Jurnal ilmiah Sistem Informasi dan Ilmu Komputer
Volume 5, Nomor 1, Maret 2025
e-ISSN: 2827-7953; p-ISSN: 2827-8135, Hal. 322-331
DOI:https://doi.org/10.55606/juisik.v5i1.1297
Available Online at: http://journal.sinov.id/index.php/juisik

Analisis Klasifikasi Pelanggaran Santri di Pesantren Menggunakan Algoritma Decision Tree

Abdullah[1*], April Lia Hananto[2], Shofa Shofia Hilabi [3] , Bayu Priyatna [4]
[1-4] Universitas Buana Perjuangan Karawang, Karawang, Indonesia
Email: si21.abdullah@mhs.ubpkarawang.ac.id[1], aprilia@ubpkarawang.ac.id[2], shofa.hilabi@ubpkarawang.ac.id[3], bayu.priyatna@ubpkarawang.ac.id[4]

Jalan Ronggo Waluyo Sirnabaya, Puseurjaya, Telukjambe Timur, Karawang, Jawa Barat 41361
Korespondensi penulis: si21.abdullah@mhs.ubpkarawang.ac.id

Abstract*. Islamic boarding schools (pesantren) are Islamic educational institutions that emphasize discipline in the daily lives of their students. However, the manual recording and evaluation of violations makes it difficult for administrators to conduct systematic analysis. This study aims to apply data mining methods, specifically the Decision Tree algorithm, to analyze student violation data for effective classification. Violation data is grouped into three categories: minor, moderate, and severe violations. A classification model is built through entropy and information gain calculations to determine the best attributes to serve as the initial nodes in the decision tree structure. The attributes used in the model include violation points, violation types, and dormitories. The results show that the Decision Tree algorithm is able to identify violation categories with excellent performance. In the minor category, the model achieved a precision of 0.99, a recall of 1.00, and an f1-score of 1.00. For the moderate category, the model achieved a precision of 0.94, a recall of 1.00, and an f1-score of 0.97, indicating that the model is able to recognize violations with a high level of accuracy and consistency. Meanwhile, in the severe category, the model demonstrated perfect precision of 1.00, recall of 0.87, and f1-score of 0.93. Overall, the model achieved 98% accuracy based on the confusion matrix evaluation, indicating that most of the data was correctly classified. The decision tree visualization also showed that the violation point, type of violation, and dormitory location were the main factors in the classification process. These results demonstrate that a data mining approach can be used to support data-driven decision-making in student guidance.*

Keywords*: Data Mining, Decision Tree, Islamic Boarding Schools, Python, Student Violations*

Abstrak. Pesantren merupakan lembaga pendidikan Islam yang menekankan pada kedisiplinan dalam kehidupan sehari-hari santri. Namun, proses pencatatan dan evaluasi pelanggaran yang masih dilakukan secara manual menyulitkan pihak pengelola dalam melakukan analisis secara sistematis. Penelitian ini bertujuan untuk menerapkan metode data mining, khususnya algoritma *Decision Tree*, dalam menganalisis data pelanggaran santri agar dapat dilakukan klasifikasi secara efektif. Data pelanggaran dikelompokkan menjadi tiga kategori, yaitu pelanggaran ringan, sedang, dan berat. Model klasifikasi dibangun melalui proses perhitungan *entropy* dan *information gain* untuk menentukan atribut terbaik yang akan dijadikan sebagai node awal dalam struktur pohon keputusan. Atribut yang digunakan dalam model antara lain poin pelanggaran, jenis pelanggaran, dan asrama. Hasil penelitian menunjukkan bahwa algoritma *Decision Tree* mampu mengidentifikasi kategori pelanggaran dengan performa yang sangat baik. Pada kategori ringan, model mencapai *precision* sebesar 0.99, *recall* 1.00, dan *f1-score* 1.00. Untuk kategori sedang, model memperoleh *precision* 0.94, *recall* 1.00, dan *f1-score* 0.97, yang menunjukkan bahwa model mampu mengenali pelanggaran dengan tingkat ketepatan dan konsistensi tinggi. Sedangkan pada kategori berat, model menunjukkan *precision* sempurna sebesar 1.00, *recall* 0.87, dan *f1-score* 0.93. Secara keseluruhan, akurasi model mencapai 98% berdasarkan hasil evaluasi *confusion matrix*, yang menunjukkan bahwa sebagian besar data terklasifikasi dengan benar. Visualisasi pohon keputusan juga memperlihatkan bahwa poin pelanggaran, jenis pelanggaran, dan lokasi asrama merupakan faktor utama dalam proses klasifikasi. Hasil ini membuktikan bahwa pendekatan data mining dapat digunakan untuk mendukung pengambilan keputusan dalam pembinaan santri secara berbasis data.

Kata Kunci: Data Minning, Decision Tree, Pelanggaran Siswa, Pesantren, Python

1. LATAR BELAKANG

Pesantren merupakan lembaga pendidikan Islam yang tidak hanya menitikberatkan pada aspek akademik, tetapi juga pada pembentukan karakter dan moral para santri. Meski demikian, pelanggaran terhadap tata tertib masih sering terjadi di lingkungan pesantren.

Received: Februari 15, 2025; Revised: Februari 30, 2025; Accepted: Maret 28, 2025; Published: Maret 30, 2025

Pelanggaran ini menjadi tantangan tersendiri bagi pengelola dalam memastikan terlaksananya proses pendidikan karakter secara menyeluruh(Sultan et al., 2025). Lebih lanjut, data pelanggaran yang ada sering kali hanya dijadikan arsip tanpa diolah pengelolaan data pelanggaran yang tidak dianalisis dapat menyebabkan terlewatnya peluang perbaikan sistem secara sistemik dan berkelanjutan(Nisa & Suharno, 2020). padahal analisis terhadap data tersebut sangat penting sebagai dasar evaluasi dan perbaikan(Putra Dhian Wahana, 2020).

Salah satu pendekatan yang dapat dimanfaatkan untuk menganalisis data pelanggaran adalah data mining. Teknologi ini memungkinkan penggalian informasi tersembunyi dari data dalam jumlah besar, termasuk dalam konteks pendidikan pesantren(Nurhadi & Ulandari, 2022). Dengan menerapkan algoritma seperti Decision Tree atau Naive Bayes, pola pelanggaran santri dapat diidentifikasi berdasarkan faktor-faktor seperti jenis pelanggaran, waktu kejadian, dan karakteristik individu santri(Romlah & Fatah, 2024). Hasil dari analisis tersebut dapat digunakan untuk merancang kebijakan yang lebih tepat sasaran, seperti pemberian pelatihan khusus bagi santri dengan kecenderungan melakukan pelanggaran atau peningkatan sistem pengawasan yang lebih efektif(Woro Mahatmi & Mediatrix Sebatubun, 2021).

Berdasarkan kondisi tersebut, penelitian ini dirancang untuk menjawab dua pertanyaan utama: bagaimana menganalisis data dan mengetahui jenis-jenis kategori pelanggaran pada santri, serta bagaimana penerapan algoritma Decision Tree dapat digunakan untuk mengklasifikasikan pelanggaran dalam rangka meningkatkan kedisiplinan dan pembinaan karakter santri(Ramadhan et al., 2025). Kedua rumusan masalah ini menjadi dasar pengembangan penelitian agar hasilnya dapat memberikan kontribusi nyata dalam pengelolaan tata tertib pesantren(Firdaus et al., 2022).

Tujuan dari penelitian ini adalah untuk menganalisis dan mengidentifikasi jenis-jenis kategori pelanggaran yang terjadi pada santri serta menerapkan algoritma Decision Tree guna mengklasifikasikan pelanggaran-pelanggaran tersebut secara sistematis. Dengan tujuan tersebut, diharapkan proses pembinaan santri dapat dilakukan secara lebih objektif, berbasis data, dan tidak semata-mata bergantung pada pendekatan subjektif atau tradisional(Sartika & Jumadi, 2023).

Manfaat dari penelitian ini dibagi menjadi dua aspek. Pertama, secara praktis, penelitian ini memberikan gambaran tentang pola pelanggaran yang terjadi, sehingga dapat dijadikan bahan evaluasi oleh pengelola pesantren dalam menyusun kebijakan disiplin yang lebih tepat dan relevan. Kedua, penelitian ini mendemonstrasikan pemanfaatan teknologi data mining khususnya algoritma decision tree dalam lingkungan pendidikan berbasis pesantren. Hal ini

diharapkan dapat membantu pengelolaan data pelanggaran secara sistematis dan mendukung pembinaan karakter santri secara lebih efektif, terarah, dan berbasis pada analisis yang dapat dipertanggungjawabkan(Santoso & Zubairi, 2022).

2. KAJIAN TEORITIS

Data minning adalah proses eksplorasi dan analisis data berskala besar untuk menemukan pola, tren, dan hubungan tersembunyi yang bermanfaat dalam pengambilan keputusan (Lia Hananto et al., n.d.) atau Data mining adalah analisis data besar untuk menemukan pola, tren, dan hubungan guna mendukung keputusan(Novia Cahya Utami et al., 2023). Penggunaan data mining dalam dunia pendidikan telah terbukti mampu membantu institusi dalam pengambilan keputusan berbasis data historis dan prediktif(Kurniawan et al., 2023). Dalam konteks pesantren, teknik ini sangat relevan digunakan untuk mengolah data pelanggaran santri yang selama ini hanya tersimpan sebagai arsip, tanpa dimanfaatkan secara maksimal untuk mencegah pelanggaran serupa di masa depan.

Python adalah bahasa pemrograman tingkat tinggi yang mudah dipahami dan banyak digunakan dalam pengolahan data dan pengembangan aplikasi berbasis kecerdasan buatan(Riziq sirfatullah Alfarizi et al., 2023). Python merupakan bahasa pemrograman yang bersifat open-source, interpretatif, dan memiliki sintaks yang sederhana(Pannadhitthana Candra, 2025). Dengan python data pelanggaran santri dapat divisualisasikan, dianalisis, dan diproses menggunakan berbagai algoritma tanpa perlu kemampuan pemrograman tingkat lanjut. Hal ini memungkinkan pengelola pesantren maupun peneliti untuk dengan cepat mengidentifikasi pola pelanggaran.

Decision Tree digunakan dalam penelitian ini sebagai algoritma utama untuk melakukan klasifikasi terhadap pelanggaran santri (Yatimah Mita Nurul, 2021). Algoritma ini memetakan data ke dalam struktur pohon keputusan(Charles zai, 2022). Pada kasus kali ini data dibagi menjadi beberapa bagian seperti jenis pelanggaran, asrama, dan waktu kejadian. Hasilnya berupa model yang mudah dipahami dan dapat menjadi dasar pertimbangan dalam menyusun strategi pembinaan karakter santri secara lebih efektif.

Dataset pelanggaran yang dianalisis mencakup 1.000 data pelanggaran santri, termasuk nama, kelas, jenis pelanggaran, asrama, dan poin sanksi. Data ini dikumpulkan dari lingkungan pesantren dan menjadi sumber utama dalam proses klasifikasi(Latifah & Ridla, 2024). Pengelolaan data yang sistematis menjadi bagian penting dalam manajemen data, yaitu serangkaian proses pengumpulan, penyimpanan, pengelolaan, dan pemanfaatan data secara efisien guna memastikan kualitas dan keamanan informasi(Shofiah Hilabi & Fauzi, 2023).

Dengan mengelola dataset ini secara sistematis, pihak pesantren dapat memperoleh gambaran menyeluruh mengenai jenis pelanggaran yang paling sering terjadi dan kelompok santri yang paling rentan melanggarnya.

Pelanggaran santri seperti ketidakhadiran, keterlambatan, atau pelanggaran etika lainnya menjadi indikator penting dalam pengelolaan disiplin di pesantren (Mulyaningsih et al., 2021). Pelanggaran siswa/santri mencakup berbagai tindakan yang tidak sesuai dengan peraturan yang telah ditetapkan oleh lembaga pendidikan(Mushafi et al., 2021). Melalui pendekatan berbasis data dan penerapan Decision Tree, penelitian ini bertujuan untuk mendukung pengambilan keputusan yang lebih objektif dan berbasis bukti. Dengan demikian, upaya pembinaan santri dapat berjalan lebih terarah dan berdampak nyata terhadap peningkatan karakter dan kedisiplinan.

3. METODE PENELITIAN

Penelitian ini diawali dengan identifikasi masalah, yaitu suatu tahapan penting dalam metode ilmiah yang bertujuan untuk merumuskan permasalahan secara jelas agar dapat ditentukan pendekatan penyelesaiannya(Lidysari et al., 2022). Dalam konteks ini, permasalahan yang diangkat adalah tingginya angka pelanggaran disiplin di lingkungan pesantren. Selanjutnya, dilakukan pengumpulan data**,** yaitu proses sistematis untuk memperoleh informasi yang relevan dari dokumentasi pelanggaran santri, seperti buku catatan pelanggaran, laporan pembina, dan data administratif lainnya(Resa Ayu Aisyah et al., 2022)

Setelah data terkumpul, dilakukan preprocessing data**,** yakni tahap persiapan data mentah menjadi format yang bersih dan dapat digunakan untuk analisis. Preprocessing mencakup pembersihan data (data cleaning), normalisasi, transformasi atribut, dan penanganan data hilang (Mardiani et al., 2023). Tahapan ini sangat krusial untuk memastikan kualitas data sehingga hasil model tidak bias atau menyesatkan.

Kemudian, penerapan algoritma Decision Tree dilakukan dengan memanfaatkan python. Decision Tree digunakan karena mampu memetakan keputusan secara hierarkis berdasarkan atribut yang paling signifikan dalam memengaruhi terjadinya pelanggaran (Jurnal et al., 2022). Model yang dibangun kemudian divalidasi menggunakan teknik cross-validation**,** yaitu metode evaluasi yang membagi data menjadi beberapa subset untuk menguji konsistensi dan akurasi model secara menyeluruh (Rifa et al., 2020)

4. HASIL DAN PEMBAHASAN

Pengumpulan data

Tahap awal dalam penelitian ini adalah pengumpulan data yang diperoleh dari dokumentasi internal pesantren, khususnya dari bagian kedisiplinan. Data yang dikumpulkan mencakup berbagai atribut penting, seperti jenis pelanggaran, kategori sanksi (ringan, sedang, berat), jenis kelamin, usia, kelas, dan identitas asrama. Total data yang berhasil dikumpulkan berjumlah 1011 entri pelanggaran santri. Data ini berasal dari catatan manual yang kemudian diolah ke dalam bentuk digital untuk memudahkan analisis menggunakan algoritma data mining. Validitas data dijaga dengan memverifikasi catatan melalui konfirmasi dengan staf pengasuhan serta pengelola data pelanggaran di pesantren.

Tujuan dari pengumpulan data ini adalah untuk menyusun dataset yang dapat digunakan dalam proses klasifikasi jenis pelanggaran menggunakan algoritma Decision Tree. Dengan atribut yang cukup beragam, diharapkan model mampu mengenali pola-pola yang tersembunyi dalam data pelanggaran santri serta membantu pihak pesantren memahami kecenderungan perilaku santri berdasarkan kategori tertentu.

Preprocesing data

Setelah data terkumpul, tahap selanjutnya adalah preprocessing atau praproses data. Tahap ini meliputi beberapa langkah penting, yaitu *data cleaning* dan *data integration*. Pada proses *data cleaning*, dilakukan penghapusan entri duplikat dan penanganan data kosong, khususnya pada kolom “Jenis Pelanggaran” dan “Kategori Sanksi”. Entri yang tidak lengkap atau mengandung kesalahan penulisan seperti “Asrama A” yang ditulis dengan variasi ejaan diperbaiki agar seragam.

Selain itu, *data integration* dilakukan dengan menggabungkan data dari berbagai sumber: data pelanggaran dari pengasuhan, asrama, dan bagian pendidikan. Penggabungan ini bertujuan untuk mendapatkan informasi yang utuh tentang profil santri yang melakukan pelanggaran. Data kemudian diolah menggunakan bahasa pemrograman Python melalui aplikasi Visual Studio Code untuk proses selanjutnya. Tahapan ini sangat krusial karena kualitas data yang baik akan sangat memengaruhi keakuratan model klasifikasi yang dibangun.

Penerapan algoritma Decision tree

Penerapan algoritma Decision Tree dimulai dengan pemilihan atribut terbaik berdasarkan perhitungan nilai entropy dan information gain. Nilai entropy awal dari pelanggaran dataset adalah 1.3022, dan yang mencerminkan tingkat penipuan atau keragaman data sebelum dilakukan verifikasi. Tetapkan kontribusi pada analisis tanpa memberikan informasi apa pun. Atribut dengan gain tertinggi dipilih sebagai akar (root node) dari pohon

keputusan. Misalnya, laporan "point" menjadi pembagi pertama karena memiliki nilai terbesar, diikuti oleh atribut lain seperti jenis kalamin, asrama, dan jenis pelanggaran. Proses sepintas ini dilakukan secara rekursif hingga data mencapai kondisi homogen atau tidak bisa dipisahkan lebih lanjut.

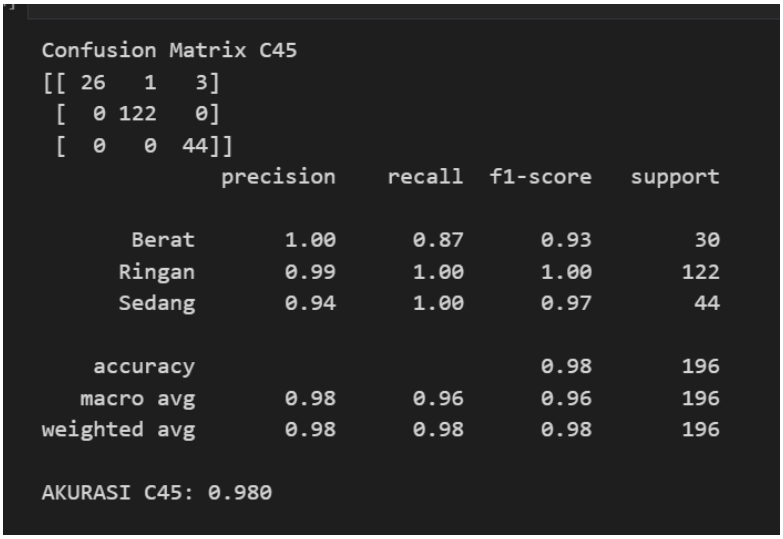
Model ini dibangun dan divisualisasikan menggunakan bahasa pemrograman Python, dengan struktur pohon biner yang menggambarkan proses pengambilan keputusan berdasarkan syarat-syarat pada setiap atribut. Menyiapkan node menggambarkan kondisi tertentu, sedangkan leaf node menunjukkan hasil klasifikasi akhir (kategori pelanggaran).

Evaluasi model

Evaluasi dilakukan dengan pendekatan konfusi matriks dan klasifikasi kinerja metrik lainnya, seperti presisi, recall, dan f1-score. Silakan bagikan yang berikut ini:

- Kategori Berat: 26 dari 30 diklasifikasi dengan benar, sisanya salah klasifikasi kelas dataran
- Kategori ringan: 122 dinar 122 diklasifikasi dengan benar (100%)
- Kategori Sedang: 44 hari 44 detik setiap tahun (100%)

Akurasi total model mencapai 98% (0.980), dengan nilai rata-rata presisi, recall, dan f1-score yang sangat tinggi (di atas 0.94 untuk semiua kelas). Hal ini menunjukkan bahwa model tidak hanya akurat secara keseluruhan, tetapi juga sangat presisi dalam mengidentifikasi masing-masing kategori pelanggaran.



```
Confusion Matrix C45
[[ 26   1   3]
 [  0 122   0]
 [  0   0  44]]
              precision    recall  f1-score   support

       Berat       1.00      0.87      0.93        30
      Ringan       0.99      1.00      1.00       122
      Sedang       0.94      1.00      0.97        44

    accuracy                           0.98       196
   macro avg       0.98      0.96      0.96       196
weighted avg       0.98      0.98      0.98       196

AKURASI C45: 0.980
```

Gambar 1. Hasil Test and Score

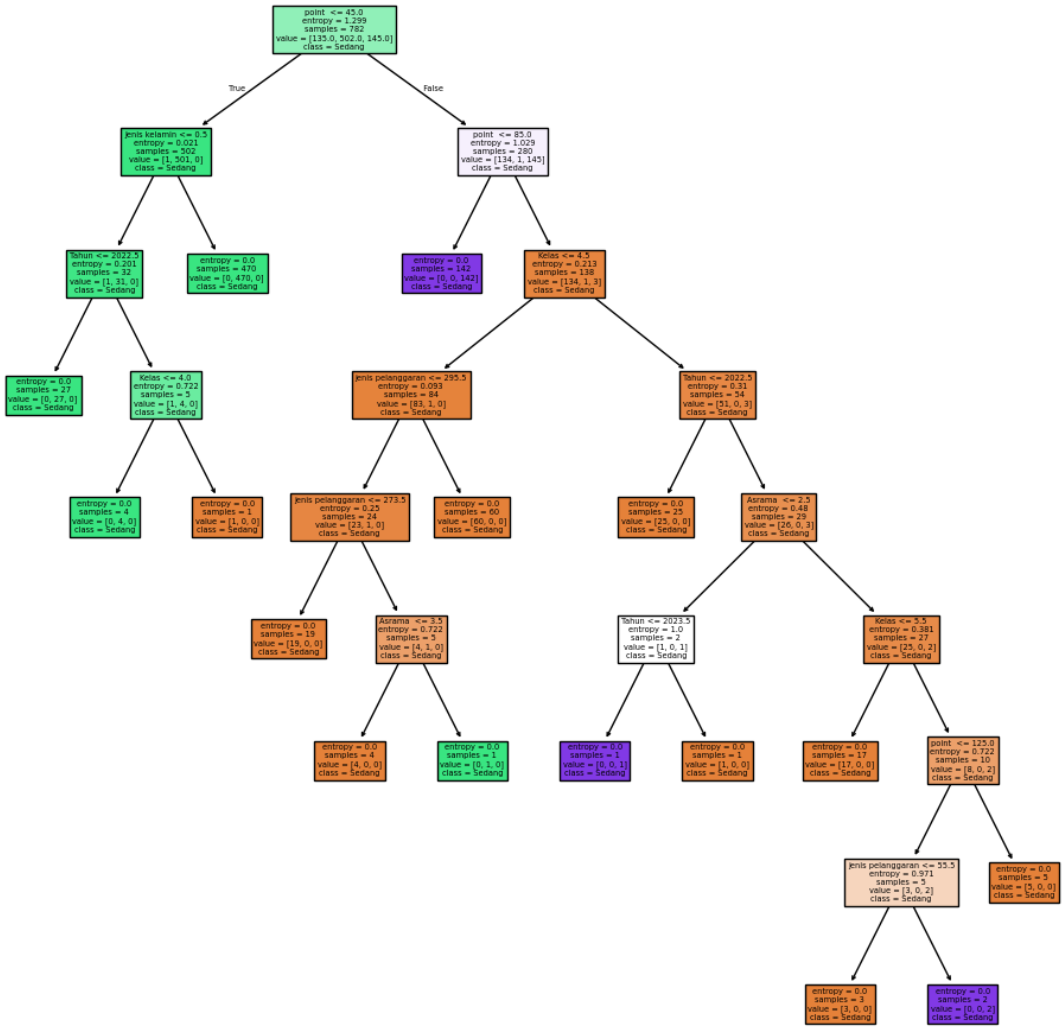
Selain itu, dilakukan pula evaluasi dengan menggunakan *confusion matrix*, yang memperlihatkan performa klasifikasi terhadap masing-masing kelas. Dari confusion matrix, terlihat bahwa pelanggaran kategori ringan sebanyak 642 dari 644 terklasifikasi dengan benar, sedangkan kategori berat dan sedang juga menunjukkan tingkat akurasi yang tinggi. Hal ini

menunjukkan bahwa model Decision Tree tidak hanya akurat secara keseluruhan, tetapi juga efektif dalam mengklasifikasikan setiap kategori pelanggaran secara spesifik.

Visualisasi hasil

Visualisasi pohon keputusan dilakukan menggunakan fungsi plot_tree dari pustaka matplotlib dalam Python. Ada banyak informasi yang dapat digunakan untuk masuk ke komputer dan sekaligus membagikannya dengan pengguna lain, di mana aman menggunakannya, dan cara membersihkannya.

Misalnya, pada cabang utama terlihat bahwa santri dengan nilai point ≤ 43.0 cenderung diklasifikasikan ke dalam kategori pelanggaran Sedang, sedangkan nilai point > 43.0 akan diproses lebih lanjut berdasarkan atribut lain seperti pelanggaran dan asrama. Struktur ini memudahkan pengguna non-teknis, seperti wali santri atau pengasuh pesantren, dalam memahami pola pelanggaran santri.



Gambar 2. Visualisasi Hasil

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian terhadap analisis pelanggaran santri di lingkungan pesantren diperoleh beberapa temuan penting sebagai berikut:

- Model klasifikasi yang dibangun memiliki tingkat efektif dan akurasi yang tinggi sebesar 98%, dengan nilai precision, recall, dan f1-score yang juga tinggi pada masing-masing kategori pelanggaran (berat, sedang, dan ringan)
- Atribut "point" merupakan variabel paling dominan dalam menentukan jenis pelanggaran, diikuti oleh atribut lain seperti jenis pelanggaran, jenis kelamin, dan asrama.
- Visualisasi pohon keputusan memberikan gambaran yang jelas dan mudah dipahami tentang alur logika pengambilan keputusan, sehingga dapat dimanfaatkan oleh pihak non-teknis.

Berdasarkan kesimpulan diatas penulis ingin memberikan saran :

- Pihak pesantren disarankan untuk memberikan perhatian lebih terhadap atribut-atribut yang terbukti memengaruhi pelanggaran, seperti nilai point dan jenis pelanggaran, khususnya pada kelompok santri yang menunjukkan kecenderungan tinggi terhadap pelanggaran berat
- Penelitian selanjutnya dapat dikembangkan dengan menambahkan variabel lain seperti riwayat pelanggaran sebelumnya, faktor psikologis, atau lingkungan keluarga untuk mendapatkan hasil analisis yang lebih komprehensif dan mendalam.

DAFTAR REFERENSI

Aisyah, R. A., Ifadah, D., Fitriani, D., & Aeni, A. N. (2022). Pengembangan video animasi motion graphic untuk mengatasi pudarnya kebudayaan Islam di kalangan mahasiswa. *Al-Fikri: Jurnal Studi dan Penelitian Pendidikan Islam*, 5(1), 1. https://doi.org/10.30659/jspi.v5i1.21132

Alfarizi, R. S., Al-Farish, Z., Taufiqurrahman, M., Ardiansah, G., & Elgar, M. (2023). Penggunaan Python sebagai bahasa pemrograman untuk machine learning dan deep learning. *Karimah Tauhid*, 2(1). https://doi.org/10.30997/karimahtauhid.v2i4.8816

Charles, Z. (2022). Implementasi data mining sebagai pengolahan data. *Portaldata.org*, 2(3).

Firdaus, M. A., Badriah, S., Arifin, B. S., & Hasanah, A. (2022). Pengembangan model pengelolaan pendidikan karakter di pesantren berbasis tradisi dan tadzkiroh. *Jurnal Basicedu*, 6(5), 8163–8174. https://doi.org/10.31004/basicedu.v6i5.3770

Hilabi, S. S., & Fauzi, A. (2023). Blockchain application on independent smart agriculture. *International Journal of Artificial Intelligence Research*, 7(1). http://ijair.id

Jurnal, H., Sandi, A. P., & Ningsih, V. W. (2022). Implementasi data mining sebagai penentu persediaan produk dengan algoritma FP-Growth pada data penjualan Sinarmart. *Jurnal Publikasi Ilmu Komputer dan Multimedia (JUPIKOM)*, 1(2). https://doi.org/10.55606/jupikom.v1i2.343

Kurniawan, M. D., Priyatna, B., & Nurapriani, F. (2023). Implementasi algoritma K-Means untuk klasterisasi data obat Puskesmas Kotabaru. *Jurnal Sains Komputer & Informatika (J-SAKTI)*, 7(2).

Latifah, T., & Ridla, M. A. (2024). Prediksi pelanggaran lalu lintas di Kabupaten Karangasem menggunakan algoritma neural network. *Jurnal Ahsana Penelitian dan Pengabdian Kepada Masyarakat*, 2(1), 21–26. https://doi.org/10.59395/ahsana.v2i1.360

Lia Hananto, A., Shofiah Hilabi, S., Hananto, A., Buana Perjuangan Karawang, U., Ronggo Waluyo, J., Timur, T., & Karawang, K. (n.d.). Strategi promosi penerapan data mining mahasiswa baru dengan metode K-Means clustering. *Dirgamaya Jurnal Manajemen dan Sistem Informasi*.

Lidysari, W., Tambunan, H. S., Qurniawan, H., & Tunas Bangsa, S. (2022). Penerapan data mining dalam menentukan kelayakan penerima bantuan sosial Pemko dengan algoritma C4.5 (Kasus Kantor Kelurahan Martoba). *Kesatria*, 3(1). https://doi.org/10.30645/kesatria.v3i1.97

Mahatmi, W. M., & Sebatubun, M. M. (2021). Strategi komunikasi stakeholders dengan pendekatan data alumni menggunakan data mining. *Public Relations Journal*, 2. https://doi.org/10.24246/precious.v2i1.5350

Mardiani, E., Rahmansyah, N., Kurniati, I., Matondang, N., Tesalonika, T., Zanitha, D. A., & Romzy, I. (2023). Membandingkan algoritma data mining dengan tools Orange untuk social economy. *Digital Transformation Technology*, 3(2), 686–693. https://doi.org/10.47709/digitech.v3i2.3256

Mulyaningsih, F., Islam, K., & Dakwah, F. (2021). Perilaku melanggar peraturan pada santri di Pondok Pesantren Darusslama Putri Utara Blokagung Banyuwangi.

Mushafi, Hidayatullah, S., & Aisyah, S. (2021). Pelanggaran siswa/santri mencakup berbagai tindakan yang tidak sesuai dengan peraturan yang telah ditetapkan oleh lembaga pendidikan.

Nisa, A. N., & Suharno, S. (2020). Penegakan hukum terhadap permasalahan lingkungan hidup untuk mewujudkan pembangunan berkelanjutan. *Jurnal Bina Mulia Hukum*, 4(2), 294. https://doi.org/10.23920/jbmh.v4i2.337

Nurhadi, A., & Ulandari, D. (2022). Manajemen pembinaan karakter disiplin santriwati di Pondok Pesantren Miftahul Qulub Polagan Pamekasan. *Evaluasi*, 6(2). https://doi.org/10.32478/evaluasi.v6i2.1030

Pannadhitthana Candra, A. (2025). Analisis data menggunakan Python: Memperkenalkan Pandas dan NumPy. *Jurnal Ilmu Sains dan Edukasi Digital (JISEd)*, 3(1), 11–16. https://doi.org/10.62386/jised.v3i1.118

Ramadhan, P., Yuhandri, & Veri, J. (2025). Eksplorasi algoritma decision tree untuk penentuan siswa berprestasi. *Bit-Tech*, 7(3), 826–833. https://doi.org/10.32877/bt.v7i3.2210

Rifa, H., Hamonangan, R., & Kurnia, D. A. (2020). Implementasi algoritma decision tree dalam klasifikasi kompetensi siswa. *KOPERTIP: Jurnal Ilmiah Manajemen Informatika dan Komputer*. http://jurnal.kopertipindonesia.or.id/

Romlah, S., & Fatah, Z. (2024). Penerapan data mining untuk memprediksi hafalan santri tahfidz menggunakan metode decision tree (Studi kasus MIK Sarina). *Global Journal of Management and Information (GJMI)*, 2(11), 160–166. https://doi.org/10.59435/gjmi.v2i11.1051

Santoso, F., & Zubairi, A. (2022). Sistem pendukung keputusan izin bepergian harian santri putra Pondok Pesantren Salafiyah Syafiiyah Sukorejo. *JUSTIFY: Jurnal Sistem Informasi Ibrahimy*, 1(1). https://doi.org/10.35316/justify.v1i1.2153

Sartika, D., & Jumadi, J. (2023). Implementasi data mining untuk pengelompokan wilayah pelanggaran lalu lintas menggunakan metode K-Means pada Polres Bengkulu. *Journal of Science and Social Research*, (2). http://jurnal.goretanpena.com/index.php/JSSR

Sultan, Ismail, Madalle, & Agil. (2025). Efektivitas tata tertib pesantren dan sanksi dalam meningkatkan. *Al-Ilmi*. https://doi.org/10.47435/Al-Ilmi

Utami, N. C., Hananto, A. L., Tukino, & Hilabi, S. S. (2023). Komparasi metode klasifikasi data mining algoritma C4.5 dan Naïve Bayes. *Journal of Social Science Research*, 3, 503–513.

Wahana, P. D. (2020). Pesantren dan pemberdayaan masyarakat (Analisis terhadap Undang-Undang Nomor 18 Tahun 2019).

Yatimah, M. N. (2021). Implementasi data mining untuk prediksi kelulusan tepat waktu mahasiswa STIMIK ESQ menggunakan decision tree C4.5. *JUMANJI*, 5(2). https://doi.org/10.26874/jumanji.v5i2.95