

**JEMSI:**
Jurnal Ekonomi Manajemen Sistem
Informasi

E-ISSN: 2686-5238
P-ISSN: 2686-4916

<https://dinastirev.org/JEMSI> dinasti.info@gmail.com +62 811 7404 455

DOI: <https://doi.org/10.38035/jemsi.v6i5>
<https://creativecommons.org/licenses/by/4.0/>

Analisis Klasifikasi Tindak Kejahatan Pencurian dengan Algoritma *K-Nearest Neighbor* dan *Naive Bayes* di Polres Buol

Nandito Marbun¹, Jarot Prianggono²

¹Sekolah Tinggi Ilmu Kepolisian PTIK, Jakarta, Indonesia, dito.marbun1805@gmail.com

²Sekolah Tinggi Ilmu Kepolisian PTIK, Jakarta, Indonesia, komputerstik@gmail.com

Corresponding Author: dito.marbun1805@gmail.com¹

Abstract: *One effective data mining technique is the use of machine learning algorithms. Machine learning, a branch of artificial intelligence (AI), enables computers to learn from data and experience without being explicitly programmed. This study explores the potential of predictive systems to assist the Indonesian National Police in addressing and preventing criminal acts, particularly theft. The research compares two algorithms—k-Nearest Neighbors (K-NN) and Naive Bayes—in predicting theft crimes in the jurisdiction of Buol Police Department. K-NN operates based on data proximity, while Naive Bayes uses a probabilistic approach based on Bayes' theorem. The results show that K-NN outperforms Naive Bayes in classifying theft cases based on location, time, and type of incident. K-NN consistently achieves higher values in accuracy, precision, recall, and F1-score. The prediction analysis also reveals that most thefts occur in the early morning hours, likely due to limited surveillance or the absence of property owners. These findings suggest the need to enhance patrols during vulnerable times, both at night and during the day in densely populated areas. Random and unpredictable night patrols are considered effective in deterring criminal activity by creating a sense of being monitored among potential offenders.*

Keyword: *K-NN Algorithm, Naïve Bayes Algorithm, Theft*

Abstrak: Salah satu teknik data mining yang efektif adalah penggunaan algoritma machine learning. Machine learning, sebagai cabang dari kecerdasan buatan (AI), memungkinkan komputer belajar dari data dan pengalaman tanpa pemrograman eksplisit. Penelitian ini tertarik mengeksplorasi potensi sistem prediksi untuk membantu kepolisian RI dalam menangani dan mencegah tindak kriminal, khususnya pencurian. Studi dilakukan dengan membandingkan dua algoritma, yaitu k-Nearest Neighbors (K-NN) dan Naive Bayes, dalam memprediksi tindak pidana pencurian di wilayah Polres Buol. K-NN bekerja berdasarkan kedekatan data, sementara Naive Bayes menggunakan pendekatan probabilistik berbasis teorema Bayes. Hasil penelitian menunjukkan bahwa K-NN memiliki kinerja lebih unggul dalam klasifikasi data pencurian berdasarkan lokasi, waktu, dan jenis kejadian. K-NN mencatat nilai akurasi, presisi, recall, dan F1-score yang lebih tinggi dibandingkan Naive Bayes. Prediksi juga mengungkapkan bahwa sebagian besar pencurian terjadi pada dini hari, diduga karena minimnya pengawasan atau ketidakhadiran pemilik. Temuan ini menyarankan perlunya peningkatan patroli pada waktu

rawan, baik malam maupun siang hari di lokasi padat penduduk. Patroli acak di malam hari dinilai efektif mencegah aksi kejahatan dengan menciptakan rasa diawasi bagi pelaku.

Kata Kunci: Algoritma k-NN, Algoritma Naïve Bayes, Pencurian

PENDAHULUAN

Perkembangan dunia saat ini telah memasuki fase revolusi industri 4.0 society 5.0, dimana revolusi industri 4.0 merupakan fenomena yang mengkolaborasikan teknologi siber dan teknologi otomatisasi atau cyber physical system. Karakteristik lain dari fase ini adalah fenomena Internet of Things (IoT), cloud computing, dan cognitive computing atau saat ini kita kenal dengan artificial intelligence. Sementara society 5.0 sendiri merupakan bentuk penyebutan evolusi peradaban manusia ketika mereka diharapkan dapat menyelesaikan berbagai tantangan dan permasalahan sosial di era revolusi industri 4.0. Pada era ini perubahan terjadi seiring dengan kemampuan manusia dalam melahirkan inovasi. Setiap periode zaman selalu saja melahirkan inovasi dahsyat yang berorientasi pada kebutuhan dan kemaslahatan kehidupan. Inovasi merupakan kunci kesuksesan di bidang industri (Kusnandar, 2019). Revolusi Industri 4.0 secara fundamental mengakibatkan berubahnya cara manusia berpikir, hidup, dan berhubungan satu dengan yang lain. Era ini akan mendisrupsi berbagai aktivitas manusia dalam berbagai bidang, tidak hanya dalam bidang teknologi saja, namun juga bidang yang lain seperti ekonomi, sosial, dan politik (Prasetyo, 2020).

Data mining adalah suatu proses semi otomatis yang memakai suatu teknik matematika, statistik, machine learning, dan kecerdasan buatan untuk mengidentifikasi dan mengekstraksi suatu informasi pengetahuan yang terkait dari macam-macam basis data yang sangat besar (Maharani, et al., 2017). Berdasarkan prinsipnya algoritma KNN melakukan klasifikasi data baru berdasarkan kemiripan atau memilih jarak terdekat. Jarak eucludien distance digunakan untuk menghitung jarak. K-Nearest Neighbor (KNN) merupakan salah satu algoritma yang berkembang khususnya dalam pengolahan citra untuk klasifikasi kematangan buah. Metode K-Nearest Neighbor (KNN) dipilih karena dalam penelitian pengolahan citra sebelumnya metode ini memiliki kinerja yang baik. Klasifikasi yang didasarkan pada sistem informasi merupakan teknik memetakan (mengklasifikasikan) data ke dalam satu atau beberapa kelas yang sudah didefinisikan sebelumnya. Ada banyak teknik klasifikasi yang dapat digunakan, diantaranya adalah K-Nearest Neighbor (KNN). Konsep penelitian dengan algoritma KNN telah banyak dilakukan oleh peneliti-peneliti sebelumnya diantaranya yaitu, algoritma KNN digunakan untuk pengenalan pola, pengenalan teks, pengolahan objek dan lain-lain. Algoritma KNN dianggap mempunyai kesederhanaan dalam pengolahan data training dan data testing dalam jumlah yang sangat besar (Tang, Jing, Li, & Atkinson, 2016).

Algoritma k-NN merupakan metode yang digunakan untuk melakukan klasifikasi data berdasarkan jarak terpendek terhadap objek data. Penentuan nilai K yang terbaik untuk algoritma ini berdasarkan pada data yang ada. Nilai K yang tinggi dapat mengurangi efek noise pada klasifikasi, bisa juga membuat batasan antara setiap klasifikasi menjadi lebih kabur (Anshori, Regasari, & Putri, 2018). Algoritma KNN merupakan algoritma berbasis contoh atau non parametric dan dianggap metode paling sederhana di dalam proses data mining (Tharwat, Mahdi, Elhoseny, & Hassanien, 2018). Algoritma KNN salah satu metode klasifikasi data yang mudah diimplementasi pada jumlah data yang kecil, tetapi jika dataset yang diolah banyak dan kompleks maka algoritma KNN memiliki kelemahan dan waktu yang tidak efisien (Yahya & Hidayanti, 2020). Metode KNN menggunakan ukuran jarak yang sesuai untuk mengklasifikasikan data baru. Jarak tetangga terdekat K dihitung dan label kelas dari tetangga terdekat diprediksi sebagai label kelas dari instance baru. Akurasi KNN sangat terpengaruh dengan memilih jumlah K tetangga terdekat. Jika terdapat nilai K yang kecil, maka akan

sensitif terhadap noise dan jika terlalu besar, dapat menyebabkan bias model (Anshori et al., 2018). Algoritma k-NN merupakan algoritma berbasis memori yang menggunakan iterasi pada data sampai atribut atau parameter data yang terdekat ditemukan. Jarak minimum data yang diproses pada data testing akan dibandingkan dengan data training dengan jarak yang terdekat (Suwirmayanti, 2017). Tahapan model satistika yang bisa mengerjakan pengkategorian yaitu klasifikasi. Banyak pilihan metode klasifikasi dan model penelitian ini dapat menggunakan metode naïve bayes. Metode naïve bayes sudah banyak digunakan pada penelitian tentang text mining, salah satu kelebihan naïve bayes yaitu algoritma sederhana tetapi mempunyai nilai akurasi yang cukup tinggi.

Kriminologi adalah seperangkat ilmu yang muncul pada abad ke-19 dan berfokus pada studi tentang kejahatan dari berbagai aspek yang melibatkan penyelidikan dan pemahaman mendalam terhadap fenomena kriminal. Menurut Edwin Sutherland dalam mempelajari kriminologi banyak menggunakan berbagai macam disiplin ilmu sehingga kriminologi dapat dikatakan ilmu yang bersifat interdisipliner. Hukum pidana, hukum acara pidana, antropologi fisik, antropologi budaya, psikologi, biologi, ekonomi, kimia, statistika, merupakan contoh disiplin ilmu yang erat kaitannya dengan kriminologi. Pencurian adalah suatu perbuatan kejahatan yang melanggar hukum, bertentangan dengan kesusilaan manusia dan norma sosial.

Tindak pidana adalah salah satu bentuk perilaku yang melenceng dari norma masyarakat dan dianggap sebagai ancaman signifikan terhadap norma dan nilai-nilai sosial. Perilaku yang melenceng ini dianggap sebagai masalah manusia dan sosial, karena dapat menimbulkan gangguan baik pada tingkat individu maupun pada tingkat sosial secara keseluruhan. Norma-norma sosial dianggap sebagai dasar bagi tatanan sosial yang teratur, dan jika perilaku yang melenceng dibiarkan tanpa penanganan, dapat mengancam keberlangsungan ketertiban sosial. tindak pidana bukan hanya merupakan masalah sosial, melainkan juga merupakan permasalahan yang berkaitan dengan kemanusiaan. Pemahaman ini menggarisbawahi kompleksitas dan dampak yang luas dari tindak pidana, tidak hanya terbatas pada tingkat sosial, tetapi juga mencakup dimensi kemanusiaan. Analisis lebih lanjut terhadap bagaimana penanganan tindak pidana dapat mencakup aspek-aspek kemanusiaan akan memberikan sudut pandang yang lebih mendalam terhadap permasalahan ini (Hutabarat et al., 2023). Pencurian dalam hukum pidana adalah suatu tindak pidana yang melibatkan pengambilan atau penggelapan barang milik orang lain tanpa izin atau tanpa hak. Tindakan pencurian ini diatur dalam berbagai sistem hukum di seluruh dunia dan umumnya dianggap sebagai suatu pelanggaran serius. Unsur-unsur utama pencurian biasanya melibatkan adanya niat untuk mengambil barang milik orang lain, pengambilan tersebut dilakukan tanpa izin atau tanpa hak, serta adanya perpindahan kepemilikan barang dari pemilik asli kepada pelaku pencurian.

Berdasarkan data Pusat Informasi Kriminal Nasional (Pusiknas) Bareskrim Polri, terdapat 394.001 tindak kejahatan yang terjadi sepanjang 2023 per bulan November. Angka ini mengalami lonjakan apabila dibandingkan dengan total tindak kejahatan yang terjadi pada 2022, yani sebanyak 322.200 kejadian. Sementara itu, pada 2021 Pusiknas mencatat bahwa terdapat 275.258 tindak kejahatan yang terjadi. Data Pusiknas menunjukkan bahwa jenis kejahatan berupa pencurian menjadi kejahatan yang paling banyak terjadi sejak Januari hingga November 2023. Terdapat hingga 155.361 tindak kejahatan berupa pencurian dengan pemberatan yang terjadi pada periode tersebut. Selain itu, terdapat total 115.324 tindak kejahatan jenis pencurian biasa.

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 asosiasi, yaitu teknik yang dapat menemukan aturan asosiasi atau hubungan antara item-item dalam data. Analisis asosiasi dapat digunakan untuk mengetahui hubungan antara hasil klasifikasi dan prediksi tindak pidana pencurian dengan keberhasilan patroli di suatu wilayah.

Dalam penelitian ini, penulis tertarik untuk mengkaji secara teoritis dan sistematis fenomena peningkatan tingkat tindak kriminal, khususnya tindak pidana pencurian, yang telah disebutkan sebelumnya. Penulis juga tertarik untuk mengeksplorasi belum adanya sistem prediksi yang mumpuni di kepolisian RI sebagai sistem yang membantu proses penanganan dan pencegahan tindak kriminal. Hasil penelitian ini akan dituangkan dalam bentuk skripsi yang berjudul: “ANALISIS KLASIFIKASI TINDAK KEJAHATAN PENCURIAN DENGAN ALGORITMA K-NEAREST NEIGHBOR DAN NAIVE BAYES DI POLRES BUOL”.

Penelitian ini bertujuan untuk menerapkan algoritma *machine learning* pada data tindak pidana pencurian di wilayah Polres Buol, untuk melakukan analisis asosiasi antara hasil klasifikasi dan prediksi tindak pidana pencurian dengan keberhasilan patroli di wilayah tersebut, Penelitian ini diharapkan dapat memberikan manfaat bagi kepolisian dalam meningkatkan kinerja dan kualitas patroli, serta memberikan rekomendasi patroli yang sesuai dengan kondisi dan kebutuhan di lapangan.

METODE

Dalam penelitian ini, peneliti menggunakan pendekatan kuantitatif sebagai aproksimasi dalam melakukan analisis. Pendekatan kuantitatif merupakan pendekatan yang mengandalkan perhitungan statistik untuk mencatat dan menganalisis data penelitian secara akurat. Sudjana (2004) mengutip Izaak Latanussa yang menyatakan bahwa penelitian kuantitatif adalah penelitian yang memakai metode angka untuk menggambarkan observasi suatu variabel atau objek, angka merupakan unsur dari pengukuran. Selain itu, pendekatan yang dipakai dalam penelitian ini juga melibatkan kombinasi antara statistik, matematik dan kecerdasan buatan untuk mengekstrak dan mengenali informasi yang berkaitan dengan *database* yang besar. Pendekatan ini biasa disebut dengan *data mining* (Efraim Turban, 2005).

Peneliti menerapkan metode algoritma *machine learning* sebagai salah satu metode *data mining* untuk melakukan klasifikasi dan prediksi data yang dapat digunakan untuk pengambilan keputusan oleh pihak terkait. Metode *data mining* jenis terawasi atau *supervised* ini menentukan kelas dari data yang digali. Data yang sudah dianalisis selanjutnya diproses dengan teknik *data mining* menggunakan *tools google colaborasion*. Beberapa kategori informasi yang diproses adalah jenis kelamin korban, umur korban, pekerjaan korban, tempat tinggal korban, tempat dan waktu terjadinya pencurian, hari pencurian, penggunaan alat oleh pelaku, sasaran pencurian, besaran kerugian, modus operandi, jenis dan pasal pencurian. Peneliti juga menambahkan informasi tambahan keputusan berupa metode pelaku dan jenis pencurian pada kolom data yang diperiksa dan mengabaikan bagian yang sudah diproses

Statistik konvensional mengenal konsep populasi dan sampel. Populasi direpresentasikan oleh sebagian kecil yang disebut sampel, yang dipilih dari populasi dengan metode tertentu agar sesuai dengan karakteristik populasi. Namun, dalam *data mining*, sampel bukanlah bagian dari populasi, melainkan populasi itu sendiri, karena kualitas analisis akan meningkat seiring dengan jumlah data yang banyak. Penelitian ini menggunakan populasi berupa Laporan Polisi (LP) Korban Pencurian yang terjadi di Wilayah Polres Buol.

HASIL DAN PEMBAHASAN

Gambaran Kasus Pencurian Kabupaten Buol

Adapun data kasus pencurian yang terjadi di Polres Buol adalah sebagai berikut:

Tabel 1. Data Kasus Tindak Pidana Pencurian Tahun 2017-2024

| NO | TINDAK PIDANA | JUMLAH | | DALAM PROSES |
|------------|-------------------------------------|--------|-----|--------------|
| | | L | S | |
| TAHUN 2017 | | | | |
| 1 | PENCURIAN BIASA | 112 | 32 | 80 |
| 2 | CURANMOR R-2 | 26 | 6 | 20 |
| 3 | CURAT | 11 | 9 | 2 |
| 4 | PENCURIAN TERNAK | 5 | 2 | 3 |
| 6 | PENCURIAN DALAM KELUARGA | 1 | 1 | 0 |
| TAHUN 2018 | | | | |
| 1 | PENCURIAN BIASA | 94 | 32 | 62 |
| 2 | CURANMOR R-2 | 43 | 6 | 37 |
| 3 | CURAT | 19 | 2 | 17 |
| 4 | PENCURIAN TERNAK | 1 | 1 | 0 |
| TAHUN 2019 | | | | |
| 1 | PENCURIAN BIASA (CUBIS) | 75 | 21 | 54 |
| 2 | PENCURIAN DENGAN PEMBERATAN (CURAT) | 40 | 4 | 36 |
| 3 | CURANMOR R-2 | 32 | 5 | 27 |
| TAHUN 2020 | | | | |
| 1 | PENCURIAN BIASA (CUBIS) | 75 | 14 | 61 |
| 2 | CURANMOR R-2 | 13 | 7 | 6 |
| 3 | PENCURIAN DENGAN PEMBERATAN (CURAT) | 10 | 3 | 7 |
| TAHUN 2021 | | | | |
| 1 | PENCURIAN BIASA (CUBIS) | 38 | 20 | 18 |
| 2 | CURANMOR R-2 | 9 | 3 | 6 |
| 3 | PENCURIAN DENGAN PEMBERATAN (CURAT) | 16 | 8 | 8 |
| TAHUN 2022 | | | | |
| 1 | PENCURIAN BIASA (CUBIS) | 49 | 22 | 27 |
| 2 | CURANMOR R-2 | 14 | 5 | 9 |
| 3 | PENCURIAN DENGAN PEMBERATAN (CURAT) | 24 | 4 | 20 |
| 4 | PENCURIAN TERNAK | 3 | 0 | 3 |
| TAHUN 2023 | | | | |
| 1 | PENCURIAN BIASA (CUBIS) | 76 | 66 | 10 |
| 2 | PENCURIAN DENGAN PEMBERATAN (CURAT) | 34 | 28 | 6 |
| 3 | PENCURIAN DENGAN KEKERASAN (CURAS) | 2 | 2 | 0 |
| 4 | PENCURIAN TERNAK | 3 | 1 | 2 |
| 5 | CURANMOR R-2 | 18 | 11 | 7 |
| TAHUN 2024 | | | | |
| 1 | PENCURIAN BIASA (CUBIS) | 115 | 30 | 26,09 |
| 2 | PENCURIAN DENGAN PEMBERATAN (CURAT) | 11 | 2 | 18,18 |
| 3 | PENCURIAN TERNAK | 7 | 1 | 14,29 |
| JUMLAH | | 976 | 348 | 628 |

Deskripsi Karakteristik Korban Tindak Pidana Pencurian

Data primer yang digunakan pada penelitian ini merupakan data korban tindak pidana pencurian di wilayah hukum Polres Buol pada rentang tahun 2017-2024. Sebelum dilakukan analisa data tersebut terlebih dahulu dilakukan pembersihan dan normalisasi agar data dapat di analisa dengan lebih mudah. Data hasil penuangan Laporan Polisi tersebut setelah dilakukan

proses pembersihan dan normalisasi menunjukkan jumlah tindak pidana pencurian sepanjang tahun 2017-2024 berjumlah 976 kasus.

Deksripsi Jenis Pencurian

Klasifikasi pencurian yang sering dijumpai dalam hukum pidana Indonesia, yang meliputi curanmor, curas, curat, curnak, pencurian biasa, dan pencurian dalam keluarga.

Tabel 2. Karakteristik Jenis Pencurian

| Jenis Pencurian | Frequency |
|--------------------------|------------|
| Curanmor R2 | 153 |
| Curas | 2 |
| Curat | 164 |
| Curnak | 21 |
| Pencurian Biasa | 635 |
| Pencurian Dalam Keluarga | 1 |
| Total | 976 |

Berdasarkan data jenis pencurian yang banyak terjadi yaitu pencurian biasa dengan presentase sebesar 65%.

Deskripsi Pasal Pencurian

Berdasarkan pasal dalam pencurian diklasifikasikan menjadi 2 yaitu Pasal 362 Kitab Undang-Undang Hukum Pidana (KUHP) mengatur tentang pencurian yang dilakukan seseorang dengan cara mengambil barang milik orang lain. Pasal ini menjelaskan bahwa seseorang yang dengan sengaja dan tanpa hak mengambil barang milik orang lain dapat dikenai sanksi pidana. Pasal 363 KUHP mengatur tentang pencurian dengan pemberatan, yaitu tindak pidana pencurian yang dilakukan dengan cara tertentu yang lebih berat, seperti dengan masuk ke dalam rumah atau bangunan lain dengan cara merusak atau tanpa izin. Pencurian jenis ini melibatkan tindakan yang lebih merugikan korban dan berisiko lebih besar.

Tabel 3. Karakteristik Tindak Pidana Pencurian Berdasarkan Pasal

| Pasal Pencurian | Frequency |
|-----------------|------------|
| 362 Kuhp | 636 |
| 363 Kuhp | 340 |
| Total | 976 |

Berdasarkan data diatas maka dapat dilihat tindak pidana pencurian berdasarkan pasal paling banyak terjadi yaitu pasal 362 KUHP dengan presentase sebesar 65%.

Deskripsi Usia Korban

Dalam konteks tindak pidana pencurian atau kejahatan lainnya, usia korban dapat menjadi faktor penting yang mempengaruhi dampak dari tindak kejahatan tersebut. Setiap kelompok usia memiliki kerentanannya masing-masing, yang juga berpengaruh pada jenis perlindungan hukum yang diberikan. Berikut adalah deskripsi mengenai data usia korban yang dibedakan berdasarkan kategori Dewasa, Manula (Lansia), dan Remaja.

Tabel 4. Karakteristik Deskripsi Usia Korban

| Usia Korban | Frequency |
|--------------|------------|
| Dewasa | 552 |
| Lansia | 108 |
| Manula | 1 |
| Remaja | 315 |
| Total | 976 |

Berdasarkan data diatas maka usia korban pencurian yaitu dewasa 57% Korban dewasa biasanya memiliki potensi untuk melindungi diri dan mengatasi dampak kejahatan lebih baik

dibandingkan kelompok usia lainnya. Namun, mereka tetap rentan terhadap kejahatan seperti pencurian, terutama jika korban tidak waspada atau berada dalam situasi yang mudah dimanfaatkan oleh pelaku (misalnya, tidak ada pengawasan, atau kondisi sosial dan ekonomi yang lemah).

Deskripsi Tempat Pencurian

Tabel 5. Karakteristik Tempat Pencurian

| Tempat Kejadian Pencurian | Frequency |
|---------------------------|------------|
| Area Parkir | 70 |
| Dalam Rumah | 402 |
| Halaman Rumah | 367 |
| Jalan Raya | 36 |
| Parkiran Rumah | 30 |
| Pedesaan | 8 |
| Pekarangan Rumah | 26 |
| Pemukiman | 20 |
| Perkotaan | 17 |
| Total | 976 |

Berdasarkan data penelitian tempat pencurian yang banyak terjadi yaitu di dalam rumah dengan presentase sebesar 41%.

Deskripsi Waktu Pencurian

Waktu pencurian merupakan gambaran terjadinya waktu yang digambarkan berdasarkan keadaan.

Tabel 6. Karakteristik Waktu Kejadian

| Waktu Kejadian | Frequency |
|----------------|------------|
| Dini Hari | 162 |
| Malam | 308 |
| Pagi | 260 |
| Sabtu | 1 |
| Siang | 131 |
| Sore | 114 |
| Total | 976 |

Berdasarkan data penelitian diatas waktu yang sering terjadi pencurian yaitu pada malam hari dengan presentase 31%.

Deskripsi Jenis Kelamin korban

Jenis kelamin korban pencurian terdiri dari laki-laki dan perempuan menyatakan bahwa korban pencurian bisa berasal dari kedua jenis kelamin, yaitu laki-laki dan perempuan. Ini menunjukkan bahwa baik laki-laki maupun perempuan dapat menjadi korban dalam kasus pencurian.

Tabel 7. Karakteristik Jenis Kelamin Korban

| Jenis Kelamin Korban | Frequency |
|----------------------|------------|
| Laki-Laki | 581 |
| Perempuan | 395 |
| Total | 976 |

Berdasarkan data diatas dapat dilihat bahwa korban yang paling banyak mengalami pencurian adalah berjenis kelamin laki-laki dengan presentase 60%.

Deskripsi Modus Operandi

Modus pencurian merujuk pada cara atau teknik yang digunakan oleh pelaku untuk melakukan aksi pencurian.

Tabel 8. Karakteristik Modus Operandi Pencurian

| MO | Frequency |
|-----------------|------------|
| Memanjat | 56 |
| Mengambil | 657 |
| Mengancam | 2 |
| Merusak Jendela | 59 |
| Merusak Kunci | 153 |
| Merusak Pintu | 49 |
| Total | 976 |

Berdasarkan data diatas maka dapat dilihat bahwa modus yang digunakan oleh pelaku paling banyak adalah dengan cara menfambil yaitu sebesar 67%. Merusak kunci, yang sering disebut sebagai pembobolan, memang merupakan salah satu modus yang banyak digunakan dalam pencurian, terutama pada kasus pencurian di rumah atau tempat usaha. Modus ini dilakukan dengan cara merusak pintu, jendela, atau alat pengaman lainnya untuk masuk ke dalam properti yang menjadi sasaran.

Deksirpi Jumlah Kerugian

Deskripsi jumlah kerugian dalam kasus pencurian merujuk pada perhitungan total nilai barang dan properti yang hilang atau rusak akibat tindakan pencurian tersebut. Jumlah kerugian ini sering kali digunakan untuk keperluan klaim asuransi, laporan polisi, atau perhitungan ganti rugi.

Tabel 9. Karakteristik Jumlah Kerugian Korban

| Jumlah Kerugian | Frequency |
|-----------------|------------|
| Besar | 327 |
| Kecil | 649 |
| Total | 976 |

Berdasarkan data dilihat maka dapat disimpulkan bahwa kerugian yang dialami korban paling banyak dengan kerugian ber skala kecil dengan presentase 66%.

Deksripsi Sasaran Pencurian

Deskripsi sasaran pencurian dapat dibagi menjadi beberapa kategori berdasarkan jenis benda yang dicuri.

Tabel 10. Karaktersitik Sasaran Pencurian

| Sasaran Pencurian | Frequency |
|----------------------|------------|
| Benda Bergerak | 185 |
| Benda Tidak Bergerak | 759 |
| Gabungan | 11 |
| Ternak/Hewan | 21 |
| Total | 976 |

Berdasarkan data diatas maka sasaran pencurian paling banyak yaitu benda tidak bergerak dengan presentase 78%.

Bagaimana melakukan prediksi waktu terjadinya tindak pidana pencurian di wilayah hukum Polres Buol menggunakan algoritma KNN dan Naive Bayes

Klasifikasi dan prediksi tindak pidana pencurian dengan algoritma *machine learning* di Polres Buol mengimplementasikan algoritma *K-Nearest Neighbor* dan algoritma *Naïve Bayes*. Proses tersebut melewati beberapa tahapan yang akan dijelaskan pada subbab selanjutnya.

1. Persiapan Dataset

Dataset yang digunakan untuk klasifikasi dan prediksi tindak pidana pencurian dengan algoritma *machine learning* di Polres Buol merupakan *dataset* primer atau *dataset* yang didapatkan dari data korban tindak pidana pencurian di wilayah hukum Polresta Buol pada rentang tahun 2017-2024 Data di luar rentang tersebut tidak akan digunakan. Berikut ini merupakan data primer yang digunakan dalam penelitian ini:

| NO | NIS KELAMIN | KORBA | KEJADIAN | KORBA | TINGGAL | KORBA | PEJADIAN | PENCURIAH | PERISTIAK | KEJADIAN | HARI KEJADIAN | AKU MENGGUNAKAN | ANSARAN KEJAHAT/MLAH KERUGM | MO | NIS PENCUR |
|----|-------------|--------|----------------|-----------|-----------------|-----------|----------|-----------|-----------|---------------------|---------------|-----------------|-----------------------------|----|------------|
| 1 | PEREMPUAN | DEWASA | HOHONER | PERKOTAAN | PEKARANAN RUMAH | PERKOTAAN | MALAM | MINGGU | TIDAK | ENDIA TIDAK BERGERA | KECIL | MENGAMBIL | PENCIAN BIAS | | |
| 2 | LAKI-LAKI | DEWASA | WIFASWASTA | PEDESAAN | HALAMAN RUMAH | PEDESAAN | MALAM | SENN | TIDAK | ENDIA TIDAK BERGERA | KECIL | MENGAMBIL | PENCIAN BIAS | | |
| 3 | LAKI-LAKI | DEWASA | TANIPKEBUN | PEDESAAN | DALAM RUMAH | PEDESAAN | MALAM | SELASA | TIDAK | ENDIA TIDAK BERGERA | KECIL | MENGAMBIL | PENCIAN BIAS | | |
| 4 | LAKI-LAKI | DEWASA | TIDAK BEKERJA | PEDESAAN | DALAM RUMAH | PEDESAAN | PAGI | RABU | TIDAK | ENDIA TIDAK BERGERA | KECIL | MENGAMBIL | PENCIAN BIAS | | |
| 5 | PEREMPUAN | DEWASA | TENAGA HOHONER | PERKOTAAN | PEKARANAN RUMAH | PERKOTAAN | MALAM | RABU | TIDAK | ENDIA TIDAK BERGERA | KECIL | MENGAMBIL | PENCIAN BIAS | | |
| 6 | PEREMPUAN | DEWASA | WIFASWASTA | PERKOTAAN | HALAMAN RUMAH | PERKOTAAN | MALAM | KAMIS | TIDAK | ENDIA TIDAK BERGERA | KECIL | MENGAMBIL | PENCIAN BIAS | | |
| 7 | PEREMPUAN | DEWASA | IRT | PERKOTAAN | DALAM RUMAH | PERKOTAAN | MALAM | MINGGU | TIDAK | ENDIA TIDAK BERGERA | KECIL | MENGAMBIL | PENCIAN BIAS | | |
| 8 | LAKI-LAKI | DEWASA | TENAGA HOHONER | PEDESAAN | DALAM RUMAH | PEDESAAN | PAGI | SELASA | TIDAK | ENDIA TIDAK BERGERA | KECIL | MENGAMBIL | PENCIAN BIAS | | |
| 9 | PEREMPUAN | DEWASA | TENAGA HOHONER | PERKOTAAN | PEKARANAN RUMAH | PERKOTAAN | MALAM | MINGGU | TIDAK | ENDIA TIDAK BERGERA | KECIL | MENGAMBIL | PENCIAN BIAS | | |
| 10 | LAKI-LAKI | DEWASA | WIFASWASTA | PEDESAAN | HALAMAN RUMAH | PEDESAAN | MALAM | RABU | TIDAK | ENDIA TIDAK BERGERA | KECIL | MENGAMBIL | PENCIAN BIAS | | |
| 11 | PEREMPUAN | DEWASA | TIDAK BEKERJA | PERKOTAAN | DALAM RUMAH | PERKOTAAN | MALAM | SELASA | TIDAK | ENDIA TIDAK BERGERA | KECIL | MENGAMBIL | PENCIAN BIAS | | |
| 12 | LAKI-LAKI | DEWASA | TIDAK BEKERJA | PERKOTAAN | DALAM RUMAH | PERKOTAAN | PAGI | RABU | TIDAK | ENDIA TIDAK BERGERA | KECIL | MENGAMBIL | PENCIAN BIAS | | |
| 13 | PEREMPUAN | DEWASA | TENAGA HOHONER | PERKOTAAN | PEKARANAN RUMAH | PERKOTAAN | MALAM | SELASA | TIDAK | ENDIA TIDAK BERGERA | KECIL | MENGAMBIL | PENCIAN BIAS | | |
| 14 | LAKI-LAKI | DEWASA | WIFASWASTA | PEDESAAN | HALAMAN RUMAH | PEDESAAN | PAGI | JUMAT | TIDAK | ENDIA TIDAK BERGERA | KECIL | MENGAMBIL | PENCIAN BIAS | | |
| 15 | LAKI-LAKI | DEWASA | TIDAK BEKERJA | PERKOTAAN | DALAM RUMAH | PERKOTAAN | PAGI | KAMIS | TIDAK | ENDIA TIDAK BERGERA | KECIL | MENGAMBIL | PENCIAN BIAS | | |
| 16 | PEREMPUAN | DEWASA | IRT | PERKOTAAN | DALAM RUMAH | PERKOTAAN | PAGI | SENN | TIDAK | ENDIA TIDAK BERGERA | KECIL | MENGAMBIL | PENCIAN BIAS | | |
| 17 | PEREMPUAN | DEWASA | HOHONER | PERKOTAAN | PEKARANAN RUMAH | PERKOTAAN | PAGI | MINGGU | TIDAK | ENDIA TIDAK BERGERA | KECIL | MENGAMBIL | PENCIAN BIAS | | |
| 18 | LAKI-LAKI | DEWASA | WIFASWASTA | PERKOTAAN | HALAMAN RUMAH | PERKOTAAN | PAGI | RABU | TIDAK | ENDIA TIDAK BERGERA | KECIL | MENGAMBIL | PENCIAN BIAS | | |
| 19 | LAKI-LAKI | DEWASA | TANIPKEBUN | PERKOTAAN | DALAM RUMAH | PERKOTAAN | MALAM | RABU | TIDAK | ENDIA TIDAK BERGERA | KECIL | MENGAMBIL | PENCIAN BIAS | | |
| 20 | LAKI-LAKI | DEWASA | TANIPKEBUN | PEDESAAN | DALAM RUMAH | PEDESAAN | MALAM | SELASA | TIDAK | ENDIA TIDAK BERGERA | KECIL | MENGAMBIL | PENCIAN BIAS | | |
| 21 | PEREMPUAN | DEWASA | HOHONER | PEDESAAN | PEKARANAN RUMAH | PEDESAAN | MALAM | MINGGU | TIDAK | ENDIA TIDAK BERGERA | KECIL | MENGAMBIL | PENCIAN BIAS | | |
| 22 | PEREMPUAN | DEWASA | WIFASWASTA | PEDESAAN | HALAMAN RUMAH | PEDESAAN | PAGI | RABU | TIDAK | ENDIA TIDAK BERGERA | KECIL | MENGAMBIL | PENCIAN BIAS | | |
| 23 | PEREMPUAN | DEWASA | TIDAK BEKERJA | PERKOTAAN | DALAM RUMAH | PERKOTAAN | PAGI | SELASA | TIDAK | ENDIA TIDAK BERGERA | KECIL | MENGAMBIL | PENCIAN BIAS | | |
| 24 | LAKI-LAKI | DEWASA | TIDAK BEKERJA | PEDESAAN | DALAM RUMAH | PEDESAAN | MALAM | MINGGU | TIDAK | ENDIA TIDAK BERGERA | KECIL | MENGAMBIL | PENCIAN BIAS | | |
| 25 | PEREMPUAN | DEWASA | HOHONER | PERKOTAAN | PEKARANAN RUMAH | PERKOTAAN | MALAM | MINGGU | TIDAK | ENDIA TIDAK BERGERA | KECIL | MENGAMBIL | PENCIAN BIAS | | |
| 26 | PEREMPUAN | DEWASA | WIFASWASTA | PEDESAAN | HALAMAN RUMAH | PEDESAAN | MALAM | RABU | TIDAK | ENDIA TIDAK BERGERA | KECIL | MENGAMBIL | PENCIAN BIAS | | |
| 27 | PEREMPUAN | DEWASA | IRT | PERKOTAAN | DALAM RUMAH | PERKOTAAN | MALAM | RABU | TIDAK | ENDIA TIDAK BERGERA | KECIL | MENGAMBIL | PENCIAN BIAS | | |
| 28 | PEREMPUAN | DEWASA | IRT | PEDESAAN | DALAM RUMAH | PEDESAAN | PAGI | SELASA | TIDAK | ENDIA TIDAK BERGERA | KECIL | MENGAMBIL | PENCIAN BIAS | | |
| 29 | PEREMPUAN | DEWASA | HOHONER | PERKOTAAN | PEKARANAN RUMAH | PERKOTAAN | MALAM | MINGGU | TIDAK | ENDIA TIDAK BERGERA | KECIL | MENGAMBIL | PENCIAN BIAS | | |

Gambar 1. Dataset

Kemudian untuk kolom yang terdapat pada *dataset* yaitu berjumlah 14 kolom dan 976 baris. Semua kolom tidak akan digunakan Pada proses Klasifikasi dan prediksi tindak pidana pencurian dengan algoritma *machine learning*, hanya kolom yang relevan yang akan digunakan untuk proses tersebut.

2. Pemilihan Atribut

Pemilihan attribut merupakan tahap pemilihan pemilihan kelas dan fitur pada kolom *dataset*. Untuk kelas menggunakan data waktu kejadian dengan jumlah total kelas 5 yaitu, dini hari, pagi, malam, siang dan sore. Kemudian untuk fitur yang akan digunakan merupakan fitur yang relevan dengan topik penelitian yang diangkat yaitu tempat kejadian, waktu kejadian dan dana pasal.

3. Preprocessing Data

Tahap *preprocessing* dilakukan yaitu melakukan penghapusan data data yang kosong serta melakukan normalisasi data untuk data yang tidak sesuai. Pertama dilakukan pengecekan data kosong pada kolom yang telah dipilih sebelumnya sebagai kelas dan fitur yaitu tempat kejadian, waktu dan pasal. Berikut hasil pengecekan data kosong untuk data tersebut. Jika tidak terdapat data yang kosong pada data yang akan digunakan untuk klasifikasi dan prediksi. Sehingga tidak perlu dilakukan penghapusan data kosong. Setelah data kosong di cek kemudian melakukan normalisasi data pada setiap kolom. Sebelum dilakukan normalisasi dilakukan perubahan nilai menjadi huruf kecil atau *lower case*.

Setelah dilakukan perubahan nilai data menjadi *lower case* kemudian dilakukan normalisasi. Untuk kolom waktu dan hari tidak perlu dilakukan normalisasi dikarenakan nilai pada kedua kolom tersebut sudah sesuai dengan yang diperlukan.

Apakah model terbaik untuk prediksi berdasarkan klasifikasi tindak pidana pencurian dari kedua algoritma tersebut

Penelitian ini bertujuan untuk membandingkan performa dua algoritma klasifikasi, yaitu K-Nearest Neighbors (KNN) dan Naive Bayes, dalam mengklasifikasikan data kriminal berdasarkan tiga atribut utama, yaitu tempat kejadian, waktu kejadian, dan jenis pencurian. Ketiga fitur ini dipilih karena merupakan elemen penting dalam pola kejadian kriminal yang dapat memberikan wawasan prediktif terhadap kemungkinan kejadian serupa di masa mendatang. Proses klasifikasi diawali dengan preprocessing data, seperti pengubahan data kategorikal (seperti nama tempat dan jenis pencurian) menjadi representasi numerik, serta normalisasi pada data numerik jika diperlukan (khusus untuk KNN yang sensitif terhadap skala data).

Model pertama yang digunakan adalah K-Nearest Neighbors (KNN). KNN merupakan algoritma non-parametrik yang mengklasifikasikan data berdasarkan kedekatan (jarak) dengan data lain yang telah diberi label. Pemilihan nilai k yang optimal dilakukan melalui proses validasi silang. Hasil klasifikasi menunjukkan bahwa KNN memiliki performa yang baik dalam mengenali pola berdasarkan kedekatan waktu dan lokasi kejadian. Namun, KNN cukup terpengaruh oleh jumlah dan distribusi data serta memerlukan waktu lebih lama dalam proses prediksi karena menggunakan pendekatan lazy learning.

Model kedua yang diuji adalah Naive Bayes, yang merupakan algoritma probabilistik berbasis Teorema Bayes dengan asumsi independensi antar fitur. Meskipun asumsi independen ini jarang sepenuhnya terpenuhi dalam kasus nyata, Naive Bayes tetap mampu memberikan hasil klasifikasi yang cukup akurat. Naive Bayes menunjukkan keunggulan dalam efisiensi waktu komputasi dan kemampuan mengatasi dataset yang besar. Dalam kasus ini, model Naive Bayes dapat dengan cepat memprediksi jenis pencurian berdasarkan kombinasi tempat dan waktu kejadian.

Tabel 11. Model k-NN Klasifikasi Tempat Pencurian

| Model | Rasio | K | Akurasi | Presisi | Recall | F1 Score |
|-------|-------|---|---------|---------|--------|----------|
| KNN | 60:40 | 3 | 0,646 | 0,573 | 0,573 | 0,573 |
| | | 5 | 0,597 | 0,441 | 0,457 | 0,449 |
| | | 7 | 0,566 | 0,493 | 0,479 | 0,486 |
| | | 9 | 0,589 | 0,450 | 0,494 | 0,471 |
| | 70:30 | 3 | 0,658 | 0,575 | 0,562 | 0,568 |
| | | 5 | 0,648 | 0,568 | 0,542 | 0,555 |
| | | 7 | 0,607 | 0,580 | 0,505 | 0,540 |
| | | 9 | 0,621 | 0,561 | 0,494 | 0,526 |
| | 80:20 | 3 | 0,646 | 0,607 | 0,567 | 0,586 |
| | | 5 | 0,635 | 0,632 | 0,571 | 0,600 |
| | | 7 | 0,600 | 0,527 | 0,537 | 0,532 |
| | | 9 | 0,600 | 0,590 | 0,485 | 0,532 |
| | 90:10 | 3 | 0,663 | 0,396 | 0,475 | 0,432 |
| | | 5 | 0,693 | 0,432 | 0,504 | 0,465 |
| | | 7 | 0,632 | 0,376 | 0,431 | 0,402 |
| | | 9 | 0,642 | 0,516 | 0,453 | 0,482 |

Tabel 12. Model Naïve Bayes Klasifikasi Tempat Pencurian

| Model | Rasio | Akurasi | Presisi | Recall | F1 Score |
|-------|---------|---------|---------|--------|----------|
| NB | 60 : 40 | 0,535 | 0,513 | 0,587 | 0,548 |
| | 70 : 30 | 0,570 | 0,575 | 0,683 | 0,624 |
| | 80 : 20 | 0,533 | 0,551 | 0,657 | 0,599 |
| | 90 : 10 | 0,489 | 0,402 | 0,475 | 0,435 |

Berdasarkan hasil tabel perbandingan klasifikasi tempat pencurian diketahui bahwa model k-NN dengan K-5 Rasio 80:20 memiliki nilai presisi dan recall paling tinggi, dan untuk naïve bayes berada pada rasio 70:30.

Tabel 13. Model k-NN Klasifikasi Waktu Pencurian

| Model | Rasio | K | Akurasi | Presisi | Recall | F1 Score |
|-------|-------|---|---------|---------|--------|----------|
| KNN | 60:40 | 3 | 0,502 | 0,426 | 0,394 | 0,409 |
| | | 5 | 0,441 | 0,346 | 0,329 | 0,337 |
| | | 7 | 0,446 | 0,367 | 0,330 | 0,347 |
| | | 9 | 0,423 | 0,347 | 0,306 | 0,325 |
| | 70:30 | 3 | 0,542 | 0,456 | 0,417 | 0,435 |
| | | 5 | 0,453 | 0,372 | 0,352 | 0,362 |
| | | 7 | 0,481 | 0,411 | 0,364 | 0,386 |
| | | 9 | 0,481 | 0,416 | 0,360 | 0,386 |
| | 80:20 | 3 | 0,579 | 0,500 | 0,456 | 0,477 |
| | | 5 | 0,461 | 0,416 | 0,367 | 0,390 |
| | | 7 | 0,466 | 0,451 | 0,371 | 0,407 |
| | | 9 | 0,461 | 0,444 | 0,356 | 0,395 |
| | 90:10 | 3 | 0,540 | 0,495 | 0,431 | 0,460 |
| | | 5 | 0,510 | 0,493 | 0,381 | 0,430 |
| | | 7 | 0,459 | 0,454 | 0,350 | 0,395 |
| | | 9 | 0,489 | 0,451 | 0,356 | 0,398 |

Tabel 14. Model Naïve Bayes Klasifikasi Waktu Pencurian

| Model | Rasio | Akurasi | Presisi | Recall | F1 Score |
|-------|---------|---------|---------|--------|----------|
| NB | 60 : 40 | 0,333 | 0,287 | 0,257 | 0,271 |
| | 70 : 30 | 0,327 | 0,276 | 0,256 | 0,265 |
| | 80 : 20 | 0,287 | 0,245 | 0,229 | 0,237 |
| | 90 : 10 | 0,326 | 0,306 | 0,270 | 0,287 |

Berdasarkan hasil tabel perbandingan klasifikasi waktu pencurian diketahui bahwa model k-NN dengan K-3 Rasio 80:20 memiliki nilai presisi dan recall paling tinggi, dan untuk naïve bayes berada pada rasio 90:10.

Tabel 15. Model k-NN Klasifikasi Jenis Pencurian

| Model | Rasio | K | Akurasi | Presisi | Recall | F1 Score |
|-------|---------|---|---------|---------|--------|----------|
| KNN | 60 : 40 | 3 | 0,967 | 0,807 | 0,798 | 0,802 |
| | | 5 | 0,969 | 0,642 | 0,637 | 0,639 |
| | | 7 | 0,969 | 0,642 | 0,636 | 0,639 |
| | | 9 | 0,961 | 0,637 | 0,628 | 0,633 |
| | 70 : 30 | 3 | 0,979 | 0,819 | 0,811 | 0,815 |
| | | 5 | 0,972 | 0,645 | 0,641 | 0,643 |
| | | 7 | 0,976 | 0,649 | 0,644 | 0,647 |
| | | 9 | 0,972 | 0,648 | 0,641 | 0,645 |
| | 80 : 20 | 3 | 0,979 | 0,654 | 0,646 | 0,650 |
| | | 5 | 0,969 | 0,643 | 0,636 | 0,639 |
| | | 7 | 0,979 | 0,654 | 0,646 | 0,650 |
| | | 9 | 0,969 | 0,643 | 0,636 | 0,639 |

| | | | | | |
|---------|---|-------|-------|-------|-------|
| 90 : 10 | 3 | 0,979 | 0,652 | 0,643 | 0,647 |
| | 5 | 0,979 | 0,652 | 0,643 | 0,647 |
| | 7 | 0,979 | 0,652 | 0,643 | 0,647 |
| | 9 | 0,979 | 0,652 | 0,643 | 0,647 |

Tabel 16. Model Naïve Bayes Klasifikasi Jenis Pencurian

| Model | Rasio | Akurasi | Presisi | Recall | F1 Score |
|-------|---------|---------|---------|--------|----------|
| NB | 60 : 40 | 0,623 | 0,540 | 0,563 | 0,551 |
| | 70 : 30 | 0,632 | 0,563 | 0,604 | 0,583 |
| | 80 : 20 | 0,606 | 0,537 | 0,496 | 0,516 |
| | 90 : 10 | 0,605 | 0,471 | 0,450 | 0,460 |

Dari hasil evaluasi menggunakan metrik akurasi, presisi, recall, dan F1-score, diperoleh bahwa performa kedua model bervariasi tergantung pada kompleksitas data. KNN cenderung lebih unggul dalam hal akurasi ketika pola distribusi data relatif teratur dan tidak terlalu beragam, sementara Naive Bayes lebih stabil dan efisien ketika menghadapi dataset yang besar dengan distribusi yang tidak merata.

Hasil dari penelitian ini menunjukkan bahwa algoritma K-Nearest Neighbors (KNN) memiliki performa yang lebih baik dibandingkan dengan algoritma Naive Bayes dalam melakukan klasifikasi data pencurian berdasarkan tempat kejadian, waktu kejadian, dan **jenis pencurian**. Dalam pengujian yang dilakukan, KNN secara konsisten menghasilkan nilai akurasi, presisi, recall, dan F1-score yang lebih tinggi dibandingkan Naive Bayes. Hal ini menunjukkan bahwa KNN lebih efektif dalam mengenali pola-pola yang terbentuk dari kombinasi waktu dan lokasi kejadian terhadap jenis pencurian yang terjadi.

Keunggulan KNN ini disebabkan oleh kemampuannya dalam memanfaatkan kedekatan antar data untuk menentukan klasifikasi. Dengan mempertimbangkan tetangga terdekat dalam ruang fitur, KNN mampu menangkap hubungan kompleks antar atribut yang mungkin tidak bisa dijelaskan secara sederhana oleh model probabilistik seperti Naive Bayes.

Sementara itu, meskipun Naive Bayes memiliki kecepatan pemrosesan yang lebih tinggi dan bekerja baik pada data berskala besar, model ini cenderung menghasilkan prediksi yang kurang akurat dalam konteks data pencurian yang memiliki keterkaitan erat antara fitur-fitur, seperti hubungan antara tempat kejadian dan waktu kejadian terhadap jenis pencurian. Dengan demikian, dapat disimpulkan bahwa KNN lebih unggul dalam memprediksi kejadian pencurian berdasarkan lokasi, waktu, dan jenis kasus, serta lebih cocok digunakan untuk sistem prediksi atau deteksi pola pencurian berbasis data historis.

Apakah analisis prediksi yang dilakukan dapat membantu Polres Buol dalam membuat pola patroli efektif berdasarkan waktu

Ilmu kepolisian sebagai cabang ilmu terapan dalam sistem hukum dan ketertiban masyarakat memiliki peranan penting dalam merespons dan mengantisipasi kejahatan, termasuk tindak pidana pencurian. Dalam konteks lokal seperti di wilayah hukum Polres Buol, penerapan ilmu kepolisian tidak hanya mencakup penegakan hukum secara represif, tetapi juga pendekatan preventif berbasis data dan analisis sosial. Ilmu kepolisian sebagai suatu disiplin multidisipliner berperan penting dalam pengembangan strategi pencegahan dan penanggulangan kejahatan secara komprehensif. Di tengah perkembangan teknologi digital dan ketersediaan big data, pendekatan konvensional dalam penegakan hukum kini diperkuat dengan metode prediktif berbasis machine learning, salah satunya adalah dengan menerapkan model K-Nearest Neighbors (KNN) dan Naive Bayes dalam menganalisis dan menangani tindak pidana pencurian.

Dalam ilmu kepolisian, upaya preventif (pencegahan) kejahatan menempati posisi strategis karena lebih efektif dan efisien daripada penindakan setelah kejahatan terjadi.

Pendekatan ini didasarkan pada prinsip bahwa keamanan masyarakat merupakan hasil kerja kolektif antara institusi kepolisian dan warga. Hal ini tercermin dalam model community policing, yang menempatkan masyarakat sebagai mitra aktif dalam mencegah tindak kriminal, termasuk pencurian. Di era digital ini, pendekatan preventif dalam ilmu kepolisian tidak lagi terbatas pada patroli rutin dan penyuluhan hukum semata, tetapi telah berkembang menuju penggunaan model analisis prediktif yang berbasis data. Dalam konteks wilayah Polres Buol, yang merupakan daerah dengan tantangan geografis dan sosial tersendiri, integrasi antara ilmu kepolisian dan teknologi prediktif menjadi penting untuk menciptakan sistem pencegahan kejahatan yang lebih akurat, efisien, dan berkelanjutan.

Fenomena pencurian sebagai salah satu bentuk kejahatan konvensional terus menjadi perhatian utama dalam kajian kriminologi maupun dalam pengembangan teknologi prediktif modern. Dalam konteks ini, analisis interdisipliner antara teori-teori kriminologi dan metode prediksi berbasis machine learning seperti K-Nearest Neighbors (KNN) dan Naive Bayes menjadi sangat relevan. Pendekatan ini tidak hanya berfungsi untuk memahami sebab-musabab dari kejahatan, tetapi juga untuk membangun sistem peringatan dini terhadap kemungkinan terjadinya pencurian berdasarkan data historis dan pola sosial tertentu.

Pencurian dapat dianalisis menggunakan berbagai teori kriminologi klasik maupun kontemporer. Dalam teori kriminologi dinyatakan bahwa individu melakukan tindak kejahatan sebagai hasil dari keputusan rasional, dengan mempertimbangkan manfaat dan risiko yang mungkin terjadi (Cornish & Clarke, 1986). Dalam konteks ini, pelaku pencurian diasumsikan memilih waktu dan tempat yang dianggap memiliki risiko rendah dan keuntungan tinggi, seperti rumah kosong atau lokasi minim pengawasan.

Dengan menggabungkan teori kriminologi dan metode prediksi machine learning, kita dapat membangun sistem yang lebih canggih dalam memahami dan mencegah pencurian. Teori kriminologi menyediakan landasan konseptual untuk memahami motif, perilaku, dan konteks sosial dari kejahatan, sementara model prediktif seperti KNN dan Naive Bayes menawarkan alat statistik yang dapat digunakan untuk mendeteksi pola dan membuat prediksi berdasarkan data aktual. Implementasi pendekatan ini dapat dilakukan dalam bentuk sistem peringatan dini berbasis lokasi, yang menggunakan data historis kejahatan dan indikator lingkungan untuk memberi tahu pihak berwenang mengenai area rawan pencurian. Ini tidak hanya meningkatkan efisiensi aparat keamanan, tetapi juga memperkuat upaya pencegahan kejahatan melalui intervensi berbasis data.

Tindak pidana pencurian merupakan salah satu bentuk kejahatan yang paling lazim terjadi di tengah masyarakat dan menjadi cerminan dari dinamika sosial yang kompleks. Dalam konteks kajian kriminologi Indonesia, Yasminah Mandasari Saragih dan Alwan Hadiyanto (2021) dalam buku *Pengantar Teori Kriminologi dan Teori dalam Hukum Pidana* menyajikan suatu pendekatan yang integratif dan sistematis dalam memahami berbagai aspek yang melingkupi kejahatan, termasuk pencurian, melalui lima elemen utama: karakteristik dan tingkat kejahatan, etiologi (faktor penyebab kejahatan), evolusi hukum pidana, tipologi pelaku, serta pola-pola kejahatan dan dampaknya terhadap perubahan sosial. Saragih dan Hadiyanto menyoroti bahwa pencurian memiliki karakteristik yang khas, yaitu dilakukan dengan niat untuk menguasai barang milik orang lain secara melawan hukum. Kejahatan ini sering kali terjadi pada wilayah-wilayah dengan tingkat kepadatan penduduk yang tinggi, sistem pengawasan sosial yang lemah, serta kesenjangan sosial yang mencolok.

Analisis prediksi yang dilakukan dapat memberikan kontribusi yang signifikan bagi Polres Buol dalam merancang pola patroli yang lebih efektif. Dengan memanfaatkan data mengenai waktu kejadian pencurian, kepolisian dapat menentukan periode waktu yang lebih rawan terhadap tindak pidana tersebut dan menyesuaikan strategi patroli.

Hasil dari analisis prediksi menunjukkan bahwa sebagian besar kejadian pencurian terjadi pada Dini hari. Hal ini memberikan informasi berharga bahwa dini hari menjadi waktu

yang lebih rentan terhadap terjadinya pencurian, mungkin karena faktor minimnya pengawasan di lingkungan sekitar atau ketidakhadiran pemilik barang di rumah atau tempat usaha.

Berdasarkan temuan ini, Polres Buol dapat mengoptimalkan pola patroli dengan meningkatkan frekuensi dan intensitas patroli pada dini hari, terutama di daerah yang lebih rawan. Kepolisian juga dapat melakukan pemantauan lebih intensif pada jam-jam tertentu yang memiliki tingkat kejadian pencurian yang lebih tinggi, serta memastikan keberadaan petugas di titik-titik strategis yang sering menjadi sasaran pencurian. Patroli malam hari memiliki peran yang sangat penting, mengingat banyak pencurian terjadi pada malam hari saat suasana lebih sepi dan banyak orang tidur. Pada malam hari, pelaku pencurian biasanya merasa lebih bebas untuk melakukan aksinya tanpa takut dilihat atau terdengar.

Waktu malam hari hingga dini hari adalah periode yang rentan untuk terjadinya pencurian. Patroli malam bisa mencegah perampokan atau pembobolan rumah dan toko yang cenderung terjadi saat pemilik tidak ada di tempat. Petugas patroli harus mengganti rute patroli secara berkala agar pelaku tidak dapat memprediksi ke mana petugas akan pergi. Memastikan area sekitar rumah, jalanan, atau tempat usaha memiliki pencahayaan yang baik, karena pencuri lebih suka beraksi di tempat yang gelap dan sepi. Pada malam hari, banyak orang yang tidur atau beristirahat, sehingga situasi di sekitar menjadi lebih sepi dan tidak banyak orang yang aktif. Ini memberi kesempatan bagi pelaku untuk bertindak tanpa terlihat atau terdeteksi oleh saksi mata.

KESIMPULAN

Berdasarkan semua penjabaran dari hasil-hasil temuan penelitian di atas berupa hasil prediksi menggunakan algoritma k-NN serta analisis pada bab sebelumnya, maka didapat kesimpulan penelitian sebagai berikut:

1. *Dataset* yang digunakan untuk klasifikasi dan prediksi tindak pidana pencurian dengan algoritma *machine learning* di Polres Buol merupakan *dataset* primer atau *dataset* yang didapatkan dari data korban tindak pidana pencurian di wilayah hukum Polresta Buol pada rentang tahun 2017-2024. Untuk pelatihan dan evaluasi model, dataset dibagi dua bagian: training set dan test set. Pembagian data dilakukan dengan rasio 80:20, yang berarti 80% dari dataset digunakan untuk melatih model dan 20% sisanya digunakan untuk mengujinya.
2. Berdasarkan hasil yang diperoleh maka model terbaik untuk prediksi yaitu menggunakan model K-NN. Hasil dari penelitian ini menunjukkan bahwa algoritma K-Nearest Neighbors (KNN) memiliki performa yang lebih baik dibandingkan dengan algoritma Naive Bayes dalam melakukan klasifikasi data pencurian berdasarkan tempat kejadian, waktu kejadian, dan jenis pencurian. Dalam pengujian yang dilakukan, KNN secara konsisten menghasilkan nilai akurasi, presisi, recall, dan F1-score yang lebih tinggi dibandingkan Naive Bayes. Hal ini menunjukkan bahwa KNN lebih efektif dalam mengenali pola-pola yang terbentuk dari kombinasi waktu dan lokasi kejadian terhadap jenis pencurian yang terjadi. Dari hasil evaluasi menggunakan metrik akurasi, presisi, recall, dan F1-score, diperoleh bahwa performa kedua model bervariasi tergantung pada kompleksitas data. KNN cenderung lebih unggul dalam hal akurasi ketika pola distribusi data relatif teratur dan tidak terlalu beragam, sementara Naive Bayes lebih stabil dan efisien ketika menghadapi dataset yang besar dengan distribusi yang tidak merata.
3. Hasil dari analisis prediksi menunjukkan bahwa sebagian besar kejadian pencurian terjadi pada dini hari. Hal ini memberikan informasi berharga bahwa dini hari menjadi waktu yang lebih rentan terhadap terjadinya pencurian, mungkin karena faktor minimnya pengawasan di lingkungan sekitar atau ketidakhadiran pemilik barang di rumah atau tempat usaha.

REFERENSI

- Afriansyah, M., et al. 2024. ALGORITMA NAIVE BAYES YANG EFISIEN UNTUK KLASIFIKASI BUAH PISANG RAJA BERDASARKAN FITUR WARNA. *Journal of Information Systems Management and Digital Business*, 1(2), 236-248.
- Anggreni, D., & Km, S. (2022). *Buku Ajar-Metodologi Penelitian Kesehatan*. E-Book Penerbit STIKes Majapahit.
- Apriyani, H., & Kurniati, K. 2020. Perbandingan Metode Naïve Bayes Dan *Support Vector Machine* Dalam Klasifikasi Penyakit Diabetes Melitus. *Journal of Information Technology Ampera*, 1(3), 133-143.
- Bawengan, G. W. (Gerson W.). 1994. *Pengantar Psikologi Kriminil / Oleh Gerson W. Bawengan*. Jakarta: Pradnya Paramitha.
- Berkowitz, L. (2006). *The Social Psychology of Crime and Justice*. Springer.
- Chun, S. A., Avinash Paturu, V., Yuan, S., Pathak, R., Atluri, V., & R. Adam, N. 2019. Crime prediction model using deep neural networks. In *Proceedings of the 20th Annual International Conference on digital government research* (pp. 512-514).
- Cornish, D. B., & Clarke, R. V. (1986). *The Reasoning Criminal: Rational Choice Perspectives on Offending*. Springer-Verlag
- Dirjosisworo. 1984. *Ruang Lingkup Kriminalogi*. Jakarta: Rajawali Press.
- Edrisy, I. F., Kamilatun, K., & Putri, A. (2023). *Kriminologi*. Bandar Lampung: Pusaka Media
- Efrain, Turban dkk. (2005). "Decicion Support Systems and Intelligent Systems." Andi Offset.
- Grataridarga, N., & Hum, M. (2019). Konsep Data, Information, Knowledge dan Wisdom (DIKW) Hierarchy pada Manajemen Kearsipan. *JIPi (Jurnal Ilmu Perpustakaan dan Informasi)*, 4(1), 117. <https://doi.org/10.30829/jipi.v4i1.4839>
- Han, J. dan M. Kamber. 2006. *Data mining Concepts and Techniques Second Edition*. San Francisco: Morgan Kaufmann.
- 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: Jhon Wiley & Sons Inc.
- Lintjewas, C. F. 2022. DELIK PENCURIAN YANG DIKUALIFIKASI (DIPERBERAT) DALAM PASAL 363 DAN PASAL 365 KUHP SEBAGAI KEJAHATAN TERHADAP HARTA KEKAYAAN. *LEX CRIMEN*, 11(2).
- Loka, S. K. P., & Marsal, A. (2023). Perbandingan Algoritma K-Nearest Neighbor dan Naïve Bayes Classifier untuk Klasifikasi Status Gizi Pada Balita: Comparison Algorithm of K-Nearest Neighbor and Naïve Bayes Classifier for Classifying Nutritional Status in Toddlers. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 3(1), 8-14.
- Loudoe, John Z. 1982. *Beberapa Aspek Hukum Material Dan Hukum Acara Dalam Praktek*. Jakarta: Bina Aksara.
- Mamluchah, L. (2020). Peningkatan angka kejahatan pencurian pada masa pandemi dalam tinjauan kriminologi dan hukum pidana islam. *Al-Jinayah: Jurnal Hukum Pidana Islam*, 6(1), 1-26.
- Moeljatno. 2008. *Kitab Undang-Undang Hukum Pidana*. Jakarta: Bumi Aksara.
- Nilamsari, Natalina. 2014. "Memahami Studi Dokumen Dalam Penelitian Kualitatif." *Wacana* 13(2): 177-81.
- Nuswantari. 1998. *Kamus Kedokteran Dorland, (Edisi 25)*. Jakarta: EGC.
- Ollenburger, Jane C. & Helen A. Moore. 1996. *Sosiologi Wanita*. Jakarta: Rineka Cipta.
- Permana, A. P., Ainiyah, K., & Holle, K. F. H. 2021. Analisis perbandingan algoritma decision tree, kNN, dan Naive Bayes untuk prediksi kesuksesan start-up. *JISKA (Jurnal Informatika Sunan Kalijaga)*, 6(3), 178-188.

- Putra, Erik Lesmana. "ANALISIS PERTANGGUNGJAWABAN PIDANA TERHADAP PELAKU TINDAK PIDANA PENCURIAN DENGAN PEMBERATAN (Studi Putusan Nomor 1043/Pid. B/2020/PN. Tjk)." (2022).
- Putra, M. Y., & Putri, D. I. (2022). Pemanfaatan Algoritma Naïve Bayes dan K-Nearest Neighbor Untuk Klasifikasi Jurusan Siswa Kelas XI. *Jurnal Tekno Kompak*, 16(2), 176-187.
- R. Soesilo. 1994. *Kitab Undang-Undang Hukum Pidana (KUHP), Serta KomentarKomentar Lengkap Pasal Demi Pasal*. Bogor: Politea.
- Rezki, M., et al. 2020. Analisis Review Pengguna Google Meet dan Zoom Cloud Meeting Menggunakan Algoritma Naïve Bayes. *Jurnal Infortech*, 2(2), 264-270.
- Rich, E. and Knight, K. 1991. *Artificial intelligence*. New York: McGrawHill.
- Saragih, Y. M., & Hadiyanto, A. (2021). *Pengantar Teori Kriminologi & Teori Dalam Hukum Pidana*. Cattleya Darmaya Fortuna.
- Sri Soedewi M. 2000. *Hukum Perdata: Hukum Benda*. Yogyakarta: Liberty.
- Sudjana, Nana. 2004. *Dasar-Dasar Proses Belajar Mengajar*. Bandung: Sinar Baru Algensido Offset.
- Wang, J., et al. 2020. Crime Analysis of spatial-temporal distribution based on KNN Algorithm. In 2020 International Conference on Information Science and Education (ICISE-IE) (pp. 150-154). IEEE.
- Wibowo, A. H., & Oesman, T. I. 2020. The comparative analysis on the *accuracy* of k-NN, Naive Bayes, and Decision Tree Algorithms in predicting crimes and criminal actions in Sleman Regency. In *Journal of Physics: Conference Series* (Vol. 1450, No. 1, p. 012076). IOP Publishing.
- Winston, P. H., & Prendergast, K. A. (1984). *From the Blocks World to the Business World*.
- Wiyli yustanti. (2012). *Algoritma K-Nearest Neighbour untuk memprediksi harga jual tanah*. ISBN:1858-1382.
- Yanto, Robi, and Riri Khoiriah. 2015. "Implementasi *Data mining* Dengan Metode Algoritma Apriori Dalam Menentukan Pola Pembelian Obat." *Creative Information Technology Journal* 2(2): 102