

## Analisis Prediksi Tindak Pidana Pencurian Dengan Metode Klasifikasi Algoritma K-Nearest Neighbor di Polresta Bengkulu, Polda Bengkulu

Firman Bagus Perdana Kusuma<sup>1</sup>, Jarot Prianggono<sup>2</sup>, Didit Bambang Wibowo. S<sup>3</sup>

<sup>1) s/d 3)</sup> Sekolah Tinggi Ilmu Kepolisian – Perguruan Tinggi Ilmu Kepolisian

email: firmanbagusp@gmail.com, komputerstik@gmail.com,

diditbws@gmail.com

### Article History

Received: 15/01/2025

Revised: 22/01/2025

Accepted: 29/01/2025

**Keywords:** K-Nearest Neighbor, crime location prediction, theft crime, machine learning, Badung Police.

**Abstract:** This research is driven by the phenomenon of high crime rates, especially theft. Data mining can be applied to crime data, especially theft crime data, to obtain information that can be used as a basis for decision making in carrying out patrols by the Police. Machine learning algorithms can be used to classify types of theft crimes based on characteristics and predict the possibility of future theft crimes based on influencing factors. Data mining is a logical process to find information that is very useful as a supporting tool in decision making. K-Nearest Neighbor (KNN) is a classification method and uses the CRISP-DM (Cross Industry Standard Process for Data Mining) framework in pulling information from a collection of datasets. This study uses a population in the form of Police Reports (LP) of Theft Victims that occurred in the Bengkulu Police Resort which occurred from 2020 to 2025. The K-Nearest Neighbor (KNN) model shows a fairly high level of reliability in predicting the time of theft, the age of the victim, and the type of item stolen at Bengkulu Police Resort. The KNN model was able to predict the time of incidents, victim age, and type of stolen items with high accuracy — 87.35%, 82.41%, and 88.43% respectively. Based on the results of this study, the application of machine learning in predictive policing can be implemented more effectively than conventional patrol methods.

### PENDAHULUAN

Penelitian ini berfokus pada penggunaan Algoritma Machine Learning K-Nearest Neighbor dalam rangka melaksanakan analisa klasifikasi dalam melaksanakan prediksi tindak pidana pencurian di Polresta Bengkulu dengan variabel kelas waktu terjadinya tindak pidana pencurian, usia korban, dan sasaran barang yang dicuri berdasarkan dataset laporan tindak pidana pencurian di Polresta Bengkulu tahun 2020-2024. Tindak pidana merupakan perbuatan yang

dilarang oleh undang-undang dan diancam dengan pidana, baik berupa pidana pokok maupun pidana tambahan (Chairunisa et al., 2021).

Tindak pidana dapat dibedakan menjadi berbagai jenis berdasarkan pelaku, motif, dampak, dan lingkungannya. Jenis-jenis tindak pidana antara lain adalah kejahatan kerah biru, kejahatan kerah putih, kejahatan tanpa korban, kejahatan terorganisir, kejahatan korporasi, kejahatan digital, kejahatan transnasional, dan kejahatan internasional. Jenis-jenis tindak pidana tersebut memiliki karakteristik, modus operandi, dan tantangan tersendiri dalam penegakan hukumnya (Muladi et al, 1992, n.d.). Oleh karena itu, penting bagi penegak hukum untuk mempelajari dan menganalisis jenis-jenis tindak pidana tersebut secara mendalam dan komprehensif. Selain itu, penelitian mengenai jenis-jenis tindak pidana juga dapat memberikan kontribusi bagi pengembangan ilmu hukum pidana, khususnya dalam rangka pembaharuan hukum pidana nasional yang sesuai dengan nilai-nilai Pancasila dan UUD 1945 (Nabilah et al., 2021).

Tindak pidana pencurian merupakan salah satu bentuk kejahatan yang sering terjadi di masyarakat dan menimbulkan kerugian bagi korban maupun masyarakat secara umum. Berdasarkan data Pusat Informasi Kriminal Nasional (Pusiknas) Bareskrim Polri, kasus tindak pidana pencurian menjadi kasus kejahatan tertinggi di Indonesia, totalnya melebihi 50.000 kasus. Pusiknas Bareskrim Polri memberikan catatan 51.312 kasus pencurian dengan pemberatan dan pencurian biasa hingga pertengahan bulan Juli 2024. Jenis kejahatan ini menjadi yang paling masif dibandingkan jenis kejahatan yang lainnya. Secara keseluruhan, hingga pertengahan Juli 2024 sudah ada 280.442 kasus kejahatan yang terjadi jika dihitung jenis tindak pidana pencurian di dalam data dari Pusiknas, Pencurian dengan pemberatan, Pencurian biasa, dan Curanmor memiliki total 60.468 dari jumlah total kasus yang masuk 280.442 atau sebesar 21,5% jadi hampir ¼ dari jumlah total kasus kejahatan yang masuk.

Pada tahun 2023 kasus tindak pidana kejahatan di Bengkulu sepanjang tahun 2023 meningkat sebesar 37,82 persen jika dibandingkan dengan tahun 2022. Tahun sebelumnya tercatat sebanyak 4225 kasus, sementara di tahun 2023 ada sebanyak 5823 kasus atau naik sebesar 1598 kasus. Dari 11 macam jenis kasus gangguan kamtibmas yang meresahkan di masyarakat, kasus pencurian dengan pemberatan (curat) mengalami peningkatan yang sangat tajam jika pada tahun 2022 ada 589 kasus, maka di tahun 2023 tercatat ada 875 kasus atau mengalami peningkatan sebesar 48,56 persen. Kasus jenis pencurian dengan kekerasan (curas) juga mengalami peningkatan dari 75 kasus di tahun 2022 menjadi 103 kasus di tahun 2023 atau mengalami peningkatan sebanyak 37,3 persen (<https://www.rri.co.id/kriminalitas/bengkulu/>, 2024). Meningkatnya tindak pidana pencurian di daerah Bengkulu tentunya menimbulkan dampak yang negatif, terutama dampak bagi korban itu sendiri.

Korban pencurian dapat mengalami kerugian materiil, yaitu kehilangan barang atau harta benda yang dicuri. Selain itu, korban pencurian juga dapat mengalami kerugian imateriel, yaitu mengalami gangguan psikologis, seperti rasa takut, marah, sedih, atau trauma. Korban pencurian juga dapat mengalami penurunan rasa aman dan nyaman dalam beraktivitas, baik di rumah, di tempat kerja, maupun di tempat umum. Beberapa faktor yang dapat menyebabkan peningkatan pencurian di daerah Kota Bengkulu meliputi faktor ekonomi, sosial, dan psikologis. Faktor ekonomi merupakan salah satu pendorong utama dalam tindakan pencurian. Kondisi ekonomi yang sulit, tingginya tingkat pengangguran, pendapatan rendah, dan kesulitan dalam memenuhi kebutuhan mendorong individu untuk mencari cara cepat dan mudah untuk mendapatkan uang, salah satunya dengan melakukan pencurian.

Untuk mencegah dan menanggulangi tindak pidana pencurian, kepolisian melakukan berbagai upaya, salah satunya adalah patroli. Patroli adalah kegiatan rutin yang dilakukan oleh anggota kepolisian untuk mengawasi, mengamankan, dan mengendalikan situasi dan kondisi di wilayah hukumnya). Patroli bertujuan untuk memberikan rasa aman dan nyaman kepada masyarakat, serta mengantisipasi dan menekan terjadinya tindak pidana (Erik Lesmana Putra et al, 2022, n.d.). Untuk meningkatkan efektivitas patroli dalam mencegah pencurian, polisi perlu mengintegrasikan teknologi, pendekatan berbasis data yang prediktif, serta berkolaborasi lebih erat dengan masyarakat.

Patroli yang dilaksanakan oleh kepolisian merupakan salah satu upaya preventif dalam menjaga keamanan dan ketertiban masyarakat. Namun, keefektifitasannya sering kali terhambat karena kurangnya pemanfaatan analisis berbasis big data dalam perencanaan dan pelaksanaan patroli. Berdasarkan penelitian yang dilakukan oleh Cyntia Ayu (2023) yang berjudul “Strategi dan Tantangan Predictive Policing di Era Big Data bagi Masyarakat Modern” diketahui bahwa di negara-negara maju, seperti Spanyol, implementasi Decision Support System (DSS) telah berhasil meningkatkan efisiensi patroli dengan memetakan wilayah berisiko tinggi kejahatan menggunakan data geospasial dan algoritma prediktif.

Sementara itu, di Indonesia, patroli kepolisian masih menggunakan pendekatan konvensional tanpa didukung oleh analisis data yang memadai, sehingga alokasi sumber daya dan waktu sering kali tidak efektif. Ketiadaan integrasi teknologi dalam patroli yang dilakukan oleh kepolisian menyebabkan kurangnya kemampuan untuk melakukan identifikasi pola kejahatan secara akurat, yang berdampak pada peningkatan kasus pencurian di berbagai wilayah. Di dalam era Revolusi Industri 4.0 yang juga ditandai dengan bertumbuhnya tren dibidang Otomasi, Internet of Things (IoT), Big Data, dan Teknologi Cloud Computing. Era Industri 4.0 memberikan peluang bagi pekerjaan data science. Data science adalah bidang yang mempelajari bagaimana data yang terstruktur maupun tidak terstruktur dapat digunakan melalui teknik analisa data sebagai pertimbangan untuk mengambil keputusan.

Banyaknya jumlah data yang dihasilkan oleh manusia dan perangkat terus mendorong perkembangan teknologi yang dapat menjawab kebutuhan untuk mengatur dan memproses informasi yang terus bertambah. Data mining (DM) adalah salah satu bidang yang berkembang pesat karena besarnya kebutuhan akan nilai tambah dari database skala besar yang makin banyak terakumulasi sejalan dengan pertumbuhan teknologi informasi. Definisi umum dari DM itu sendiri adalah serangkaian proses untuk menggali nilai tambah berupa pengetahuan yang selama ini tidak diketahui secara manual dari suatu kumpulan data (Pramudiono, 2003).

Oleh karena itu, diperlukan suatu metode yang dapat membantu kepolisian dalam merencanakan dan melaksanakan patroli yang lebih optimal dan tepat sasaran. Salah satu metode yang dapat digunakan adalah menggunakan data mining, yaitu proses penggalian dan analisis data untuk menemukan pola, hubungan, dan pengetahuan. Data mining dapat diterapkan pada data kejahatan, khususnya data tindak pidana pencurian, untuk mendapatkan informasi yang dapat digunakan sebagai dasar pengambilan keputusan dalam melaksanakan patroli oleh Kepolisian (Data-driven Patrolling).

Salah satu teknik data mining yang dapat digunakan adalah algoritma machine learning. Machine learning adalah cabang dari kecerdasan buatan (AI) yang mempelajari bagaimana membuat sistem komputer yang dapat belajar dari data dan pengalaman, tanpa harus diprogram secara eksplisit (Rahman, n.d.). Machine learning memanfaatkan algoritma - algoritma yang dapat

menemukan pola, hubungan, dan aturan yang tersembunyi dalam data, dan kemudian menggunakannya untuk membuat prediksi, klasifikasi, atau rekomendasi. Algoritma machine learning dapat digunakan untuk mengklasifikasikan jenis tindak pidana pencurian berdasarkan karakteristiknya, serta memprediksi kemungkinan terjadinya tindak pidana pencurian di masa depan berdasarkan faktor-faktor yang mempengaruhinya. Selain itu, algoritma machine learning juga dapat digunakan untuk melakukan analisis klasifikasi yaitu teknik yang dapat menemukan aturan atau hubungan antara item-item dalam data.

Berdasarkan uraian latar belakang di atas, maka rumusan masalah pada penelitian ini adalah:

1. Bagaimana melakukan prediksi tindak pidana pencurian dengan sasaran waktu kejadian, usia korban, dan jenis barang yang dicuri di wilayah hukum Polresta Bengkulu menggunakan algoritma KNN?
2. Apakah tingkat akurasi model prediksi KNN yang dievaluasi dengan uji terhadap data TP Pencurian terbaru dapat menunjukkan keandalan yang tinggi?
3. Apakah model hasil prediksi KNN dapat dimanfaatkan secara efektif oleh Polri dalam melakukan patroli berbasis data?

## **METODE PENELITIAN**

Jenis penelitian yang digunakan adalah penelitian kuantitatif yang memandang fenomena/gejala itu dapat diklasifikasikan, relatif tetap, kongkrit, teramati, terukur, dan hubungan yang bersifat sebab akibat. Penelitian ini menerapkan algoritma K-Nearest Neighbor (KNN) sebagai metode supervised dalam data mining untuk melakukan klasifikasi dan prediksi data kejahatan guna mendukung pengambilan keputusan oleh kepolisian. Proses analisis dilakukan menggunakan tools RapidMiner karena antarmuka yang intuitif serta kemampuannya menangani dataset yang kompleks. Model prediktif dibangun berdasarkan data historis kejahatan untuk menghasilkan wawasan berbasis data yang mendukung strategi kepolisian. Seluruh proses penelitian mengikuti kerangka kerja CRISP-DM sebagai standar dalam implementasi analisis big data untuk kebijakan prediksi kejahatan.

Dalam data mining, sampel dianggap sebagai keseluruhan populasi karena analisis menjadi lebih akurat dengan jumlah data yang besar. Penelitian ini menggunakan sekitar 1.200 data laporan polisi (LP) kasus pencurian di wilayah Polresta Bengkulu selama periode 2020 hingga 2025 sebagai populasi sekaligus dataset utama.

Teori Routine Activity menekankan pentingnya waktu, tempat, dan ketidakhadiran penjaga dalam menciptakan peluang kejahatan, sementara Rational Choice Theory menjelaskan bahwa pelaku kejahatan bertindak rasional dalam memilih target berdasarkan keuntungan dan risiko. Oleh karena itu, peneliti memilih variabel-variabel seperti lokasi kejadian, hari, waktu kejadian, wilayah kejadian (kelurahan dan kecamatan), jenis pencurian sesuai KUHP, jenis kelamin korban, usia korban, pekerjaan korban, serta benda atau objek yang dicuri.

## **HASIL PENELITIAN DAN PEMBAHASAN**

### **Hasil Penelitian**

#### **1. Profil Data Atribut / Variabel dari Tindak Pidana Jenis Pencurian**

Pengambilan data primer dalam penelitian ini difokuskan pada profil data atribut atau variabel dari tindak pidana jenis pencurian, yang diperoleh langsung dari laporan polisi (LP) di

---

Polresta Bengkulu pada kurun waktu 2020–2025. Data tersebut mencakup variabel-variabel penting seperti waktu, hari kejadian, lokasi atau tempat kejadian, jenis tindak pidana pencurian, wilayah kejadian, usia, jenis kelamin korban, pekerjaan korban, serta objek atau barang yang menjadi target pencurian. Pemilihan variabel ini didasarkan pada teori *Routine Activity* dan *Rational Choice*. Dengan mengolah data primer ini, peneliti dapat membangun model klasifikasi dan prediksi yang lebih akurat untuk mendukung strategi pencegahan dan penanggulangan kejahatan secara tepat sasaran.

No	Jenis Pencurian	Jumlah
1.	Curas	26
2.	Curat	723
3.	Curbis	453
No	Kecamatan	Jumlah
1.	Gading Cempaka	216
2.	Kampung Melayu	12
3.	Muara Bangkahulu	76
4.	Ratu Agung	387
5.	Ratu Samban	115
6.	Selebar	67
7.	Singaran Pati	177
8.	Sungai Serut	40
9.	Teluk Segara	112
Total		1.202

No	Hari Terjadinya Pencurian	Jumlah
1.	Minggu	173
2.	Senin	169
3.	Selasa	187
4.	Rabu	158
5.	Kamis	161
6.	Jumat	197
7.	Sabtu	157
Total		1.202

No	Tempat Kejadian	Jumlah
1.	Café	28
2.	Jalan Raya	231
3.	Kantor	87
4.	Kost	130
5.	Pasar	27
6.	Rumah	489
7.	Sekolah	28
8.	Tempat Ibadah	26
9.	Tempat Umum	44
10.	Toko	112
Total		1.202

No	Waktu Pencurian	Jumlah
1.	Malam Awal	149
2.	Malam Larut	161
3.	Menjelang Siang	131
4.	Menjelang Subuh	104
5.	Pagi	96
6.	Siang	132
7.	Sore	129
8.	Tengah Malam	300
Total		1.202

No	Usia Korban Pencurian	Jumlah
1.	Masa Dewasa Akhir	354
2.	Masa Dewasa Awal	298
3.	Masa Lansia Akhir	35
4.	Masa Lansia Awal	89
5.	Masa Manula	6
6.	Masa Remaja Akhir	411
7.	Masa Remaja Awal	9
Total		1.202

No	Jenis Kelamin Korban Pencurian	Jumlah
1.	Laki-laki	732
2.	Perempuan	470
Total		1.202

No	Jenis Barang Yang Dicuri	Jumlah
1.	Alat Bekerja	29
2.	Hewan Ternak	13
3.	HP	389
4.	Kotak Amal Tempat Ibadah	9
5.	Laptop	37
6.	Motor	542
7.	Peralatan Rumah Tangga	83
8.	Perhiasan	44
9.	Uang	56
Total		1.202

No	Pekerjaan Korban Pencurian	Jumlah
1.	BUMN	6
2.	Buruh	52
3.	Pedagang	20
4.	Dokter	6
5.	Dosen	5
6.	Guru	13
7.	Hakim	1
8.	Honoror	37
9.	IRT	131
10.	Mahasiswa	283
11.	Nelayan	2
12.	Pengacara	37
13.	PNS / Polri / TNI	102
14.	Satpam	17
15.	Sopir	2
16.	Petani	307
17.	Wiraswasta	14
18.	Tidak Bekerja	169
Total		1.202

**Tabel 1 – Data Deskriptif**

## 2. Data Untuk Mengevaluasi Model Prediksi KNN

Evaluasi model prediksi K-Nearest Neighbor (KNN) dalam penelitian ini dilakukan menggunakan data baru yang tidak termasuk dalam dataset pelatihan, yaitu laporan tindak pidana pencurian bulan Januari 2025 dari Sat Reskrim Polresta Bengkulu. Fokus evaluasi ini adalah untuk menguji keandalan model dalam memprediksi pola kejahatan secara akurat di luar data latih. Data laporan polisi yang digunakan memiliki atribut yang sama dengan dataset pelatihan dan dibandingkan dengan hasil prediksi model untuk mengukur akurasi pada tiga kelas target: waktu kejadian, usia korban, dan jenis barang yang dicuri.

NO	NO LAPORAN POLISI	TANGGAL LP	KECAMATAN	USIA KORBAN	WAKTU	PEKERJAAN KORBAN	BARANG YANG DICURI
1.	LP/B-2/I/2025/SPKT/RES.BKL/POLDA.BKL	01/01/2025	Ratu Samban	29	02:00	Swasta	Motor
2.	LP/B-7/I/2025/SPKT/RES.BKL/POLDA.BKL	03/01/2025	Ratu Agung	31	01:25	Swasta	Motor
3.	LP/B-8/I/2025/SPKT/RES.BKL/POLDA.BKL	03/01/2025	Sungai Serut	35	17:00	IRT	Motor
4.	LP/B-17/I/2025/SPKT/RES.BKL/POLDA.BKL	09/01/2025	Ratu Samban	39	11:20	WiraSwasta	HP
5.	LP/B-18/I/2025/SPKT/RES.BKL/POLDA.BKL	10/01/2025	Sungai Serut	23	15:00	Mahasiswa	Motor
6.	LP/B-27/I/2025/SPKT/RES.BKL/POLDA.BKL	15/01/2025	Muara Bangkahulu	26	16:19	Honoror	HP
7.	LP/B-28/I/2025/SPKT/RES.BKL/POLDA.BKL	15/01/2025	Muara Bangkahulu	37	21:00	Swasta	Alat Bekerja
8.	LP/B-32/I/2025/SPKT/RES.BKL/POLDA.BKL	19/01/2025	Ratu Samban	30	23:00	WiraSwasta	Uang
9.	LP/B-34/I/2025/SPKT/RES.BKL/POLDA.BKL	21/01/2025	Gading Cempaka	21	14:00	Mahasiswa	Laptop
10.	LP/B-45/I/2025/SPKT/RES.BKL/POLDA.BKL	28/01/2025	Selebar	40	13:31	IRT	Motor
11.	LP/B-47/I/2025/SPKT/RES.BKL/POLDA.BKL	29/01/2025	Muara Bangkahulu	34	11:40	PNS	HP
12.	LP/B-51/I/2025/SPKT/RES.BKL/POLDA.BKL	31/01/2025	Selebar	40	13:20	IRT	HP

Tabel 2 – Laporan Polisi

### 3. Data Skema Patroli di Polresta Bengkulu

Untuk memahami sistem patroli di Polresta Bengkulu, peneliti menelusuri skema patroli yang melibatkan personel Reskrim dan Samapta, khususnya dalam kegiatan KRYD (Kegiatan Rutin yang Ditingkatkan). Meskipun KRYD bertujuan meningkatkan kehadiran polisi di lapangan guna menekan angka kriminalitas, hasil temuan menunjukkan bahwa pelaksanaannya masih bersifat reaktif dan belum berbasis data. Perencanaan patroli belum mempertimbangkan secara terstruktur faktor-faktor prediktif seperti waktu kejadian, karakteristik korban, dan jenis barang yang menjadi sasaran, sehingga efektivitas pencegahan kejahatan belum optimal.

KEPOLISIAN NEGARA REPUBLIK INDONESIA  
 DAERAH BENGKULU  
 RESORTA BENGKULU

**JADWAL PIKET PATROLI SAT SAMAPTA  
 BULAN SEPTEMBER TAHUN 2024**

NO	NAMA	PANGKAT	NRP	TANGGAL																														KET
				1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	
<b>REGU IA</b>																																		
1	IRENDY ADITYA	BRIPKA	8/051220	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	
2	M RIZKI	BRIPDA	02091160	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	
3	WAHYU ALPADIR	BRIPDA	03011136	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	
4	PENDO MARCELINO	BRIPDA	02041145	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	
<b>REGU IB</b>																																		
1	HERI SUSANTO	BRIPKA	81060915	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	
2	HANDRI SETIAWAN	BRIPDA	93120988	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	
3	FEDRI REPALDIANSYAH	BRIPDA	05040187	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	
4	AL AZIZ ROBI PUTRA	BRIPDA	01111357	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	P	M	L	

KEPOLISIAN NEGARA REPUBLIK INDONESIA  
 DAERAH BENGKULU  
 RESOR KOTA BENGKULU

LAMPIRAN SURAT PERINTAH KAPOLRESTA  
 NOMOR : SPRIN/95a/IX/OPS.1.3/2024  
 TANGGAL : 27 SEPTEMBER 2024

**JADWAL KEGIATAN KRYD POLRESTA BENGKULU BULAN SEPTEMBER S/D OKTOBER 2024**

NO	UKL	TANGGAL 28 SEPTEMBER S/D 2 NOVEMBER 2024																														KET					
		28	29	30	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27		28	29	30	1	2
1	I	X																																			
2	II						X																														
3	III												X																								
4	IV																							X													
5	V																													X							
6	VI																																				X

Tabel 3 – Jadwal Patroli

Pembahasan

**1. Bagaimana melakukan prediksi tindak pidana pencurian dengan sasaran waktu kejadian, usia korban, dan jenis barang yang dicuri di wilayah hukum Polresta Bengkulu menggunakan algoritma KNN ?**

Penelitian ini menggunakan algoritma K-Nearest Neighbor (KNN) untuk membangun model klasifikasi prediksi tindak pidana pencurian berdasarkan tiga kelas utama: waktu kejadian, usia korban, dan jenis barang yang dicuri. KNN dipilih karena mampu mengklasifikasikan data nominal berdasarkan kesamaan karakteristik historis. Pengembangan model mengikuti kerangka kerja CRISP-DM (Cross-Industry Standard Process for Data Mining) yang mencakup enam tahapan: pemahaman bisnis, pemahaman data, persiapan data, pemodelan, evaluasi, dan implementasi, guna memastikan hasil prediksi yang akurat dan aplikatif bagi sistem patroli kepolisian.

**A. Tahap Business Understanding.**

Dalam penelitian ini bertujuan untuk mengatasi permasalahan patroli kepolisian yang masih bersifat reaktif dalam mencegah tindak pidana pencurian di Polresta Bengkulu. Dengan menggunakan algoritma K-Nearest Neighbor (KNN), penelitian ini berupaya memprediksi pola kejahatan berdasarkan waktu kejadian, usia korban, dan jenis barang yang dicuri. Pendekatan berbasis data ini diharapkan dapat membantu kepolisian dalam menyusun jadwal patroli yang lebih efektif dan proaktif, sehingga alokasi sumber daya menjadi lebih terarah dan mampu menekan angka kejahatan.

**B. Tahap Data Understanding.**

Bertujuan untuk memahami struktur dan kualitas data yang digunakan dalam membangun model prediksi tindak pidana pencurian. Data diperoleh dari laporan polisi di Satreskrim Polresta Bengkulu tahun 2020–2024, sebanyak 1.202 data, mencakup variabel seperti waktu kejadian, usia korban, jenis barang yang dicuri, dan lokasi kejadian. Dataset ini disusun berdasarkan teori Routine Activity dan Rational Choice, yang menjelaskan bahwa kejahatan dipengaruhi oleh pertemuan antara pelaku, target, dan ketiadaan penjaga, serta pertimbangan rasional pelaku dalam memilih target berdasarkan risiko dan keuntungan.

**C. Tahap Data Preparation**

Dilakukan untuk mempersiapkan dataset agar siap digunakan dalam pemodelan prediksi dengan algoritma KNN. Proses ini mencakup pembersihan data dari duplikasi dan missing values, pengecekan inkonsistensi, serta transformasi variabel numerik menjadi kategorikal, seperti pengelompokan waktu kejadian. Normalisasi juga dilakukan untuk memastikan konsistensi data. Hasil dari tahap ini adalah dataset yang lebih terstruktur dan siap menghasilkan model prediktif yang akurat dan andal.

**D. Tahap Modeling.**

Bertujuan membangun model prediksi tindak pidana pencurian menggunakan algoritma K-Nearest Neighbor (KNN) untuk tiga variabel utama: waktu kejadian, usia korban, dan jenis barang

---

yang dicuri. Proses dimulai dengan menentukan nilai K terbaik guna memperoleh akurasi optimal, lalu dataset dibagi menjadi data latih dan data uji menggunakan teknik split data.

### **E. Tahap Evaluation.**

Bertujuan menilai kinerja model KNN dalam memprediksi waktu kejadian, usia korban, dan jenis barang yang dicuri menggunakan data laporan pencurian bulan Januari 2025 yang tidak termasuk dalam data latih. Hasil evaluasi digunakan untuk menentukan keandalan model dan apakah diperlukan optimasi lebih lanjut sebelum diterapkan dalam strategi patroli berbasis data.

### **F. Tahap Deployment**

Bertujuan mengimplementasikan hasil prediksi model KNN ke dalam operasional kepolisian, khususnya untuk menyusun jadwal patroli berbasis data di Polresta Bengkulu. Prediksi waktu kejadian, usia korban, dan jenis barang yang dicuri digunakan untuk menentukan lokasi dan waktu patroli yang lebih efektif. Model yang telah divalidasi ini membantu memprioritaskan daerah rawan dan mengoptimalkan alokasi sumber daya guna meningkatkan pencegahan kejahatan.

## **Membangun Model Prediksi KNN Waktu Kejadian Pencurian, Usia Korban Pencurian, dan Jenis Barang yang dicuri**

Pada tahap ini, peneliti membangun model prediksi untuk waktu kejadian pencurian, Usia Korban Pencurian, dan Jenis Barang yang dicuri menggunakan algoritma K-Nearest Neighbor (KNN) dengan memanfaatkan tools RapidMiner sebagai platform analisis data. KNN dipilih karena data yang dimiliki semuanya akan berbentuk nominal sehingga peneliti menggunakan KNN untuk melakukan klasifikasi yang dimanfaatkan untuk prediksi.

Langkah pertama dalam membangun model adalah menentukan atribut yang paling berpengaruh/melakukan *feature selection* dalam memprediksi waktu kejadian pencurian. Untuk itu, penulis menggunakan metode *Weight by Information Gain*, yang berfungsi untuk menganalisis dan memberikan bobot pada setiap atribut dalam dataset. Metode Information Gain digunakan melalui operator Weight by Information Gain untuk menentukan atribut yang paling berpengaruh dalam klasifikasi dengan menghitung selisih entropy sebelum dan sesudah pemisahan atribut (Much. Rifqi Maulana & M. Adib Al Karomi, 2015).

Hasil analisis ini digunakan untuk mengidentifikasi atribut spesial yang memiliki dampak signifikan terhadap variabel target, dan membagi kelas sebagai berikut :

1. Waktu yang sebelumnya ada 8 bagian waktu menjadi 5 bagian waktu yaitu : Pagi, Malam, Siang, Sore, dan Dini Hari sehingga model KNN dapat melakukan klasifikasi dengan lebih akurat. Dengan pembagian Kelas waktu adalah sebagai berikut Pagi (04.00–10.59), Siang (11.00–13.59), Sore (14.01–17.59), Malam (18.00–00.00), dan Dini Hari (00.01–03.59).
2. Usia korban yang sebelumnya ada 7 bagian menjadi 3 bagian yaitu : Remaja (dibawah 24 tahun), Dewasa (25 – 59 tahun) , dan Lansia (diatas 60 tahun)
3. Barang curian menjadi 4 bagian yaitu : Kendaraan Bermotor (Motor/Mobil), Barang Bernilai Tinggi (Perhiasan, Uang, Dokumen Berharga), Barang Elektronik (HP, Laptop, Tablet, Gadget), Barang Pendukung dan Peralatan Rumah Tangga sehingga (Alat yang mendukung untuk melakukan pekerjaan, Hewan Ternak, Peralatan-peralatan rumah tangga)

Setelah mendapatkan atribut spesial lalu Penulis menggunakan operator *Optimize Parameters (Grid Search)* untuk menemukan nilai K yang menghasilkan akurasi tertinggi dan

Untuk memastikan model memiliki generalisasi yang baik, dataset dibagi menjadi dua bagian dengan rasio 70:30, dimana 70% data digunakan sebagai training set untuk melatih model, sedangkan 30% data digunakan sebagai testing set untuk mengukur kinerja prediksi model. Setelah pemilihan spesial atribut, menentukan nilai K terbaik dan pemilihan rasio sebesar 70:30 selanjutnya peneliti melakukan running model menggunakan algoritma KNN.

No	Model KNN	Spesial Atribut	Nilai K
1.	Prediksi Waktu	Hari, Jenis Pencurian, Kelurahan, Klasifikasi Pekerjaan, Tempat	1
2.	Prediksi Usia Korban	Klasifikasi Barang, Klasifikasi Pekerjaan, Kelurahan, Tempat	16
3.	Prediksi Jenis Barang Curian	Hari, Kelurahan, Klasifikasi Waktu, Tempat	1

Hasil Running Model KNN						
<b>Akurasi Model KNN Prediksi Waktu</b>						
accuracy: 87.35%						
	true Malam	true Siang	true Sore	true Pagi	true Dini_Hari	class precision
pred. Malam	47	9	6	3	4	68.12%
pred. Siang	1	144	6	2	0	94.12%
pred. Sore	2	2	122	1	0	96.06%
pred. Pagi	1	6	6	77	1	84.62%
pred. Dini_Hari	4	6	2	3	59	79.73%
class recall	85.45%	86.23%	85.92%	89.53%	92.19%	
classification_error: 12.65%						
kappa: 0.836						
weighted_mean_recall: 87.86%, weights: 1, 1, 1, 1, 1						
weighted_mean_precision: 84.53%, weights: 1, 1, 1, 1, 1						
<b>Akurasi Model KNN Prediksi Usia Korban</b>						
accuracy: 82.41%						
	true Dewasa	true Lansia	true Remaja	class precision		
pred. Dewasa	201	18	52	74.17%		
pred. Lansia	15	216	9	90.00%		
pred. Remaja	19	1	117	85.40%		
class recall	85.53%	91.91%	65.73%			
classification_error: 17.59%						
kappa: 0.732						
weighted_mean_recall: 81.06%, weights: 1, 1, 1						
weighted_mean_precision: 83.19%, weights: 1, 1, 1						
<b>Akurasi Model KNN Prediksi Jenis Barang Curian</b>						

	true Kendaraan B...	true Barang Elektr...	true Barang Bernil...	true Barang Pend...	class precision
pred. Kendaraan ...	93	6	13	17	72.09%
pred. Barang Elek...	1	63	2	26	68.48%
pred. Barang Bern...	2	1	84	5	91.30%
pred. Barang Pen...	4	0	1	356	98.61%
class recall	93.00%	90.00%	84.00%	88.12%	

accuracy: 88.43%

classification\_error: 11.57%

kappa: 0.812

weighted\_mean\_recall: 88.78%, weights: 1, 1, 1, 1

weighted\_mean\_precision: 82.62%, weights: 1, 1, 1, 1

No	Evaluasi Model KNN	Akurasi Model	Jumlah Data Testing	Hasil Benar	Hasil Salah	Hasil Akurasi Evaluasi Model
1.	Prediksi Waktu	87.35%	12	11	1	91,67%
2.	Prediksi Usia Korban	82,41%	12	12	0	100%
3.	Prediksi Jenis Barang Curian	88,43%	12	11	1	91,67%

**Tabel 4 - Hasil Running KNN**

## 2. Apakah model hasil prediksi KNN dapat dimanfaatkan secara efektif oleh Polri dalam melakukan patroli berbasis data ?

Penelitian ini telah memasuki tahap Deployment dalam kerangka CRISP-DM, di mana model prediksi K-Nearest Neighbor (KNN) yang telah terbukti handal diterapkan untuk mendukung strategi patroli berbasis data. Berdasarkan tiga sasaran utama—waktu kejadian, usia korban, dan jenis barang yang dicuri—penelitian ini menyusun prototype jadwal patroli yang lebih optimal dan ilmiah guna meningkatkan efektivitas pencegahan pencurian. Temuan menunjukkan bahwa jadwal patroli di Polresta Bengkulu masih bersifat konvensional dan belum berbasis prediksi. Oleh karena itu, implementasi jadwal patroli berbasis KNN diharapkan dapat meningkatkan efisiensi alokasi sumber daya dan efektivitas patroli melalui pendekatan data-driven policing yang lebih terukur dan ilmiah.

### Penyusunan Prototype Jadwal Patroli dengan menggunakan hasil Prediksi model KNN

Setiap model KNN memiliki atribut khusus yang disesuaikan dengan target prediksinya. Untuk model prediksi waktu, atribut seperti hari, jenis pencurian, kelurahan, klasifikasi pekerjaan, dan tempat digunakan untuk menentukan klasifikasi waktu kejadian pencurian (pagi, siang, sore, malam, dini hari). Model prediksi usia korban menggunakan atribut klasifikasi barang, klasifikasi pekerjaan, dan tempat untuk memperkirakan usia korban (remaja, dewasa, lansia). Sementara itu, model prediksi jenis barang curian menggunakan atribut hari, kelurahan, klasifikasi waktu, dan tempat untuk memprediksi jenis barang yang sering menjadi target kejahatan.



**Gambar 1 - Kerangka Penyusunan Prototype Jadwal Patroli dengan Model Prediksi KNN**

Atribut tempat dan kelurahan dipilih sebagai penghubung antara ketiga model prediksi, selain juga dikarenakan memiliki weight score yang tinggi dalam mempengaruhi hasil prediksi. Setelah itu proses Deployment dilakukan peneliti dengan cara memasukkan data testing dengan predictor yang sudah disiapkan sesuai dengan kelas atribut masing-masing. Lalu dimulai dari proses deployment klasifikasi waktu yang menghasilkan kelas (Klasifikasi Waktu) digunakan sebagai input predictor dalam proses deployment klasifikasi jenis barang curian yang menghasilkan kelas (Klasifikasi Barang), dan hasil dari proses deployment jenis barang curian digunakan sebagai input predictor proses deployment Usia Korban yang menghasilkan kelas (Klasifikasi Usia) sehingga hasil prediksi tersebut terhubung satu sama lain sesuai dengan predictor yang dimasukkan. Dengan demikian, prototype jadwal patroli yang dihasilkan dari proses deployment model KNN tersebut akan lebih efektif karena didasarkan pada pola kejahatan yang terprediksi secara ilmiah, memungkinkan kepolisian untuk mengalokasikan sumber daya secara efisien, meningkatkan kewaspadaan di area rawan, dan mencegah tindak pidana pencurian dengan lebih tepat sasaran.

Contoh Proses Deployment Model Prediksi KNN									
Hasil Deployment KNN Prediksi Waktu									
HARI	KECAMATAN	KELURAHAN	TEMPAT	prediction(KLA...	confidence(...	confidence(...	confidence(...	confidence(...	confidence...
senin	Ratu_Agung	Sawah_Lebar	Cafe	Dini_Hari	0	0	0	0	1
senin	Ratu_Agung	Sawah_Lebar	Jalan_Raya	Siang	0	1	0	0	0
senin	Ratu_Agung	Sawah_Lebar	Kantor	Siang	0	1	0	0	0
senin	Ratu_Agung	Sawah_Lebar	Kost	Dini_Hari	0	0	0	0	1
senin	Ratu_Agung	Sawah_Lebar	Pasar	Malam	1	0	0	0	0
senin	Ratu_Agung	Sawah_Lebar	Rumah	Malam	1	0	0	0	0
senin	Ratu_Agung	Sawah_Lebar	Sekolah	Siang	0	1	0	0	0
senin	Ratu_Agung	Sawah_Lebar	Tempat_Ibadah	Pagi	0	0	0	1	0
senin	Ratu_Agung	Sawah_Lebar	Tempat_Umum	Malam	1	0	0	0	0
senin	Ratu_Agung	Sawah_Lebar	Toko	Malam	1	0	0	0	0

  

Hasil Deployment KNN Prediksi Usia Korban				
TEMPAT ↑	prediction(KLASIFIKASI_USIA(3))	confidence(Dewasa)	confidence(Lansia)	confidence(Remaja)
Cafe	Dewasa	0.733	0.167	0.100
Cafe	Dewasa	0.800	0	0.200

  

Hasil Deployment KNN Prediksi Jenis Barang Curian				
HARI	TEMPAT	KELURAHAN	KLASIFIKASI_WAKTU(5) ↑	prediction(KLASIFIKASI_BARANG)
senin	Cafe	Sawah_Lebar	Dini_Hari	Kendaraan Bermotor
senin	Cafe	Sawah_Lebar	Dini_Hari	Kendaraan Bermotor
senin	Cafe	Sawah_Lebar	Dini_Hari	Kendaraan Bermotor

## Gambar 2- Hasil Deployment KNN Prediksi Waktu

Gambar di atas adalah proses deployment model prediksi yang telah melalui serangkaian proses evaluasi untuk memastikan keandalannya. Dari ketiga model tersebut, penulis akan menyusun contoh skema prototipe jadwal patroli harian yang dilaksanakan di kecamatan dan kelurahan tertentu. Skema jadwal ini mengintegrasikan hasil prediksi terkait waktu kejadian, usia korban, dan jenis barang yang dicuri, sehingga memungkinkan kepolisian mengalokasikan personel dan sumber daya secara optimal. Dengan demikian, jadwal patroli berbasis data-driven ini diharapkan dapat meningkatkan efektivitas pengawasan dan pencegahan tindak pidana pencurian di wilayah yang rawan kejahatan.

KEPOLISIAN NEGARA REPUBLIK INDONESIA  
 DAERAH BENGKULU  
 RESOR KOTA BENGKULU

**Jadwal Patroli Harian  
 Polres Kota Bengkulu**

Hari : Senin, 24 Februari 2025  
 Kecamatan : Ratu Agung  
 Kelurahan : Sawah Lebar

Tempat :

	Sasaran Waktu	Sasaran Barang	Sasaran Orang (Korban)
Cafe	Dini_Hari	Kendaraan Bermotor	Dewasa
Kost	Dini_Hari	Kendaraan Bermotor	Dewasa
Pasar	Malam	Kendaraan Bermotor	Dewasa
Rumah	Malam	Kendaraan Bermotor	Dewasa
Tempat Umum	Malam	Kendaraan Bermotor	Dewasa
Toko	Malam	Barang Pendukung dan Peralatan Rumah Tangga	Dewasa
Tempat Ibadah	Pagi	Kendaraan Bermotor	Dewasa
Jalan Raya	Siang	Kendaraan Bermotor	Dewasa
Kantor	Siang	Barang Bermilai Tinggi	Dewasa
Sekolah	Siang	Kendaraan Bermotor	Remaja

Keterangan :

- Waktu :**
  - Pagi (04.00–10.59)
  - Siang (11.00–13.59)
  - Sore (14.01–17.59)
  - Malam (18.00–00.00)
  - Dini Hari (00.01–03.59)
- Barang :**
  - Kendaraan Bermotor (Motor/Mobil).
  - Barang Bermilai Tinggi (Perhiasan, Uang, Dokumen Berharga).
  - Barang Elektronik (HP, Laptop, Tablet, Gadget).
  - Barang Pendukung dan Peralatan Rumah Tangga sehingga (Alat yang mendukung untuk melakukan pekerjaan, Hewan Ternak, Peralatan-peralatan rumah tangga)
- Orang :**
  - Remaja (dibawah 24 tahun)
  - Dewasa (25 – 59 tahun)
  - Lansia (diatas 60 tahun)

**Tabel 5 – Contoh Prototype Jadwal Patroli Harian dari Proses Deployment Model Prediksi KNN**

### KESIMPULAN

- Berdasarkan hasil penelitian, model prediksi dengan algoritma K-Nearest Neighbor (KNN) yang dipakai dalam penelitian ini untuk melakukan klasifikasi yang digunakan untuk prediksi jenis tindak pidana pencurian, terbukti dapat mengidentifikasi pola kejadian tindak pidana pencurian berdasarkan waktu kejadian, usia korban, dan jenis barang yang dicuri di wilayah hukum Polresta Bengkulu. Penarikan suatu informasi dari sejumlah dataset yang dikumpulkan mengenai jenis pidana pencurian menghasilkan suatu hubungan antar atributnya yang dapat di manfaatkan untuk melakukan prediksi dengan kerangka kerja CRISP-DM (Cross Industry Standard Process for Data Mining). Dari hasil klasifikasi, model prediksi waktu kejadian pencurian memiliki akurasi sebesar 87,35%, prediksi usia korban 82,41%, dan prediksi jenis barang curian 88,43%, menunjukkan tingkat akurasi yang cukup tinggi. Model ini dapat memberikan wawasan berbasis data untuk memahami pola kejahatan dan dapat digunakan sebagai dasar dalam perencanaan strategi pencegahan tindak pidana pencurian.

2. Evaluasi model prediksi KNN dilakukan dengan menggunakan data laporan polisi terbaru pada Januari 2025 (diluar dataset latih) untuk mengukur keandalan model dalam kondisi riil. Hasil evaluasi menunjukkan bahwa model prediksi waktu pencurian mengalami peningkatan akurasi dari 87,35% menjadi 91,67%, model prediksi usia korban mencapai akurasi sempurna 100%, dan model prediksi jenis barang yang dicuri meningkat dari 88,43% menjadi 91,67%. Konsistensi antara hasil prediksi awal dan hasil evaluasi menunjukkan bahwa model tidak mengalami overfitting dan tetap mampu menggeneralisasi pola kejahatan dengan baik. Hasil ini membuktikan bahwa model KNN dapat dikategorikan sebagai model yang robust (tangguh) dalam melakukan prediksi kejahatan, dengan adanya konsistensi antara akurasi model dan akurasi evaluasi mengindikasikan bahwa kejahatan yang terjadi membentuk suatu pola dan terkonsentrasi pada data-data tertentu sehingga dapat di manfaatkan dalam menarik suatu kebijakan dalam strategi pencegahan pencurian, serta model tersebut dapat diandalkan untuk mendukung strategi preventif kepolisian dalam mencegah tindak pidana pencurian.
3. Hasil dari akurasi model dan analisa evaluasi model menunjukkan bahwa model prediksi KNN dapat dimanfaatkan oleh Polri dalam mendukung patroli berbasis data (data-driven policing). Prototype Jadwal patroli yang dihasilkan dari model prediksi yang dibuat dalam tahap atau proses deployment memberikan fokus pada sasaran waktu kejadian, kelompok usia korban, dan jenis barang curian yang paling rentan terhadap pencurian sehingga pelaksanaan patroli dapat berjalan lebih efisien. Dibandingkan dengan skema patroli konvensional, yang masih dilakukan secara manual dan kurang mempertimbangkan pola kejahatan berbasis data, patroli berbasis prediksi dengan model KNN memungkinkan strategi yang lebih tepat sasaran dan berbasis ilmiah. Dengan hasil prediksi yang akurat, Polri dapat melakukan pencegahan lebih efektif, seperti menyesuaikan intensitas patroli pada waktu dan lokasi dengan risiko tinggi serta memberikan sosialisasi keamanan, seperti himbauan pemasangan kunci ganda untuk kendaraan di area rawan curanmor. Berdasarkan hasil penelitian ini, penerapan machine learning dalam predictive policing / pemolisian prediktif dapat dilaksanakan dengan lebih efektif dibandingkan metode patroli konvensional. Oleh karena itu, model prediksi KNN yang dikembangkan dalam penelitian ini direkomendasikan untuk diterapkan dalam sistem kepolisian terutama di Polresta Bengkulu guna meningkatkan efektivitas pengawasan dan pencegahan kejahatan di wilayah hukum Kota Bengkulu.

## DAFTAR PUSTAKA

- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). "Crisp-Dm 1.0," *CRISP-DM Consortium*, 76. <https://doi.org/10.1109/ICETET.2008.239>
- Chazawi, Adami (2002). *Pelajaran Hukum Pidana*. Jakarta: PT Raja Grafindo Persada.
- Cornish, D. B., & Clarke, R. V. (1986). *Reasoning Criminal - Rational Choice Perspectives on Offending*. United States: Springer.
- Cynthia Ayu Windani (2023). "Strategi dan Tantangan Predictive Policing di Era Big Data bagi Masyarakat Modern," *Deviance Jurnal Kriminologi* Volume 7 Nomor 2 Desember 2023 Hal: 101-120 DOI: <http://dx.doi.org/10.36080/djk.2385>
- Daniel T. Larose and Chantal D. Larose (2014). *Discovering Knowledge in Data: An Introduction in Data mining Second Edition*. New Jersey: Wiley
- Dirjosisworo (1984). *Ruang Lingkup Kriminalogi*. Jakarta: Rajawali Press.
- Efraim Turban, Dkk (2005). *Decision Support Systems and Intelligent Systems*. Yogyakarta: Andi.

- G. Sivapriya, B. Vijay Ganesh, U.G. Pradeeshwar, Vishnu Dharshini, Muhammad Al-Amin (2023) "Crime Prediction and Analysis Using Data Mining," *Transactions on Sustainable Computer Letters and Machine Learning: A Simple Approach that Helps Predictive Policing*
- Ganesh, M. S. S., (et. al) (2023). "Crime Prediction Using Machine Learning Algorithms," *Advances in Science and Technology* journal, 124, 457-461.
- Hadmanto, Aditya, Jarot Prianggono (2024) "Prediction of Theft with Machine Learning Technology at Police Station. Building of Informatics," *Technology and Science (BITS)* Volume 6, No 1, June 2024 Page: 550-558 ISSN 2684-8910 (media cetak) ISSN 2685-3310 (media online) DOI 10.47065/bits.v6i1.5107
- Han, J. dan M. Kamber (2006). *Data mining Concepts and Techniques Second Edition*. San Francisco: Morgan Kaufmann.
- Hermawan, Febry dan Jarot Prianggono (2023). "Crime of theft prediction using Machine Learning KNearest Neighbour Algorithm at Polresta Bengkulu," *Jurnal Teknik Informatika Sinkron*. Vol 7 no 3 (2023). DOI: 10.33395/sinkron. v8i3.12422.
- Hoetomo, M. A (2005). *Kamus Lengkap Bahasa Indonesia*. Surabaya: Mitra Belajar.
- Koentjaraningrat (1991). *Metode-metode Penelitian Masyarakat*. Jakarta: Gramedia Pustaka Umum
- Larose, D. T (2005). *An introduction to data mining. Traduction et adaptation de Thierry Vallaud*. New Jersey: John Wiley & Sons Inc.
- Moeljatno (2008). *Kitab Undang-Undang Hukum Pidana*. Jakarta: Bumi Aksara.
- N. Naga Swathi, Sarah Vineela Cole, Polani Nikhil Manikanta, Shaik Ashmiya Jafreen, Shaik Salma, Talari Manoj (2022). "Crime Prediction Using K-Nearest Neighboring Algorithm. International Research Journal of Modernization in Engineering Technology and Science," Peer-Reviewed, Open Access, Fully Refereed International Journal, Volume:04/Issue:04/April-2022 Impact Factor- 6.752.
- R. Soesilo (1994). *Kitab Undang-Undang Hukum Pidana (KUHP), Serta Komentar-Komentar Lengkap Pasal Demi Pasal*. Bogor: Politea.
- Vijay, Kotu and Bala Deshpande (2019). *Data Science: Concepts and Practice Second Edition*. Cambridge: Morgan Kaufmann