

Penerapan Data Mining untuk Mengukur Prestasi Kinerja Dosen dengan Menggunakan Algoritma C4.5

Ester Ulina Sihotang¹, Gogor Abiezer Ginting², Icha Dahlia¹, Rizka Amanda³, David JM Sembiring^{1,*}, Sinek Mehuli Br Peranginangin⁴

¹Program Studi Teknik Informatika, Fakultas Sains dan Teknologi, Institut Teknologi dan Bisnis Indonesia, Medan, Indonesia

²Program Studi Teknologi Informasi, Fakultas Sains dan Teknologi, Institut Teknologi dan Bisnis Indonesia, Medan, Indonesia

³Program Studi Rekayasa Perangkat Lunak, Fakultas Sains dan Teknologi, Institut Teknologi dan Bisnis Indonesia, Medan, Indonesia

⁴Program Studi Manajemen Informatika, Fakultas Sains dan Teknologi, Institut Teknologi dan Bisnis Indonesia, Medan, Indonesia
Jl. Sei Batang Hari No.84a, Babura Sunggal, Kec. Medan Sunggal, Kota Medan, Sumatera Utara, Indonesia

Email: ^{1,*}davidjms Sembiring@itbi.ac.id, ²sinekmehuli19@gmail.com

(* : coresponding author)

Abstrak—Kinerja dosen merupakan salah satu faktor penting dalam peningkatan mutu pendidikan tinggi, karena dosen tidak hanya berperan sebagai pendidik, tetapi juga sebagai peneliti dan pengabdian kepada masyarakat. Namun, penilaian kinerja dosen seringkali menghadapi kendala, seperti belum adanya standar evaluasi yang seragam, kecenderungan subjektivitas dalam penilaian, serta keterbatasan instrumen evaluasi yang mampu menilai kinerja secara komprehensif. Untuk mengatasi permasalahan tersebut, diperlukan pendekatan berbasis data yang mampu memberikan hasil analisis objektif dan terukur. Salah satu metode yang dapat digunakan adalah data mining dengan algoritma C4.5, yang merupakan salah satu algoritma klasifikasi berbasis pohon keputusan. Penelitian ini bertujuan untuk menerapkan algoritma C4.5 dalam mengukur prestasi kinerja dosen berdasarkan data historis yang mencakup berbagai indikator tridharma perguruan tinggi. Tahapan penelitian meliputi identifikasi masalah, studi literatur, pengumpulan data, pemilihan teknik analisis, implementasi algoritma C4.5 dengan bantuan perangkat lunak RapidMiner, serta analisis hasil pengujian. Model klasifikasi yang dihasilkan divisualisasikan dalam bentuk pohon keputusan sehingga mudah dipahami dan dapat digunakan sebagai dasar evaluasi. Hasil pengujian menunjukkan bahwa algoritma C4.5 mampu menghasilkan model klasifikasi dengan tingkat akurasi sebesar 86,67%. Hal ini membuktikan bahwa C4.5 efektif dalam mengolah data kinerja dosen dan menghasilkan evaluasi yang lebih objektif, transparan, serta dapat mengurangi potensi subjektivitas dalam penilaian. Penelitian ini memberikan kontribusi strategis dalam mendukung pengambilan keputusan manajerial di perguruan tinggi, khususnya dalam merumuskan kebijakan peningkatan mutu pendidikan dan pengembangan profesional dosen secara berkelanjutan.

Kata Kunci: Data Mining; Pengukuran; Prestasi; Kinerja; Algoritma C4.5

Abstract—Lecturer performance is an important factor in improving the quality of higher education, because lecturers not only act as educators, but also as researchers and community service. However, lecturer performance assessment often faces obstacles, such as the lack of uniform evaluation standards, a tendency for subjectivity in assessments, and limited evaluation instruments capable of assessing performance comprehensively. To overcome these problems, a data-driven approach is needed that can provide objective and measurable analysis results. One method that can be used is data mining with the C4.5 algorithm, which is a decision tree-based classification algorithm. This study aims to apply the C4.5 algorithm to measure lecturer performance achievements based on historical data that includes various indicators of the tridharma of higher education. The research stages include problem identification, literature review, data collection, selection of analysis techniques, implementation of the C4.5 algorithm with the help of RapidMiner software, and analysis of test results. The resulting classification model is visualized in the form of a decision tree so that it is easy to understand and can be used as a basis for evaluation. The test results show that the C4.5 algorithm is able to produce a classification model with an accuracy level of 86.67%. This demonstrates that C4.5 is effective in processing lecturer performance data and producing more objective and transparent evaluations, while also reducing the potential for subjectivity in assessments. This research provides a strategic contribution to supporting managerial decision-making in higher education, particularly in formulating policies for improving the quality of education and sustainable professional development of lecturers.

Keywords: Data Mining; Measurement; Achievement; Performance; C4.5 Algorithm

1. PENDAHULUAN

Kinerja seorang dosen tidak hanya diukur dari kemampuan menyampaikan materi perkuliahan, tetapi juga dari bagaimana dosen mampu menjalankan perannya sebagai pendidik, peneliti, dan pengabdian kepada masyarakat. Dalam konteks tridharma perguruan tinggi, kinerja dosen mencakup aspek pengajaran yang berkualitas, penelitian yang relevan dengan kebutuhan ilmu pengetahuan maupun industri, serta pengabdian yang memberi dampak positif bagi masyarakat luas[1]. Selain itu, dosen juga dituntut untuk terus meningkatkan kompetensi akademik dan profesional melalui pengembangan diri serta pemanfaatan teknologi pembelajaran. Dengan kinerja yang baik, dosen tidak hanya berkontribusi pada peningkatan mutu pendidikan, tetapi juga mencetak lulusan yang kompeten, berdaya saing, dan memiliki karakter yang sesuai dengan kebutuhan bangsa[2].

Meskipun kinerja dosen menjadi aspek penting dalam peningkatan mutu pendidikan tinggi, proses penilaian kinerja seringkali menghadapi berbagai permasalahan. Salah satu kendala utama adalah belum adanya standar penilaian yang seragam dan komprehensif, sehingga penilaian kinerja cenderung subjektif dan bervariasi antar institusi. Selain itu, sistem penilaian yang ada terkadang lebih menekankan pada aspek administratif, seperti jumlah publikasi atau laporan kegiatan, daripada menilai kualitas sebenarnya dari proses pembelajaran dan dampak penelitian[3]. Beban kerja dosen yang meliputi tridharma perguruan tinggi juga seringkali tidak seimbang, sehingga sulit mengukur kontribusi dosen secara proporsional pada setiap bidang[4]. Permasalahan lain adalah keterbatasan instrumen evaluasi yang mampu menangkap

kinerja dosen secara holistik, termasuk soft skills, inovasi pembelajaran, dan kemampuan adaptasi terhadap perkembangan teknologi. Hal ini membuat penilaian kinerja dosen berpotensi tidak mencerminkan kondisi sebenarnya, sehingga kurang efektif sebagai dasar untuk pengembangan kualitas pendidikan.

Urgensi penelitian terkait penilaian kinerja dosen sangatlah penting karena hasilnya dapat menjadi dasar dalam merumuskan strategi peningkatan mutu pendidikan tinggi secara berkelanjutan. Melalui penelitian, dapat diperoleh model evaluasi yang lebih objektif, terukur, dan sesuai dengan tuntutan tridarma perguruan tinggi serta perkembangan era digital. Tanpa adanya penelitian yang mendalam, sistem penilaian kinerja berisiko tetap bersifat subjektif, tidak konsisten, dan kurang mencerminkan kualitas sebenarnya dari kontribusi dosen. Kondisi ini dapat menimbulkan dampak negatif, seperti rendahnya motivasi dosen untuk meningkatkan kualitas pengajaran dan penelitian, terhambatnya inovasi dalam pembelajaran, serta menurunnya daya saing institusi pendidikan di tingkat nasional maupun internasional. Oleh karena itu, penelitian dalam bidang ini tidak hanya mendesak dilakukan, tetapi juga memiliki nilai strategis untuk menjamin keberlangsungan peningkatan kualitas pendidikan tinggi.

Salah satu pendekatan yang dapat digunakan untuk menyelesaikan permasalahan dalam penilaian kinerja dosen adalah penerapan teknik data mining. Data mining memungkinkan pengolahan data kinerja dosen yang kompleks, besar, dan beragam menjadi informasi yang lebih terstruktur dan bermakna[5]. Melalui algoritma klasifikasi, clustering, atau decision tree, data kinerja dosen dapat dianalisis untuk mengidentifikasi pola, kecenderungan, serta faktor-faktor utama yang memengaruhi kualitas kinerja. Dengan demikian, evaluasi yang dihasilkan menjadi lebih objektif, transparan, dan berbasis data, sehingga mampu mengurangi subjektivitas dalam penilaian[6]. Selain itu, hasil analisis data mining dapat dijadikan dasar untuk memberikan rekomendasi strategis dalam peningkatan kompetensi dosen, perbaikan metode pembelajaran, maupun penyesuaian beban tridarma secara lebih proporsional[7].

Data mining pada hakikatnya merupakan proses untuk menggali pengetahuan baru dari kumpulan data yang besar dengan memanfaatkan teknik statistik, matematika, dan kecerdasan buatan. Salah satu kekuatan utama data mining adalah kemampuannya dalam menganalisis data historis atau data masa lampau yang tersimpan dalam basis data. Melalui pengolahan data masa lampau, data mining dapat menemukan pola tersembunyi, hubungan antar variabel, serta tren yang mungkin tidak terlihat melalui analisis konvensional[8]. Informasi yang diperoleh dari data historis tersebut dapat digunakan sebagai dasar prediksi maupun pengambilan keputusan yang lebih akurat di masa depan. Dalam konteks penilaian kinerja dosen, analisis data masa lampau dapat memberikan gambaran mengenai konsistensi kinerja, efektivitas metode pembelajaran yang pernah diterapkan, serta kontribusi dosen dalam tridarma perguruan tinggi dari waktu ke waktu. Dengan demikian, data mining bukan hanya membantu memahami kondisi saat ini, tetapi juga berperan strategis dalam merancang kebijakan peningkatan mutu pendidikan di masa mendatang[9].

Salah satu algoritma data mining yang relevan digunakan dalam penilaian kinerja dosen adalah algoritma C4.5. Algoritma ini termasuk dalam metode klasifikasi yang bekerja dengan membangun sebuah decision tree berdasarkan data historis[10]. C4.5 mampu membagi data ke dalam kategori tertentu dengan menggunakan konsep information gain ratio, sehingga hasil klasifikasi lebih akurat dibandingkan algoritma pendahulunya seperti ID3. Dalam konteks penilaian kinerja dosen, algoritma C4.5 dapat digunakan untuk mengklasifikasikan dosen ke dalam kelompok kinerja tertentu, misalnya “sangat baik”, “baik”, “cukup”, atau “kurang”, berdasarkan variabel-variabel penilaian seperti publikasi ilmiah, metode pembelajaran, keterlibatan dalam pengabdian masyarakat, maupun penilaian mahasiswa[11]. Keunggulan algoritma ini adalah kemampuannya menangani data dengan nilai yang bervariasi, mengatasi atribut yang hilang, serta menghasilkan model yang mudah dipahami dalam bentuk pohon keputusan. Dengan demikian, penerapan algoritma C4.5 dapat memberikan evaluasi kinerja dosen yang lebih objektif, transparan, serta menjadi dasar pengambilan keputusan manajerial dalam peningkatan mutu pendidikan tinggi[12].

Algoritma C4.5 merupakan salah satu metode data mining yang banyak digunakan dalam teknik klasifikasi untuk membangun decision tree. Prinsip utama dari C4.5 adalah memilih atribut yang paling relevan dalam memisahkan data ke dalam kelas tertentu dengan menggunakan ukuran gain ratio. Proses ini dilakukan secara rekursif hingga terbentuk pohon keputusan yang merepresentasikan pola atau aturan dari data yang dianalisis[13]. Keunggulan algoritma C4.5 dibandingkan metode lainnya adalah kemampuannya menangani data kontinu maupun diskrit, mengelola nilai atribut yang hilang, serta menghasilkan model klasifikasi yang relatif mudah dipahami oleh pengguna karena divisualisasikan dalam bentuk pohon keputusan[14]. Dalam konteks penelitian penilaian kinerja dosen, algoritma C4.5 dapat dimanfaatkan untuk mengklasifikasikan kinerja berdasarkan variabel-variabel yang telah ditentukan, sehingga menghasilkan informasi yang objektif dan dapat digunakan sebagai dasar dalam pengambilan kebijakan peningkatan mutu pendidikan[15].

Beberapa penelitian terdahulu berkaitan dengan algoritma C4.5 sebagai pendukung terhadap pelaksanaan algoritma C4.5 dapat membantu mengklasifikasikan Masyarakat Penerima Bantuan Sosial[16], menganalisis kepuasan konsumen[17], untuk memprediksi pre-eklampsia dalam kehamilan[18], penentuan strategi promosi[19], klasifikasi produk[20] dan penilaian kelayakan kredit[21].

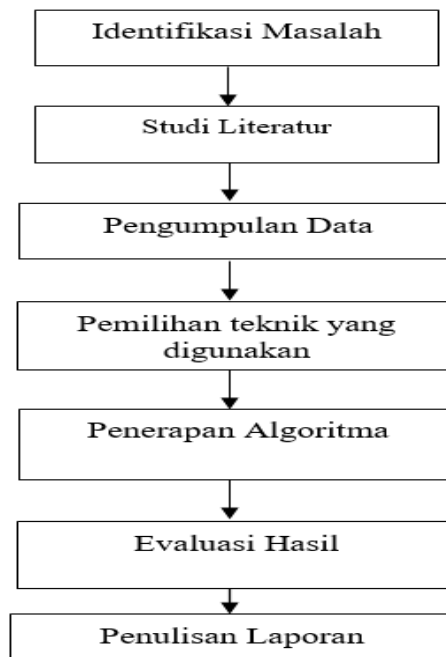
Penelitian ini memiliki perbedaan dengan penelitian sebelumnya karena fokus pada penerapan algoritma C4.5 dalam mengevaluasi kinerja dosen berdasarkan data historis. Jika penelitian terdahulu cenderung menitikberatkan pada penilaian administratif atau hanya berfokus pada satu indikator tertentu, maka penelitian ini berupaya menghadirkan pendekatan berbasis data mining untuk menghasilkan klasifikasi kinerja yang lebih objektif, terukur, dan transparan. Tujuan utama penelitian ini adalah merancang model evaluasi yang mampu mengolah data kinerja dosen secara akurat, mengurangi subjektivitas dalam proses penilaian, serta memberikan gambaran mengenai faktor-faktor utama yang memengaruhi efektivitas kinerja dosen. Hasil yang diharapkan dari penelitian ini adalah tersusunnya sistem evaluasi berbasis algoritma C4.5 yang dapat membantu institusi pendidikan dalam mengidentifikasi dosen dengan kinerja optimal,

merumuskan strategi pengembangan profesional, serta meningkatkan kualitas proses pembelajaran secara menyeluruh.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Merupakan model kerja dengan mempergunakan pendekatan dalam memecahkan masalah. Berikut ini merupakan framework penulis gambarkan.



Gambar 1. Kerangka Kerja Penelitian

Berdasarkan gambar 1, Sesuai dengan kerangka riset yang telah dijelaskan di atas, maka pembahasannya dalam riset ini, antara lain :

- a. Identifikasi Masalah
Pada tahapan ini mendefinisikan suatu permasalahan yang dihadapi dalam mengukur prestasi
- b. Studi Literatur
Pada tahapan ini mempelajari literatur yang berkaitan pada penelitian yang dijadikan referensi seperti buku, jurnal maupun situs web
- c. Pengumpulan Data
Pada kegiatan ini mengumpulkan data yang dibutuhkan pada riset dimana teknik mengumpulkan informasi dengan pengamatan langsung ke lapangan
- d. Pemilihan Teknik Yang Digunakan
Kegiatan ini digunakan untuk melakukan dan memiliki teknik yang tepat dalam melakukan proses pengukuran prestasi.
- e. Implementasi
Pada tahapan ini dilakukan untuk mengetahui proses dari metode tsukamoto dan Algoritma C.45
- f. Hasil Analisa
Pada tahapan ini digunakan untuk mengetahui hasil metode tsukamoto dan algoritma C4.5 dalam menentukan mengukur prestasi.
- g. Penulisan Laporan
Pada kegiatan ini adalah pembuatan laporan kerja dari riset yang telah dilaksanakan.

2.2 Data Mining

Data mining adalah suatu proses pengolahan data dalam jumlah besar untuk menemukan pola, hubungan, atau informasi yang sebelumnya tersembunyi dan belum diketahui. Proses ini merupakan bagian penting dari Knowledge Discovery in Database (KDD) yang bertujuan menghasilkan pengetahuan baru yang dapat digunakan sebagai dasar pengambilan keputusan. Data mining tidak hanya berfokus pada pengumpulan data, tetapi lebih kepada bagaimana data tersebut dianalisis dan diekstraksi sehingga mampu memberikan nilai tambah. Secara umum, data mining melibatkan tahapan pembersihan data, integrasi, seleksi, transformasi, hingga proses analisis yang kemudian dievaluasi dan disajikan dalam bentuk yang mudah dipahami[22], [23].

Dalam penerapannya, data mining memiliki beberapa teknik utama yang dapat digunakan sesuai dengan tujuan analisis. Teknik-teknik tersebut antara lain: klasifikasi yang digunakan untuk mengelompokkan data ke dalam kategori

tertentu, klasterisasi untuk mengelompokkan data berdasarkan kemiripan karakteristik tanpa label tertentu, asosiasi yang bertujuan menemukan hubungan antar variabel dalam suatu basis data, serta prediksi yang digunakan untuk memperkirakan nilai atau kecenderungan di masa depan berdasarkan data historis. Selain itu, terdapat pula teknik estimasi untuk memperkirakan nilai numerik dan analisis outlier untuk mendeteksi data yang menyimpang dari pola umum. Dengan berbagai teknik tersebut, data mining mampu membantu organisasi atau lembaga dalam menggali informasi berharga yang bermanfaat bagi pengambilan keputusan strategis[24], [25].

2.3 Algoritma C4.5

Algoritma C4.5 merupakan algoritma yang digunakan untuk membentuk pohon keputusan. Pohon keputusan merupakan metode klasifikasi dan prediksi yang sangat kuat dan terkenal. Metode pohon keputusan mengubah fakta yang sangat besar menjadi pohon keputusan yang merepresentasikan aturan. Aturan dapat dengan mudah dipahami dengan bahasa alami. Pohon Keputusan juga berguna untuk mengeksplorasi data, menemukan hubungan tersembunyi antara sejumlah calon variabel input dengan sebuah variabel target. Karena pohon keputusan memadukan antara eksplorasi data dan pemodelan, dia sangat bagus sebagai langkah awal dalam proses pemodelan bahkan ketika dijadikan sebagai model akhir dari beberapa teknik lain[26], [27].

Dalam pembuatan decision tree menggunakan metode C4.5 digunakan untuk membangun sebuah pohon keputusan dimulai dengan pemilihan variabel ataupun atribut sebagai akar, untuk membangun cabang pada setiap nilai, pembagian kasus pada cabang serta melakukan pengulangan proses di setiap cabang sampai keseluruhan kasus pada cabang memiliki kelas yang sama. Decision tree adalah teknik pengelompokan terhadap record, meliputi dari kumpulan decision node dan terhubung ke cabang, mundur ke bawah dari root node sampai berakhir di leaf node. Data pada decision tree biasanya dapat dinyatakan ke dalam bentuk tabel ataupun atribut serta record. Pada teknik decision tree dengan memakai teknik “membagi dan menaklukkan” dalam pembagian ruang himpunan dengan pencarian masalah. Berikut ini adalah tahapan yang terdapat pada algoritma C4.5 antara lain[28]:

a. Mempersiapkan data training

Pada proses ini data training biasanya data dapat diambil dari sejarah sebelumnya telah terjadi serta telah diklasifikasi pada kelas tertentu.

b. Menghitung akar dari pohon

Sebelum menghitung gain pada atribut, maka proses pertama adalah menghitung nilai entropy dengan menggunakan rumus [29]:

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

Pada rumus fungsi diatas dapat dijelaskan untuk tahapan sebagai berikut Entropy (S) menggambarkan terhadap banyaknya kasus yang terdapat pada data. Nilai n merupakan jumlah dari setiap pembagian nilai pada masing-masing kelas. P_i merupakan pembagian terhadap kelas dengan masing-masing nilai atribut.

c. Perhitungan nilai gain menggunakan rumus[30]:

$$\text{Gain (S, A)} = \text{entropy (S)} - \sum_{i=1}^n \frac{|S_i|}{|S|} \text{entropy (S}_i) \quad (2)$$

Secara naratif, rumus tersebut menjelaskan bahwa nilai information gain diperoleh dengan cara menghitung selisih antara entropy awal dari seluruh data (entropy(S)) dengan rata-rata tertimbang entropy dari setiap subset data (S_i) yang terbentuk setelah data dibagi berdasarkan atribut A. Bobot rata-rata tersebut ditentukan oleh perbandingan ukuran subset $|S_i|/|S|$ terhadap ukuran keseluruhan data $|S|/|S|$. Dengan demikian, semakin besar penurunan nilai entropy setelah pembagian data menggunakan atribut A, maka semakin besar pula nilai gain-nya. Artinya, atribut tersebut dianggap lebih informatif dan lebih baik untuk digunakan dalam proses pembentukan cabang pada pohon keputusan.

d. Ulangi langkah ke 2 dan 3 sampai semua register terpecah

e. Keputusan proses partisi pohon akan berhenti jika:

1. Semua record ada di node N memiliki kelas yang sama
2. Atribut untuk record yang dipartisi tidak ada lagi
3. Catatan cabang kosong tidak ada lagi

3. HASIL DAN PEMBAHASAN

3.1 Analisa Masalah

Analisa masalah merupakan tahap penting dalam penelitian karena berfungsi untuk mengidentifikasi akar persoalan yang dihadapi, sehingga dapat ditentukan pendekatan yang tepat dalam mencari solusi. Dalam konteks penilaian kinerja dosen, permasalahan utama yang muncul adalah belum adanya metode evaluasi yang objektif, terukur, dan komprehensif. Penilaian kinerja sering kali dilakukan secara administratif, terbatas pada jumlah publikasi, laporan kegiatan, atau indikator yang bersifat kuantitatif semata, tanpa mempertimbangkan kualitas pengajaran, kontribusi penelitian, maupun dampak pengabdian kepada masyarakat. Hal ini mengakibatkan hasil penilaian kinerja dosen cenderung subjektif, tidak seragam antar institusi, serta berpotensi tidak mencerminkan kondisi sebenarnya.

Selain itu, data kinerja dosen pada umumnya sangat beragam, meliputi aspek pengajaran, penelitian, pengabdian, hingga pengembangan diri. Data tersebut bersifat kompleks, dalam jumlah besar, dan tersebar pada berbagai sumber sehingga sulit dianalisis secara manual. Kondisi ini menimbulkan kesulitan dalam menemukan pola atau faktor dominan yang memengaruhi kualitas kinerja dosen. Tanpa adanya teknik pengolahan data yang tepat, informasi yang dihasilkan dari penilaian kinerja dosen menjadi kurang bermakna dan sulit digunakan sebagai dasar kebijakan peningkatan mutu pendidikan. Oleh karena itu, dibutuhkan pendekatan berbasis data mining yang mampu mengolah data kinerja dosen secara sistematis untuk menghasilkan analisis yang objektif, transparan, dan dapat dijadikan landasan dalam pengambilan keputusan manajerial di lingkungan perguruan tinggi.

3.2 Penerapan Algoritma C4.5

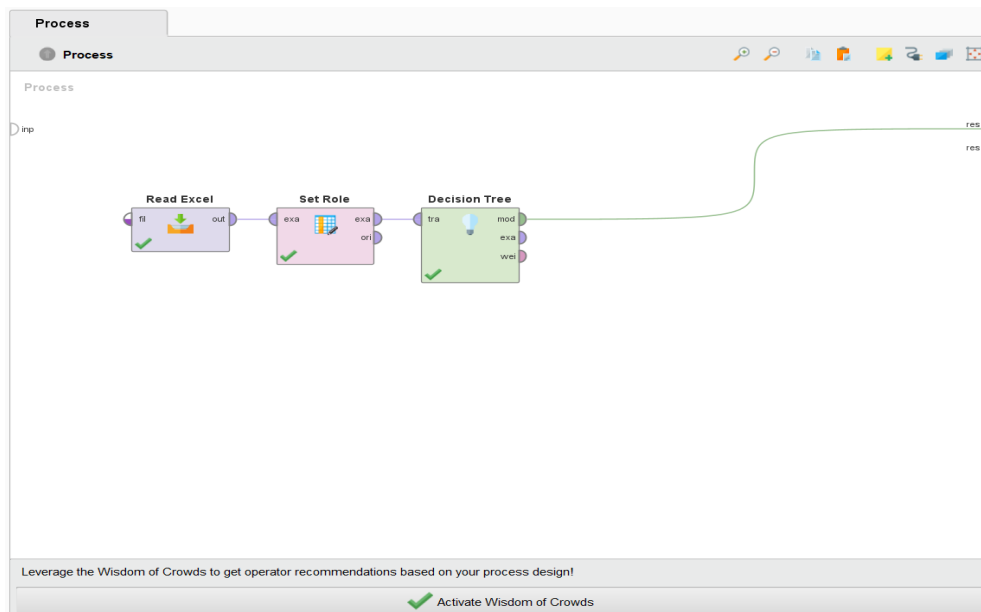
Algoritma C4.5 merupakan salah satu metode klasifikasi dalam data mining yang banyak digunakan untuk membangun pohon keputusan (decision tree). Pohon keputusan berfungsi untuk mengubah data yang kompleks menjadi model klasifikasi yang lebih sederhana, terstruktur, dan mudah dipahami. Proses klasifikasi dengan C4.5 diawali dengan pemilihan atribut yang paling relevan berdasarkan nilai gain ratio sebagai akar pohon, kemudian dilanjutkan dengan pembentukan cabang-cabang yang merepresentasikan kondisi dari setiap atribut. Proses ini dilakukan secara rekursif hingga semua data dapat terklasifikasi ke dalam kelas yang sesuai.

Dalam penerapannya, algoritma C4.5 memiliki keunggulan karena mampu mengolah data baik yang bersifat diskrit maupun kontinu, menangani atribut yang hilang, serta menghasilkan aturan klasifikasi yang dapat divisualisasikan dalam bentuk pohon keputusan. Hal ini menjadikan C4.5 sebagai salah satu algoritma yang efektif untuk menghasilkan model analisis yang jelas dan mudah dipahami. Pada penelitian ini, penerapan algoritma C4.5 digunakan untuk mengolah data kinerja dosen sehingga dapat diperoleh model klasifikasi yang menggambarkan kategori kinerja “baik”, “cukup baik”, dan “perlu ditingkatkan”. Dengan demikian, hasil yang diperoleh diharapkan lebih objektif, transparan, dan dapat digunakan sebagai dasar pertimbangan dalam evaluasi serta peningkatan mutu pendidikan tinggi. Sebelum dilakukan penelitian maka selanjutnya yang dilakukan adalah data penelitian, adapun data penelitian yang digunakan adalah sebagai berikut:

Tabel 1. Sampel Data Penelitian

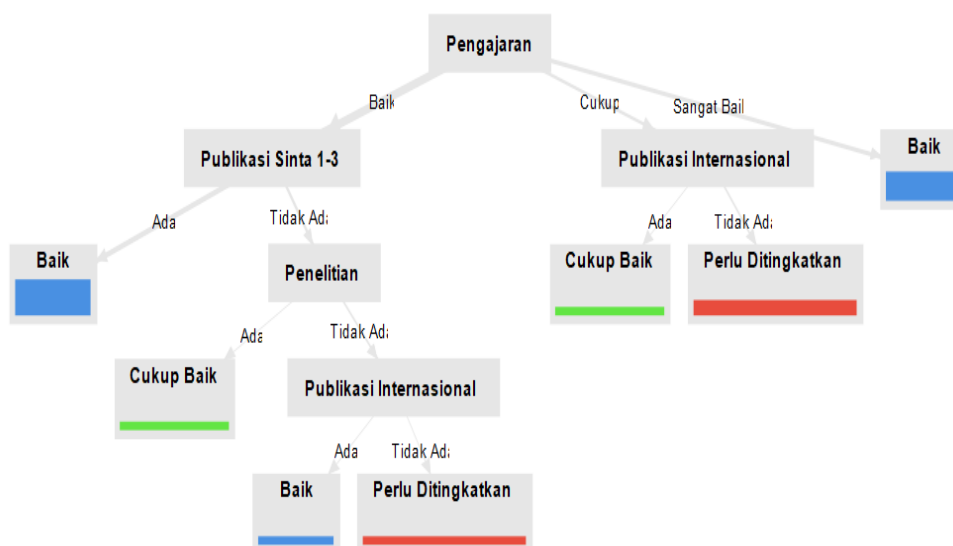
Dosen	Pengajaran	Penelitian	Publikasi Internasional	Publikasi Sinta 1-3	Publikasi Sinta 4-5	PkM	Pelatihan	Penilaian
Dosen 1	Sangat Baik	Tidak Ada	Ada	Ada	Ada	Tidak Ada	Ada	Baik
Dosen 2	Baik	Ada	Tidak Ada	Ada	Tidak Ada	Ada	Tidak Ada	Baik
Dosen 3	Baik	Tidak Ada	Ada	Tidak Ada	Ada	Ada	Tidak Ada	Baik
Dosen 4	Sangat Baik	Ada	Ada	Tidak Ada	Ada	Ada	Tidak Ada	Baik
Dosen 5	Cukup	Tidak Ada	Ada	Tidak Ada	Ada	Ada	Ada	Cukup Baik
Dosen 6	Baik	Tidak Ada	Ada	Ada	Ada	Ada	Tidak Ada	Baik
Dosen 7	Cukup	Tidak Ada	Tidak Ada	Tidak Ada	Ada	Tidak Ada	Ada	Perlu Ditingkatkan
Dosen 8	Sangat Baik	Ada	Ada	Tidak Ada	Ada	Ada	Tidak Ada	Baik
Dosen 9	Cukup	Ada	Tidak Ada	Ada	Tidak Ada	Ada	Tidak Ada	Perlu Ditingkatkan
Dosen 10	Baik	Ada	Tidak Ada	Ada	Ada	Ada	Ada	Baik
Dosen 11	Baik	Tidak Ada	Tidak Ada	Tidak Ada	Ada	Ada	Tidak Ada	Perlu Ditingkatkan
Dosen 12	Baik	Tidak Ada	Ada	Ada	Ada	Ada	Tidak Ada	Baik
Dosen 13	Baik	Ada	Tidak Ada	Tidak Ada	Ada	Tidak Ada	Ada	Cukup Baik
Dosen 14	Sangat Baik	Ada	Tidak Ada	Ada	Tidak Ada	Ada	Tidak Ada	Baik
Dosen 15	Baik	Tidak Ada	Tidak Ada	Ada	Ada	Tidak Ada	Ada	Baik

Pada Tabel 1. Sampel Data Penelitian menampilkan sebagian data yang digunakan sebagai objek penelitian dalam proses analisis kinerja dosen. Data tersebut berisi sejumlah atribut yang merepresentasikan indikator penilaian kinerja, seperti aktivitas pengajaran, jumlah publikasi ilmiah, keterlibatan dalam kegiatan pengabdian kepada masyarakat, serta hasil evaluasi mahasiswa terhadap proses pembelajaran. Setiap atribut dalam tabel memiliki peran penting karena akan menjadi dasar dalam proses klasifikasi menggunakan algoritma C4.5. Selanjutnya, dilakukan proses penyelesaian penelitian. Pada penelitian dilakukan penyelesaian dengan melakukan pengujian data pada Algoritma C4.5 dengan menggunakan Tools pengujian Rapidminer. Adapun gambaran pengujian dapat dilihat berikut:



Gambar 2. Operator Pengujian C4.5 Dengan Rapidminer

Pada Gambar 2. Operator Pengujian C4.5 dengan RapidMiner menunjukkan rancangan proses yang digunakan dalam penelitian untuk menguji algoritma C4.5. Pada gambar tersebut terlihat bahwa proses diawali dengan memasukkan dataset penelitian yang berisi atribut-atribut kinerja dosen, kemudian dilakukan pembagian data menjadi training set dan testing set untuk keperluan klasifikasi. Operator Decision Tree (C4.5) digunakan sebagai inti dari proses analisis, di mana sistem membangun model klasifikasi berdasarkan data training dan selanjutnya melakukan pengujian terhadap data testing untuk mengevaluasi tingkat akurasi model. Dengan adanya visualisasi alur kerja pada Gambar 2, dapat memahami bagaimana algoritma C4.5 diimplementasikan dalam perangkat lunak RapidMiner secara sistematis. Setelah dilakukan proses pengujian dapat dilihat untuk hasil pengujian. Pada algoritma C4.5 hasil yang didapatkan berupa sebuah pohon keputusan. Adapun hasil dari keputusan dapat dilihat berikut:



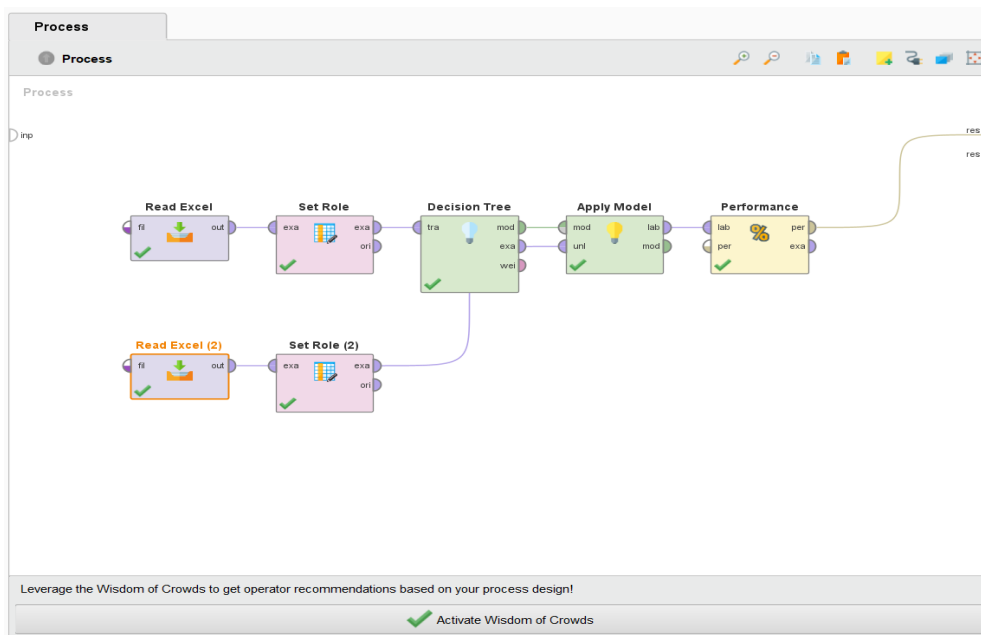
Gambar 3. Pohon Keputusan Hasil Pengujian Algoritma C4.5

Pada Gambar 3. Pohon Keputusan Hasil Pengujian Algoritma C4.5 memperlihatkan struktur model klasifikasi yang terbentuk setelah data penelitian diolah menggunakan algoritma C4.5. Pohon keputusan tersebut menggambarkan hubungan antar atribut kinerja dosen dan hasil klasifikasi yang diperoleh. Setiap node pada pohon merepresentasikan atribut yang digunakan sebagai dasar pemisahan data. Adapun dari hasil pohon keputusan dapat juga dilihat untuk aturan rule yang didapatkan sebagai berikut:

- Pengajaran = Baik
 - | Publikasi Sinta 1-3 = Ada: Baik {Baik=5, Cukup Baik=0, Perlu Ditingkatkan=0}
 - | Publikasi Sinta 1-3 = Tidak Ada
 - | | Penelitian = Ada: Cukup Baik {Baik=0, Cukup Baik=1, Perlu Ditingkatkan=0}
 - | | Penelitian = Tidak Ada
 - | | | Publikasi Internasional = Ada: Baik
 - | | | Publikasi Internasional = Tidak Ada: Perlu Ditingkatkan
- Pengajaran = Cukup
 - | Publikasi Internasional = Ada: Cukup Baik
 - | Publikasi Internasional = Tidak Ada: Perlu Ditingkatkan
- Pengajaran = Sangat Baik
 - | Baik

- | | | Publikasi Internasional = Ada: Baik {Baik=1, Cukup Baik=0, Perlu Ditingkatkan=0}
- | | | Publikasi Internasional = Tidak Ada: Perlu Ditingkatkan {Baik=0, Cukup Baik=0, Perlu Ditingkatkan=1}
- Pengajaran = Cukup
- | Publikasi Internasional = Ada: Cukup Baik {Baik=0, Cukup Baik=1, Perlu Ditingkatkan=0}
- | Publikasi Internasional = Tidak Ada: Perlu Ditingkatkan {Baik=0, Cukup Baik=0, Perlu Ditingkatkan=2}
- Pengajaran = Sangat Baik: Baik {Baik=4, Cukup Baik=0, Perlu Ditingkatkan=0}

Aturan rule yang terbentuk dari proses pengujian algoritma c4.5 dipergunakan sebagai dasar dalam proses pengambilan keputusan. Untuk melihat terhadap kinerja dari algoritma c4.5 maka selanjutnya dilakukan pengukuran kinerja akurasi dari algoritma c4.5 dengan menggunakan operator berikut:



Gambar 4. Operator Pengukuran Kinerja Algoritma C4.5

Pada gambar 4 dapat dilihat terhadap operator pengukuran kinerja dari algoritma c4.5, dari proses tersebut didapatkan untuk hasil kinerja algoritma C4.5. Hasil kinerja algoritma C4.5 yang didapatkan dapat dilihat berikut:

Tabel 2. Hasil Kinerja Algoritma C4.5

	True Cukup Baik	True Baik	True Perlu Ditingkatkan	Class Precision
Pred Cukup Baik	4	1	0	80%
Pred Baik	0	6	0	100%
Pred Perlu Ditingkatkan	0	1	3	75%
	100%	75%	100%	86,67%

Pada tabel 2 dapat dilihat hasil kinerja terhadap algoritma C4.5. Dari hasil tersebut dapat dilihat bahwasannya hasil kinerja algoritma C.45 sebesar 86,67%.

4. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa penerapan data mining dengan algoritma C4.5 mampu memberikan pendekatan yang lebih objektif dan terukur dalam mengevaluasi kinerja dosen. Melalui proses klasifikasi, data kinerja yang kompleks dapat diolah menjadi pohon keputusan yang mudah dipahami serta mampu menunjukkan faktor-faktor dominan yang memengaruhi kualitas kinerja dosen. Hal ini membantu mengurangi subjektivitas dalam penilaian serta memberikan dasar yang lebih kuat dalam perumusan strategi peningkatan mutu pendidikan. Hasil pengujian yang dilakukan dengan menggunakan perangkat lunak RapidMiner menunjukkan bahwa model klasifikasi yang dibangun memiliki tingkat akurasi sebesar **86,67%**. Nilai ini mengindikasikan bahwa algoritma C4.5 cukup efektif dalam mengklasifikasikan data kinerja dosen ke dalam kategori yang telah ditentukan. Dengan akurasi yang tinggi, hasil analisis dapat dijadikan rujukan dalam mengevaluasi performa dosen secara lebih transparan dan proporsional. Secara keseluruhan, penelitian ini menunjukkan bahwa data mining, khususnya dengan algoritma C4.5, berpotensi menjadi solusi strategis dalam mendukung sistem evaluasi kinerja dosen di perguruan tinggi. Sebagai tindak lanjut, penelitian selanjutnya disarankan untuk menggunakan dataset yang lebih besar dan bervariasi agar hasil analisis

lebih representatif. Selain itu, dapat dilakukan perbandingan dengan algoritma lain seperti Random Forest, Naïve Bayes, atau Support Vector Machine untuk menemukan metode terbaik dalam evaluasi kinerja dosen. Penelitian mendatang juga sebaiknya menambahkan indikator penilaian yang lebih komprehensif, termasuk kualitas pengajaran dan inovasi pembelajaran, serta mengembangkan sistem evaluasi berbasis aplikasi agar hasil analisis dapat dimanfaatkan secara langsung oleh manajemen perguruan tinggi. Dengan demikian, model evaluasi kinerja dosen yang dihasilkan dapat semakin akurat, praktis, dan bermanfaat bagi peningkatan mutu pendidikan tinggi.

REFERENCES

- [1] L. S. Nasution, N. S. Wahyuni, Julkarnain, and Syafrizal, "Pengaruh Kepuasan Kerja dan Motivasi Kerja Terhadap Kinerja Dosen di Fakultas Ekonomi dan Bisnis Universitas Islam Sumatera Utara," *J. Ilm. Manaj. dan Bisnis*, vol. 6, no. 1, pp. 125–129, 2025, doi: 10.31289/jimbi.v6i1.6151.
- [2] N. S. Wahyuni, R. A. Pohan, and Supriadi, "Pengaruh Lingkungan Kerja dan Disiplin Kerja Terhadap Kinerja Dosen Di Fakultas Ekonomi dan Bisnis Universitas Islam Sumatera Utara," *J. Ilm. Manaj. dan Bisnis*, vol. 6, no. 1, pp. 118–124, 2025, doi: 10.31289/jimbi.v6i1.6150.
- [3] S. Sufina and L. Wati, "Implementation of Random Forest Algorithm in Lecturer Performance Evaluation System," *J. Artif. Intell. Softw. Eng.*, vol. 5, no. 2, pp. 790–798, 2025, doi: 10.30811/jaise.v5i2.6975.
- [4] S. Fatmawati, M. R. A. Prasetya, S. E. Prastyta, and S. P. Cipta, "Penerapan Naïve Bayes dalam Analisis Sentimen untuk Evaluasi Kinerja Pengajaran Dosen," *J. SAINTEKOM*, vol. 15, no. 1, pp. 82–93, 2025, doi: 10.33020/saintekom.v15i1.765.
- [5] D. P. Indini, Mesran, and Dito Putro Utomo, "Penerapan Data Mining Dalam Pengelompokan Data Reseller di Telkomsel Authorized Partner (TAP) Deli Tua Dengan Algoritma K-Means," *J. Ilm. Media Sisfo*, vol. 17, no. 2, pp. 189–202, 2023, doi: 10.33998/mediasisfo.2023.17.2.1391.
- [6] M. Mesran, M. Syahrizal, S. Sarwandi, S. Aripin, D. P. Utomo, and A. Karim, "A comparison of the performance of data mining classification algorithms on medical datasets with the application of data normalization," *AIP Conf. Proc.*, vol. 3048, no. 1, 2024, doi: 10.1063/5.0207994.
- [7] E. Bu'ulolo, "Strategi Inisialisasi Centroid K-Means Berbasis Mean Global untuk Peningkatan Akurasi Clustering Data," *JIKTEKS J. Ilmu Komput. dan Teknol. Inf.*, vol. 3, no. 02, pp. 69–75, 2025, doi: 10.70404/jikteks.v3i02.177.
- [8] K. D. A. Sastrawan, D. Suranti, and I. Y. Beti, "Implementation Of Data Mining Using Algorithms A Priori in Determining The Pattern Of Product Purchases Perfume Sold Case Study On (Rafflesia Aromatic Professional Perfume) Implementasi Data Mining Menggunakan Algoritma A Priori Dalam Penentuan Pola Pembe," *J. Media Comput. Sci.*, vol. 1, no. 2, pp. 163–174, 2022.
- [9] J. Dongga, A. Sarungallo, N. Koru, and G. Lante, "Implementasi Data Mining Menggunakan Algoritma Apriori Dalam Menentukan Persediaan Barang (Studi Kasus: Toko Swapen Jaya Manokwari)," *G-Tech J. Teknol. Terap.*, vol. 7, no. 1, pp. 119–126, 2023, doi: 10.33379/gtech.v7i1.1938.
- [10] J. Teknik, I. C. I. T. Medicom, E. Bu, R. Syahputra, E. Sabrina, and A. Simorangkir, "Outlier detection in the clustired data," *J. Tek. Inform. C.I.T Medicom*, vol. 16, no. 6, pp. 394–404, 2025.
- [11] N. K. Juliani, N. M. S. Iswari, and N. W. Utami, "Penerapan Data Mining Untuk Menentukan Kelayakan Kendaraan Sepeda Motor Bekas Menggunakan Algoritma C4.5," *INFOTECH J.*, vol. 11, no. 2, pp. 1–5, 2025.
- [12] Z. Nur Saputra and Zaehol Fatah, "Pengunaan Data Minig Untuk Mengidentifikasi Pelanggan Beresiko Tinggi Dalam Penjualan Mengunakan Algoritma Decition Tree C4.5," *J. Ilm. Multidisiplin Ilmu*, vol. 2, no. 1, pp. 46–51, 2025, doi: 10.69714/s91z1k09.
- [13] A. R. Hakim, "Penerapan Algoritma C4.5 untuk Memprediksi Profit Perusahaan di PT JMS Batam," *J. Sains Inform. Terap.*, vol. 4, no. 2, pp. 16–20, 2025.
- [14] M. W. Martadiansyah, A. Ghufiron, R. A. Hidayah, D. Salzabila, and L. Amanda, "Perbandingan Algoritma C4.5 dengan Naive Bayes Untuk Klasifikasi Curah Hujan," *J. Komput. Antart.*, vol. 3, no. 1, pp. 8–17, 2025, doi: 10.70052/jka.v3i1.648.
- [15] P. E. Putra, M. A. Amrullah, Y. H. Fauzi, R. F. Saputri, and L. C. R. Fatma, "Penerapan Algoritma C4.5 untuk Klasifikasi Tingkat Korban Banjir di Indonesia," *J. Komput. Antart.*, vol. 3, no. 1, pp. 1–7, 2025, doi: 10.70052/jka.v3i1.647.
- [16] J. Bramanda, "Klasifikasi Masyarakat Penerima Bantuan Sosial dari Pemerintah dengan Metode Algoritma C4.5," *J. Komput. Antart.*, vol. 3, no. 1, pp. 34–41, 2025, doi: 10.70052/jka.v3i1.234.
- [17] N. M. Wasilah, "Penerapan Data Mining Untuk Menganalisis Kepuasan Konsumen Menggunakan Metode C4.5," *Intern. (Information Syst. Journal)*, vol. 8, no. 1, pp. 85–90, 2025.
- [18] D. A. Octaviani, Dhias Widiastuti, Rizky Amelia, and A. Salam, "Implementasi Data Mining Untuk Memprediksi Pre-Eklampsia Dalam Kehamilan Menggunakan Algoritma C4.5," *J. Kesehatan*, vol. 18, no. 1, pp. 29–38, 2025, doi: 10.32763/ps5qjp75.
- [19] S. Dewi, A. Kresnawati, and S. H. Salsabil, "Implementasi Data Mining Untuk Menentukan Strategi Promosi DiSmk AI-Amah Cimanggung Menggunakan Algoritma C4.5," *IJRSE Indones. J. Inform. Res. Softw. Eng.*, vol. 5, no. 1, pp. 20–28, 2025.
- [20] N. S. Sitorusa and D. Leman, "Implementasi Data Mining Dalam Klasifikasi Produk Laris Dan Tidak Laris Dengan Algoritma C4.5," *J. Rekayasa Sist.*, vol. 3, no. 2, pp. 423–426, 2025.
- [21] O. Pahlevi, H. Rianto, and C. Author, "Analisa Komparasi Model Data Mining Algoritma C4.5, CHAID, dan Random Forest Untuk Penilaian Kelayakan Kredit," *Comput. Sci. (CO-SCIENCE)*, vol. 5, no. 1, pp. 49–57, 2025, [Online]. Available: <http://jurnal.bsi.ac.id/index.php/co-science>
- [22] J. A. Noyari, A. Aprillia, R. G. Munthe, and A. Sutarman, "Optimasi Kinerja Sistem Informasi Manajemen Kampus Menggunakan Teknik Data Mining," *J. MENTARI Manajemen, Pendidik. dan Teknol. Inf.*, vol. 3, no. 1, pp. 52–63, 2024.
- [23] C. S. Octiva, T. I. Fajri, E. B. Sulistiarini, S. Suharjo, and U. W. Nuryanto, "Penggunaan Teknik Data Mining untuk Analisis Perilaku Pengguna pada Media Sosial," *J. Minfo Polgan*, vol. 13, no. 1, pp. 1074–1078, 2024, doi: 10.33395/jmp.v13i1.13936.
- [24] E. Lutfina et al., "Identifikasi Teknik Data Mining Metode Asosiasi: Systematic Literature Review," *Sci. Technol. Manag. J.*, vol. 4, no. 2, pp. 57–61, 2024, doi: 10.53416/stmj.v4i2.257.
- [25] M. F. Haryanti et al., "Pengaruh Data Mining, Strategi Metode Perusahaan, Terhadap Laporan Kinerja Perusahaan," *J. Portofolio J. Manaj. dan Bisnis*, vol. 3, no. 1, pp. 71–90, 2024, doi: 10.70704/jpjmb.v3i1.285.
- [26] A. Ernawati and S. Wahyuni, "Analisis Data Mining Pola Penggunaan Seluler dan Klasifikasi Perilaku Pengguna di Berbagai

- Perangkat Menggunakan Metode C4 . 5,” Bull. Inf. Technol., vol. 5, no. 4, pp. 162–168, 2024, doi: 10.47065/bit.v5i2.1689.
- [27] I. J. N. Ayu, N. R. Putri, R. P. Nugraha, and R. Febriansyah, “Klasifikasi Kualitas Udara di Jakarta Pada Bulan Agustus 2024 Menggunakan Algoritma C4.5,” JDMIS J. Data Min. Inf. Syst., vol. 3, no. 1, pp. 1–8, 2025, doi: 10.54259/jdmis.v3i1.3745.
- [28] D. Mayangsari and T. Wahyudi, “Klasifikasi Seleksi Penerimaan Karyawan Baru Menggunakan Algoritma C4.5,” MALCOM Indones. J. Mach. Learn. Comput. Sci., vol. 5, no. 1, pp. 26–34, 2024, doi: 10.57152/malcom.v5i1.1525.
- [29] A. Z. D. Nur Adiya, A. F. Desvita, A. Fidela, D. Amelia, and T. Astuti, “Penerapan Data Mining Untuk Klasifikasi Kualitas Udara di Daerah Istimewa Yogyakarta Menggunakan Algoritma C4.5,” JDMIS J. Data Min. Inf. Syst., vol. 2, no. 2, pp. 59–65, 2024, doi: 10.54259/jdmis.v2i2.2800.
- [30] J. R. Daulai and E. Panggabean, “Penerapan Data Mining Untuk Memprediksi Penjualan Aquarium,” J. Data Sci. Penusa, vol. 1, no. 1, pp. 1–12, 2024.