

Submit : 06 Juli 2024

# Analisis Dan Implementasi Algoritma Logistik Regresi Dalam Mengklasifikasi Calon Konsumen Barbershop Di Kota Medan Sebagai Ide Pengembangan Bisnis Aplikasi Pangkasin

<sup>1</sup> Farhan Al Ikhsan, <sup>2</sup> Halim Maulana,  
Universitas Muhammadiyah Sumatera Utara  
Medan, Indonesia

[farhansan47@gmail.com](mailto:farhansan47@gmail.com), [halimmaulana@umsu.ac.id](mailto:halimmaulana@umsu.ac.id)

## ABSTRAK

Perkembangan teknologi digital mendorong transformasi di berbagai sektor, termasuk jasa pangkas rambut (barbershop). Startup Pangkasin hadir sebagai aplikasi yang menawarkan layanan reservasi dan pangkas rambut berbasis digital. Namun, tantangan muncul dalam mengidentifikasi konsumen potensial yang berminat menggunakan aplikasi ini. Penelitian ini bertujuan mengimplementasikan algoritma Logistik Regresi untuk mengklasifikasikan calon konsumen barbershop di Kota Medan berdasarkan data observasi yang meliputi umur, jenis kelamin, dan frekuensi cukur rambut. Penelitian ini menggunakan metode observasi langsung di delapan barbershop serta analisis data dengan Python. Data dibagi menjadi data latih (training) dan data uji (testing) dengan perbandingan 70:30. Hasil penelitian menunjukkan model Logistik Regresi yang dikembangkan mampu mengklasifikasikan konsumen potensial dengan tingkat akurasi 80%, recall 91,8%, presisi 85%, dan F1-Score 88%. Karakteristik pengguna potensial aplikasi Pangkasin cenderung berumur 20–35 tahun, pria, dan memiliki frekuensi cukur 2–3 kali per bulan. Hasil penelitian ini diharapkan dapat menjadi acuan dalam pengembangan strategi bisnis digital pada sektor jasa pangkas rambut.

**Kata Kunci:** logistik regresi, klasifikasi, barbershop, pangkasin, machine learning, python

## PENDAHULUAN

Perkembangan teknologi digital telah membawa perubahan signifikan dalam berbagai aspek kehidupan, termasuk sektor jasa. Industri barbershop, yang masih mengandalkan sistem antrian manual tradisional, menghadapi tantangan dalam mengadopsi digitalisasi. Sistem manual ini sering mengakibatkan waktu tunggu yang tidak pasti, menurunkan tingkat kepuasan pelanggan, dan mengurangi loyalitas terhadap layanan.

Data menunjukkan bahwa penetrasi pengguna internet di Indonesia terus meningkat. Pengguna internet mencapai 88 juta jiwa dari total populasi 258 juta jiwa hingga Juni 2016. Pada periode 2019-2020, penetrasi meningkat menjadi 73,7%. Hal ini menunjukkan peluang besar dalam penerapan teknologi digital di berbagai sektor, termasuk jasa pangkas rambut.

Startup Pangkasin hadir sebagai solusi aplikasi jasa cukur rambut yang menawarkan layanan hairstylist ke rumah pelanggan atau penjadwalan di barbershop mitra.

Namun, tantangan utama yang dihadapi adalah kesulitan dalam mengidentifikasi dan mengelompokkan calon pengguna aplikasi potensial secara otomatis berdasarkan karakteristik mereka.

Dalam menghadapi tantangan identifikasi dan pengelompokan calon pengguna potensial, Pangkasin memerlukan pendekatan berbasis data yang sistematis dan terukur. Penerapan teknik customer segmentation menggunakan algoritma machine learning menjadi solusi strategis untuk mengklasifikasikan calon pengguna berdasarkan karakteristik demografis, perilaku, dan preferensi layanan mereka. Metode seperti K-Means Clustering atau algoritma klasifikasi lainnya dapat menganalisis pola data pengguna untuk membentuk segmen-segmen yang homogen, sehingga memudahkan tim pemasaran dalam merancang strategi komunikasi dan promosi yang tepat sasaran. Dengan segmentasi yang akurat, Pangkasin dapat mengoptimalkan alokasi sumber daya pemasaran, meningkatkan efektivitas akuisisi pengguna baru, serta mempersonalisasi pengalaman layanan sesuai dengan kebutuhan spesifik setiap segmen pelanggan. Implementasi sistem ini tidak hanya meningkatkan efisiensi operasional startup, tetapi juga berkontribusi pada peningkatan kepuasan pelanggan melalui layanan yang lebih relevan dan responsif terhadap ekspektasi pasar yang beragam.

## TINJAUAN PUSTAKA

### Logistik Regresi Biner

Regresi logistik biner adalah metode statistik yang digunakan untuk menganalisis hubungan antara variabel independen dan variabel dependen biner. Dalam konteks ini, variabel dependen adalah niat beli ulang pelanggan (biner: ya/tidak), variabel independen adalah karakteristik produk. Metode ini cocok untuk situasi di mana hasil yang diprediksi hanya memiliki dua kemungkinan, dan memberikan estimasi probabilitas bahwa pelanggan akan melakukan pembelian ulang berdasarkan karakteristik produk.

Selain itu Metode Logistic Regression dipilih di penelitian ini karena sederhana namun efektif untuk klasifikasi biner. Algoritma ini termasuk yang paling luas digunakan dalam tugas klasifikasi biner karena kemudahannya dalam implementasi dan interpretasi. Model logistik regresi juga menghasilkan output probabilitas yang dapat membantu memahami seberapa yakin model terhadap prediksinya – hal ini penting untuk pengambilan keputusan bisnis.

Model regresi logistik biner ini dimanfaatkan agar menguji hubungan dari satu variabel dependen dengan beberapa variabel yang independen. Apabila variabel dependen menghasilkan dua kategori dengan nilai 0 dan 1, digunakan model regresi logistik biner, model ini mengikuti distribusi Bernoulli.

### Klasifikasi

Klasifikasi merupakan cara pengelompokan benda berdasarkan ciri – ciri yang dimiliki oleh objek klasifikasi. Dalam prosesnya, klasifikasi dapat dilakukan dengan banyak cara baik secara manual ataupun dengan bantuan teknologi. Klasifikasi yang dilakukan secara manual adalah klasifikasi yang dilakukan oleh manusia tanpa adanya bantuan dari algoritma cerdas komputer. Klasifikasi yang dilakukan dengan bantuan teknologi, memiliki beberapa algoritma, diantaranya Naïve Bayes, Support Vector Machine, Decision Tree, Fuzzy dan Jaringan Saraf Tiruan.

Klasifikasi adalah tipe analisis data yang dapat membantu orang menentukan kelas label dari sampel yang ingin di klasifikasi. Klasifikasi merupakan Metode supervised

learning, metode yang mencoba menemukan hubungan antara atribut masukan dan atribut target. Tujuan klasifikasi untuk meningkatkan kehandalan hasil yang diperoleh dari data.

Dalam analisis data, klasifikasi merupakan tipe analisis yang dapat membantu menentukan kelas label dari sampel yang ingin diklasifikasi, di mana metode ini termasuk dalam supervised learning yang berusaha menemukan hubungan antara atribut masukan dan atribut target dengan tujuan untuk meningkatkan kehandalan hasil yang diperoleh dari data. Proses klasifikasi melibatkan beberapa tahapan penting, antara lain mendefinisikan kelas, menentukan hubungan antara kelas seperti hubungan hierarki, membuat sistem klasifikasi, dan menetapkan elemen atau dokumen untuk sebuah kelas dalam suatu sistem klasifikasi. Proses ini analog dengan mendefinisikan konsep-konsep, menentukan hubungan semantik antara konsep-konsep, dan menentukan elemen-elemen yang termasuk dalam konsep tertentu. Sebagai ilustrasi, seperti yang dikemukakan oleh Ingetraut Dahlberg, ketika sebuah konsep dipilih untuk sebuah kelas, konsep tersebut merujuk pada sesuatu yang harus dimiliki bersama oleh semua dokumen dari kelas tersebut, misalnya dalam sistem Dewey Decimal Classification (DDC), label "011" merujuk pada kelas dokumen yang memiliki fitur umum berupa bibliografi dan bukan tentang bibliografi itu sendiri.

### **Customer Relationship Management (CRM)**

Setelah era digital yang ditandai dengan penggunaan awalan "e-" pada berbagai istilah seperti e-commerce, e-book, e-sales, dan e-mail, muncul akronim baru yang menjadi sangat populer di kalangan profesional sales dan marketing, yaitu Customer Relationship Management atau yang disingkat CRM. Jika diterjemahkan ke dalam bahasa Indonesia, istilah ini dapat dipahami sebagai Manajemen Hubungan Pelanggan (MHP). Untuk memahami konsep CRM secara menyeluruh, perlu dilakukan telaah terhadap setiap kata yang membentuknya. Kata pertama, "pelanggan" atau "customer", dalam kamus bahasa Inggris didefinisikan sebagai seseorang yang berulang kali atau secara teratur melakukan pembelian kepada seorang pedagang. Definisi ini berfokus pada orangnya sebagai subjek, tanpa menyinggung aspek-aspek lain seperti kepuasan, mahal atau murahnya pembelian, dan sebagainya.

Hubungan atau relationship, adalah bentuk komunikasi dua arah antara pembeli dan penjual. Manajemen, kata terakhir, artinya pengelolaan (secara luas tanpa perlu menjabarkan detail bagaimana mengelola sesuatu). Jadi definisi diatas kalau digabungkan kira-kira menjadi pengelolaan hubungan dua arah antara suatu perusahaan dengan orang yang menjadi pelanggan di perusahaan tersebut.

Kata kedua, "hubungan" atau "relationship", merujuk pada bentuk komunikasi dua arah yang terjalin antara pembeli dan penjual. Komunikasi ini tidak bersifat satu arah, melainkan melibatkan interaksi aktif dari kedua belah pihak yang saling memberikan informasi, feedback, dan respons. Sementara itu, kata terakhir, "manajemen", memiliki arti pengelolaan dalam cakupan yang luas tanpa perlu menjabarkan detail spesifik tentang bagaimana cara mengelola sesuatu. Konsep manajemen di sini mencakup berbagai aktivitas seperti perencanaan, pengorganisasian, pelaksanaan, dan pengawasan. Ketika ketiga kata tersebut digabungkan, CRM dapat didefinisikan sebagai pengelolaan hubungan dua arah antara suatu perusahaan dengan orang-orang yang menjadi pelanggan di perusahaan tersebut. Dengan kata lain, CRM merupakan strategi bisnis yang berfokus pada pembangunan dan pemeliharaan hubungan jangka panjang dengan pelanggan melalui komunikasi yang efektif dan pengelolaan yang terstruktur

## METODE PENELITIAN

### Desain Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode observasi lapangan dan analisis data sekunder. Data dikumpulkan melalui observasi langsung di delapan barbershop di Kota Medan yang telah menyetujui menjadi sampel penelitian.

### Pra-pemrosesan data

Pra-pemrosesan data merupakan tahapan fundamental yang menentukan kualitas output dari model segmentasi pelanggan. Pada tahap awal, dilakukan pembersihan data secara menyeluruh untuk menghilangkan nilai kosong dan data duplikat yang dapat menurunkan akurasi analisis serta menimbulkan bias dalam hasil prediksi. Proses pembersihan ini memastikan bahwa setiap record dalam dataset memiliki informasi yang lengkap dan unik, sehingga algoritma dapat bekerja dengan data yang konsisten dan reliable. Selanjutnya, dilakukan transformasi variabel kategorik menjadi format numerik menggunakan teknik encoding, di mana variabel jenis kelamin dikonversi dengan nilai 0 untuk pria dan 1 untuk wanita, memungkinkan algoritma machine learning untuk menginterpretasi dan memproses data kategorikal secara matematis. Tahap terakhir adalah penyiapan struktur data yang mencakup pengorganisasian dataset ke dalam format yang sesuai untuk pelatihan model, termasuk pemisahan fitur dan target variable, serta pengaturan dimensi data yang optimal untuk proses komputasi. Keseluruhan tahapan pra-pemrosesan ini dilakukan secara sistematis untuk memastikan bahwa data yang digunakan dalam pelatihan model telah memenuhi standar kualitas yang diperlukan, sehingga menghasilkan model segmentasi yang akurat dan dapat diandalkan dalam mengklasifikasikan calon pengguna aplikasi Pangkasin.

### Logistik Regresi

Model Logistik Regresi diimplementasikan menggunakan bahasa pemrograman Python dengan memanfaatkan library scikit-learn yang menyediakan framework lengkap untuk pengembangan model machine learning. Dalam proses pembangunan model, dataset dibagi menjadi dua bagian menggunakan metode stratified sampling, yaitu 70% untuk data latih yang digunakan dalam proses pembelajaran pola dan parameter model, serta 30% untuk data uji yang berfungsi mengevaluasi performa model terhadap data yang belum pernah dilihat sebelumnya. Penggunaan stratified sampling sangat penting untuk memastikan distribusi kelas tetap seimbang pada kedua subset data, sehingga model tidak mengalami bias terhadap kelas tertentu. Model ini bekerja dengan menghitung fungsi linier  $z = b + w_1x_1 + w_2x_2 + w_3x_3$ , di mana  $b$  merupakan bias atau intercept,  $w_1$ ,  $w_2$ , dan  $w_3$  adalah bobot untuk setiap fitur input  $x_1$ ,  $x_2$ , dan  $x_3$  yang merepresentasikan karakteristik calon pengguna. Nilai  $z$  yang dihasilkan dari perhitungan linier tersebut kemudian ditransformasikan menjadi probabilitas menggunakan fungsi sigmoid  $\sigma(z) = 1/(1 + e^{-z})$ , yang mengkonversi nilai  $z$  dari rentang tak terbatas menjadi nilai probabilitas antara 0 dan 1, memungkinkan model untuk mengklasifikasikan apakah seorang calon pengguna termasuk dalam kategori pengguna potensial atau bukan berdasarkan threshold probabilitas yang telah ditentukan.

## HASIL DAN PEMBAHASAN

Berdasarkan proses pelatihan model yang telah dilakukan, diperoleh beberapa nilai parameter penting yang menggambarkan hubungan antara variabel-variabel dalam model. Nilai bias atau konstanta ( $b$ ) diperoleh sebesar 2,44, yang merepresentasikan nilai dasar

dari model ketika semua variabel prediktor bernilai nol. Untuk bobot umur ( $w_1$ ), diperoleh nilai -0,08 yang menunjukkan adanya hubungan negatif, artinya setiap penambahan satu satuan umur akan menurunkan nilai prediksi sebesar 0,08 satuan. Sementara itu, bobot jenis kelamin ( $w_2$ ) memiliki nilai 0,62, yang mengindikasikan bahwa variabel jenis kelamin memberikan kontribusi positif terhadap prediksi model. Adapun bobot frekuensi cukur ( $w_3$ ) menunjukkan nilai tertinggi yaitu 0,92, yang berarti variabel ini memiliki pengaruh paling besar terhadap output model dibandingkan dengan variabel lainnya, di mana setiap peningkatan satu satuan frekuensi cukur akan meningkatkan nilai prediksi sebesar 0,92 satuan.

Model Logistik Regresi yang dikembangkan menunjukkan performa yang baik dalam mengklasifikasikan calon pengguna potensial:

Tabel 4.1 Hasil Evaluasi Model

Metrik	Nilai	Interpretasi
Akurasi	0,80	Model memprediksi dengan benar 80% dari data uji
Recall	0,918	Model dapat mengidentifikasi 91,8% dari konsumen yang benar-benar minat
Precision	0,85	85% dari prediksi "minat" benar-benar akurat
F1-Score	0,88	Keseimbangan yang baik antara precision dan recall

Berdasarkan hasil confusion matrix yang diperoleh dari evaluasi model, terdapat empat kategori prediksi yang menggambarkan performa model dalam mengklasifikasikan data. Kategori True Positive (TP) menunjukkan nilai 31, yang berarti terdapat 31 data yang memang benar-benar memiliki minat dan berhasil diprediksi dengan tepat oleh model sebagai minat. Untuk kategori True Negative (TN), diperoleh nilai 7, yang mengindikasikan bahwa terdapat 7 data yang memang tidak memiliki minat dan model berhasil memprediksinya dengan benar sebagai tidak minat. Sementara itu, terdapat kesalahan prediksi yang tercermin dalam dua kategori lainnya. Kategori False Positive (FP) memiliki nilai 6, yang berarti ada 6 data yang seharusnya tidak memiliki minat namun keliru diprediksi oleh model sebagai minat. Sedangkan kategori False Negative (FN) juga menunjukkan nilai 6, yang menandakan bahwa terdapat 6 data yang sebenarnya memiliki minat tetapi gagal diidentifikasi oleh model dan diprediksi sebagai tidak minat. Hasil confusion matrix ini memberikan gambaran komprehensif tentang akurasi dan kesalahan yang terjadi dalam proses prediksi model.

Berdasarkan hasil Exploratory Data Analysis (EDA) yang telah dilakukan, diperoleh pemahaman mendalam mengenai karakteristik pengguna potensial aplikasi. Secara umum, profil responden menunjukkan bahwa rata-rata umur pengguna adalah 31 tahun dengan dominasi pengguna pria sebesar 81,2%, sementara rata-rata frekuensi cukur mencapai 2 kali per bulan. Yang menarik, persentase minat terhadap aplikasi cukup tinggi yaitu mencapai 83%, mengindikasikan potensi pasar yang cukup besar.

Ketika dianalisis lebih lanjut, terdapat perbedaan karakteristik yang signifikan antara pengguna yang berminat dan tidak berminat terhadap aplikasi. Pengguna yang menunjukkan minat cenderung lebih muda dengan rata-rata umur 29 tahun dan memiliki frekuensi cukur yang lebih tinggi yaitu 2,31 kali per bulan. Kelompok ini didominasi oleh pria dengan persentase mencapai 20,16% dari total pengguna yang berminat. Sebaliknya, pengguna yang tidak berminat terhadap aplikasi memiliki karakteristik yang berbeda, dengan rata-rata umur yang lebih tua yaitu 36 tahun dan frekuensi cukur yang lebih rendah

sebesar 1,64 kali per bulan. Menariknya, kelompok ini menunjukkan komposisi jenis kelamin yang lebih seimbang dengan persentase wanita yang lebih tinggi mencapai 12%. Temuan ini mengindikasikan bahwa target pengguna potensial aplikasi adalah pria berusia lebih muda dengan kebiasaan cukur yang lebih rutin, sementara kelompok dengan umur lebih tua dan frekuensi cukur lebih rendah cenderung kurang tertarik dengan aplikasi tersebut.

Hasil penelitian menunjukkan bahwa algoritma Logistik Regresi efektif dalam mengklasifikasikan calon pengguna potensial aplikasi Pangkasin dengan akurasi 80%. Nilai recall yang tinggi (91,8%) menunjukkan bahwa model sangat baik dalam mengidentifikasi konsumen yang benar-benar berminat, yang penting untuk strategi pemasaran yang efektif.

Karakteristik pengguna potensial yang teridentifikasi (umur 20-35 tahun, pria, frekuensi cukur 2-3 kali per bulan) memberikan insight berharga untuk Tim Pangkasin dalam mengarahkan strategi pemasaran dan pengembangan fitur aplikasi.

Dibandingkan dengan penelitian terdahulu, hasil akurasi 80% ini sejalan dengan temuan Ndangi et al. (2019) yang mencapai akurasi 80% pada data testing, dan sedikit lebih rendah dari Tran et al. (2023) yang mencapai 86-87%. Perbedaan ini dapat dijelaskan oleh perbedaan kompleksitas data dan jumlah fitur yang digunakan.

## KESIMPULAN

Penelitian ini berhasil membuktikan efektivitas algoritma Logistik Regresi dalam mengklasifikasikan calon pengguna aplikasi Pangkasin dengan performa yang baik, ditunjukkan oleh akurasi 80%, recall 91,8%, precision 85%, dan F1-Score 88%. Model yang dibangun menggunakan tiga variabel prediktor yaitu umur, jenis kelamin, dan frekuensi cukur, berhasil mengidentifikasi karakteristik pengguna potensial sebagai pria berusia 20-35 tahun dengan frekuensi cukur 2-3 kali per bulan. Untuk pengembangan lebih lanjut, disarankan untuk menambahkan variabel lain seperti pekerjaan, lokasi geografis, dan penghasilan, memperluas area observasi ke wilayah lain, serta membandingkan dengan algoritma klasifikasi alternatif seperti Random Forest atau Support Vector Machine. Secara praktis, model ini dapat diintegrasikan ke dalam backend aplikasi Pangkasin untuk klasifikasi real-time dan pengembangan sistem rekomendasi layanan, serta perlu dilakukan evaluasi dan pembaruan model secara berkala agar tetap relevan dengan pola perilaku pengguna aktual, sehingga dapat menjadi indikator strategis dalam penentuan arah marketing dan pengembangan bisnis Startup Pangkasin.

## REFERENSI

- Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression*. John Wiley & Sons.
- Telkomuniversity.ac.id. (2024, Oktober 28). *Memahami Apa Itu Aplikasi dan Fungsinya di Era Digital*. Diakses pada 7 April 2025, dari <https://dif.telkomuniversity.ac.id/memahami-apa-itu-aplikasi-dan-fungsinya-di-era-digital/>
- Bisono, A.T., & Zulherry, A. (2025). *Analisis Sentimen Game Genshin Impact untuk Mengetahui Reaksi dan Harapan Pemain Menggunakan Metode Naïve Bayes*. *sudo Jurnal Teknik Informatika* 4 (2), 183-193

- Alfayulanda, P., & Murni, D. (2024). Penerapan Analisis Regresi Logistik Biner Pada Faktor-Faktor Yang Mempengaruhi Keputusan Mahasiswa Berbelanja Online Di Tiktok. *Jurnal Lebesgue : Jurnal Ilmiah Pendidikan Matematika, Matematika Dan Statistika*, 5(1), 178–189. <https://doi.org/10.46306/lb.v5i1.547>
- Alwin, M. I., & Aditiawan, F. P. (2022). Aplikasi Layanan Jasa Tukang Berbasis Website. *INTEGER: Journal of Information Technology*, 7(2), 93–99. <https://doi.org/10.31284/j.integer.2022.v7i2.3298>
- Basri, M., & Zulherry, A. (2025). Analysis of the Impact of Gambling and Online Loans in the Perspective of Informatics, Islam, and Kemuhammadiyah. *AR-RASYID: Jurnal Pendidikan Agama Islam* 5 (1)
- Sari, I.P., Hariani, P.P., Al-Khowarizmi, A., Ramadhani, F., Sulaiman, O.K., Satria, A., & Manurung, A.A. (2024). CLUSTERING HIV/AIDS DISEASE USING K-MEANS CLUSTERING ALGORITHM. *Proceeding International Seminar on Islamic Studies* 5 (1), 1668-1676
- Sari, I.P., Ramadhani, F., Satria, A., & Sulaiman, O.K. Leukocoria Identification: A 5-Fold Cross Validation CNN and Adaboost Hybrid Approach. *2023 6th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*, 486-491
- Manurung, A.A., Nasution, M.D., & Sari, I.P. (2023). Implementation of Fuzzy K-Nearest Neighbor Method in Dengue Disease Classification. *2023 11th International Conference on Cyber and IT Service Management (CITSM)*, 1-4
- Ameliya, A., Nainggolan, D. O., & Pinem, H. G. (2024). Analisis Kemampuan Mahasiswa Matematika FMIPA Unimed dalam Menyelesaikan Pertidaksamaan Nilai Mutlak dengan Berbantuan Python. 2(5).
- Badan Pengembangan dan Pembinaan Bahasa, K. P. dan K. R. I. (2016). Implementasi. <https://kbbi.kemdikbud.go.id/entri/Implementasi>
- Ichsan, A., Zulherry, A., Lubis, T.A., & Shahnaz, B.A.Z. (2025). Utilization of Mobile Applications to Speed Up The Search for Android-Based Index Places. *IJATCoS: Indonesian Journal of Applied Technology, Computer and Science* 2 (1)
- Bogdanchikov, A., Zhaparov, M., & Suliyev, R. (2013). Python to learn programming. *Journal of Physics: Conference Series*, 423(1). <https://doi.org/10.1088/1742-6596/423/1/012027>
- Sari, I.P., Ramadhani, F., Satria, A., & Apdilah, D. (2023). Implementasi Pengolahan Citra Digital dalam Pengenalan Wajah menggunakan Algoritma PCA dan Viola Jones. *Hello World Jurnal Ilmu Komputer* 2 (3), 146-157
- Danardatu, A. H. (2019). Pengenalan Customer Relationship Management ( CRM ). *Jurnal Ilmiah Teknologi Informasi Asia*, 1–5.
- Zulherry, A. (2023) Decision making for network security with simple additive weighting method. *Journal of Intelligent Decision Support System (IDSS)* 6 (3), 155-159
- Faisal, F., Sari, I. R., Saraswati, I., & Heikal, J. (2024). Analisis Pengaruh Karakteristik Produk terhadap Niat Beli Ulang Pelanggan menggunakan Metode Regresi Logistik Biner. *AKADEMIK: Jurnal Mahasiswa Humanis*, 4(3), 1182–1190. <https://doi.org/10.37481/jmh.v4i3.1064>
- Gunawan, R., Aulia, S., Supeno, H., Wijanarko, A., Uwiringiyimana, J. P., & Mahayana, D. (2021). Adiksi Media Sosial dan Gadget bagi Pengguna Internet di Indonesia. *Techno-Socio Ekonomika*, 14(1), 1–14. <https://doi.org/10.32897/techno.2021.14.1.544>

- Hendrian, S. (2018). Algoritma Klasifikasi Data Mining Untuk Memprediksi Siswa Dalam Memperoleh Bantuan Dana Pendidikan. *Faktor Exacta*, 11(3), 266–274. <https://doi.org/10.30998/faktorexacta.v11i3.2777>
- Irma Devi Lestari. (2016). 196927-ID-klasifikasi-online-dan-google. *Iqra'*, 83–94.
- Sari, I.P., Al-Khowarizmi, A, Sulaiman, O.K., & Apdilah, D. (2023). Implementation of Data Classification Using K-Means Algorithm in Clustering Stunting Cases. *Journal of Computer Science, Information Technology and Telecommunication Engineering* 4 (2), 402-412
- Sulaiman, O.K & Batubara, I.H. (2021). Implementation Data Mining For Level Analysis Traffic Violation By Algorithm Association Rule. *Al'adzkiya International of Computer Science and Information Technology (AIOCSIT) Journal* 2 (2), 128-135
- Zulherry, A., Gunawan, T.S., & Wanayumini, W. (2021). Analisis Hasil Pendukung Keputusan Mendapatkan Rumah Dinas Perusahaan Menggunakan Metode Analytical Hierarchy Process (AHP) dan Technique for Order Preference by Similarity to Ideal Solution (TOPSIS). *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 2021
- Khotimah, K., & Febriansyah, F. (2018). Pengaruh kemudahan penggunaan, kepercayaan konsumen dan kreativitas iklan terhadap minat beli konsumen online-shop. *Jurnal Manajemen Strategi Dan Aplikasi Bisnis*, 1(1), 19–26. <https://doi.org/10.36407/jmsab.v1i1.16>
- Sari, I.P., Batubara, I.H., & Al-Khowarizmi, A. (2021). Sensitivity Of Obtaining Errors In The Combination Of Fuzzy And Neural Networks For Conducting Student Assessment On E-Learning. *International Journal of Economic, Technology and Social Sciences (Injects)* 2 (1), 331-338
- Sari, I.P., Al-Khowarizmi, A., & Batubara, I.H. (2021). Cluster Analysis Using K-Means Algorithm and Fuzzy C-Means Clustering For Grouping Students' Abilities In Online Learning Process. *Journal of Computer Science, Information Technology and Telecommunication Engineering* 2 (1), 139-144
- Apdilah, D., & Sari, I.P. (2021). Optimization Of The Fuzzy C-Means Cluster Center For Credit Data Grouping Using Genetic Algorithms. *Al'adzkiya International of Computer Science and Information Technology (AIOCSIT) Journal* 2 (2), 156-163
- Kurniasih, D., Rusfiana, Y., Subagyo, A., & Nuradhawati, R. (2021). Teknik Analisa. Alfabeta Bandung, 1–119. [www.cvalfabeta.com](http://www.cvalfabeta.com)
- Naïve, A., Classifier, B., & Penyakit, P. (2024). Analisis klasifikasi menggunakan regresi logistik biner dan algoritma naïve bayes classifier pada penyakit hipertensi 1,2,3. *13(2007)*, 319–327. <https://doi.org/10.14710/j.gauss.13.2.319-327>
- Ndangi, W. R. A., Resmawan, R., & Djakaria, I. (2019). Perbandingan Analisis Diskriminan dan Regresi Logistik Multinomial. *Jambura Journal of Mathematics*, 1(2), 54–63. <https://doi.org/10.34312/jjom.v1i2.2100>
- Zulherry, A., Siregar, F.A., Gultom, Z.A., & Raihan, E.A. (2023). Optimalisasi Website untuk Monitoring Jaringan OPD di Dinas Kominfo Kota Medan dengan Metode Triangulasi. *Bulletin of Computer Science Research* 3 (5), 357-363
- Tran, H., Le, N., & Nguyen, V. H. (2023). Customer Churn Prediction in the Banking Sector Using Machine Learning-Based Classification Models. *Interdisciplinary Journal of Information, Knowledge, and Management*, 18(May), 87–105. <https://doi.org/10.28945/5086>
- Van Rossum, G., & Drake, F. L. . (2006). An introduction to Python, release 2.5.

- Wibawa, A. P., Guntur, M., Purnama, A., Fathony Akbar, M., & Dwiyanto, F. A. (2018). Metode-metode Klasifikasi. *Prosiding Seminar Ilmu Komputer Dan Teknologi Informasi*, 3(1), 134–138.
- Indah Purnama Sari. *Algoritma dan Pemrograman*. Medan: UMSU Press, 2023, pp. 290.
- Andi Zulherry, Muhammad Basri, Muhammad Haris, Ferdy Riza, Zuli Agustina Gultom, Farid Akbar Siregar, Okvi Nugroho, Mahardika Abdi Prawira Tanjung. *Komunikasi Data dan Jaringan Komputer*. Medan: UMSU Press, 2025, pp. 202.