

PENERAPAN ALGORITMA C4.5 UNTUK KLASIFIKASI LAJU PERTUMBUHAN PENDUDUK DI KOTA BINJAI

Putri Mawarni Rojalian Nst¹, Yusfrizal², Indah Ambarita³

Sistem Informasi, STMIK Kaputama Binjai^{1,3}, Politeknik Gihon²

E-mail: putrimawarni164@gmail.com¹, yusfrizal80@gmail.com², yesnovanda@yahoo.com³

ABSTRAK

Pertumbuhan penduduk merupakan salah satu aspek penting dalam perencanaan pembangunan karena perubahan jumlah penduduk pada setiap kecamatan dipengaruhi oleh kelahiran, penduduk datang, dan penduduk pindah. Banyaknya data kependudukan yang tersedia membuat proses pengolahan dan pengklasifikasian secara manual menjadi kurang efektif. Penelitian ini bertujuan untuk menerapkan algoritma C4.5 dalam mengklasifikasikan laju pertumbuhan penduduk di Kota Binjai serta menghasilkan sistem berbasis web yang dapat menampilkan proses perhitungan dan pohon keputusan. Penelitian menggunakan pendekatan kuantitatif dengan metode data mining. Data yang digunakan merupakan data sekunder dari BPS Kota Binjai sebanyak 300 record periode 2018–2026 yang mencakup lima kecamatan. Tahapan penelitian meliputi identifikasi masalah, studi literatur, pengumpulan data, preprocessing, penerapan algoritma C4.5, pengujian dan analisis hasil, serta implementasi sistem. Algoritma C4.5 diterapkan dengan menghitung entropy dan information gain untuk menentukan atribut pembentuk decision tree. Hasil penelitian menunjukkan bahwa 300 data terdiri atas 92 data kategori Tinggi, 148 data kategori Sedang, dan 60 data kategori Rendah. Nilai entropy total yang diperoleh sebesar 1,490. Atribut Jumlah Kelahiran memperoleh nilai gain tertinggi sebesar 0,211 sehingga ditetapkan sebagai root node, sedangkan Jumlah Penduduk Datang memperoleh gain 0,046, Kecamatan 0,034, dan Jumlah Penduduk Pindah 0,012 sebagai atribut pembentuk cabang. Hasil tersebut menunjukkan bahwa algoritma C4.5 dapat diterapkan untuk mengklasifikasikan laju pertumbuhan penduduk Kota Binjai ke dalam kategori Tinggi, Sedang, dan Rendah. Implementasi berbasis web juga dapat menampilkan data, data latih, perhitungan C4.5, serta struktur pohon keputusan. Dengan demikian, penerapan algoritma C4.5 dapat membantu proses klasifikasi data pertumbuhan penduduk secara lebih terstruktur berdasarkan pola atribut kependudukan yang digunakan dalam penelitian.

Kata kunci

Data Mining, Algoritma C4.5, Decision Tree, Laju Pertumbuhan Penduduk, Kota Binjai.

ABSTRACT

Population growth is a crucial aspect of development planning, as changes in population size within each district are influenced by births, in-migration, and out-migration. The vast amount of available demographic data renders manual processing and classification ineffective. This study aims to apply the C4.5 algorithm to classify population growth rates in Binjai City and to develop a web-based system capable of displaying the calculation process and the resulting decision tree. A quantitative approach utilizing data mining methods was employed. The study utilized secondary data from the Binjai City Central Statistics Agency (BPS), comprising 300 records covering the 2018–2026 period across five districts. The research stages included problem identification, literature review, data collection, preprocessing, C4.5 algorithm application, testing and analysis of results, and system implementation. The C4.5 algorithm was applied by calculating entropy and information gain to determine the attributes forming the decision tree. The results indicate that the 300 data points comprised 92 records in the "High" category, 148 in the "Medium" category, and 60 in the "Low" category. The total entropy value obtained was 1.490. The "Number of

Births" attribute yielded the highest gain value (0.211) and was thus designated as the root node, while "Number of In-migrants" (gain: 0.046), "District" (0.034), and "Number of Out-migrants" (0.012) served as attributes forming the branches. These findings demonstrate that the C4.5 algorithm can effectively classify Binjai City's population growth rates into High, Medium, and Low categories. The web-based implementation also enables the display of raw data, training data, C4.5 calculations, and the decision tree structure. Consequently, the application of the C4.5 algorithm facilitates a more structured classification process for population growth data based on the demographic attribute patterns utilized in this study.

Keywords

Data Mining, C4.5 Algorithm, Decision Tree, Population Growth Rate, Binjai City.

1. PENDAHULUAN

Badan Pusat Statistik (BPS) merupakan lembaga pemerintah yang menyediakan data statistik sebagai dasar dalam perencanaan, evaluasi, dan pengambilan kebijakan pembangunan. BPS Kota Binjai secara rutin menyediakan berbagai data kependudukan yang menggambarkan kondisi dan perkembangan jumlah penduduk di wilayah Kota Binjai. Data kependudukan tersebut dapat menjadi sumber informasi penting dalam memahami perubahan jumlah penduduk dari waktu ke waktu. Pertumbuhan penduduk dipengaruhi oleh beberapa faktor, seperti kelahiran, kematian, penduduk yang datang, dan penduduk yang pindah. Perubahan tersebut menyebabkan laju pertumbuhan penduduk pada setiap wilayah dapat berbeda. Informasi mengenai pertumbuhan penduduk sangat diperlukan untuk mendukung perencanaan pembangunan, penyediaan fasilitas umum, pelayanan kesehatan, pendidikan, serta kebijakan sosial dan ekonomi. Namun, semakin bertambahnya jumlah data kependudukan menyebabkan proses pengolahan dan analisis menjadi lebih kompleks apabila dilakukan secara manual. Pengolahan data secara manual membutuhkan waktu yang relatif lama dan berpotensi menimbulkan kesalahan dalam proses pengelompokan maupun penentuan kategori pertumbuhan penduduk (Lusiana et al., 2021).

Data mining merupakan proses untuk memperoleh pola, informasi, dan pengetahuan yang bermanfaat dari sekumpulan data dalam jumlah besar. Data mining dapat digunakan untuk mengolah data sehingga menghasilkan informasi yang dapat mendukung proses analisis dan pengambilan keputusan (Pramudiono, 2006). Data mining menjadi salah satu bagian penting dalam proses Knowledge Discovery in Database (KDD), yaitu suatu proses yang dilakukan untuk menemukan pengetahuan dari data melalui beberapa tahapan pengolahan dan analisis (Lusiana et al., 2021). Pemanfaatan data mining dapat membantu proses pengolahan data menjadi lebih efektif karena mampu mengidentifikasi pola tertentu yang sulit ditemukan melalui analisis manual. Proses penggalian pengetahuan dari kumpulan data juga dapat memberikan nilai tambah terhadap data yang sebelumnya hanya berupa kumpulan informasi sehingga dapat dimanfaatkan untuk menghasilkan pengetahuan baru (Purwadi et al., 2019).

Dalam penerapannya, data mining memiliki berbagai teknik, salah satunya adalah klasifikasi yang digunakan untuk mengelompokkan data berdasarkan karakteristik atau atribut tertentu (Hermawati, 2013). Salah satu algoritma klasifikasi yang dapat digunakan adalah algoritma C4.5. Algoritma C4.5 merupakan algoritma pembentukan decision tree atau pohon keputusan yang menentukan atribut terbaik sebagai dasar pemisahan data melalui perhitungan entropy dan information gain. Pohon keputusan yang dihasilkan dapat digunakan untuk membentuk aturan klasifikasi sehingga hasil analisis lebih mudah

dipahami dan diinterpretasikan. Algoritma C4.5 banyak digunakan dalam berbagai permasalahan klasifikasi karena mampu menghasilkan model berupa aturan keputusan berdasarkan atribut yang terdapat dalam dataset (Quinlan, 1996).

Penerapan algoritma C4.5 telah dilakukan dalam berbagai bidang permasalahan. Rohman et al. (2020) melakukan komparasi algoritma C4.5 berbasis PSO dan GA untuk diagnosis penyakit stroke. Penelitian tersebut menunjukkan bahwa algoritma C4.5 dapat diterapkan dalam proses klasifikasi dengan memanfaatkan sejumlah atribut sebagai dasar pembentukan keputusan. Istiawan (2019) menerapkan algoritma C4.5 untuk melakukan prediksi lahan kritis pada kawasan budidaya pertanian di Daerah Aliran Sungai Pemali Jratun. Selain itu, Helmi (2025) menerapkan algoritma C4.5 untuk mendeteksi potensi putus sekolah pada siswa SMP di Giligenting. Penerapan data mining juga telah digunakan dalam bidang hukum untuk membantu proses analisis terhadap data yang tersedia. Simanjuntak (2024) menerapkan data mining dalam proses hukum pidana bagi pelaku kekerasan pada wanita. Berbagai penelitian tersebut menunjukkan bahwa algoritma C4.5 dan data mining dapat digunakan dalam berbagai bidang dengan karakteristik data yang berbeda untuk menghasilkan informasi yang mendukung proses klasifikasi dan pengambilan keputusan. Hal ini menjadi salah satu dasar pemilihan algoritma C4.5 dalam penelitian untuk mengklasifikasikan data laju pertumbuhan penduduk Kota Binjai.

Berdasarkan permasalahan tersebut, penelitian ini berfokus pada Penerapan Algoritma C4.5 untuk Klasifikasi Laju Pertumbuhan Penduduk di Kota Binjai. Penerapan algoritma tersebut diharapkan dapat menghasilkan model decision tree berdasarkan faktor-faktor kependudukan yang digunakan dalam penelitian. Selain menghasilkan klasifikasi, penelitian ini juga bertujuan untuk mengembangkan sistem berbasis web yang dapat menampilkan data kependudukan, proses perhitungan algoritma C4.5, hasil klasifikasi, serta visualisasi pohon keputusan. Dengan adanya sistem tersebut, proses pengolahan dan analisis data laju pertumbuhan penduduk diharapkan dapat dilakukan secara lebih cepat, sistematis, dan mudah dipahami sehingga dapat mendukung pemanfaatan data kependudukan sebagai informasi dalam proses analisis dan pengambilan keputusan.

2. METODE PENELITIAN

2.1 Data Penelitian

Penelitian menggunakan pendekatan kuantitatif dengan metode data mining. Data penelitian merupakan data sekunder yang diperoleh dari BPS Kota Binjai. Dataset terdiri atas 300 record dan mencakup lima kecamatan, yaitu Binjai Kota, Binjai Barat, Binjai Selatan, Binjai Timur, dan Binjai Utara. Periode data yang digunakan adalah 2018–2026.

Tabel 1. Definisi Operasional Atribut Data Kependudukan

Atribut	Keterangan
Tahun	Periode pengamatan data kependudukan
Kecamatan	Wilayah administratif objek pengamatan
Penduduk Awal	Jumlah penduduk pada awal periode
Kelahiran	Jumlah penduduk yang lahir
Penduduk Datang	Jumlah penduduk yang masuk ke wilayah
Penduduk Pindah	Jumlah penduduk yang keluar dari wilayah
Total Penduduk	Jumlah penduduk setelah perubahan
Kategori Pertumbuhan	Label klasifikasi: Tinggi, Sedang, Rendah

2.2 Tahapan Penelitian

Tahapan penelitian meliputi identifikasi masalah, studi literatur, pengumpulan data, preprocessing, penerapan algoritma C4.5, pengujian dan analisis hasil, serta penarikan kesimpulan. Preprocessing dilakukan melalui seleksi data, pemeriksaan data, transformasi, dan pengelompokan data sesuai kebutuhan algoritma.

- Identifikasi masalah: menentukan kebutuhan analisis laju pertumbuhan penduduk.
- Studi literatur: mengkaji data mining, algoritma C4.5, decision tree, dan pertumbuhan penduduk.
- Pengumpulan data: mengambil data kependudukan yang digunakan sebagai dataset penelitian.
- Preprocessing: menyeleksi dan menyesuaikan atribut serta kategori target.
- Penerapan C4.5: menghitung entropy dan gain untuk menentukan node.
- Pengujian dan analisis: mengevaluasi hasil klasifikasi dan struktur decision tree.
- Implementasi: menampilkan pengolahan data melalui sistem berbasis web.

2.3 Algoritma C4.5

Algoritma C4.5 membangun decision tree secara bertahap. Atribut dengan information gain tertinggi dipilih sebagai root atau node pemisah. Tingkat ketidakpastian data dihitung menggunakan entropy dengan persamaan:

$$Entropy(S) = - \sum_{i=1}^n p_i \log_2 p_i \quad (1)$$

Setelah entropy diperoleh, gain dihitung menggunakan persamaan:

$$Gain(S,A) = Entropy(S) - \sum_{i=1}^n \left| \frac{S_i}{S} \right| Entropy(S_i) \quad (2)$$

Pada penelitian ini, atribut numerik dikelompokkan menggunakan batas yang terdapat pada skripsi, yaitu Jumlah Kelahiran ≥ 3000 dan < 3000 , Jumlah Penduduk Datang ≥ 2500 dan < 2500 , serta Jumlah Penduduk Pindah ≥ 1500 dan < 1500 . Atribut Kecamatan digunakan sebagai atribut kategorikal.

3. HASIL DAN PEMBAHASAN

3.1 Istribusi Data

Dari 300 record yang digunakan, terdapat 92 data berkategori Tinggi, 148 data berkategori Sedang, dan 60 data berkategori Rendah. Kategori Sedang menjadi kelompok dengan jumlah data terbesar, sedangkan kategori Rendah merupakan kelompok dengan jumlah paling sedikit.

Tabel 2. Distribusi Kategori Pertumbuhan Penduduk

Kategori	Jumlah	Persentase
Tinggi	92	30,67%
Sedang	148	49,33%
Rendah	60	20,00%
Total	300	100%

3.2 Perhitungan Entropy Total

Entropy total dihitung berdasarkan tiga kelas target. Dengan 92 data Tinggi, 148 data Sedang, dan 60 data Rendah, nilai entropy total yang diperoleh adalah 1,490. Nilai ini digunakan sebagai dasar untuk membandingkan kemampuan setiap atribut dalam mengurangi ketidakpastian klasifikasi.

$$\begin{aligned} \text{entropy}(S) &= -\left(\frac{92}{300}\right) \log_2 \left(\frac{148}{300}\right) \log_2 \left(\frac{60}{300}\right) \\ &= (0,306)(-1,705) - (0,493)(-1,019) - (0,2)(-2,321) = 1,490 \end{aligned}$$

3.3 Perhitungan Information Gain

Perhitungan gain dilakukan pada atribut Kecamatan, Jumlah Kelahiran, Jumlah Penduduk Datang, dan Jumlah Penduduk Pindah. Hasil perhitungan menunjukkan bahwa Jumlah Kelahiran memberikan nilai gain tertinggi.

Tabel 3. Perhitungan Nilai Gain Atribut pada Algoritma C4.5

No.	Atribut	Nilai Gain	Keterangan
1	Kecamatan	0,034	Atribut pembentuk cabang
2	Jumlah Kelahiran	0,211	Gain tertinggi; root node
3	Jumlah Penduduk Datang	0,046	Atribut pembentuk cabang
4	Jumlah Penduduk Pindah	0,012	Atribut pembentuk cabang

Nilai gain Jumlah Kelahiran sebesar 0,211 merupakan nilai tertinggi dibandingkan atribut lainnya. Dengan demikian, Jumlah Kelahiran dipilih sebagai root node. Hasil tersebut menunjukkan bahwa pada dataset penelitian, atribut Jumlah Kelahiran memberikan pengurangan ketidakpastian terbesar pada tahap awal pembentukan pohon. Pada atribut Kecamatan, data terbagi menjadi Binjai Kota, Binjai Barat, Binjai Selatan, Binjai Timur, dan Binjai Utara. Nilai gain yang diperoleh adalah 0,034. Untuk Jumlah Penduduk Datang, data dikelompokkan pada batas ≥ 2500 dan < 2500 dengan gain sebesar 0,046. Jumlah Penduduk Pindah menggunakan batas ≥ 1500 dan < 1500 dengan gain sebesar 0,012.

3.4 Pembentukan Decision Tree

Berdasarkan hasil perhitungan nilai Gain pada setiap atribut, selanjutnya dilakukan pembentukan node dan cabang pada pohon keputusan menggunakan algoritma C4.5. Proses ini dilakukan dengan menentukan atribut yang memiliki nilai Gain tertinggi sebagai node utama, kemudian menentukan kondisi atau cabang berdasarkan nilai atribut yang digunakan. Hasil dari setiap node dan cabang tersebut digunakan untuk menentukan kategori klasifikasi pertumbuhan penduduk, yaitu Tinggi, Sedang, dan Rendah. Adapun hasil pembentukan node dan cabang pohon keputusan menggunakan algoritma C4.5 dapat dilihat pada Tabel 4.

Tabel 4. Hasil Pembentukan Node dan Cabang Pohon Keputusan Algoritma C4.5

Node	Atribut	Kondisi/Cabang	Hasil Klasifikasi
Node 1	Kecamatan	Binjai Kota; Binjai Barat; Binjai Selatan; Binjai Timur; Binjai Utara	Tinggi/Sedang/Rendah
Node 2	Jumlah Kelahiran	≥ 3000 ; < 3000	Tinggi/Sedang/Rendah
Node 3	Penduduk Datang	≥ 2500 ; < 2500	Tinggi/Sedang/Rendah
Node 4	Penduduk Pindah	≥ 1500 ; < 1500	Tinggi/Sedang/Rendah

Struktur decision tree menggunakan beberapa atribut secara bertahap. Berdasarkan hasil penelitian, jumlah kelahiran menjadi atribut yang paling dominan pada pemilihan root. Setelah node tersebut, proses klasifikasi dapat menggunakan jumlah

penduduk datang, jumlah penduduk pindah, dan kecamatan untuk mempersempit hasil keputusan. Setiap cabang menunjukkan kondisi tertentu dan berakhir pada kategori pertumbuhan yang sesuai dengan pola data latih. Pada kondisi jumlah kelahiran ≥ 3000 , model dalam skripsi menunjukkan kecenderungan menuju kategori Tinggi. Sementara pada jumlah kelahiran < 3000 , hasil masih dapat berada pada kategori Sedang atau Rendah. Jumlah penduduk datang dan jumlah penduduk pindah digunakan sebagai atribut lanjutan ketika kondisi sebelumnya belum menghasilkan satu kelas yang pasti.

3.5 Implementasi Sistem

Model klasifikasi kemudian diimplementasikan dalam sistem berbasis web. Sistem menyediakan Dashboard, Data Penduduk, Data Latih, Perhitungan C4.5, dan Pohon Keputusan. Dashboard menampilkan informasi umum sistem dan distribusi data. Data Penduduk berfungsi mengelola master data, sedangkan Data Latih menyiapkan data yang digunakan dalam proses klasifikasi. Menu Perhitungan C4.5 menampilkan entropy total dan perbandingan gain antaratribut. Berdasarkan hasil sistem, entropy total sebesar 1,49 dan Jumlah Kelahiran memiliki gain tertinggi sebesar 0,211. Menu Pohon Keputusan menampilkan struktur node dan kondisi yang digunakan sebagai dasar klasifikasi. Implementasi tersebut membuat proses pengolahan data lebih terstruktur. Data dapat dimasukkan, dikelola, dipilih sebagai data latih, kemudian diproses menggunakan C4.5 hingga menghasilkan decision tree. Aturan yang terbentuk dapat digunakan untuk menentukan kategori Tinggi, Sedang, atau Rendah pada data yang belum diketahui kategorinya.

3.6 Pembahasan

Hasil penelitian memperlihatkan bahwa algoritma C4.5 dapat diterapkan untuk mengklasifikasikan data pertumbuhan penduduk Kota Binjai. Keunggulan utama model pada penelitian ini adalah bentuk aturan keputusan yang dapat ditelusuri melalui node dan cabang, sehingga hasil klasifikasi tidak hanya berupa label tetapi dapat dijelaskan berdasarkan kondisi atribut. Dominannya Jumlah Kelahiran pada root menunjukkan bahwa atribut tersebut memberikan kontribusi paling besar terhadap pemisahan kelas pada dataset yang digunakan. Namun, hasil klasifikasi tetap melibatkan atribut lain karena beberapa kondisi pada satu atribut belum cukup untuk menghasilkan satu kelas. Kombinasi Jumlah Penduduk Datang, Jumlah Penduduk Pindah, dan Kecamatan membantu mempersempit keputusan pada cabang berikutnya. Penelitian ini belum mencantumkan nilai akurasi akhir model secara eksplisit dalam dokumen skripsi. Oleh sebab itu, artikel ini tidak menyatakan persentase akurasi tertentu. Pengujian yang terdokumentasi berfokus pada pembentukan model, distribusi kelas, perhitungan entropy dan gain, serta struktur decision tree. Pengukuran akurasi dengan confusion matrix atau cross-validation dapat menjadi pengembangan untuk penelitian berikutnya.

4. KESIMPULAN

4.1 Kesimpulan

Algoritma C4.5 dapat diterapkan untuk mengklasifikasikan laju pertumbuhan penduduk Kota Binjai berdasarkan data kependudukan. Dataset penelitian terdiri atas 300 record periode 2018–2026 dan memiliki tiga kelas target, yaitu Tinggi, Sedang, dan Rendah. Hasil perhitungan menunjukkan entropy total sebesar 1,490. Jumlah Kelahiran memperoleh nilai gain tertinggi sebesar 0,211, diikuti Jumlah Penduduk Datang sebesar 0,046, Kecamatan sebesar 0,034, dan Jumlah Penduduk Pindah sebesar 0,012. Jumlah Kelahiran ditetapkan sebagai root node pada pembentukan decision tree.

Model menggunakan Kecamatan, Jumlah Kelahiran, Jumlah Penduduk Datang, dan Jumlah Penduduk Pindah untuk membentuk aturan klasifikasi. Model tersebut dapat mengelompokkan laju pertumbuhan penduduk ke dalam kategori Tinggi, Sedang, dan Rendah. Sistem berbasis web yang dikembangkan menyediakan fitur pengelolaan data, data latih, perhitungan C4.5, dan visualisasi pohon keputusan. Penelitian selanjutnya disarankan menggunakan jumlah data dan periode yang lebih panjang, menambahkan atribut seperti tingkat kematian, kepadatan penduduk, kondisi ekonomi, pendidikan, dan faktor sosial, serta membandingkan C4.5 dengan *Random Forest*, *Naive Bayes*, atau *K-Nearest Neighbor*. Optimasi preprocessing dan parameter model juga dapat dilakukan untuk memperoleh evaluasi performa yang lebih lengkap.

4.2 Saran

Penelitian selanjutnya disarankan menggunakan jumlah data yang lebih banyak dan periode pengamatan yang lebih panjang serta menambahkan atribut yang berkaitan dengan pertumbuhan penduduk, seperti jumlah kematian, kepadatan penduduk, kondisi ekonomi, pendidikan, dan faktor sosial. Evaluasi model juga perlu dilakukan menggunakan metode pengukuran kinerja seperti confusion matrix dan cross-validation untuk memperoleh gambaran performa klasifikasi yang lebih objektif. Selain itu, algoritma C4.5 dapat dibandingkan dengan metode lain, seperti Random Forest, Naive Bayes, dan K-Nearest Neighbor, serta sistem dapat dikembangkan menjadi aplikasi yang lebih interaktif agar hasil klasifikasi dapat dimanfaatkan secara lebih mudah oleh instansi terkait.

5. DAFTAR PUSTAKA

- Helmi, F. (2025). Deteksi Potensi Putus Sekolah Menggunakan Algoritma C4.5: Studi Kasus SMP di Giligenting. *Jurnal Penelitian Pendidikan*.
- Hermawati, F. A. (2013). *Data Mining*. Jakarta/Yogyakarta: Penerbit Terkait.
- Istiawan, D. (2019). Implementation of C4.5 Algorithm for Critical Land Prediction in Agricultural Cultivation Areas in Pemali Jratun Watershed. *Journal of Data Mining and Agricultural Cultivation*.
- Lusiana, F. O., Fatma, I., & Windarto, A. P. (2021). Penerapan Data Mining dalam Knowledge Discovery in Database (KDD). *Jurnal Data Mining*, 1(2).
- Pramudiono. (2006). *Pengantar Data Mining dan Analisis Perilaku*. Penerbit Terkait.
- Purwadi, Ramadhan, P. S., & Safitri, N. (2019). Penerapan Data Mining untuk Menggali Nilai Tambah Kumpulan Data Eksplorasi Pengetahuan. *Jurnal Kecerdasan Buatan dan Machine Learning*.
- Quinlan, J. R. (1996). *Induction of Decision Trees and the C4.5 Algorithm Concept*. *Machine Learning Journal*.
- Rohman, R. S., Saputra, R. A., & Firmansaha, D. A. (2020). Komparasi Algoritma C4.5 Berbasis PSo dan GA untuk Diagnosa Penyakit Stroke. *Jurnal Klasifikasi Medis*, 5(1).
- Simanjuntak, L. P. (2024). Penerapan Data Mining Dalam Proses Hukum Pidana Bagi Pelaku Kekerasan Pada Wanita. *Jurnal Ilmu Hukum dan Data Mining*, 1(2).
- Yatini, B. (2010). *Pengantar Flowchart dan Representasi Grafis Pemecahan Masalah*. Penerbit Terkait.