E - ISSN 2829- 2529

RJTI
Riau Jurnal Teknik Informatika

Riau Jurnal Teknik Informatika (https://journal.upp.ac.id/index.php/rjti/index)
Dipublikasikan oleh **Universitas Pasir Pengaraian**
Jl.Tuanku Tambusai Jl. Raya Kumu, Rambah, Kec. Rambah Hilir, Kab. Rokan Hulu, Provinsi Riau, 28558

RIAU JURNAL TEKNIK INFORMATIKA

SURAT KETERANGAN PENERIMAAN ARTIKEL JURNAL

No. 019/RJTI/INF-FILKOM/VI/2026

Dewan penyunting Jurnal Riau Jurnal Teknik Informatika telah menerima artikel,

Nama : Fahriyadi Purnama Thaib[1,*], Frangky Tupamahu[2], Alter Lasarudin[3], Hilmansyah Gani[4]

Judul : **Analisis Pelanggaran Disiplin dan Kode Etik Anggota Polri Menggunakan Decision Tree C4.5**

Asal Instansi : [1,2,3,4] Universitas Muhammadiyah Gorontalo, Indonesia

Menyatakan bahwa artikel tersebut **telah diproses** sesuai Prosedur Penulisan Jurnal **Riau Jurnal Teknik Informatika dan akan diterbitkan pada** Volume 5 Nomor 2 Bulan Juli Tahun 2026. Demikian surat keterangan ini dibuat dan harap dipergunakan dengan sebaik-baiknya.

Riau, 26 Juni 2026
Ketua Dewan Penyunting

Imam Rangga Bakti, M.Kom



Published online on the page: https://journal.upp.ac.id/index.php/rjti/index

R J T I

Riau Jurnal Teknik Informatika

| E - ISSN (*Online*) 2829-2529 |

Analisis Pelanggaran Disiplin dan Kode Etik Anggota Polri Menggunakan Decision Tree C4.5

Fahriyadi Purnama Thaib[1,*], Frangky Tupamahu[2], Alter Lasarudin[3], Hilmansyah Gani[4]

[1,2,3,4]Sistem Informasi, Fakultas Sains dan Ilmu Komputer, Universitas Muhammadiyah Gorontalo, Indonesia

Informasi Artikel

Sejarah Artikel:
Submit: 01 Januari 2026
Revisi: 04 Februari 2026
Diterima: 17 Maret 2026
Diterbitkan: 30 Maret 2026

Kata Kunci

Data Mining, Decision Tree, C4.5, Ethical Violations, Police Discipline.

Korespondensi

E-mail: fahriyadithaib2@gmail.com

A B S T R A C T

Disciplinary and ethical violations committed by police officers can affect organizational professionalism and reduce public trust in law enforcement institutions. The management of violation records at Polresta Gorontalo Kota is still administrative, making it difficult to identify violation patterns and support data-driven decision making. This study aims to analyze disciplinary and ethical violation data using the Decision Tree C4.5 algorithm to generate a classification model and decision rules. The dataset used consists of police disciplinary and ethical violation records from 2022 to 2026. The results show that the violation category attribute becomes the root node of the decision tree with the highest Gain Ratio value of 0.694. The generated model is capable of classifying violations into three sanction levels, namely minor, moderate, and severe sanctions, while producing interpretable decision rules. Model evaluation using a confusion matrix achieved an accuracy rate of 70.8%. These findings indicate that the C4.5 algorithm can effectively identify patterns between violation types and sanction levels and can be utilized as a decision-support tool for managing disciplinary and ethical violations within the Indonesian National Police institution.

Abstrak

Pelanggaran disiplin dan kode etik anggota Polri dapat memengaruhi profesionalisme institusi serta menurunkan kepercayaan masyarakat. Pengelolaan data pelanggaran di Polresta Gorontalo Kota yang masih bersifat administratif menyebabkan kesulitan dalam mengidentifikasi pola pelanggaran dan mendukung pengambilan keputusan berbasis data. Penelitian ini bertujuan menganalisis data pelanggaran disiplin dan kode etik menggunakan algoritma Decision Tree C4.5 untuk menghasilkan model klasifikasi dan aturan pengambilan keputusan. Data yang digunakan merupakan data pelanggaran anggota Polri periode 2022–2026. Hasil penelitian menunjukkan bahwa atribut kategori pelanggaran menjadi akar pohon keputusan dengan nilai Gain Ratio tertinggi sebesar 0,694. Model yang dihasilkan mampu mengklasifikasikan pelanggaran ke dalam kategori sanksi ringan, sedang, dan berat serta menghasilkan aturan keputusan yang mudah diinterpretasikan. Evaluasi model menggunakan confusion matrix memperoleh tingkat akurasi sebesar 70,8%. Hasil penelitian menunjukkan bahwa algoritma C4.5 mampu mengidentifikasi pola hubungan antara jenis pelanggaran dan tingkat sanksi sehingga dapat dimanfaatkan sebagai pendukung pengambilan keputusan dalam pengelolaan pelanggaran disiplin dan kode etik anggota Polri.

This is an open access article under the CC–BY-SA license

Creative Commons Attribution-ShareAlike 4.0 International License. *Some rights reserved*



Published online on the page: https://journal.upp.ac.id/index.php/rjti/index

R J T I

Riau Jurnal Teknik Informatika

| E - ISSN (*Online*) 2829-2529 |

1. Pendahuluan

Perkembangan teknologi informasi dan komunikasi telah mendorong transformasi digital pada berbagai sektor, termasuk pemerintahan dan organisasi publik. Pemanfaatan teknologi tidak lagi terbatas pada fungsi penyimpanan data, tetapi telah berkembang menjadi sarana untuk menghasilkan informasi strategis yang mendukung proses pengambilan keputusan. Seiring meningkatnya volume data yang dihasilkan organisasi, diperlukan metode yang mampu mengolah data menjadi informasi yang bernilai. Salah satu pendekatan yang banyak digunakan adalah data mining, yaitu proses mengekstraksi pola dan pengetahuan tersembunyi dari kumpulan data untuk menghasilkan informasi yang dapat digunakan dalam mendukung pengambilan keputusan secara lebih efektif dan objektif [1].

Kepolisian Negara Republik Indonesia (Polri) merupakan institusi yang memiliki tanggung jawab dalam menjaga keamanan dan ketertiban masyarakat, menegakkan hukum, serta memberikan perlindungan dan pelayanan kepada masyarakat. Pelaksanaan tugas tersebut menuntut setiap anggota Polri untuk menjunjung tinggi profesionalisme, integritas, dan akuntabilitas. Untuk menjamin terciptanya standar perilaku yang sesuai dengan nilai-nilai organisasi, pemerintah menetapkan Peraturan Pemerintah Republik Indonesia Nomor 2 Tahun 2003 tentang Peraturan Disiplin Anggota Polri serta Peraturan Kepolisian Negara Republik Indonesia Nomor 7 Tahun 2022 tentang Kode Etik Profesi Polri dan Komisi Kode Etik Polri [2], [3]. Kedua regulasi tersebut menjadi landasan dalam pembinaan, pengawasan, dan penegakan disiplin guna menjaga kepercayaan masyarakat terhadap institusi kepolisian.

Meskipun perangkat regulasi telah tersedia, pelanggaran disiplin dan kode etik masih ditemukan dalam pelaksanaan tugas kepolisian. Data pelanggaran anggota Polri di Polresta Gorontalo Kota periode 2022–2026 menunjukkan adanya berbagai bentuk pelanggaran, mulai dari mangkir, ketidakpatuhan terhadap perintah pimpinan, perjudian, penyalahgunaan narkotika, konsumsi minuman keras, perbuatan asusila, hingga perilaku yang tidak mencerminkan profesionalisme dalam menjalankan tugas. Fenomena tersebut menunjukkan bahwa upaya pembinaan dan pengawasan masih menghadapi berbagai tantangan. Selain berdampak pada efektivitas organisasi, pelanggaran yang dilakukan anggota Polri berpotensi menurunkan citra institusi dan tingkat kepercayaan publik.

Permasalahan lain yang dihadapi adalah pengelolaan data pelanggaran yang masih berorientasi pada kebutuhan administrasi. Data yang tersimpan selama bertahun-tahun umumnya hanya digunakan sebagai arsip tanpa dianalisis lebih lanjut untuk menemukan pola hubungan antara jenis pelanggaran dan tingkat sanksi yang diberikan. Kondisi ini menyebabkan informasi yang terkandung dalam data historis belum dimanfaatkan secara optimal sebagai dasar penyusunan kebijakan pembinaan maupun evaluasi disiplin anggota. Oleh karena itu, diperlukan pendekatan analitis yang mampu mengubah data pelanggaran menjadi informasi yang dapat mendukung proses pengambilan keputusan secara lebih sistematis.

Dalam bidang Sistem Informasi, teknik data mining telah banyak digunakan untuk menyelesaikan berbagai permasalahan klasifikasi dan prediksi. Salah satu metode yang populer adalah algoritma Decision Tree C4.5. Algoritma ini membangun model klasifikasi dalam bentuk pohon keputusan dengan memanfaatkan perhitungan entropy, information gain, dan gain ratio untuk menentukan atribut terbaik sebagai pemisah data [4]. Selain memiliki kemampuan klasifikasi yang baik, algoritma C4.5 menghasilkan aturan keputusan (decision rules) yang mudah dipahami sehingga

Creative Commons Attribution-ShareAlike 4.0 International License. *Some rights reserved*

banyak digunakan dalam berbagai bidang yang memerlukan transparansi dalam proses pengambilan keputusan.

Sejumlah penelitian telah membuktikan efektivitas algoritma tersebut. Penelitian yang dilakukan oleh Liputo, Tupamahu, dan rekan-rekan menunjukkan bahwa algoritma Decision Tree mampu menghasilkan model klasifikasi yang efektif dalam memprediksi kesehatan mental siswa sekolah dasar serta memiliki tingkat stabilitas yang baik melalui proses validasi silang [5]. Pada sektor layanan publik, Lasarudin, Hasyim, dan Dumako menerapkan algoritma C4.5 untuk menganalisis kepuasan masyarakat terhadap layanan siaran RRI Gorontalo. Temuan penelitian tersebut memperlihatkan bahwa pohon keputusan yang dihasilkan mampu mengidentifikasi faktor-faktor yang memengaruhi tingkat kepuasan masyarakat secara jelas dan mudah diinterpretasikan [6].

Penelitian lain yang dilakukan oleh Lasarudin, Gani, dan Tomayahu membandingkan algoritma Naïve Bayes dan C4.5 pada klasifikasi status gizi balita. Hasil pengujian menunjukkan bahwa algoritma C4.5 mampu memberikan performa yang kompetitif serta menghasilkan model yang lebih mudah dipahami oleh pengguna [7]. Pada kajian yang berbeda, Lasarudin, Handayani, dan Yane mengembangkan sistem pendukung keputusan untuk pemberian sanksi terhadap siswa yang melakukan pelanggaran menggunakan metode SMART. Penelitian tersebut membuktikan bahwa pendekatan berbasis data dapat membantu proses penentuan sanksi secara lebih objektif dan terstruktur [8].

Efektivitas algoritma C4.5 juga dilaporkan pada penelitian Nugraha dkk. yang memanfaatkan metode tersebut untuk memprediksi kebutuhan bimbingan siswa dan menghasilkan aturan klasifikasi yang mudah diterapkan dalam proses pengambilan keputusan [9]. Santoso dkk. menggunakan algoritma yang sama untuk mengklasifikasikan risiko penyakit jantung dan memperoleh model yang mampu menjelaskan hubungan antarvariabel secara sistematis [10]. Sementara itu, penelitian yang dilakukan oleh Sari dan Jannah menunjukkan bahwa algoritma C4.5 memiliki tingkat akurasi, presisi, dan recall yang lebih baik dibandingkan metode K-Nearest Neighbor pada klasifikasi anemia [11].

Kajian-kajian terdahulu memperlihatkan bahwa algoritma C4.5 telah berhasil diterapkan pada berbagai bidang, seperti kesehatan, pendidikan, layanan publik, dan sistem pendukung keputusan. Namun, pemanfaatan algoritma tersebut untuk menganalisis data pelanggaran disiplin dan kode etik anggota Polri masih sangat terbatas. Sebagian besar penelitian terdahulu berfokus pada klasifikasi kondisi kesehatan, evaluasi layanan, maupun penentuan bantuan sosial. Sementara itu, penelitian yang mengkaji hubungan antara kategori pelanggaran dan tingkat sanksi dalam lingkungan kepolisian belum banyak ditemukan. Kesenjangan inilah yang menjadi dasar perlunya penelitian lebih lanjut pada domain penegakan disiplin dan kode etik anggota Polri.

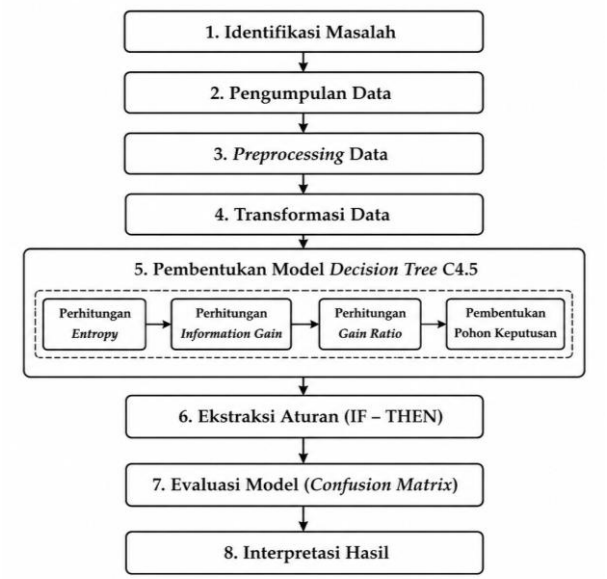
Kontribusi utama penelitian ini terletak pada penerapan algoritma Decision Tree C4.5 untuk mengidentifikasi pola hubungan antara jenis pelanggaran dan tingkat sanksi yang dijatuhkan kepada anggota Polri. Penelitian ini tidak hanya menghasilkan model klasifikasi, tetapi juga menghasilkan aturan keputusan yang dapat digunakan sebagai referensi dalam proses pembinaan, pengawasan, dan penegakan disiplin di lingkungan Polresta Gorontalo Kota. Dengan demikian, hasil penelitian diharapkan mampu memberikan dukungan terhadap pengambilan keputusan yang lebih objektif, konsisten, dan berbasis data.

Penelitian ini bertujuan untuk menganalisis data pelanggaran disiplin dan kode etik anggota Polri di Polresta Gorontalo Kota menggunakan algoritma Decision Tree C4.5 serta membangun model klasifikasi yang dapat mengungkap pola hubungan antara jenis pelanggaran dan tingkat sanksi. Hasil penelitian diharapkan dapat menjadi sumber informasi bagi fungsi Propam dalam meningkatkan efektivitas pembinaan dan pengawasan anggota sekaligus mendukung penerapan tata kelola organisasi yang lebih profesional.

2. Metode Penelitian

Penelitian ini dilaksanakan di Polresta Gorontalo Kota dengan memanfaatkan data pelanggaran disiplin dan kode etik anggota Polri periode 2022–2026. Penelitian menggunakan pendekatan kuantitatif dengan metode data mining untuk menganalisis pola pelanggaran serta hubungan antara jenis pelanggaran dan tingkat sanksi yang diberikan. Pendekatan kuantitatif dipilih karena penelitian berfokus pada pengolahan data historis untuk menghasilkan model klasifikasi yang dapat digunakan sebagai dasar pengambilan keputusan. Metode klasifikasi yang digunakan adalah algoritma Decision Tree C4.5 karena mampu menghasilkan model yang mudah dipahami serta direpresentasikan dalam bentuk pohon keputusan (decision tree) dan aturan keputusan (decision rules) [4], [7].

Tahapan penelitian diawali dengan identifikasi masalah dan pengumpulan data pelanggaran disiplin serta kode etik anggota Polri. Data yang diperoleh kemudian melalui proses preprocessing dan transformasi data untuk menghasilkan dataset yang siap diolah. Selanjutnya dilakukan pembentukan model klasifikasi menggunakan algoritma Decision Tree C4.5 melalui perhitungan entropy, information gain, dan gain ratio untuk membangun pohon keputusan. Model yang terbentuk kemudian diekstraksi menjadi aturan keputusan (IF–THEN), dilanjutkan dengan evaluasi menggunakan confusion matrix, serta interpretasi hasil untuk mengidentifikasi pola hubungan antara jenis pelanggaran dan tingkat sanksi yang diberikan. Alur penelitian ditunjukkan pada Gambar 1.



Gambar 1. Tahapan Penelitian Analisis Data Pelanggaran Disiplin dan Kode Etik Menggunakan Algoritma C4.5

ahapan penelitian yang dilakukan pada penelitian ini ditunjukkan pada Gambar 1. Proses penelitian diawali dengan identifikasi masalah dan pengumpulan data, kemudian dilanjutkan dengan preprocessing dan transformasi data sebelum dilakukan pembentukan model klasifikasi menggunakan algoritma Decision Tree C4.5. Model yang dihasilkan dievaluasi menggunakan confusion matrix dan diinterpretasikan untuk memperoleh pola hubungan antara jenis pelanggaran dan tingkat sanksi.

2.1 Tahapan Penelitian

Tahapan penelitian yang dilakukan pada penelitian ini ditunjukkan pada Gambar 1. Penelitian diawali dengan identifikasi masalah untuk memahami kondisi pengelolaan data pelanggaran disiplin dan kode etik anggota Polri di Polresta Gorontalo Kota serta mengidentifikasi kebutuhan analisis yang diperlukan. Tahap selanjutnya adalah pengumpulan data melalui studi dokumentasi dan wawancara untuk memperoleh data pelanggaran disiplin dan kode etik periode 2022–2026.

Data yang telah terkumpul kemudian melalui tahap preprocessing untuk meningkatkan kualitas data sebelum proses analisis. Tahap ini meliputi pemeriksaan kelengkapan data, standarisasi atribut, serta penyesuaian format data agar konsisten dan siap digunakan dalam proses klasifikasi. Setelah itu dilakukan transformasi data dengan mengelompokkan atribut yang relevan sesuai kebutuhan algoritma Decision Tree C4.5.

Tahap berikutnya adalah pembentukan model klasifikasi menggunakan algoritma Decision Tree C4.5. Proses ini dilakukan melalui perhitungan entropy, information gain, dan gain ratio untuk menentukan atribut terbaik sebagai pembentuk pohon keputusan (decision tree) [4], [6]. Pohon keputusan yang dihasilkan kemudian diekstraksi menjadi aturan keputusan (IF–THEN) yang dapat digunakan untuk menjelaskan hubungan antara kategori pelanggaran dan tingkat sanksi yang diberikan.

Model yang terbentuk selanjutnya dievaluasi menggunakan confusion matrix untuk mengukur performa klasifikasi berdasarkan nilai akurasi, presisi, recall, dan F1-score [9]–[11]. Tahap terakhir adalah interpretasi hasil untuk mengidentifikasi pola pelanggaran serta hubungan antara jenis pelanggaran dan tingkat sanksi yang dapat dimanfaatkan sebagai pendukung pengambilan keputusan dalam proses pembinaan dan penegakan disiplin anggota Polri.

2.2 Algoritma Decision Tree C4.5

Penelitian ini menggunakan algoritma Decision Tree C4.5 sebagai metode klasifikasi untuk menganalisis data pelanggaran disiplin dan kode etik anggota Polri. Algoritma C4.5 dipilih karena mampu menghasilkan model klasifikasi dalam bentuk pohon keputusan (decision tree) yang mudah dipahami dan diinterpretasikan oleh pengguna [4], [6]. Proses pembentukan pohon keputusan dilakukan dengan memilih atribut terbaik sebagai node akar (root node) berdasarkan nilai Gain Ratio tertinggi. Pemilihan atribut dilakukan melalui perhitungan Entropy, Information Gain, dan Gain Ratio pada setiap atribut yang digunakan dalam penelitian [7]. Proses tersebut dilakukan secara berulang hingga seluruh data berada pada kelas yang homogen atau tidak terdapat atribut yang dapat digunakan untuk melakukan pemisahan lebih lanjut.

Hasil dari proses klasifikasi berupa pohon keputusan (decision tree) yang kemudian diekstraksi menjadi aturan keputusan (IF–THEN rules). Aturan yang dihasilkan digunakan untuk mengidentifikasi pola hubungan antara kategori pelanggaran dan tingkat sanksi pada data pelanggaran disiplin dan kode etik anggota Polri di Polresta Gorontalo Kota.

2.3 Evaluasi Model

Evaluasi model dilakukan untuk mengetahui kemampuan algoritma Decision Tree C4.5 dalam mengklasifikasikan tingkat sanksi pelanggaran disiplin dan kode etik anggota Polri. Proses evaluasi menggunakan confusion matrix, yaitu metode yang umum digunakan untuk mengukur performa model klasifikasi berdasarkan hasil prediksi dan data aktual [9]–[11]. Parameter evaluasi yang digunakan meliputi accuracy, precision, recall, dan F1-score. Nilai accuracy digunakan untuk mengukur tingkat ketepatan model dalam melakukan klasifikasi secara keseluruhan. Precision menunjukkan tingkat ketepatan prediksi pada setiap kelas yang dihasilkan oleh model, sedangkan recall digunakan untuk mengukur kemampuan model dalam mengidentifikasi seluruh data pada kelas yang sebenarnya. Adapun F1-score merupakan kombinasi antara precision dan recall yang digunakan untuk memberikan gambaran performa model secara lebih seimbang [10], [11].

Hasil evaluasi digunakan untuk menilai tingkat performa model klasifikasi yang dihasilkan serta mengetahui kemampuan algoritma Decision Tree C4.5 dalam mengidentifikasi pola hubungan antara kategori pelanggaran dan tingkat sanksi pada data pelanggaran disiplin dan kode etik anggota Polri di Polresta Gorontalo Kota. Nilai evaluasi yang diperoleh selanjutnya digunakan sebagai dasar untuk

menganalisis efektivitas model dalam mendukung proses pengambilan keputusan pada fungsi pembinaan dan penegakan disiplin anggota Polri.

3. Hasil dan Pembahasan

3.1 Hasil Preprocessing Data

Data pelanggaran disiplin dan kode etik anggota Polri yang diperoleh dari Polresta Gorontalo Kota berjumlah 65 kasus. Setelah dilakukan proses preprocessing, seluruh data dinyatakan valid dan dapat digunakan dalam proses klasifikasi karena tidak ditemukan data kosong (missing value) maupun inkonsistensi atribut yang dapat memengaruhi proses pembentukan model. Selanjutnya dilakukan transformasi atribut Putusan Sidang menjadi tiga kategori tingkat sanksi, yaitu ringan, sedang, dan berat. Hasil transformasi menghasilkan distribusi data sebagaimana ditunjukkan pada Tabel 1.

Tabel 1. Distribusi Tingkat Sanksi Hasil Transformasi Data

Tingkat Sanksi	Jumlah	Persentase
Ringan	16	24,6%
Sedang	17	26,2%
Berat	32	49,2%
Total	**65**	**100%**

Sumber: Hasil Perhitungan Penelitian, 2026

Hasil transformasi menunjukkan bahwa kategori sanksi berat memiliki jumlah kasus tertinggi yaitu 32 kasus (49,2%), diikuti kategori sanksi sedang sebanyak 17 kasus (26,2%), dan kategori sanksi ringan sebanyak 16 kasus (24,6%). Distribusi data tersebut selanjutnya digunakan sebagai dasar dalam proses pembentukan model klasifikasi menggunakan algoritma Decision Tree C4.5.

3.2 Pembentukan Model Decision Tree C4.5

Pembentukan model klasifikasi dilakukan menggunakan algoritma Decision Tree C4.5 dengan menghitung nilai Information Gain dan Gain Ratio pada setiap atribut yang digunakan dalam penelitian. Proses ini bertujuan untuk menentukan atribut yang memiliki kemampuan terbaik dalam memisahkan data berdasarkan tingkat sanksi yang diberikan kepada anggota Polri. Hasil perhitungan menunjukkan bahwa atribut Kategori Pelanggaran memiliki nilai Information Gain sebesar 0,690 dan Gain Ratio sebesar 0,694. Nilai tersebut lebih tinggi dibandingkan atribut Wujud Perbuatan yang memiliki Information Gain sebesar 0,622 dan Gain Ratio sebesar 0,171. Ringkasan hasil perhitungan ditunjukkan pada Tabel 2.

Tabel 2. Hasil Perhitungan Gain Ratio

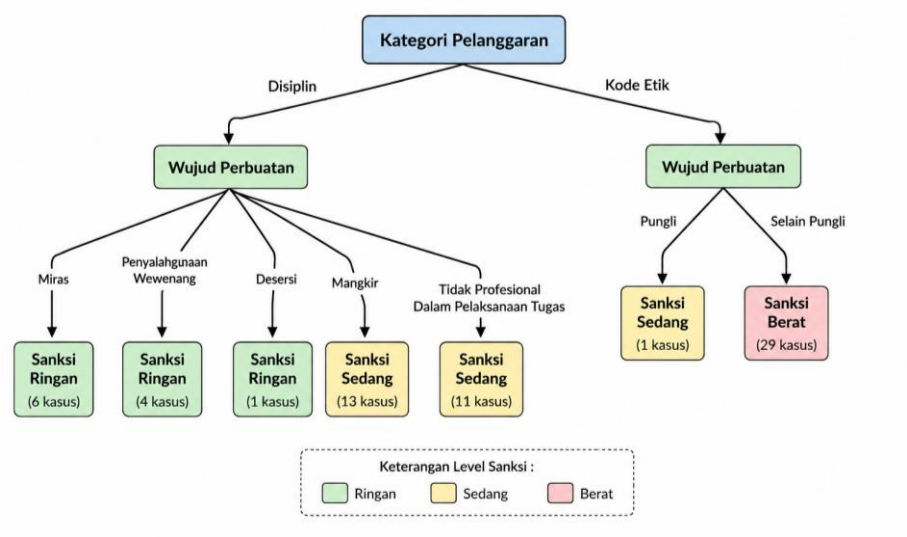
Atribut	Information Gain	Split Info	Gain Ratio
Kategori Pelanggaran	0,690	0,995	**0,694**
Wujud Perbuatan	0,622	3,623	0,171

Sumber: Hasil Perhitungan Penelitian, 2026

Hasil pada Tabel 2 menunjukkan bahwa atribut Kategori Pelanggaran memiliki nilai Gain Ratio tertinggi sehingga dipilih sebagai akar pohon (root node) dalam pembentukan model klasifikasi. Nilai Gain Ratio yang lebih tinggi menunjukkan bahwa atribut tersebut memiliki kemampuan yang lebih baik dalam memisahkan data ke dalam kategori tingkat sanksi yang berbeda. Pemilihan atribut Kategori Pelanggaran sebagai root node menunjukkan bahwa klasifikasi pelanggaran ke dalam kategori Disiplin atau Kode Etik merupakan faktor yang paling berpengaruh dalam menentukan tingkat sanksi yang diberikan. Setelah root node terbentuk, algoritma C4.5 melakukan proses pemisahan data secara berulang pada setiap cabang menggunakan atribut yang masih tersedia hingga diperoleh pohon keputusan yang mampu mengklasifikasikan data secara optimal. Hasil pembentukan pohon keputusan selanjutnya disajikan pada Subbab 3.3.

3.3 Pohon Keputusan dan Aturan Klasifikasi

Berdasarkan hasil perhitungan Gain Ratio, atribut Kategori Pelanggaran dipilih sebagai akar pohon (root node) karena memiliki nilai Gain Ratio tertinggi sebesar 0,694. Selanjutnya, algoritma C4.5 melakukan proses pemisahan data secara bertahap hingga diperoleh pohon keputusan yang mampu mengklasifikasikan tingkat sanksi berdasarkan kategori pelanggaran dan wujud perbuatan. Hasil pembentukan pohon keputusan ditunjukkan pada Gambar 2.



Gambar 2. Pohon Keputusan Hasil Algoritma Decision Tree C4.5

Gambar 2 menunjukkan bahwa atribut Kategori Pelanggaran menjadi faktor utama dalam menentukan tingkat sanksi yang diberikan kepada anggota Polri. Pada cabang Kode Etik, sebagian besar data langsung mengarah pada kategori sanksi berat. Dari 30 kasus pelanggaran kode etik, sebanyak 29 kasus menghasilkan sanksi berat, sedangkan satu kasus lainnya menghasilkan sanksi sedang. Kondisi ini menunjukkan bahwa pelanggaran yang masuk dalam ranah kode etik cenderung mendapatkan penanganan yang lebih tegas dibandingkan pelanggaran disiplin.

Sebaliknya, pada cabang Disiplin, tingkat sanksi yang diberikan lebih bervariasi sehingga diperlukan pemisahan lebih lanjut berdasarkan atribut Wujud Perbuatan. Hasil klasifikasi menunjukkan bahwa beberapa jenis pelanggaran memiliki kecenderungan terhadap tingkat sanksi tertentu. Pelanggaran berupa mangkir dan tidak profesional dalam pelaksanaan tugas cenderung menghasilkan sanksi sedang, sedangkan pelanggaran seperti miras lebih banyak menghasilkan sanksi ringan. Variasi tersebut menunjukkan bahwa dalam ranah disiplin, jenis perbuatan memiliki pengaruh yang cukup besar terhadap tingkat sanksi yang dijatuhkan. Pohon keputusan yang terbentuk kemudian diekstraksi menjadi sejumlah aturan klasifikasi (IF–THEN rules) yang dapat digunakan sebagai dasar pendukung pengambilan keputusan. Beberapa aturan utama yang dihasilkan ditunjukkan pada Tabel 3.

Tabel 3. Aturan Klasifikasi Hasil Decision Tree C4.5

No	Aturan Klasifikasi
1	IF Kategori Pelanggaran = Kode Etik AND Wujud Perbuatan ≠ Pungli THEN Level Sanksi = Berat
2	IF Kategori Pelanggaran = Kode Etik AND Wujud Perbuatan = Pungli THEN Level Sanksi = Sedang
3	IF Kategori Pelanggaran = Disiplin AND Wujud Perbuatan = Miras THEN Level Sanksi = Ringan
4	IF Kategori Pelanggaran = Disiplin AND Wujud Perbuatan = Mangkir THEN Level Sanksi = Sedang
5	IF Kategori Pelanggaran = Disiplin AND Wujud Perbuatan = Tidak Profesional Dalam Pelaksanaan Tugas THEN Level Sanksi = Sedang

Sumber: Hasil Perhitungan Penelitian, 2026

Aturan-aturan yang dihasilkan menunjukkan bahwa kategori pelanggaran merupakan variabel yang paling dominan dalam menentukan tingkat sanksi. Selain itu, model juga mampu

mengidentifikasi pola historis yang terjadi pada data pelanggaran di Polresta Gorontalo Kota. Informasi tersebut dapat dimanfaatkan sebagai dasar pendukung pengambilan keputusan dalam proses pembinaan, pengawasan, dan penegakan disiplin anggota Polri sehingga proses penjatuhan sanksi dapat dilakukan secara lebih objektif dan berbasis data.

3.4 Evaluasi Model

Evaluasi model dilakukan untuk mengukur kemampuan algoritma Decision Tree C4.5 dalam mengklasifikasikan tingkat sanksi pelanggaran disiplin dan kode etik anggota Polri. Pengujian dilakukan menggunakan confusion matrix dengan parameter evaluasi berupa accuracy, precision, recall, dan F1-score [9]–[11]. Hasil pengujian menunjukkan bahwa model klasifikasi yang dihasilkan memperoleh tingkat akurasi sebesar 67,69%. Nilai tersebut menunjukkan bahwa algoritma C4.5 mampu mengklasifikasikan sebagian besar data pelanggaran sesuai dengan kelas tingkat sanksi yang sebenarnya. Ringkasan hasil evaluasi model ditunjukkan pada Tabel 4.

Tabel 4. Hasil Evaluasi Model

Kelas	Precision (%)	Recall (%)	F1-Score (%)
Ringan	62,5	62,5	62,5
Sedang	52,9	52,9	52,9
Berat	78,1	78,1	78,1

Sumber: Hasil Perhitungan Penelitian, 2026

Hasil evaluasi menunjukkan bahwa performa terbaik diperoleh pada kategori sanksi berat dengan nilai precision sebesar 78,1%, recall sebesar 78,1%, dan F1-score sebesar 78,1%. Tingginya nilai tersebut menunjukkan bahwa model memiliki kemampuan yang baik dalam mengidentifikasi kasus pelanggaran yang berpotensi menghasilkan sanksi berat. Kondisi ini sejalan dengan hasil pembentukan pohon keputusan yang menunjukkan bahwa sebagian besar pelanggaran pada kategori kode etik memiliki kecenderungan menghasilkan sanksi berat. Pada kategori sanksi ringan, model memperoleh nilai precision sebesar 62,5%, recall sebesar 62,5%, dan F1-score sebesar 62,5%. Hasil tersebut menunjukkan bahwa model cukup baik dalam mengenali pelanggaran yang menghasilkan sanksi ringan, meskipun masih terdapat beberapa data yang diklasifikasikan ke dalam kategori lain. Sementara itu, kategori sanksi sedang menunjukkan performa terendah dengan nilai precision, recall, dan F1-score masing-masing sebesar 52,9%. Nilai tersebut mengindikasikan bahwa karakteristik pelanggaran yang menghasilkan sanksi sedang memiliki kemiripan dengan kategori sanksi lainnya sehingga lebih sulit dipisahkan oleh model klasifikasi.

Secara keseluruhan, hasil evaluasi menunjukkan bahwa algoritma Decision Tree C4.5 mampu menghasilkan model klasifikasi dengan performa yang cukup baik untuk mengidentifikasi hubungan antara kategori pelanggaran, wujud perbuatan, dan tingkat sanksi pada data pelanggaran disiplin dan kode etik anggota Polri di Polresta Gorontalo Kota. Tingkat akurasi yang diperoleh menunjukkan bahwa model dapat digunakan sebagai alat bantu dalam mendukung proses pengambilan keputusan pada fungsi pembinaan dan penegakan disiplin anggota Polri.

3.5 Pembahasan

Hasil penelitian menunjukkan bahwa algoritma Decision Tree C4.5 mampu mengidentifikasi pola hubungan antara kategori pelanggaran, wujud perbuatan, dan tingkat sanksi pada data pelanggaran disiplin dan kode etik anggota Polri di Polresta Gorontalo Kota. Temuan utama penelitian ini menunjukkan bahwa atribut Kategori Pelanggaran merupakan faktor yang paling berpengaruh dalam proses klasifikasi, yang ditunjukkan oleh nilai Gain Ratio tertinggi sebesar 0,694. Hasil tersebut mengindikasikan bahwa pengelompokan pelanggaran ke dalam kategori disiplin dan kode etik memiliki peran dominan dalam menentukan tingkat sanksi yang diberikan kepada anggota Polri. Pohon keputusan yang dihasilkan menunjukkan bahwa sebagian besar pelanggaran yang termasuk dalam kategori kode etik berujung pada sanksi berat. Temuan ini menunjukkan adanya konsistensi dalam penerapan ketentuan disiplin dan kode etik di lingkungan Polresta Gorontalo Kota,

di mana pelanggaran yang berkaitan dengan integritas, penyalahgunaan wewenang, tindak pidana, maupun perilaku yang mencederai citra institusi cenderung dikenakan sanksi yang lebih tegas. Sebaliknya, pada kategori pelanggaran disiplin ditemukan variasi tingkat sanksi yang lebih beragam, sehingga diperlukan atribut tambahan berupa wujud perbuatan untuk menentukan tingkat sanksi yang sesuai. Kondisi tersebut menunjukkan bahwa karakteristik pelanggaran disiplin memiliki tingkat kompleksitas yang lebih tinggi dibandingkan pelanggaran kode etik dalam proses klasifikasi.

Hasil evaluasi model menunjukkan bahwa algoritma Decision Tree C4.5 memperoleh tingkat akurasi sebesar 67,69%. Nilai tersebut menunjukkan bahwa model mampu mengklasifikasikan sebagian besar data pelanggaran dengan baik. Performa terbaik diperoleh pada kategori sanksi berat dengan nilai F1-score sebesar 78,1%, sedangkan performa terendah terdapat pada kategori sanksi sedang dengan nilai F1-score sebesar 52,9%. Perbedaan performa ini menunjukkan bahwa karakteristik data pada kategori sanksi berat lebih mudah dikenali karena memiliki pola yang lebih konsisten dibandingkan kategori sanksi sedang. Sebaliknya, kategori sanksi sedang memiliki karakteristik yang lebih beragam sehingga sebagian data memiliki kemiripan dengan kategori ringan maupun berat. Temuan penelitian ini sejalan dengan penelitian yang dilakukan oleh Liputo, Tupamahu, dan rekan-rekan yang menunjukkan bahwa algoritma Decision Tree mampu menghasilkan model klasifikasi yang mudah dipahami dan efektif dalam mengidentifikasi pola pada data kategorikal [5]. Kesamaan tersebut terlihat dari kemampuan pohon keputusan dalam menghasilkan aturan klasifikasi yang dapat diinterpretasikan secara langsung oleh pengguna. Hasil penelitian ini juga mendukung penelitian Lasarudin, Hasyim, dan Dumako yang menyimpulkan bahwa algoritma C4.5 mampu menghasilkan aturan keputusan yang jelas sehingga dapat digunakan sebagai dasar evaluasi dan pengambilan keputusan pada suatu organisasi [6].

Perbedaan penelitian ini dengan penelitian terdahulu terletak pada objek dan konteks penerapannya. Penelitian sebelumnya lebih banyak diterapkan pada bidang pendidikan, kesehatan, dan layanan publik [5]–[11], sedangkan penelitian ini mengimplementasikan algoritma C4.5 pada data pelanggaran disiplin dan kode etik anggota Polri. Penerapan pada domain kepolisian menghasilkan informasi baru mengenai pola hubungan antara kategori pelanggaran dan tingkat sanksi yang diberikan. Dengan demikian, penelitian ini memberikan kontribusi dalam pemanfaatan teknik data mining sebagai alat bantu analisis pada bidang pembinaan dan penegakan disiplin aparatur negara. Dari perspektif praktis, aturan klasifikasi yang dihasilkan dapat dimanfaatkan sebagai pendukung pengambilan keputusan pada fungsi Profesi dan Pengamanan (Propam). Informasi yang diperoleh dari model dapat digunakan untuk mengidentifikasi kecenderungan tingkat sanksi berdasarkan karakteristik pelanggaran yang dilakukan anggota. Selain itu, hasil penelitian juga dapat menjadi bahan evaluasi dalam penyusunan strategi pembinaan dan pengawasan guna meminimalkan terjadinya pelanggaran serupa di masa mendatang. Temuan penelitian menunjukkan bahwa pelanggaran dalam kategori kode etik memiliki kecenderungan menghasilkan sanksi berat, sedangkan pelanggaran disiplin menunjukkan pola yang lebih bervariasi dan dipengaruhi oleh wujud perbuatan yang dilakukan. Pada ranah disiplin, pelanggaran seperti mangkir dan tidak profesional memiliki kecenderungan menghasilkan sanksi sedang, sedangkan beberapa bentuk pelanggaran lainnya dapat menghasilkan tingkat sanksi yang berbeda sesuai karakteristik kasus yang terjadi.

Meskipun demikian, penelitian ini masih memiliki keterbatasan. Dataset yang digunakan hanya berasal dari satu satuan kerja, yaitu Polresta Gorontalo Kota, dengan jumlah data sebanyak 65 kasus. Selain itu, atribut yang digunakan dalam proses klasifikasi masih terbatas pada kategori pelanggaran dan wujud perbuatan. Penelitian selanjutnya dapat menggunakan jumlah data yang lebih besar serta menambahkan atribut lain, seperti riwayat pelanggaran, masa dinas, pangkat, atau satuan fungsi anggota, sehingga model yang dihasilkan memiliki tingkat akurasi dan kemampuan generalisasi yang lebih baik.

4. Kesimpulan

Penelitian ini berhasil menerapkan algoritma Decision Tree C4.5 untuk menganalisis data pelanggaran disiplin dan kode etik anggota Polri di Polresta Gorontalo Kota periode 2022–2026. Hasil penelitian menunjukkan bahwa atribut Kategori Pelanggaran merupakan faktor yang paling berpengaruh dalam proses klasifikasi dengan nilai Gain Ratio tertinggi sebesar 0,694, sehingga ditetapkan sebagai akar pohon keputusan (root node). Pohon keputusan yang dihasilkan mampu mengidentifikasi pola hubungan antara kategori pelanggaran, wujud perbuatan, dan tingkat sanksi yang diberikan. Hasil klasifikasi menunjukkan bahwa pelanggaran yang termasuk dalam kategori kode etik cenderung menghasilkan sanksi berat, sedangkan pelanggaran disiplin memiliki variasi tingkat sanksi yang dipengaruhi oleh jenis perbuatan yang dilakukan. Pada kategori pelanggaran disiplin, wujud perbuatan menjadi faktor penentu dalam proses klasifikasi sehingga menghasilkan tingkat sanksi yang berbeda sesuai karakteristik pelanggaran yang terjadi.

Evaluasi model menunjukkan bahwa algoritma Decision Tree C4.5 memperoleh tingkat akurasi sebesar 67,69% dengan performa terbaik pada kategori sanksi berat yang memiliki nilai F1-score sebesar 78,1%. Hasil tersebut menunjukkan bahwa model mampu mengidentifikasi pola pelanggaran dan tingkat sanksi dengan cukup baik serta dapat digunakan sebagai alat bantu pendukung pengambilan keputusan dalam proses pembinaan, pengawasan, dan penegakan disiplin anggota Polri di lingkungan Polresta Gorontalo Kota. Penelitian selanjutnya disarankan menggunakan jumlah data yang lebih besar dan menambahkan atribut lain yang relevan, seperti riwayat pelanggaran, masa dinas, pangkat, dan satuan fungsi anggota, sehingga dapat meningkatkan akurasi model dan menghasilkan pola klasifikasi yang lebih komprehensif.

5. Daftar Pustaka

[1] I. K. A. Subagio and A. M. N. Limbong, "Dampak Teknologi Informasi dan Komunikasi terhadap Aktivitas Pendidikan," *Journal of Learning and Technology*, vol. 2, no. 1, pp. 43–52, 2023, doi: 10.33830/jlt.v2i1.5844.

[2] Republik Indonesia, *Peraturan Pemerintah Republik Indonesia Nomor 2 Tahun 2003 tentang Peraturan Disiplin Anggota Kepolisian Negara Republik Indonesia*. Jakarta, Indonesia, 2003.

[3] Kepolisian Negara Republik Indonesia, *Peraturan Kepolisian Negara Republik Indonesia Nomor 7 Tahun 2022 tentang Kode Etik Profesi Polri dan Komisi Kode Etik Polri*. Jakarta, Indonesia, 2022.

[4] A. Srirahayu and L. Setya Pribadie, "Review Paper Data Mining Klasifikasi Data Mining," *Jurnal Ilmiah Informatika Global*, vol. 13, no. 2, 2022, doi: 10.36982/jiig.v13i2.2307.

[5] B. S. D. Nugraha, S. Defit, and G. W. Nurcahyo, "Algoritma C4.5 untuk Prediksi Bimbingan Siswa Berdasarkan Tipologi Hippocrates-Galenus," *Jurnal Teknoif Teknik Informatika Institut Teknologi Padang*, vol. 11, no. 1, pp. 1–8, 2023, doi: 10.21063/jtif.2023.v11.1.1-8.

[6] A. Lasarudin, W. Hasyim, and R. Dumako, "Analisis Kepuasan Masyarakat terhadap Layanan Siaran RRI Gorontalo Menggunakan Algoritma C4.5," *Jurnal Repositor*, 2022.

[7] A. Lasarudin, H. Gani, and M. Tomayahu, "Perbandingan Metode Naïve Bayes dan C4.5 Klasifikasi Status Gizi Bayi Balita," *SPECTA Journal of Technology*, vol. 6, no. 3, pp. 273–283, 2023, doi: 10.35718/specta.v6i3.789.

[8] N. A. A. Sari and U. M. Jannah, "Komparasi Algoritma C4.5 dan K-Nearest Neighbor (K-NN) pada Klasifikasi Penyakit Anemia," vol. 9, no. 2, 2024.

[9] M. Santoso, B. Al-Akbar, H. Nurjaya, S. A. Ramadhan, N. A. Rizky, and A. Fadillah, "Klasifikasi Potensi Penyakit Jantung Menggunakan Algoritma C4.5," *Jurnal INSAN (Journal of Information Systems Management Innovation)*, vol. 3, no. 2, 2023.

[10] M. B. Rahman and Z. Fatah, "Analisis Pola Kehadiran Mahasiswa Menggunakan Algoritma Decision Tree," *Jurnal Riset Sistem Informasi*, vol. 2, no. 1, pp. 60–66, 2025, doi: 10.69714/6z8kc143.

[11] K. J. Wijaya and B. A. Wijaya, "Implementasi Algoritma Clustering dan Classification dalam Data Mining: Systematic Literature Review terhadap Tren dan Tantangan Terkini," *Jurnal Publikasi Sistem Informasi dan Manajemen Bisnis*, vol. 4, no. 3, pp. 372–386, 2025, doi: 10.55606/jupsim.v4i3.5240.