

Analisis Pola Kriminalitas Menggunakan Data Science Dengan Metode Catboost Classifier Di Polres Kota Bengkulu

¹ Santiya Saputri, ² Devi Sartika,, ³ Abdussalam Al Akbar²

¹ Mahasiswa, Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Dehasen Bengkulu
Alamat : Jl. Cendrawasih no. 36 Rt. 03 Rw 01 kebun geran
email : saputrisantiya@gmail.com

^{2,3} Dosen Tetap, Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Dehasen Bengkulu
Jl. Meranti Raya No. 32 Kota Bengkulu 38228 Telp. (0736) 22027, 26957 Fax. (0736) 341139;
e-mail: devisartika@unived.ac.id akbarabenk@unived.ac.id

(Received: Juni 2026, Revised: Agustus 2026, Accepiend: Oktober 2026)

Abstract-Crime data management at the Bengkulu City Police Station currently faces several operational constraints, including underutilized historical data and conventional analytical methods that prolong the crime-pattern mapping process. This study aimed to analyze crime patterns through a data science approach using the CatBoost Classifier method to support data-driven tactical decision-making within the police department. A quantitative methodology using an Applied Data Science framework was executed following the Cross-Industry Standard Process for Data Mining (CRISP-DM) framework. The study analyzed crime report data extracted from the Bengkulu City Police database, focusing on temporal (time of occurrence) and spatial (district/kecamatan) variables. Model performance was comprehensively evaluated using a Confusion Matrix. Findings revealed that Motor Vehicle Theft (Curanmor) was the most dominant crime category, with peak vulnerability observed in Selebar District and the majority of incidents concentrated during the early morning hours. Feature Importance analysis indicated that the temporal dimension (Time Category) contributed the most (30.37%), followed by geographical location (District) at 29.30% identifying them as the primary predictors shaping crime patterns. Model evaluation for the CatBoost Classifier yielded an overall accuracy of 49%, reflecting performance limitations due to the small dataset size and imbalanced class distribution. Nevertheless, the model demonstrated a high recall value (89%) specifically in detecting Motor Vehicle Theft, maintaining its militly as an early warning signal for street crimes. Conversely, the model completely failed to predict fraud offenses (0% accuracy) because fraud characteristics are not tied to spatial-temporal patterns. The application of the CatBoost Classifier in this study ts best suited as a tactical intelligence support instrument for street crime mitigation, enabling police administrators to optimize directed patrol strategies in high-risk zones and time windows

Keywords: Crime. Data Science CatBoost Classifier. CRISP-DM. Strategic Decision Making

Intisari-Pengelolaan data kriminalitas di Polres Kota Bengkulu saat ini masih menghadapi kendala, di antaranya pemanfaatan data historis yang belum optimal dan analisis yang masih konvensional, sehingga proses pemetaan pola kejahatan membutuhkan waktu yang lama.

Penelitian ini bertujuan untuk menganalisis pola kriminalitas dengan pendekatan data science menggunakan metode CatBoost Classifier guna mendukung pengambilan keputusan taktis kepolisian yang berbasis data. Penelitian ini menggunakan metode kuantitatif dengan pendekatan Applied Data Science melalui kerangka kerja Cross-Industry Standard Process for Data Mining (CRISP-DM). Subjek penelitian merupakan data laporan kriminal dari database Polres Kota Bengkulu, yang difokuskan pada pengolahan variabel temporal (waktu) dan spasial (kecamatan). Performa model dievaluasi secara komprehensif menggunakan matriks pengujian Confusion Matrix. Hasil penelitian mengungkapkan bahwa Pencurian Kendaraan Bermotor (Curanmor) merupakan tindak pidana yang paling mendominasi, dengan titik kerawanan tertinggi berada di Kecamatan Selebar dan mayoritas kejadian terpusat pada rentang waktu dini hari. Berdasarkan analisis Feature Importance, dimensi temporal atau Kategori Waktu memberikan kontribusi paling besar (30,37%) yang disusul oleh letak geografis atau Kecamatan (29,30%) sebagai prediktor utama pembentuk pola kejahatan. Evaluasi model CatBoost Classifier mencatatkan tingkat akurasi global sebesar 49%, yang mengindikasikan keterbatasan performa akibat ukuran dataset yang kecil dan distribusi kelas yang tidak seimbang. Meskipun demikian, model menunjukkan nilai Recall yang tinggi (89%) secara khusus dalam mendeteksi Pencurian Kendaraan Bermotor (Curanmor), sehingga tetap berpotensi digunakan sebagai sinyal kewaspadaan pada jenis kejahatan tersebut. Sebaliknya, model gagal total dalam mendeteksi kejahatan penipuan (Akurasi 0%) karena karakteristiknya yang tidak terikat pola ruang dan waktu. Sebagai kesimpulan, implementasi CatBoost Classifier pada penelitian ini lebih tepat diposisikan sebagai instrumen pendukung intelijen taktis dalam mitigasi kejahatan jalanan, yang dapat dimanfaatkan untuk mengoptimalkan strategi directed patrol (patroli terarah) pada wilayah dan waktu rawan.

Kata Kunci: Kriminalitas, Data Science, CatBoost Classifier, CRISP-DM, Keputusan Strategis

I. PENDAHULUAN

Perkembangan teknologi informasi dan komputasi dalam satu dekade terakhir menunjukkan akselerasi

yang sangat signifikan, khususnya pada bidang *data science* dan *machine learning*. Pemanfaatan teknologi ini memungkinkan pengolahan data berukuran besar (*big data*) secara cepat dan akurat untuk menghasilkan informasi yang bernilai strategis. Dalam konteks keamanan dan ketertiban masyarakat, teknologi informasi berperan penting dalam membantu institusi penegak hukum melakukan analisis pola kriminalitas secara lebih sistematis dan berbasis data.

Terjadinya kegiatan kriminal dalam wilayah hukum Polres Kota Bengkulu menunjukkan perubahan variabel yang bergantung pada jenis kejahatan, waktu kejadian, serta lokasi kejadian. Polres Kota Bengkulu mempunyai peran penting dalam menjaga keamanan dan ketertiban masyarakat melalui tindakan pencegahan dan hukuman. Kumpulan data historis yang berhubungan dengan kriminal memiliki karakteristik kompleks sehingga membutuhkan pendekatan analisis berbasis komputasi. Visualisasi dan pengelompokan data kriminal terbukti membantu aparat penegak hukum dalam memahami distribusi kejahatan dan menentukan strategi pengamanan wilayah (Rosiana et al., 2023).

Pada Polres Kota Bengkulu, pengelolaan data kriminal masih menghadapi beberapa kendala seperti kurang optimalnya pemanfaatan data histori, analisis data yang masih konvensional serta keterbatasan sistem klasifikasi untuk wilayah yang rentang terhadap aksi kriminal. Kondisi ini menyebabkan proses identifikasi pola kriminalitas memerlukan waktu yang lama. Padahal pemetaan wilayah rawan kejahatan sangat penting untuk mendukung kebijakan pengamanan yang tepat sasaran (Saputra et al., 2024). Beberapa penelitian terkait analisis data kriminal telah menunjukkan hasil yang efektif. Yoandito et al (2025) mengimplementasikan algoritma *Random Forest* untuk prediksi tingkat kejahatan dan memperoleh akurasi tinggi hingga 98%. Implementasi *K-Means* untuk mengelompokkan jenis kriminalitas efektif dalam visual tetapi tidak mampu melakukan klasifikasi prediktif (Rosiana et al., 2023). Sementara itu penelitian yang dilakukan Aditya et al (2025) membuktikan *CatBoost* memiliki performa baik dalam klasifikasi indeks standar pencemaran udara. Sakmar et al (2026) menjelaskan algoritma *XGBoost* dan *CatBoost* efektif dan baik terhadap dataset yang tidak seimbang. Dari beberapa studi tersebut bahwa penggunaan *CatBoost* pada data kriminalitas masih sangat sedikit dilakukan sehingga menjadi *research gap* pada penelitian ini. Berdasarkan penelitian terdahulu yang telah dijabarkan, penelitian ini akan mengimplementasikan metode *CatBoost Classifier* untuk menganalisa pola kriminalitas di Polres Kota Bengkulu. *CatBoost Classifier* merupakan algoritma *gradient boosting* yang dirancang untuk mengatasi data kategorikal

secara optimal serta mampu meningkatkan akurasi model (Razi et al., 2025). Metode ini akan diterapkan melalui tahapan preprocessing, pembentukan model, serta evaluasi performa menggunakan metrik klasifikasi (Pahlevi et al., 2024). Hasil analisis diharapkan mampu memberikan rekomendasi prioritas pengamanan wilayah berbasis data.

Berdasarkan latar belakang tersebut, penelitian analisis pola kriminalitas menggunakan *data science* dengan metode *CatBoost Classifier* menjadi penting dilakukan. Penelitian ini diharapkan memberikan kontribusi akademik dalam pengembangan penerapan *data science* serta manfaat praktis bagi Polres Kota Bengkulu dalam mendukung pengambilan keputusan strategis yang lebih akurat dan berbasis data.

II. TINJAUAN PUSTAKA

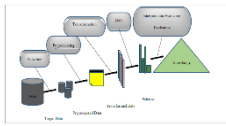
Beberapa penelitian berikut akan menggambarkan dan memperkuat arah penelitian yang akan dilakukan. Selain itu melalui penelitian terdahulu ini juga penulis dapat melakukan pendekatan terhadap tema penelitian.

No	Judul, Penulis dan Tahun	Metode yang Digunakan	Algoritma yang Digunakan	Hasil Penelitian	Limitasi Penelitian
1	Klasifikasi Pola Kriminalitas Berbasis Data (Razi et al., 2025)	Analisis regresi linier, analisis deskriptif, dan analisis korelasi	Model regresi linier	Hasil analisis menunjukkan bahwa tipe data waktu memiliki pengaruh signifikan terhadap tingkat kejahatan	Pada penelitian ini, data yang digunakan adalah data sekunder, sehingga mungkin terdapat bias dalam analisis
2	Penerapan Sistem Prediksi Kriminalitas Berbasis Data (Sakmar et al., 2026)	Analisis deskriptif, analisis korelasi, dan analisis regresi linier	Model regresi linier	Hasil analisis menunjukkan bahwa tipe data waktu memiliki pengaruh signifikan terhadap tingkat kejahatan	Pada penelitian ini, data yang digunakan adalah data sekunder, sehingga mungkin terdapat bias dalam analisis
3	Penerapan Sistem Prediksi Kriminalitas Berbasis Data (Yoandito et al., 2025)	Analisis deskriptif, analisis korelasi, dan analisis regresi linier	Model regresi linier	Hasil analisis menunjukkan bahwa tipe data waktu memiliki pengaruh signifikan terhadap tingkat kejahatan	Pada penelitian ini, data yang digunakan adalah data sekunder, sehingga mungkin terdapat bias dalam analisis
4	Penerapan Sistem Prediksi Kriminalitas Berbasis Data (Aditya et al., 2025)	Analisis deskriptif, analisis korelasi, dan analisis regresi linier	Model regresi linier	Hasil analisis menunjukkan bahwa tipe data waktu memiliki pengaruh signifikan terhadap tingkat kejahatan	Pada penelitian ini, data yang digunakan adalah data sekunder, sehingga mungkin terdapat bias dalam analisis
5	Penerapan Sistem Prediksi Kriminalitas Berbasis Data (Saputra et al., 2024)	Analisis deskriptif, analisis korelasi, dan analisis regresi linier	Model regresi linier	Hasil analisis menunjukkan bahwa tipe data waktu memiliki pengaruh signifikan terhadap tingkat kejahatan	Pada penelitian ini, data yang digunakan adalah data sekunder, sehingga mungkin terdapat bias dalam analisis

Gambar 1. Penelitian Terdahulu

Kriminalitas adalah segala bentuk tindakan atau perbuatan yang merugikan orang lain, baik secara ekonomi maupun secara psikologis, serta melanggar hukum, norma sosial, dan norma agama yang berlaku di Indonesia. Artinya, tindakan kriminal adalah perbuatan yang tidak sesuai dengan aturan dan nilai-nilai yang dianut masyarakat, sehingga dianggap salah dan ditolak oleh masyarakat (Yazid & Mulya, 2025). *Data science* merupakan gabungan ilmu komputer, statistika, dan pengetahuan domain bisnis untuk ekstraksi tumpukan data yang besar menjadi pengetahuan sehingga mendapatkan pattern atau pola-pola yang dapat memudahkan pengambil keputusan (Hairani & Amrullah, 2021). *Data mining* adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar (Dongga et al., 2023). Selanjutnya menurut Rahmat et al., (2021) mengungkapkan bahwa data mining merupakan pendekatan sistematis yang memanfaatkan algoritma komputasi untuk mengolah data berskala besar sehingga menghasilkan informasi strategis yang mendukung perencanaan dan evaluasi kebijakan.

Data mining merupakan salah satu dari rangkaian *Knowledge Discovery in Database* (KDD). KDD berhubungan dengan teknik integrasi dan penemuan ilmiah, interpretasi dan visualisasi dari pola-pola sejumlah kumpulan data (Naldy & Andri, 2021)



Gambar 2. Tahapan Data Mining
(Sumber : Naldy & Andri, (2021))

Klasifikasi merupakan salah satu metode dalam pengolahan data yang merupakan pengelompokkan sekumpulan data sesuai dengan karakteristik atau ciri-ciri data itu sendiri satu sama lain. Klasifikasi dilakukan dengan mengumpulkan sekumpulan data latih dengan beberapa kelas yang telah ditentukan. Pengklasifikasian ini bisa dilakukan dengan dua cara, yaitu dengan cara yang manual atau dengan memanfaatkan teknologi yang ada. Pada pemanfaatan teknologi di metode klasifikasi dapat dilakukan dengan menerapkan algoritma yang ditujukan untuk klasifikasi seperti *Decision Tree*, *Naive Bayes*, dan *K-Nearest Neighbor* (Biyantoro & Prasetyo, 2024).

CatBoost merupakan algoritma boosting yang terurut, dan diimplementasikan berdasarkan algoritma peningkatan gradien dengan menggunakan decision tree yang digunakan untuk regresi, dan dari masing-masing tree sesuai dengan partisi sebuah feature space dan output value. CatBoost ini menangani data kategorikal secara otomatis menggunakan metode statistik. CatBoost ini dapat menghindari overfitting data dengan mengoptimalkan berbagai parameter masukan, hal tersebut tidak berurusan dengan karakteristik kategori selama pemrosesan. CatBoost ini akan melakukan eksekusi permutasi secara acak pada data kategori alih-alih penggantian biner dan tiap-tiap label dihitung rata-ratanya (Istianto et al., 2024). CatBoost melatih *ensemble* pohon keputusan simetris (oblivious trees) dengan kerangka boosting berbasis gradien; setiap iterasi menambahkan pohon yang meminimalkan loss menggunakan informasi gradien yang dihitung tanpa kebocoran target.

Keunggulan utama *CatBoost* terletak pada kemampuannya memproses fitur kategorikal tanpa memerlukan prapemrosesan *One-Hot Encoding* secara manual. Algoritma ini menerapkan teknik *Ordered Target Statistics* untuk mengubah nilai kategorikal menjadi angka berdasarkan statistik target. Nguyen & Ngo (2025) menjelaskan bahwa pendekatan ini memungkinkan model mempertahankan informasi spasial dari kategori yang memiliki kardinalitas tinggi tanpa meningkatkan dimensi data secara signifikan. Penggunaan CatBoost dalam keamanan publik sangat efektif karena data kriminalitas sering kali mengandung banyak variabel kategorikal, seperti jenis pelanggaran, lokasi

kejadian, dan identitas waktu. Penelitian terbaru menunjukkan bahwa CatBoost mampu menangani dataset yang tidak seimbang (*imbalanced data*) dengan lebih baik dibandingkan model *Random Forest* atau *XGBoost* (Deswandi & Hastomo, 2024). Python merupakan bahasa pemrograman tingkat tinggi yang dibuat oleh Guido Van Rossum dan dirilis pada tahun 1991. Python juga merupakan bahasa yang sangat populer belakangan ini. Selain itu python juga merupakan bahasa pemrograman yang multi fungsi contohnya python dapat digunakan untuk Machine Learning dan Deep Learning. Python dipilih sebagai penelitian karena python memiliki penulisan sintaksis yang mudah selain itu python juga memiliki library yang lengkap dan memiliki dukungan komunitas yang kuat karena python bersifat open source. Untuk menuliskan source code python anda dapat menggunakan IDE seperti Visual Studio Code, Sublime Text, PyCharm atau anda juga dapat menggunakan IDE online seperti Jupyter notebook dan google colab (Alfarizi et al., 2023).

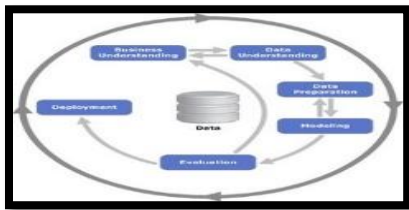
Google Collaboratory, juga dikenal sebagai *Google Colab*, adalah perangkat lunak gratis yang berbasis *cloud* yang digunakan untuk tujuan pendidikan dan penelitian. perangkat ini dapat diakses secara gratis melalui internet (Octarina et al., 2025). *Google Colab* adalah *platform cloud computing* dari Google yang memungkinkan pengguna menjalankan kode Python langsung di browser tanpa perlu instalasi lokal. Platform ini banyak digunakan untuk pemrosesan data, pelatihan model kecerdasan buatan, dan analisis data (Irsabmalfi & Rahmadewi, 2025). Flowchart yang sering disebut dengan diagram alir adalah suatu jenis diagram yang merepresentasikan algoritma atau langkah-langkah instruksi yang berurutan dalam sistem. flowchart dibuat dengan tujuan agar melihatkan setiap proses yang harus dilalui dalam suatu sistem (Febryana et al., 2024).

III. METODOLOGI PENELITIAN

Subjek penelitian ini adalah data laporan kriminal yang tercatat pada database Polres Kota Bengkulu. Fokus utama penelitian ini adalah menganalisa klasifikasi pola kriminalitas untuk mendukung pengambilan keputusan. Waktu penelitian dimulai dari bulan November 2025 sampai dengan Mei 2026. Penelitian dilakukan di Polres Kota Bengkulu yang beralamat di Jl. A. Yani, Malabero, Bengkulu, Kota Bengkulu, Bengkulu 38116. Penelitian ini menggunakan jenis penelitian kuantitatif dengan pendekatan *Applied Data Science*. Pendekatan ini berfokus pada penggunaan data untuk menghasilkan pengetahuan yang dapat ditindaklanjuti (*actionable insights*). Tujuannya bukan sekadar mendeskripsikan data, melainkan memprediksi pola kriminalitas melalui algoritma *CatBoost Classifier* agar dapat

memberikan solusi bagi manajemen Polres Kota Bengkulu.

Metode penelitian yang digunakan dalam penelitian ini adalah pendekatan data mining untuk menemukan pola tersembunyi pada data transaksi penjualan. Pada penelitian ini, digunakannya metodologi data mining *Cross-Industry Standard Process for Data Mining* (CRISP-DM) sebagai pemecah masalah yang umum untuk bisnis dan penelitian. Metodologi ini terdiri dari enam tahapan yaitu *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modelling*, *Evaluation*, dan *Deployment* (Hasanah et al., 2021).



Gambar 3. Tahapan CRISP-DM
(Sumber : (Hasanah et al., 2021))

Berdasarkan observasi awal yang dilakukan, diperoleh keterangan bahwa setiap laporan tindak kriminal yang masuk melalui Sentra Pelayanan Kepolisian Terpadu (SPKT) dicatat ke dalam buku register atau diinput ke dalam format spreadsheet sederhana (seperti Microsoft Excel). Sistem yang akan dibangun membutuhkan data histori transaksi penjualan sebagai input utama yang akan diolah dengan tahapan data mining menggunakan algoritma Catboost. Hasil dari analisis ini diharapkan menghasilkan informasi berupa pola transaksi atau rule keputusan yang dapat digunakan sebagai dasar dalam menentukan target promosi. Arsitektur sistem ini dirancang dengan prinsip modularitas, di mana setiap komponen memiliki peran spesifik untuk memastikan transformasi data dari laporan kriminal mentah menjadi rekomendasi strategis yang valid. Secara garis besar.



Gambar 4. Perancangan Arsitektur Analisis Pola Kriminalitas

Pada perancangan evaluasi dilakukan dengan melihat tingkat performa dari pola yang dihasilkan oleh algoritma CatBoost. Parameter yang digunakan untuk evaluasi komparasi algoritma adalah *Confusion Matrix* (Hasanah et al., 2021).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\%$$

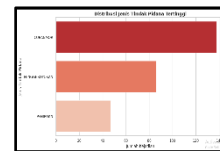
$$Precision = \frac{TP}{FP + TP} \times 100\%$$

$$Recall = \frac{TP}{TP + FN} \times 100\%$$

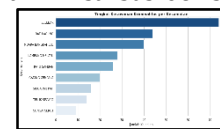
IV. HASIL DAN PEMBAHASAN

Penelitian analisa pola Kriminalitas dengan metode CatBoost Classifier di Polres Kota Bengkulu berdasarkan pengolahan dataset histori tindak pidana tidak hanya menunjukkan tingkat kemampuan model dalam melakukan prediksi jenis tindak pidana, tetapi juga memberikan gambaran mengenai pola kriminalitas berdasarkan aspek waktu dan wilayah. Informasi tersebut diharapkan dapat menjadi dasar pendukung dalam proses pengambilan keputusan operasional di lingkungan Polres.

Gambar 5. Hasil Ekstraksi Dataset Data_Kriminal.xlsx

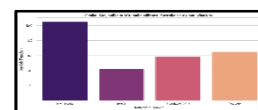


Gambar 6. Grafik Distribusi Jenis Tindak Pidana



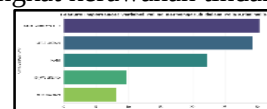
Gambar 7. Grafik Kerawanan Kriminalitas Per Kecamatan

Terlihat jelas bahwasanya tindak kriminal paling tinggi adalah curanmor yang kemudian diikuti dengan pengeroyokan dan terakhir adalah penipuan. Selain grafik distribusi jenis tindak pidana, proses EDA juga menghasilkan grafik daerah rawan kriminal per kecamatan.



Gambar 8. Grafik Pola Kejadian Kriminalitas Berdasarkan Waktu

Kecamatan Seblar merupakan daerah yang mempunyai tingkat kerawanan tindak kriminalitas.



Gambar 9. Grafik Feature Importance

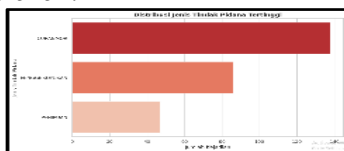
Feature importance merupakan tahapan yang sangat penting untuk dilakukan dalam proses data science, dengan tahapan ini akan mengetahui kontribusi kriteria yang mempunyai hubungan dalam pola tindak kriminal.



Gambar 10. Source Code dan Hasil Modelling CatBoost Pada tahap modelling CatBoost, dataset dibagi menjadi dua proporsi, yaitu 80% sebagai data latih (*training data*) dan 20% sebagai data uji (*testing data*). menggunakan metode train-test split dengan parameter stratify yang memastikan rasio kelas target pada data uji sebanding dengan data latih, sehingga pengujian berlangsung secara representatif dan seimbang

Gambar 11. Source Code dan Hasil Evaluasi Tahapan evaluasi dan pengujian dilakukan untuk mengukur performa, reliabilitas, dan kemampuan generalisasi model CatBoost Classifier terhadap data yang belum pernah diproses sebelumnya (*unseen data*). Evaluasi ini merupakan fase penting untuk memvalidasi apakah model yang telah dilatih mampu bekerja secara objektif di luar ruang lingkup data latih. Pengujian tidak dilakukan hanya berdasarkan metrik akurasi global (*accuracy*), melainkan menggunakan pendekatan multi-metrik yang lebih representatif, yaitu *Confusion Matrix*, *Precision*, *Recall*, dan *F1-Score*

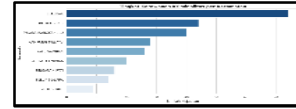
Distribusi jenis tindak pidana menunjukkan frekuensi kejadian untuk setiap kategori tindak pidana yang terdapat pada dataset. Hasil visualisasi memperlihatkan bahwa **Curanmor** merupakan jenis tindak pidana yang paling dominan dengan jumlah kejadian sekitar **138 kasus**. Selanjutnya diikuti oleh tindak pidana **Pengeroyokan** sebanyak sekitar **86 kasus**, sedangkan **Penipuan** memiliki jumlah kejadian paling sedikit, yaitu sekitar **47 kasus**. Berdasarkan hasil tersebut, fokus pencegahan kriminalitas mengharuskan institusi penegak hukum memprioritaskan rancangan pengamanan preventif untuk menekan angka kejahatan properti tanpa mengesampingkan mitigasi responsif terhadap kejahatan jalanan.



Gambar 12. Jenis Tindak Pidana

Visualisasi tingkat kerawanan kriminalitas berdasarkan kecamatan menunjukkan adanya perbedaan jumlah kejadian pada masing-masing wilayah. Kecamatan **Selebar** memiliki jumlah kasus tertinggi dengan sekitar **74 kejadian**, diikuti oleh **Ratu Agung** sebanyak **44 kejadian** dan **Muara Bangkahulu** sebanyak **40 kejadian**. Sebaliknya,

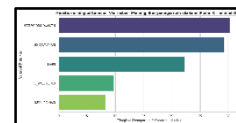
Kecamatan **Sungai Serut** merupakan wilayah dengan jumlah kejadian paling rendah, yaitu sekitar **9 kasus**.



Gambar 13. Kecamatan Rawan Tindak Kriminal Perbedaan jumlah kasus tersebut menunjukkan bahwa distribusi kriminalitas tidak terjadi secara merata pada seluruh wilayah hukum Polres Kota Bengkulu. Konsentrasi tindak pidana pada Kecamatan Selebar mengindikasikan bahwa wilayah tersebut memiliki tingkat kerawanan yang lebih tinggi dibandingkan kecamatan lainnya.

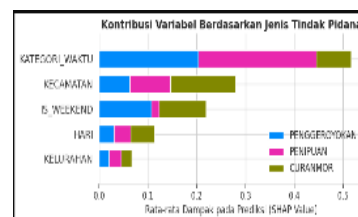


Gambar 14. Pola Kejadian Berdasarkan Waktu Analisis terhadap kategori waktu bertujuan untuk mengetahui kecenderungan terjadinya tindak pidana berdasarkan rentang waktu tertentu. Hasil visualisasi menunjukkan bahwa kejadian kriminalitas paling banyak terjadi pada kategori **Dini Hari**, yaitu sekitar **105 kejadian**. Selanjutnya kategori **Malam** memiliki sekitar **65 kejadian**, sedangkan kategori **Siang/Sore** mencapai sekitar **59 kejadian**. Jumlah kejadian paling sedikit terjadi pada kategori **Pagi**, yaitu sekitar **42 kejadian**.



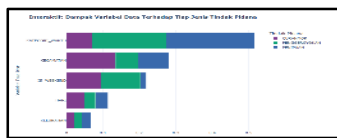
Gambar 15. Variabel Paling Berpengaruh Dalam Pola Kriminalitas

Berdasarkan hasil analisa dengan metode CatBoost, variabel **Kategori Waktu** memiliki nilai kontribusi terbesar, yaitu sekitar **30%**. Posisi berikutnya ditempati oleh variabel **Kecamatan** dengan kontribusi sekitar **29%**, kemudian **Hari** sebesar **22%**, **Is_Weekend** sebesar **10%**, dan **Kelurahan** sekitar **8%**.



Gambar 16. Kontribusi Variabel Berdasarkan Jenis Tindak Pidana

Hasil visualisasi menunjukkan bahwa **Kategori Waktu** memiliki kontribusi terbesar terhadap seluruh jenis tindak pidana. Variabel ini memberikan pengaruh yang cukup dominan pada klasifikasi **Pengeroyokan**, **Penipuan**, maupun **Curanmor**. Kondisi tersebut memperkuat hasil analisis sebelumnya bahwa dimensi waktu merupakan faktor utama yang membedakan karakteristik setiap jenis tindak pidana.



Gambar 17. Dampak Variabel Terhadap Jenis Tindak Pidana

Visualisasi interaktif dampak variabel terhadap jenis tindak kriminal memberikan gambaran mengenai besarnya pengaruh setiap variabel prediktor terhadap masing-masing kategori tindak pidana secara lebih rinci. Hasil analisis menunjukkan bahwa *Kategori Waktu* merupakan variabel dengan dampak paling besar terhadap ketiga jenis tindak pidana, yaitu *Curanmor*, *Pengeroyokan* dan *Penipuan*. Untuk memperjelas pola kriminalitas per jenis tindak pidana, dilakukan analisis crosstabulation antara jenis tindak pidana dengan kategori waktu kejadian. Hasil analisis menunjukkan bahwa Curanmor dan Penggeroyokan memiliki pola waktu yang serupa, yaitu terkonsentrasi pada dini hari (masing-masing 44,2% dan 45,3% dari total kejadian), yang mengindikasikan lemahnya pengawasan pada rentang waktu tersebut. Sebaliknya, tindak pidana Penipuan menunjukkan pola yang berbeda, dengan 42,6% kejadian terjadi pada siang/sore hari, sesuai karakteristik modus penipuan yang umumnya memerlukan interaksi langsung dengan korban pada jam aktif.

Jenis Tindak Pidana	Dini Hari	Malam	Pagi	Siang/Sore
Curanmor (n=138)	61 (44,2%)	29 (21,0%)	26 (18,8%)	22 (15,9%)
Pengeroyokan (n=86)	39 (45,3%)	25 (29,1%)	5 (5,8%)	17 (19,8%)
Penipuan (n=47)	5 (10,6%)	11 (23,4%)	11 (23,4%)	20 (42,6%)

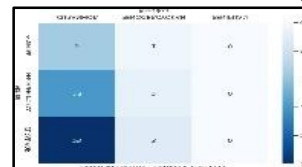
Tabel 1. Crosstabulation Jenis Tindak Pidana terhadap Kategori Waktu

Jenis Tindak Pidana	Kecamatan Dominan (5 teratas)
Curanmor	Selebar (44), Muara Bangkahulu (28), Kampung Melayu (14), Ratu Agung (12), Ratu Samban (12)
Pengeroyokan	Ratu Agung (24), Selebar (17), Ratu Samban (11), Gading Cempaka (8), Muara Bangkahulu (8)
Penipuan	Selebar (13), Ratu Agung (8), Kampung Melayu (7), Singaran Pati (5), Muara Bangkahulu (4)

Tabel 2. Kecamatan Dominan per Jenis Tindak Pidana
Tahap evaluasi dilakukan untuk mengukur kemampuan model **Categorical Boosting (Catboost) Classifier** dalam mengklasifikasikan pola kriminalitas tindak pidana berdasarkan dataset yang digunakan. Evaluasi dilakukan menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score*.

No	Kelas Penyebab	Pesisi (%)	Recall (%)	F1-Score (%)	Jumlah Data (support)
1	Curanmor	0.51	0.89	0.65	28
2	Pengeroyokan	0.33	0.12	0.17	17
3	Penipuan	0	0	0	10
DATA UNTUK RINGKASAN PERSENTASE					
Akurasi Keseluruhan		49%			
Macro Average F1-Score		27%			
Weighted Average F1-Score		38%			

Tabel 3. Hasil Evaluasi dan Pengujian



Gambar 18. Confusion Matrix Model CatBoost Classifier
Analisis metrik evaluasi memperlihatkan adanya variasi performa yang baik antar-jenis tindak pidana, yang secara langsung merepresentasikan karakteristik *imbalanced dataset* (ketidakseimbangan jumlah data antar kelas). Berikut adalah penjabaran analitis terhadap kapabilitas prediktif model pada masing-masing tindak kejahatan:

1. Kapabilitas Prediksi Tindak Pidana Curanmor

Model menunjukkan performa prediktif paling optimal pada kelas kejahatan Pencurian Kendaraan Bermotor (Curanmor). Nilai *Recall* (sensitivitas) menyentuh angka 0.89, yang mengartikan bahwa model berhasil mendeteksi dan memprediksi dengan benar 89% dari seluruh insiden Curanmor aktual yang terjadi di lapangan. Di sisi lain, nilai Presisi sebesar 0.51 menunjukkan bahwa setiap kali sistem mengeluarkan peringatan bahaya Curanmor, probabilitas ketepatannya adalah 51%. Dari perspektif taktis kepolisian, nilai *Recall* yang sangat tinggi ini bersifat fatal dan menguntungkan; sistem dirancang untuk memiliki sensitivitas maksimal agar tidak ada potensi kejahatan properti yang luput dari radar patroli (*minimizing false negatives*), menjadikan model ini sangat layak dijadikan acuan kewaspadaan Curanmor.

2. Kapabilitas Prediksi Tindak Pidana Pengeroyokan

Pada kelas kejahatan Penggeroyokan, model mengalami penurunan performa dengan mencatatkan nilai Presisi 0.33 dan *Recall* 0.12. Rendahnya sensitivitas ini mengindikasikan bahwa algoritma kesulitan menangkap pola deterministik yang memicu terjadinya penggeroyokan. Fenomena ini dapat dijustifikasi secara kriminologis: berbeda dengan Curanmor yang sangat terikat pada lokasi sepi dan jam rawan, eskalasi penggeroyokan cenderung bersifat

reaktif, insidental, dan sering kali dipicu oleh gesekan sosial sporadis yang tidak selalu terikat pada parameter ruang dan waktu secara absolut

3. Kapabilitas Prediksi Tindak Pidana Penipuan Model mencatatkan nilai absolut 0.00 pada seluruh metrik (Presisi, *Recall*, dan F1-Score) untuk kelas Penipuan. Kegagalan algoritma dalam memprediksi kelas ini bukanlah cacat komputasional, melainkan pembuktian dari teori anomali spasial. Merujuk pada tabel kontribusi variabel, model ini bertumpu pada Kategori Waktu (30.37%) dan **Kecamatan (29.30%)** sebagai parameter pengambil keputusan utama. Faktanya, kejahatan penipuan (terutama di era modern) merupakan bentuk *white-collar crime* atau kejahatan siber yang bersifat *borderless* (tidak bersinggungan dengan batas geografis fisik) dan tidak mengenal jam rawan. Karena penipuan tidak terdistribusi berdasarkan kaidah ruang (Kecamatan/Kelurahan) dan waktu (Dini hari/Malam), wajar jika model yang dilatih menggunakan fitur spasio-temporal gagal memetakannya.

Dari hasil pengujian ini membuktikan bahwa algoritma CatBoost yang dibangun sangat andal dan direkomendasikan untuk memetakan jenis kejahatan konvensional jalanan (Curanmor) yang memiliki akar kausalitas kuat terhadap ruang wilayah dan jam kejadian. Di sisi lain, jarak nilai antara *macro average* (0.27) dan *weighted average* (0.38) menegaskan bahwa model ini sangat dipengaruhi oleh distribusi mayoritas data latih.

V. PENUTUP

A. Kesimpulan

Berdasarkan seluruh tahapan analisis, pemodelan, dan evaluasi yang telah dilakukan, rumusan konklusi dari penelitian ini adalah sebagai berikut

1. Implementasi algoritma CatBoost Classifier pada penelitian ini berperan utama sebagai instrumen pengukur kontribusi variabel kategorikal (Kecamatan, Kelurahan, Hari) dan temporal terhadap pola kriminalitas, melalui feature importance dan SHAP value, dengan kemampuan prediksi jenis tindak pidana sebagai bukti pendukung tambahan. Model berhasil memproses himpunan data kriminalitas yang didominasi oleh variabel kategorikal tanpa memerlukan encoding manual, dengan akurasi prediksi 49% pada data uji (dijelaskan lebih lanjut pada poin 5).
2. Ekstraksi Feature Importance menunjukkan bahwa parameter temporal dan spasial merupakan prediktor paling berpengaruh dalam pola kriminalitas pada dataset ini. Kategori Waktu menyumbangkan pengaruh terbesar (30,37%), diikuti oleh variabel Kecamatan (29,30%). Temuan ini sejalan dengan teori kriminologi bahwa tindak kejahatan konvensional cenderung

terikat pada dimensi ruang (wilayah dengan pengawasan minim) dan waktu (momen kelengahan masyarakat), meskipun hubungan ini bersifat korelatif pada data yang tersedia dan belum tentu kausal.

3. Analisis pola per jenis tindak pidana (subbab 4.2.7) menunjukkan bahwa karakteristik pola berbeda antar kelas: Curanmor dan Penggeroyokan sama-sama terkonsentrasi pada dini hari (masing-masing 44,2% dan 45,3% dari total kejadian), namun terjadi di kecamatan yang berbeda — Selebar untuk Curanmor, Ratu Agung untuk Penggeroyokan. Sebaliknya, Penipuan justru dominan pada siang/sore hari (42,6%), menunjukkan pola yang berkebalikan dari dua jenis kejahatan lainnya. Temuan ini membuktikan bahwa pola kriminalitas di wilayah hukum Polres Kota Bengkulu tidak seragam dan perlu didekati secara spesifik per jenis kejahatan, bukan sebagai satu pola umum.
4. Kriminalitas di wilayah hukum Polres Kota Bengkulu didominasi oleh tindak pidana Pencurian Kendaraan Bermotor (Curanmor) dengan 138 dari 271 kasus, dengan titik kerawanan tertinggi berada di Kecamatan Selebar (74 kasus) dan Ratu Agung (44 kasus). Kejadian kriminalitas juga terkonsentrasi pada rentang waktu Dini Hari (pukul 00.00–05.59), yang mencatatkan 105 dari 271 kejadian. Berdasarkan analisis SHAP, variabel *Is_Weekend* memberikan kontribusi yang relatif lebih tinggi khususnya pada tindak pidana Penggeroyokan dibandingkan jenis tindak pidana lainnya.
5. Hasil pengujian metrik evaluasi menunjukkan performa model yang masih timpang antar kelas: model mencapai *Recall* 89% dalam mendeteksi Curanmor, namun *Precision* hanya 0,33 dan *Recall* 0,12 untuk Penggeroyokan, serta gagal total (0% pada seluruh metrik) untuk kelas Penipuan. Ketimpangan ini konsisten dengan kondisi data yang belum seimbang (*imbalanced*) antar kelas dan jumlah data yang masih terbatas (271 baris untuk satu tahun).

Kegagalan model pada kelas Penipuan justru memperkuat temuan pada poin 3, bahwa pola spasio-temporal memang tidak berlaku untuk jenis kejahatan tersebut, sehingga hal ini merupakan bagian dari hasil analisis itu sendiri dan bukan sekadar kelemahan model. Dengan demikian, model CatBoost pada penelitian ini lebih tepat diposisikan sebagai alat bantu eksplorasi awal untuk memetakan pola kejahatan berbasis wilayah dan waktu (khususnya Curanmor), dan belum dapat diandalkan sebagai alat prediksi operasional untuk seluruh jenis tindak pidana, terutama Penipuan yang polanya tidak banyak bergantung pada faktor lokasi dan waktu.

B.Saran

Guna menyempurnakan kualitas mitigasi keamanan serta pengembangan penelitian di ranah sains data kriminologi pada masa mendatang, direkomendasikan sejumlah langkah strategis berikut:

1. Optimalisasi Strategi Operasional Kepolisian (Polres) dengan mempertimbangkan transisi dari pola patroli acak menjadi directed patrol (patroli terarah) berbasis data, khususnya untuk kejahatan Curanmor yang pola spasial-temporalnya cukup jelas pada penelitian ini. Alokasi personel Sabhara dan Tim Patroli dapat dipertimbangkan untuk lebih difokuskan di Kecamatan Selebar dan Muara Bangkahulu pada rentang waktu Dini Hari, sebagai salah satu masukan pendukung keputusan bagi pihak Polres.
2. Untuk meningkatkan kualitas data dan performa model Machine Learning di masa depan, sistem pelaporan di Sentra Pelayanan Kepolisian Terpadu (SPKT) perlu mengakomodasi pencatatan data yang lebih granular dan dalam rentang waktu yang lebih panjang (lebih dari satu tahun). Penambahan variabel spesifik seperti titik koordinat (Latitude/Longitude), kondisi cuaca saat kejadian, hingga tingkat keramaian lokasi berpotensi memberikan fitur yang lebih kaya bagi model.
3. Penelitian selanjutnya disarankan mengintegrasikan teknik penyeimbangan data seperti SMOTE (Synthetic Minority Over-sampling Technique) atau class weighting sebelum tahap klasifikasi untuk mengurangi bias terhadap kelas mayoritas (Curanmor), serta menerapkan validasi silang (cross-validation) agar hasil evaluasi lebih dapat dipertanggungjawabkan secara statistik. Untuk kasus seperti Penipuan yang sulit dijelaskan hanya dari faktor lokasi dan waktu, dapat dipertimbangkan penambahan variabel deskriptif lain atau pendekatan Natural Language Processing (NLP) pada narasi laporan kejadian.
4. Cakupan data pada penelitian ini difokuskan pada tiga jenis tindak pidana yang paling banyak dilaporkan selama periode satu tahun, sesuai dengan ruang lingkup data yang diberikan dan disepakati untuk keperluan penelitian ini. Untuk penelitian selanjutnya, disarankan memperluas cakupan data baik dari sisi jenis tindak pidana maupun rentang waktu pengamatan (multi-tahun), sehingga pola kriminalitas yang terpetakan dapat lebih representatif terhadap kondisi keamanan di wilayah hukum Polres Kota Bengkulu secara keseluruhan.

DAFTAR PUSTAKA

- [1] A'la, F. Y., Hartatik, & Riasti, B. K. (2025). A Hybrid Approach for Recommender Systems Based on Alternating Least Squares and CatBoost. *Jurnal Teknik Informatika (Jutif)*, 6(4), 2825–2836. <https://doi.org/10.52436/1.jutif.2025.6.4.5002>
- [2] Adawiyah, Q. (2024). Penerapan Algoritma K-Means Clustering untuk Mengelompokkan Rekomendasi Jenis KB Berbasis Machine Learning (Studi Kasus di Puskesmas Pegang Baru). Universitas Putra Indonesia YPTK Padang.
- [3] Ade, H., Auti, P. A. H., Bansod, A., Bhaisare, S., & Dutta, S. (2025). CDAPS – Crime Data Analysis & Prediction System. *International Journal of Scientific Research in Engineering and Management*, 09(11), 1–9. <https://doi.org/10.55041/ijserem53669>
- [4] Aditya, A., Umbara, F. R., & Sabrina, P. N. (2025). Klasifikasi Indeks Standar Pencemaran Udara Menggunakan Algoritma Catboost Dengan Teknik Balancing Data Random UnderSampling. *Jurnal Algoritma*, 22(2), 1751–1760. <https://doi.org/10.33364/algoritma/v.22-2.2971>
- [5] Aditya, M. A., Mulyana, R. D., Eka, I. P., & Widiyanto, S. R. (2020). Penggabungan Teknologi Untuk Analisa Data Berbasis Data Science. *Seminar Nasional Teknologi Komputer & Sains (SAINTEKS)*, 51–56.
- [6] Ashby, M. (2023). Forecasting crime trends to support police strategic decision making. *CrimRxiv*, 1–28. <https://doi.org/10.21428/cb6ab371.8c79f146>
- [7] Biyantoro, A. S., & Prasetyo, B. (2024). Penerapan Decision Tree untuk Klasifikasi Status Kesehatan dengan perbandingan KNN dan Naive Bayes. *Indonesian Journal of Informatic Research and Software Engineering*, 4(1), 47–55. <https://journal.irpi.or.id/index.php/ijirse/article/view/1342>
- [8] Deswandi, A., & Hastomo, W. (2024). Predicting Crime Time Intervals Using Machine Learning Models. *Journal of Information Systems and Informatics*, 6(4), 2397–2418. <https://doi.org/10.51519/journalisi.v6i4.881>
- [9] Dongga, J., Koru, N., & Lante, G. (2023). Implementasi Data Mining Menggunakan Algoritma Apriori Dalam Menentukan Persediaan Barang (Studi Kasus: Toko Swapen Jaya Manokwari). *G-Tech: Jurnal Teknologi Terapan*, 7(1), 119–126.
- [10] Hairani, & Amrullah, A. Z. (2021). PELATIHAN PENGENALAN DATA SCIENCE UNTUK MENINGKATKAN KEMAMPUAN DALAM PENGOLAHAN DATA. *Jurnal Abdidas*, 2(5), 1092–1098.
- [11] Hasan, Z., AS, D. A., Febriyanti, A., & Mariska, S. (2023). Kriminalitas Pencurian Sepedah Motor Di Desa Gandri Kecamatan Penengahan Kabupaten Lampung Selatan. *JURNAL RECTUM: Tinjauan Yuridis Penanganan Tindak Pidana*, 5(3), 245. <https://doi.org/10.46930/jurnalrectum.v5i3.3501>
- [12] He, Y., Yang, B., & Chu, C. (2024). GA-CatBoost-Weight Algorithm for Predicting Casualties in Terrorist Attacks: Addressing Data Imbalance and Enhancing Performance. *Mathematics*, 12(6). <https://doi.org/10.3390/math12060818>
- [13] Insiyah, I., Khasanah, M., & Hendarsyah, T. P. (2023). Penerapan Metode Ward Clustering Untuk

- Pengelompokkan Daerah Rawan Kriminalitas Di Jawa Timur Tahun 2021. *Jurnal Statistika Dan Komputasi*, 2(1), 44–54. <https://doi.org/10.32665/statkom.v2i1.1664>
- [14] Istianto, A. F., Hadiana, A. I., & Umbara, F. R. (2024). Prediksi Curah Hujan Menggunakan Metode Categorical Boosting (Catboost). *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(4), 2930–2937. <https://doi.org/10.36040/jati.v7i4.7304>
- [15] Nahjan, M. R., Heryana, N., & Voutama, A. (2023). Implementasi Rapidminer Dengan Metode Clustering K-Means Untuk Analisa Penjualan Pada Toko Oj Cell. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(1), 101–104.
- [16] Naldy, E. T., & Andri, A. (2021). Penerapan Data Mining Untuk Analisis Daftar Pembelian Konsumen Dengan Menggunakan Algoritma Apriori Pada Transaksi Penjualan Toko Bangunan MDN. *Jurnal Nasional Ilmu Komputer*, 2(2), 89–101. <https://doi.org/10.47747/jurnalnik.v2i2.525>
- [17] Nguyen, N., & Ngo, D. (2025). Comparative analysis of boosting algorithms for predicting personal default. *Cogent Economics and Finance*, 13(1). <https://doi.org/10.1080/23322039.2025.2465971>
- [18] Pahlevi, O., Wulandari, D. A. N., Rahayu, L. K., Leidiyana, H., & Handrianto, Y. (2024). Model Klasifikasi Risiko Stunting Pada Balita Menggunakan Algoritma CatBoost Classifier. *Bulletin of Computer Science Research*, 4(6), 414–421. <https://doi.org/10.47065/bulletincsr.v4i6.373>
- [19] Rahmat, A., Hardi, H., Syam, F. A., Zamzami, Z., Febriadi, B., & Windarto, A. P. (2021). Utilization of the field of data mining in mapping the area of the Human Development Index (HDI) in Indonesia. *Journal of Physics: Conference Series*, 1783(1). <https://doi.org/10.1088/1742-6596/1783/1/012035>
- [20] Ramadhon, R. N., Ogi, A., Agung, A. P., Putra, R., Febrihartina, S. S., & Firdaus, U. (2024). Implementasi Algoritma Decision Tree untuk Klasifikasi Pelanggan Aktif atau Tidak Aktif pada Data Bank. *Karimah Tauhid*, 3(2), 1860–1874. <https://doi.org/10.30997/karimahtauhid.v3i2.11952>
- [21] Razi, S. M. A., Octariadi, B. C., & Siregar, A. C. (2025). Klasifikasi Penyakit Hepatitis Menggunakan Algoritma Catboost. *Kohesi: Jurnal Multidisiplin Saintek*, 8(3), 1–11. <https://doi.org/10.8734/Kohesi.v1i2.365>
- [22] Rosiana, P. S., Mohsa, A. A., Fadila, M. A., & Haerul, J. (2023). Visualisasi Data Karawang Dengan Menggunakan Algoritma Clustering K-Means. *JITET (Jurnal Informatika Dan Teknik Elektro Terapan)*, 11(3), 822–828.
- [23] Sabili, N. L., Umbara, F. R., & Melina. (2024). Klasifikasi Penyakit Diabetes Menggunakan Algoritma Categorical Boosting Dengan Faktor Risiko Diabetes. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(6), 11391–11398. <https://doi.org/10.36040/jati.v8i6.11447>
- [24] Sakmar, M., Kadir, N. T., Shoho, P. A., & Darmawan, A. (2026). Efektivitas Xgboost Lightghm Dan Catboost Pada Dataset Imbalanced Predictive Maintenance. *SINTA Jurnal Sistem Informasi Dan Teknologi Komputasi*, 3, 36–44. <https://doi.org/10.61124/sinta.v3i1.145>
- [25] Saputra, D. C., Herlambang, B. A., & Anam, A. K. (2024). Sistem Informasi Geografis Tindak Kriminal Di Jawa Tengah Tahun 2019 – 2021. *Jurnal Informatika Upgris*, 10(1), 47–55. <https://doi.org/10.26877/jiu.v10i1.18133>
- [26] Sartika, D., & Jumadi, J. (2023). Implementasi Data Mining Untuk Pengelompokkan Wilayah Pelanggaran Lalu Lintas Menggunakan Metode K-Means Pada Polres Bengkulu. *Journal of Science and Social Research*, 4307(2), 345–349. <http://jurnal.goretanpena.com/index.php/JSSR>
- [27] Yazid, F., & Mulya, E. S. (2025). Tipologi Kriminal: Memahami Kategori, Bentuk, Karakteristik Dan Faktor-Faktor Penyebab Terjadinya Tindakan Kriminal. *Jurnal Psikologi Dan Bimbingan Konseling*, 11(2), 1–10.
- [28] Yoandito, N., Rosyidah, U. A., & Abdurrahman, G. (2025). Penerapan Algoritma Random Forest Dalam Analisis Dan Prediksi Tingkat Kejahatan Berdasarkan Data Kriminalitas Polres Jember. *Jurnal Smart Teknologi*, 6(6), 2774–1702. <http://jurnal.unmuhjember.ac.id/index.php/JST>
- [29] Zhao, X., Fan, W., Liu, H., & Tang, J. (2022). Multi-Type Urban Crime Prediction. *Proceedings : Conference on Artificial Intelligence*, 36(1), 4388–4396. <https://doi.org/10.1609/aaai.v36i4.20360>