang paling efektif untuk konteks restoran di Indonesia. Sebagian besar penelitian lebih berfokus pada *clustering* dan analisis prediktif, sementara perbandingan kinerja algoritma klasifikasi *Random Forest* dan *Decision Tree* dalam kategori menu cepat saji masih jarang dilakukan. Oleh karena itu, penelitian ini bertujuan untuk mengevaluasi kinerja kedua metode klasifikasi tersebut dalam mengelompokkan menu makanan cepat saji, dengan mempertimbangkan metrik evaluasi seperti akurasi, presisi, *recall*, dan *F1-score*. Pemahaman metode yang lebih optimal diharapkan dapat membantu industri makanan cepat saji meningkatkan efisiensi operasional, profitabilitas, dan kepuasan pelanggan.

B. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan menerapkan metode klasifikasi berbasis pembelajaran mesin. Alur penelitian dimulai dengan dengan pengumpulan dan pemrosesan data, diikuti dengan penerapan algoritma *Random Forest* dan *Decision Tree* untuk mengklasifikasikan menu cepat saji. Selanjutnya, hasil klasifikasi dianalisis menggunakan berbagai metrik evaluasi guna menentukan metode yang paling optimal dalam konteks penelitian ini. Alur penelitian klasifikasi menu fast food dapat dilihat pada Gambar 1.



Gambar 1. Alur Penelitian Klasifikasi Menu *Fast Food*

Gambar 1 menjelaskan proses klasifikasi dengan Decision Tree diawali dengan analisis data menggunakan penghitungan entropy dan information gain untuk menentukan struktur pohon keputusan. Decision Tree dipilih karena kemampuannya dalam mengidentifikasi pola berdasarkan atribut tertentu secara konseptual dan jelas. Sementara itu, Random Forest digunakan untuk meningkatkan akurasi model dengan menggabungkan hasil dari beberapa pohon keputusan, sehingga mengurangi kemungkinan overfitting pada dataset pelatihan. Evaluasi dan validasi model dilakukan untuk mengukur efektivitas metode yang digunakan. Penilaian dilakukan berdasarkan parameter seperti akurasi, presisi, recall, dan F1-score untuk memastikan kualitas hasil klasifikasi. Model yang dikembangkan diharapkan dapat diimplementasikan dalam industri restoran guna mendukung proses otomatisasi dalam pengelolaan menu makanan cepat saji secara lebih efisien.

C. HASIL DAN PEMBAHASAN

Pada bagian ini akan dibahas dari proses pengujian algoritma *Random Forest* dan *Decision Tree* pada dataset makanan siap saji. Berikut proses yang dilakukan mulai dari pengumpulan dataset hingga evaluasi dan validasi model:

1. Dataset

Dataset yang digunakan dalam penelitian ini diambil dari sumber database publik, terdiri dari 17 kolom dan 516 baris data. Dataset terdiri dari variabel restaurant, item (menu), calories, cal_fat, total_fat, sat_fat, trans_fat, kolesterol, sodium, total_carb, fiber, sugar, protein, vit_a, vit_c, calcium dan salad.

Tabel 1. Dataset Fast Food

Restaurant	Item	Calories	Cal_fat	...	Calcium	Salad
Mcdonalds	Artisan Grilled Chicken Sandwich	380	60	...	20	Other
Mcdonalds	Single Bacon Smokehouse Burger	840	410	...	20	Other
Mcdonalds	Double Bacon Smokehouse Burger	1130	600	...	50	Other
...
Subway	Pepperoni Pizza	790	290	...	60	Other
Subway	Sausage Pizza	820	310	...	60	Other

2. Pembersihan dan Pra-Proses Data

Tahapan selanjutnya adalah pembersihan dan pra-proses data sebelum masuk ke proses pemodelan dengan *Decision Tree* dan *Random Forest*. Berikut langkah-langkah pembersihan dan pra proses data pada dataset makanan cepat saji.

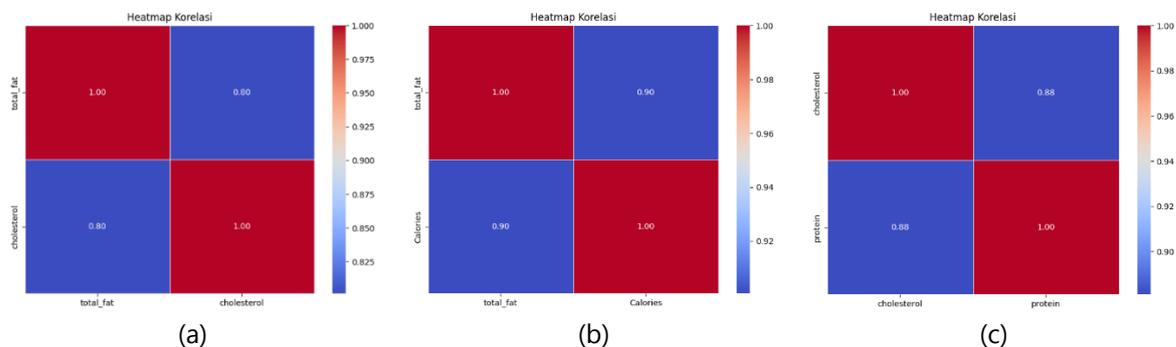
a. Menghapus data duplikat

Pada proses *cleaning data*, menghapus data duplikat adalah salah satu tahapan penting. Data duplikat merupakan baris yang memiliki nilai sama persis di setiap kolom, dalam penelitian ini dataset yang digunakan terdiri dari 516 baris dengan 17 kolom, setelah dilakukan pemeriksaan ditemukan 2 baris yang memiliki duplikat. Setelah penghapusan duplikat, jumlah data yang tersisa yaitu 514 baris, artinya data yang teridentifikasi telah berhasil dihapus. Langkah ini membantu memastikan dataset yang digunakan memiliki kualitas baik dan tidak mengandung informasi yang redundan, yang dapat mempengaruhi keakuratan model yang dibangun.

b. Mendeteksi *Missing Value*

Setelah tahapan menghapus data duplikat maka akan dilanjutkan dengan mendeteksi nilai yang hilang atau missing value. Data yang hilang bisa disebabkan oleh beberapa faktor seperti kesalahan input, dan kehilangan data pada saat proses pengumpulan (Saputra, 2023). Terdapat beberapa variabel yang memiliki *missing value* seperti fiber, protein, vit_a, vit_c dan calcium.

c. Korelasi antar variabel

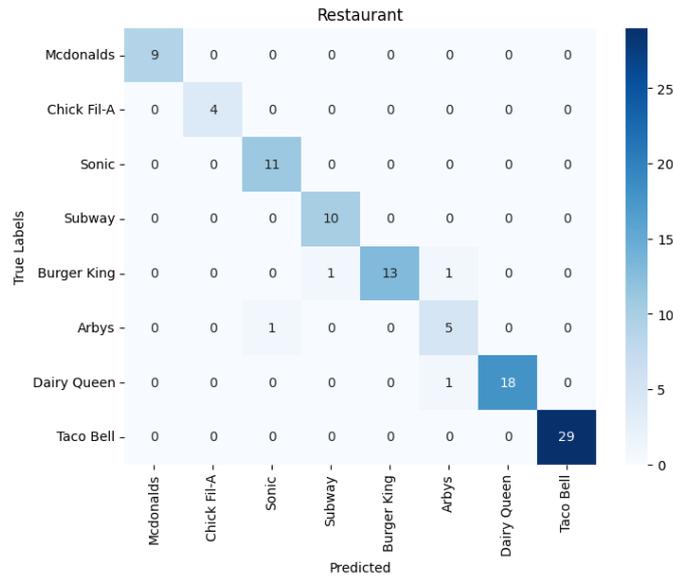


Gambar 2. Korelasi antar variabel, total_fat dan kolesterol (a), total_fat dan calories (b), kolesterol dan protein (c)

Berdasarkan gambar pada point (a) menunjukkan nilai korelasi 0.80, artinya bahwa terdapat hubungan positif yang akurat antara lemak total (*total_fat*) dan kadar kolesterol (*cholesterol*). Semakin tinggi total lemak dalam suatu makanan, semakin tinggi juga kadar kolesterol dan sebaliknya. Hal ini dapat memicu penyakit jantung koroner dengan kombinasi lemak, kolesterol dan trigliserida salah satu komponen kadar lemak dalam darah (Intan Sahara et al., 2021). Pada point (b), terlihat antara variabel *total_fat* dan *calories* menunjukkan nilai korelasi positif yang sangat kuat yaitu 0.90. Hal ini menunjukkan bahwa lemak merupakan kontributor utama dalam jumlah kalori suatu makanan. Mengurangi jumlah total lemak pada makanan akan menjadi salah satu strategi efektif (Kamilah et al., 2024). Terakhir pada point (c) menunjukkan korelasi antara variabel *cholesterol* dan *protein*. Pada heatmap tersebut nilai korelasi sebesar 0.88, artinya bahwa makanan yang memiliki kadar kolesterol tinggi cenderung memiliki kandungan protein yang tinggi. Berdasarkan informasi yang diperoleh melalui penerapan pendekatan korelasi *Pearson*, hubungan antara variabel ini memberikan wawasan baru tentang peran penting kandungan nutrisi dalam makanan cepat saji dalam menciptakan pilihan yang lebih sehat untuk dikonsumsi (Rakhmawati & Sulistyoningih, 2020).

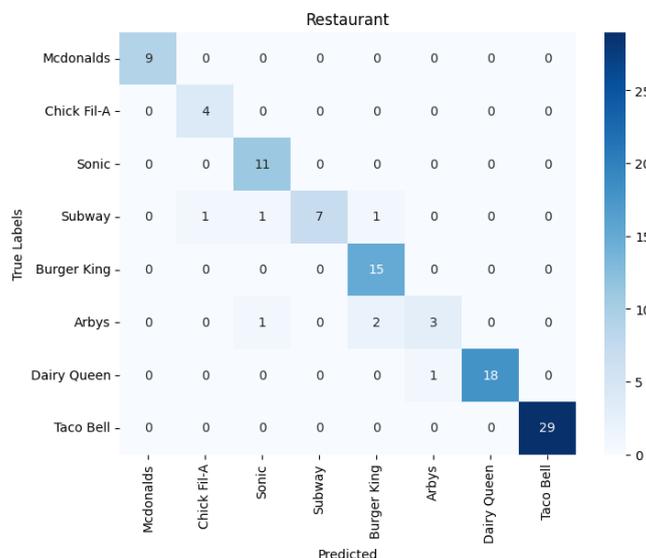
3. Klasifikasi dengan Algoritma Decision Tree dan Random Forest

Pada penelitian ini, dilakukan perbandingan antara dua algoritma klasifikasi yaitu *Decision Tree* dan *Random Forest* pada dataset makanan cepat saji dari berbagai restoran di Indonesia. Berdasarkan hasil pengujian diperoleh akurasi 64% dengan model *Decision Tree* dan *Random Forest* sebesar 93%. Berikut hasil pengujian yang ditunjukkan dengan confusion matrix pada algoritma *Decision Tree* dan *Random Forest*.



Gambar 3. Confusion Matrix Klasifikasi dengan *Decision Tree*

Terlihat pada Gambar 3, bahwa model mengklasifikasikan beberapa kategori dengan benar, namun jumlahnya terbatas. *Taco Bell* diprediksi dengan benar, namun beberapa kategori lain memiliki jumlah yang lebih sedikit.



Gambar 4. Confusion Matrix Klasifikasi dengan *Decision Tree*

Pada Gambar 4 terlihat bahwa kategori restoran memiliki banyak prediksi yang benar. Kelas *Taco Bell*(29), *Dairy Queen* (18) dan *Burger King* (15) diklasifikasikan dengan baik walaupun Subway mengalami sedikit kesalahan prediksi, tetapi memiliki jumlah klasifikasi yang cukup baik. Melalui pendekatan ensemble yang menggabungkan banyak pohon keputusan, *Random Forest* mampu

menangkap lebih banyak pola dari data. *Random Forest* dapat menangani variasi dalam fitur dengan baik, menghasilkan prediksi lebih akurat untuk setiap restoran. *Random Forest* adalah pilihan yang lebih baik dalam mengklasifikasikan makanan cepat saji berdasarkan restoran asalnya. Berdasarkan hasil yang diperoleh dapat membantu dalam memahami bagaimana setiap restoran dapat membedakan menu mereka dari kompetitor.

D. SIMPULAN DAN SARAN

Berdasarkan hasil penelitian yang dilakukan dalam klasifikasi menu makanan cepat saji di restoran Indonesia, dapat disimpulkan bahwa kinerja metode *Random Forest* lebih baik dibandingkan dengan kinerja metode *Decision Tree*. Hasil evaluasi menggunakan metrik akurasi, metode *Random Forest* mencapai akurasi sebesar 93%, sedangkan dengan metode *Decision Tree* memperoleh akurasi 64%. Perbedaan signifikan ini menunjukkan bahwa *Random Forest* lebih andal dalam menangani kompleksitas data serta mampu menghasilkan prediksi yang lebih akurat. Kenggulan *Random Forest* dapat dikaitkan dengan kemampuannya dalam menggabungkan banyak pohon keputusan (*ensemble learning*), untuk meningkatkan stabilitas model. Oleh sebab itu, dalam konteks klasifikasi menu makanan cepat saji, metode *Random Forest* merupakan pilihan yang lebih optimal untuk di terapkan dalam industri restoran dalam meningkatkan efisiensi operasional, strategi pemasaran, serta pemahaman pola konsumsi pelanggan.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Bapak Ibu Nani Sulistianingsih, S.Kom.,M.Eng selaku pembimbing yang senantiasa memberikan bimbingan, nasehat, dan motivasi kepada penulis sehingga penelitian ini selesai dengan baik. Penulis juga mengucapkan terima kasih kepada seluruh pihak yang telah mendukung dan berkontribusi dalam penyusunan paper ini.

REFERENSI

- Andriyani, A., Lee, Y. Z., Win, K. K., Tan, C. K., Amini, F., Tan, E. S. S., Thiagarajah, S., Ng, E. S. C., & Ahmad Bustami, N. (2024). Fast food consumption, obesity and nutrient intake among adults in Indonesia. *Food Research*, 8, 55–65. [https://doi.org/10.26656/fr.2017.8\(S3\).5](https://doi.org/10.26656/fr.2017.8(S3).5)
- Azwanti, N., & Elisa, E. (2020). Analisa Kepuasan Konsumen Menggunakan Algoritma C4.5. *SNISTEK*.
- Bagus, I., Dharma Putra, W., Astina, A., Made, G., & Arnawa, S. (n.d.). Analisis Strategi Menu Engineering dalam Upaya Meningkatkan Pendapatan Restaurant Menu Engineering Strategy Analysis in an Effort to Increase Restaurant Revenue. *Jurnal Pariwisata Dan Bisnis*, 02(6), 1444–1457. <https://doi.org/10.22334/paris.v2i6>
- Fahmi, M., Gozali, G., Syaquy, D., & Kurniawan, W. (2025). *Klasifikasi Kelayakan Konsumsi Susu Kambing Etawa Pasteurisasi Berdasarkan Warna, Aroma, dan PH dengan Metode Random Forest Berbasis Arduino* (Vol. 9, Issue 2). <http://j-ptiik.ub.ac.id>
- Firmansyach, W. A., Hayati, U., & Wijaya, Y. A. (2023). Analisa Terjadinya Overfitting Dan Underfitting Pada Algoritma Naive Bayes Dan Decision Tree Dengan Teknik Cross Validation. In *Jurnal Mahasiswa Teknik Informatika* (Vol. 7, Issue 1).
- Intan Sahara, L., Adelina, R., Gizi Poltekkes Kemenkes Malang, J., & Sarjana Terapan Gizi dan Dietetika, P. (2021). *Analisis Asupan Lemak Terhadap Profil Lemak Darah Berkaitan Dengan Kejadian Penyakit Jantung Koroner (Pjk) Di Indonesia: Studi Literatur Analysis of Fat Intake on Lipid Profile Related to Coronary Heart Disease (CHD) in Indonesia: Literature Study* (Vol. 1, Issue 2). <http://journal.binawan.ac.id/JAKAGI>
- Ivanenko, V. O., Kaschuck, K. M., Botsian, T. V., & Klimova, I. O. (2022). Menu Analysis as an Effective Marketing Tool for Increasing the Restaurant Establishments' Profitability. *Business Inform*, 12(539), 258–263. <https://doi.org/10.32983/2222-4459-2022-12-258-263>
- Kamilah, W., Hartani, P., Arifyan, M., Hartanto, N., Jamil, D. F., Rinaket, J., Abdillah, A. S., Fawaz Andana, S., Scrowbaneg, L., Kesehatan, P. J., Rekreasi, D., & Semarang, U. N. (2024). Pentingnya Mengonsumsi Makanan Rendah Lemak dan Kalori untuk Menjaga dan Meningkatkan Energi dalam Tubuh Sebelum Melakukan Olahraga terhadap Mahasiswa Fakultas Ilmu Keolahragaan di Universitas Negeri Semarang. In *Jurnal Analis* (Vol. 3, Issue 2). <http://jurnalilmiah.org/journal/index.php/Analis>
- Kumala Sari, P., & Randy Suryono, R. (2024). *Komparasi Algoritma Support Vector Machine Dan Random Forest Untuk Analisis Sentimen Metaverse* (Vol. 7, Issue 1).
- Rakhmawati, R., & Sulistyoningsih, M. (2020). Kandungan Kolesterol Darah pada Berbagai Jenis Ayam Konsumsi. *Titian Ilmu: Jurnal Ilmiah Multi Sciences*, 12(1), 31–34. <https://doi.org/10.30599/jti.v12i1.590>

- Ramadhon, R. N., Ogi, A., Agung, A. P., Putra, R., Febrihartina, S. S., & Firdaus, U. (2024). *Implementasi Algoritma Decision Tree untuk Klasifikasi Pelanggan Aktif atau Tidak Aktif pada Data Bank* (Vol. 3).
- Rombe, Y., Thamrin, S. A., & Lawi, A. (2022). Application of Adaptive Synthetic Nominal and Extreme Gradient Boosting Methods in Determining Factors Affecting Obesity: A Case Study of Indonesian Basic Health Research Survey 2013. *Indonesian Journal of Statistics and Its Applications*, 6(2), 309–317. <https://doi.org/10.29244/ijisa.v6i2p309-317>
- Saputra, I. (2023). *Belajar Mudah Data Mining Untuk Pemula*. Informatika Bandung.
- Sumiyati, I., Anggriyani, A., & Mukhsin, A. (2022). Hubungan Antara Konsumsi Makanan Fast Food Dengan Kejadian Obesitas Pada Remaja. *JUMANTIK (Jurnal Ilmiah Penelitian Kesehatan)*, 7(3), 242. <https://doi.org/10.30829/jumantik.v7i3.11485>
- Supriyadi, R., Gata, W., Maulidah, N., Fauzi, A., Komputer, I., & Nusa Mandiri Jalan Margonda Raya No, S. (2020). *Penerapan Algoritma Random Forest Untuk Menentukan Kualitas Anggur Merah*. 13(2), 67–75. <http://journal.stekom.ac.id/index.php/E-Bisnis/page67>
- Syahputra, H. (2022). Clustering Tingkat Penjualan Menu (Food and Beverage) Menggunakan Algoritma K-Means. *Jurnal KomtekInfo*, 29–33. <https://doi.org/10.35134/komtekinfo.v9i1.274>
- Yaman, N. I., Juwita, A. R., Lestari, S. A. P., & Faisal, S. (2024). Perbandingan Kinerja Algoritma Decision Tree dan Random Forest untuk Klasifikasi Nutrisi pada Makanan Cepat Saji. *Jurnal Algoritma*, 21(2), 184–196. <https://doi.org/10.33364/algoritma/v.21-2.1649>
- Zebua, R. S. Y., Hendriyani, C., Sukmadewi, R., Thaha, A. R., Tahir, R., Purbasari, R., Novel, N. J. A., Dewintari, P., Paramita, C. C. P., Hierdwati, T., & Subagja, A. D. (2023). *Bisnis Digital : Startegi Administrasi Bisnis Digital Untuk Menghadapi Masa Depan* (E. Efitra & S. Sepriano, Eds.).