

ANALISIS PERBANDINGAN ALGORITMA ANN DAN KNN DALAM PREDIKSI TINDAK PIDANA PENCURIAN KENDARAAN BERMOTOR DI POLRESTA MALANG KOTA

Widya Bhakti Dira, Jarot Prianggono

Sekolah Tinggi Ilmu Kepolisian PTIK, Kota Jakarta Selatan

Penulis korespondensi: bhakti.dira10@gmail.com

ABSTRAK

Penelitian ini bertujuan untuk membandingkan performa dua algoritma *machine learning*, yaitu *Artificial Neural Network* (ANN) dan *K-Nearest Neighbours* (KNN), dalam memprediksi dengan mengklasifikasi tindak pidana pencurian kendaraan bermotor di Kota Malang. Tingkat kejahatan jalanan, khususnya pencurian kendaraan bermotor, telah menjadi perhatian utama pemerintah setempat, seperti yang terlihat dari catatan Polresta Malang Kota. Dataset yang berhasil dicatat dan dirapihkan sejak tahun 2018 hingga 2023 sejumlah 606. Penelitian ini memanfaatkan dua algoritma *machine learning*, yaitu ANN dan KNN, untuk menganalisis dan memprediksi tindak pidana pencurian kendaraan bermotor (curanmor) di wilayah hukum Polresta Malang Kota. Penelitian ini memanfaatkan prediksi kejahatan yang menggunakan pendekatan *research and development* (R&D) dengan algoritma *machine learning*, seperti ANN dan KNN, untuk mengidentifikasi pola kejahatan, lokasi kejadian, dan faktor-faktor yang mempengaruhi tingkat kejahatan. Metode penelitian melibatkan pengumpulan data, *preprocessing*, pemilihan fitur dan model, pengembangan model, serta validasi dan evaluasi. Data yang digunakan berasal dari catatan Polresta Malang Kota tentang kasus curanmor dari tahun 2018 hingga 2023 sebagai data primer dan menggunakan kuesioner yang diberikan kepada 10 penyidik polisi di Polresta Malang Kota sebagai data sekunder. Untuk memprediksi tingkat kejahatan tindak pidana pencurian kendaraan bermotor di Kota Malang, dapat menggunakan data jenis kelamin, usia korban, kecamatan TKP, waktu kejadian, hari kejadian dan tindak pidana sebagai atribut fitur dan data tingkat kejahatan sebagai atribut target. Kemudian dilakukan normalisasi *preprocessing* dengan melakukan pembersihan dan normalisasi *dataset*. Tingkat kejahatan ditentukan berdasarkan gabungan dari 6 faktor yang berbeda yang digunakan sebagai fitur. Kemudian dihitung jumlah duplikasi data atau berapa kali kejadian dengan 6 faktor itu terjadi. Tingkat kejahatan dianggap tinggi jumlah kejahatan terjadi lebih dari 2 kali. Sementara jika kejahatan terjadi kurang dari 2 kali maka tingkat kejahatan dianggap rendah. Hasil dari pembuatan tingkat kejahatan membuat jumlah *dataset* menyusut dari yang awalnya berjumlah 1005 menjadi 606 karena pada *dataset* yang baru tersebut tidak terdapat duplikasi data untuk kolom yang digunakan sebagai fitur. Dengan menggunakan data yang sudah diolah, penulis dapat memprediksi tingkat kejahatan pencurian kendaraan bermotor di Kota Malang menggunakan dua metode: *Artificial Neural Network* dan *K-Nearest Neighbors*. Hal ini menghasilkan dataset yang lebih kecil dari awalnya, yaitu 606 dari 1005 data. Dalam penelitian ini, eksperimen dilakukan untuk menerapkan nilai $k=7$ dan $k=9$ pada algoritma KNN. Rasio perbandingan antara data *training* dan *testing* adalah 0.6:0.4, 0.7:0.3, 0.8:0.2, dan 0.9:0.1, kemudian dibandingkan untuk mencari model prediksi terbaik antara ANN dan KNN dalam memprediksi tindak kejahatan curanmor. Metode terbaik teridentifikasi sebagai ANN, yang menunjukkan *recall* yang lebih tinggi dengan rasio pembagian data *training* dan *testing* 0.9:0.1. Pada rasio tersebut, dengan 545 data *training* dan 61 data *testing*, nilai akurasi mencapai 96.72%, *recall* 97.50%, presisi 95.65%, dan *f1-score* 96.44%. Menggunakan hasil dapat membantu pihak Polresta Malang Kota dalam perencanaan

dan pelaksanaan tindakan pencegahan dengan alokasi sumber daya yang lebih efektif dan efisien.

Kata Kunci : Prediksi tindak pidana pencurian kendaraan bermotor, *Machine learning*, *Artificial Neural Network*, *K-Nearest Neighbours*

1 PENDAHULUAN

Kota Malang, sebagai pusat perekonomian, menghadapi tingkat kejahatan jalanan yang tinggi, termasuk pencurian kendaraan bermotor (curanmor). Melalui analisis dokumen, diketahui bahwa terdapat sekitar 1005 kasus curanmor yang tercatat dari tahun 2018 hingga 2023. Penanganan dalam kasus curanmor seringkali hanya sebatas penanganan secara reaktif, di mana pihak berwenang hanya mendapatkan laporan dari korban selepas kejadian sudah terjadi. Sehingga dibutuhkan cara pencegahan untuk mengurangi tindak kejahatan tersebut.

Penelitian telah menyoroti perlunya pendekatan preventif dalam menangani masalah ini. Metode machine learning telah terbukti efektif dalam menganalisis dan memprediksi tren kejahatan. Studi oleh Tamir et al. (2021) menggunakan *algoritma Neural Network*, yang berhasil memprediksi kejahatan dengan akurasi 90,77%. Di sisi lain, penelitian oleh Zhang et al. (2020) menunjukkan bahwa algoritma *Recurrent Neural Network* (RNN) memiliki kinerja terbaik dalam analisis deret waktu kejahatan. Penelitian lainnya oleh He & Zheng (2021) mengimplementasikan algoritma *GAN neural network* untuk memetakan lokasi kejahatan di Philadelphia, dengan hasil memprediksi area konsentrasi kejahatan yang akan datang.

Dari penelitian terdahulu, model neural network terbukti unggul dalam memprediksi tren kejahatan, terutama ANN. Meskipun model ini membutuhkan sumber daya lebih besar karena kompleksitasnya, algoritma machine learning seperti *K-Nearest Neighbours* (KNN) juga dapat digunakan, seperti yang dilakukan dalam studi Tamir et al. (2021). Dengan pengetahuan ini, dapat dibangun model *machine learning* yang optimal untuk menangani tindak kejahatan curanmor di Kota Malang, sehingga penegak hukum dapat bertindak secara efektif dan efisien dengan menggunakan data yang relevan. Firmansyah dan Prianggono (2023), dalam penelitiannya yang berjudul *Prediksi Tindak Pidana UU Ite Dengan Metode K-Nn Oleh Personel Polres Metro Bekasi Era Police 4.0*, menggunakan algoritma KNN dengan nilai $K=3, 5, 7$, dan 9 didapat nilai rata-rata performa akurasinya 98.67%. Selanjutnya Hermawan dan Prianggono (2023). Dalam penelitiannya yang berjudul *Crime of theft prediction using Machine Learning K-Nearest Neighbour Algorithm at Polresta Bandar Lampung*, melaporkan dengan menggunakan $K=3$ dan rasio dataset training dan testing adalah 70% berbanding 30%, didapat performa akurasinya 75.71%.

Oleh karena itu, penelitian perbandingan algoritma ANN dan KNN dalam prediksi tindak pidana pencurian kendaraan bermotor di Polresta Malang Kota menjadi penting untuk mengetahui model machine learning yang optimal untuk diterapkan pada kasus penelitian ini. Sehingga menggunakan model yang terbaik dapat meningkatkan keamanan terkait tindak pidana curanmor. Dengan penelitian ini diharapkan tercipta kerangka kerja yang lebih sistematis dalam pencegahan dan penanganan tindak kejahatan curanmor. Hal ini juga akan memberikan kontribusi signifikan bagi pengembangan sistem informasi kepolisian yang lebih fleksibel dan responsif.

2 METODE

2.1.2. Pendekatan Penelitian

Penelitian ini bertujuan untuk memprediksi tindak pidana pencurian kendaraan bermotor di Polresta Malang Kota menggunakan pendekatan yang terstruktur dengan algoritma *machine learning*. Tahap awal melibatkan pengumpulan data kejahatan yang terdokumentasi, termasuk waktu, lokasi kejadian, jenis kejahatan, serta detail pelaku dan korban. Data ini kemudian diolah dan dipersiapkan untuk analisis. Pemilihan fitur-fitur relevan seperti waktu kejadian, koordinat lokasi, pola kejadian, dan karakteristik tindak pidana pencurian kendaraan bermotor dilakukan. Algoritma *machine learning*, khususnya *Artificial Neural Networks* (ANN) dan *K-Nearest Neighbours* (KNN), diterapkan pada data yang telah disiapkan. Evaluasi model menggunakan metrik seperti *precision*, *recall*, dan *F1-score* untuk mengukur kinerja dan akurasi model. Analisis lebih lanjut dilakukan untuk memahami prediksi lokasi kejahatan dan mengidentifikasi pola atau tren khusus dari hasil prediksi tersebut.

2.1.2. Metode Penelitian

Penelitian mengenai perbandingan algoritma ANN dan KNN dalam prediksi tindak pidana pencurian kendaraan bermotor di Polresta Malang Kota mengadopsi metode yang beragam dan terstruktur. Berikut tahapan yang dilakukan dalam metode penelitian ini.

- 1) Pengumpulan data:
- 2) Persiapan data
- 3) Analisis data
- 4) Pengembangan model
- 5) Validasi dan evaluasi model
- 6) Interpretasi hasil
- 7) Pelaporan hasil

2.2. Operasionalisasi Variabel

Penelitian mengenai perbandingan algoritma ANN dan KNN dalam prediksi tindak pidana pencurian kendaraan bermotor di Polresta Malang Kota memerlukan sumber data yang komprehensif dan terstruktur. Sehingga operasional variabel menjadi penting untuk penelitian ini karena berfungsi sebagai fondasi untuk melakukan analisis dan prediksi tindak pidana pencurian kendaraan bermotor. Tujuan operasionalisasi variabel adalah agar variabel tersebut dapat diuji secara empiris sehingga memungkinkan pengumpulan data yang dapat dianalisis untuk menguji hipotesis atau teori yang terkait. Hal ini membantu memastikan kejelasan dalam pengukuran serta memungkinkan peneliti untuk membuat kesimpulan yang didasarkan pada data yang dapat diamati dan diukur secara konkret.

Operasionalisasi variabel pada penelitian ini dijabarkan sebagai berikut.

- a. Jenis Kelamin Korban
- b. Usia Korban

Menurut Kementerian Kesehatan dalam Al Amin usia dapat diklasifikasikan sebagai berikut (Hakim, L. N. 2020).

- i) Masa balita mulai dari umur 0 sampai umur 5 tahun.
- ii) Masa kanak – kanak mujlai dari umur 5 sampai umur 10 tahun.
- iii) Masa remaja awal mulai dari umur 12 sampai umur 16 tahun.
- iv) Masa remaja akhir mulai dari umur 17 sampai umur 25 tahun.
- v) Masa dewasa awal mulai dari umur 26 sampai umur 35 tahun.
- vi) Masa dewasa akhir mulai dari umur 36 sampai umur 45 tahun.
- vii) Masa lansia awal mulai dari umur 46 sampai umur 55 tahun.
- viii) Masa lansia akhir mulai dari umur 56 sampai umur 65 tahun.

- ix) Masa manula mulai dari umur lebih dari 65 tahun.
- c. Pekerjaan Korban
 Pengelompokan jenis pekerjaan di Indonesia telah disusun dalam KBJI 2014 baik pekerjaan formal ataupun informal. Berikut ringkasan golongan pokok dalam KBJI 2014.
- i) Tentara Nasional Indonesia (TNI) dan Kepolisian Negara Republik Indonesia (POLRI).
 - ii) Manajer
 - iii) Professional
 - iv) Teknisi dan Asisten Profesional
 - v) Tenaga Tata Usaha
 - vi) Tenaga Usaha Jasa dan Tenaga Penjualan
 - vii) Pekerjaan Terampil Pertanian, Kehutanan dan Perikanan
 - viii) Pekerja Pengolahanm, Kerajinan, dan Tbd
 - ix) Operator dan Perakit Mesin
 - x) Pekerja Kasar
- d. Tempat Kejadian Perkara
 Menurut petunjuk lapangan No. Pol: Skep/1205/IX/2000 Tempat Kejadian Perkara atau yang disingkat TKP dibagai menjadi dua yaitu (Sinaga, L. V., & Simatupang, M. Y. M. 2020):
- Tempat dimana tindak pidana terjadi atau menimbulkan akibatnya.
 - Tempat lain yang berhubungan dengan tindak pidana tersebut seperti lokasi barang bukti, lokasi penangkapan tersangka, atau lokasi korban ditemukan.
- Menurut Peraturan Kepala Kepolisian Negara Republik Indonesia Nomor 6 Tahun 2010 Pasal 1 ayat (19) Tentang Manajemen Penyidikan Oleh Penyidik Pegawai Negeri Sipil, TKP yaitu “tempat dimana suatu tindak pidana dilakukan/terjadi dan tempat-tempat lain, dimana tersangka dan/atau korban dan/atau barang bukti yang berhubungan dengan tindak pidana tersebut dapat ditemukan.”
- e. Tempat Tinggal Korban
 Tempat tinggal korban dapat diklasifikasikan berdasarkan tingkat hunian sesuai dengan Peraturan Menteri Negara Perumahan Rakyat Nomor: 11/PERMEN/M/2008 tentang Pedoman Keserasian Kawasan Perumahan dan Permukiman. Klasifikasi tingkat hunian atau kepadatan penduduk ditunjukkan pada Tabel 2.1 sebagai berikut.

Tabel 2.1 Klasifikasi tingkat hunian

No.	Klasifikasi Tingkat Hunian	Jumlah Penduduk Dalam Jiwa/Ha	Jumlah Rumah Dalam Unit/Ha
1	Zona lindung	0	0
2	Zona pedesaan	<50	≤15
3	Zona pinggiran kota	51-100	≤25
4	Zona perkotaan	101-300	≤75
5	Zona pusat kota	301-500	≤125
6	Zona pusat metro	>501	≤300
7	Zona preservasi	Sesuai ketentuan daerah masing – masing	

- f. Waktu Kejadian Peristiwa
 Penelitian ini menggunakan indikator waktu kejadian peristiwa terjadi. Berdasarkan Kamus Besar Bahasa Indonesia waktu dikasifikasikan Pagi (04.00-10.59), Siang (11.00-13.59), Sore / Petang (14.01-17.59), Malam (18.00-00.00), dan Dini hari (00.01-03.59).

- g. Hari Kejadian Peristiwa
Penelitian ini menggunakan variabel hari kejadian peristiwa. Berdasarkan Kamus Besar Bahasa Indonesia satu hari merupakan waktu dari pagi sampai pagi lagi atau 24 jam. Sementara dalam periode satu minggu terdapat tujuh hari.
- h. Modus Operandi
Menurut Modus operandi merupakan metode operasi atau cara prosedur; terutama pola perilaku kriminal yang perilaku kriminal yang sangat khas sehingga para penyelidik mengaitkannya dengan pekerjaan orang yang sama (Supriyadi, A. 2022).
- i. Akibat dari Peristiwa
Akibat yang pasti terjadi dari peristiwa kendaraan bermotor yaitu kerugian bagi korban. Selain berdampak pada harta benda korban tindak pidana pencurian kendaraan bermotor juga berdampak keresahaan bagi masyarakat sekitar (Mufidah, A. U. 2021).
- j. Jenis Tindak Pidana
Jenis tindak pidana pada penelitian ini merupakan variabel yang berisi keterangan jenis tindak pidana yang dilakukan oleh pelaku. Variabel ini condong ke jenis kendaraan bermotor yang dicuri oleh pelaku.

2.3. Populasi dan Sampel

Populasi dari penelitian tentang perbandingan algoritma ANN dan KNN dalam prediksi tindak pidana pencurian kendaraan bermotor di Polresta Malang Kota berupa data 1005 data tindak pidana pencurian kendaraan bermotor yang bersumber dari data internal dan eksternal Polresta Malang Kota. Data diambil mulai dari tahun 2018 sampai dengan tahun 2023. Sampel yang dimaksud dalam penelitian ini merupakan sebagian atau semua populasi yang memiliki korelasi tertinggi serta relevansi terhadap tujuan penelitian sehingga data yang digunakan dapat menghasilkan model machine learning yang optimal.

2.4. Teknik Pengumpulan Data

Dalam penelitian tentang perbandingan algoritma ANN dan KNN dalam prediksi tindak pidana pencurian kendaraan bermotor di Polresta Malang Kota, teknik pengumpulan data menjadi kunci. Beberapa teknik pengumpulan data yang bisa digunakan antara lain:

- a) Pengumpulan data dari sumber internal Polresta Malang Kota
 - 1) Catatan Kejahatan: Data kejahatan dari catatan internal polisi menjadi sumber utama. Ini mencakup jenis kejahatan, waktu, lokasi, dan detail tentang pelaku dan korban.
 - 2) Laporan Polisi: Laporan resmi mengenai kejahatan dan insiden tertentu yang terjadi di wilayah Polresta Malang Kota.
- b) Pengumpulan data dari sumber eksternal:
 - a. Sistem Pengadilan: Data dari sistem pengadilan seperti putusan pengadilan terkait kasus-kasus kejahatan tertentu.
 - b. Sumber Media dan Sosial: Informasi dari berita atau platform media sosial yang dapat memberikan gambaran lebih luas tentang kejahatan di wilayah tersebut.
- c) Teknik pengumpulan data secara langsung yang dilakukan melalui survei langsung atau wawancara dengan aparat kepolisian untuk mendapatkan wawasan yang lebih mendalam tentang kasus-kasus spesifik atau strategi penanganan kejahatan.

2.5. Teknik Analisis Data

Teknik analisis data dalam penelitian tersebut melibatkan serangkaian langkah penting yang dapat digunakan untuk memproses, menginterpretasi, dan mengidentifikasi pola atau tren terkait tindak pidana pencurian kendaraan bermotor di Kota Malang. Berikut adalah rincian teknik analisis data yang dapat digunakan dalam penelitian ini ini:

- a. Pembersihan data
Langkah awal adalah membersihkan data dari potensi kesalahan, duplikasi, atau nilai yang hilang. Ini termasuk menangani missing values, standardisasi format data, dan memastikan keakuratan informasi.
- b. Transformasi data
Proses ini mencakup pengubahan format data ke format yang sesuai untuk analisis machine learning. Misalnya, mengonversi data waktu dan lokasi menjadi format yang dapat dipahami oleh algoritma.
- c. Eksplorasi data
Analisis eksploratori digunakan untuk memahami karakteristik data. Grafik, ringkasan statistik, dan visualisasi digunakan untuk mengidentifikasi pola, outlier, atau tren yang menarik perhatian.
- d. *Feature engineering*
Merupakan langkah penting untuk mengekstrak fitur-fitur yang relevan dari data. Misalnya, mengidentifikasi fitur-fitur kunci dari data waktu dan lokasi, seperti pola kejadian di jam tertentu atau daerah dengan tingkat kejahatan tinggi.
- e. Model *machine learning*
Penggunaan algoritma *Artificial Neural network* dan *K-Nearest Neighbours* untuk prediksi kejahatan. Ini termasuk pemilihan model, pembagian data untuk pelatihan dan pengujian, serta evaluasi kinerja model menggunakan metrik seperti akurasi, presisi, recall, dan F1-score.
- f. Analisis spasial dan temporal
Analisis spasial memungkinkan identifikasi pola spasial kejahatan di Kota Malang, sementara analisis temporal membantu dalam memahami pola kejahatan seiring waktu.
- g. Identifikasi pola dan tren
Setelah penerapan model *machine learning*, analisis dilakukan untuk mengidentifikasi pola atau tren yang terungkap dari hasil prediksi. Misalnya, apakah terdapat peningkatan kejahatan pada jam tertentu atau pola spesifik di area geografis tertentu.

3. Hasil Penelitian

3.1. Karakteristik Data Penelitian

Karakteristik data penelitian merujuk pada atribut yang digunakan dalam penelitian tersebut. Dalam metode machine learning tahap preprocessing merupakan tahap yang harus dilakukan agar kualitas data yang digunakan untuk pembuatan model dalam kondisi optimal. Salah satu yang dilakukan pada *preprocessing* adalah pemilihan atribut. Atribut atau fitur adalah ciri-ciri dari data yang dipakai untuk meramalkan kelas atau nilai output. Pemilihan atribut yang paling relevan sangat penting agar prediksi menjadi akurat. Sementara itu, kelas atau nilai output merupakan data yang akan diklasifikasikan dan diprediksi. Atribut dan kelas yang dimanfaatkan dalam penelitian ini ditunjukkan pada Tabel 3.1.

Atribut Kelas/Target dihasilkan dari jumlah setiap kombinasi dari 6 Atribut/Fitur. Jika jumlah suatu kombinasi diatas 2 maka tingkat kejahatan akan diberi label tinggi dan jika suatu kombinasi dibawah 2 maka akan diberi label rendah.

Tabel 3.1 Atribut dan kelas data penelitian

No	Atribut/Fitur	Kelas/Target
1	Jenis Kelamin	Tingkat
2	Usia Korban	Kejahatan
3	Kecamatan TKP	
4	Waktu Kejadian	
5	Hari Kejadian	
6	Tindak Pidana	

3.2.1 Data Hasil Penelitian

Data yang digunakan pada penelitian ini berasal dari laporan korban tindak kejahatan curanmor di Poresta Malang Kota tahun 2018 - 2023 dengan jumlah 1005 data/kasus dan memiliki rincian untuk setiap tahunnya yaitu, 173, 155, 208, 167 dan 153 data/kasus dengan rata – rata data/kasus setiap tahunnya yaitu 167.5 data/kasus. Puncak kasus curanmor di wilayah hukum Polresta Malang Kota terdapat pada tahun 2020 dengan jumlah 208 data/kasus memiliki selisih 40.5 data/kasus jika dibandingkan rata - rata per-tahun. Berikut penjelasan karakteristik setiap atribut dari data laporan korban tindak kejahatan curanmor di Poresta Malang Kota tahun 2018 -2023.

a. Karakteristik data jenis kelamin korban

Dari Tabel 3.2, dapat disimpulkan jika pada rentang tahun 2018 – 2023 mayoritas korban memiliki jenis kelamin laki – laki dengan jumlah 691 kasus dari 1005 kasus. Hal tersebut kemungkinan disebabkan mayoritas pengguna kendaraan bermotor adalah laki – laki

Tabel 3.2 Data jenis kelamin korban

No	Jenis Kelamin Korban	Tahun						Total
		2018	2019	2020	2021	2022	2023	
1	Laki - Laki	122	118	134	114	92	111	691
2	Perempuan	51	37	74	53	57	42	314
	Total	173	155	208	167	149	153	1005

Tabel 3.3 Data usia korban

No	Usia Korban	Tahun						Total
		2018	2019	2020	2021	2022	2023	
1	Balita	0	0	0	0	0	1	1
2	Kanak - Kanak	0	0	0	0	0	0	0
3	Remaja Awal	1	33	0	1	0	1	36
3	Remaja Akhir	77	22	109	62	75	87	432
4	Dewasa Awal	49	0	40	54	33	34	210
5	Dewasa Akhir	21	69	32	25	19	10	176
6	Lansia Awal	21	24	16	14	16	11	102
7	Lansia Akhir	4	7	10	9	6	8	44
8	Manula	0	0	1	2	0	1	4
	Total	173	155	208	167	149	153	1005

b. Karakteristik data usia korban

Pada Tabel 3.3, dapat disimpulkan jika pada rentang tahun 2018 – 2023 mayoritas korban memiliki usia remaja akhir dengan jumlah 432 kasus dari 1005 kasus. Dari tabel tersebut mayoritas korban berusia remaja akhir yaitu rentang 17 – 25 tahun dengan persentase 48%, disusul dengan dewasa akhir awal 24% atau setengah dari data remaja akhir. Hal tersebut

dikarenakan untuk usia tersebut korban sudah dapat melakukan pembuatan surat ijin mengemudi. Sehingga mayoritas korban yang menggunakan kendaraan bermotor adalah orang yang memiliki surat ijin mengemudi.

c. Karakteristik data kecamatan TKP

Pada Tabel 3.4, dapat disimpulkan jika pada rentang tahun 2018 – 2023 mayoritas kejadian curanmor terjadi di Kecamatan Sukun dengan 286 kasus dari 1005 kasus. Jika durutkan dari kasus yang terbanyak adalah Kecamatan Sukun, Kecamatan Blimbing, dan Kecamatan Lowokwaru dengan persentase diatas 20%.

Tabel 3.4 Data kecamatan TKP

No	Kecamatan TKP	Tahun						Total
		2018	2019	2020	2021	2022	2023	
1	Blimbing	34	43	39	30	19	34	199
2	Kedungkandang	26	15	31	28	33	24	157
3	Klojen	30	26	56	20	6	26	164
4	Lowokwaru	46	28	33	41	16	35	199
5	Sukun	37	43	49	48	75	34	286
	Total	173	155	208	167	149	153	1005

d. Karakteristik data waktu kejadian

Dalam data yang ditunjukkan pada Tabel 3.5, dapat disimpulkan jika pada rentang tahun 2018-2023 mayoritas kejadian curanmor terjadi pada malam hari dengan 267 kasus dari 1005 kasus. Apabila kasus pada tahun 2018-2023 di Polresta Malang Kota durutkan dari yang terbanyak adalah malam dan dini hari dengan presentase yaitu 28% dan 24%. Hal tersebut dikarenakan pada rentang waktu tersebut mayoritas penduduk dan korban saat dalam kondisi beristirahat. Sehingga pelaku dapat dengan mudah melakukan pencurian kendaraan bermotor dikarenakan kemungkinana besar pada rentang waktu tersebut sedikit atau bahkan tidak ada orang yang melakukan aktivitas yang dapat mengganggu pelaku melakukan aksi pencurian kendaraan bermotor.

e. Karakterisitik data hari kejadian

Pada Tabel 3.6, dapat disimpulkan jika pada rentang tahun 2018 – 2023 mayoritas kejadian curanmor terjadi pada hari senin dengan 164 kasus dari 1005 kasus. Dari data pada tabel, dapat dilihat hari kejadian paling sering terjadi pada hari senin. Namun untuk hari lainnya hanya memiliki selisihh jumlah yang sedikit dengan presentase disekitar 13% - 16%. Sehingga untuk hari kejadian tindak pidana pencurian kendaraan bermotor dapat terjadi pada hari apa aja.

f. Karakteristik data jenis tindak pidana

Penelitian menggunakan data laporan korban tindak pidana pencurian kendaraan bermotor di Polresta Malang Kota pada tahun 2018 – 2023. Sehingga untuk jenis tindak pidana yang terjadi hanya tindak pidana pencurian kendaraan bermotor (curnamor).

Tabel 3.5 Data waktu kejadian

No	Waktu Kejadian	Tahun						Total
		2018	2019	2020	2021	2022	2023	
1	Pagi	41	29	50	32	6	36	194
2	Siang	14	17	20	28	16	12	107
3	Sore	22	32	28	26	75	15	198
3	Malam	50	32	52	51	33	49	267
4	Dini Hari	46	45	58	30	19	41	239
	Total	173	155	208	167	149	153	1005

Tabel 3.6 Data hari kejadian

No	Waktu Kejadian	Tahun						Total
		2018	2019	2020	2021	2022	2023	
1	Senin	25	29	29	29	28	24	164
2	Selasa	25	24	30	20	20	25	144
3	Rabu	31	22	38	32	15	18	156
4	Kamis	21	19	25	25	20	21	131
5	Jum'at	24	18	26	22	21	22	133
6	Sabtu	22	13	30	25	30	23	143
7	Minggu	25	30	30	14	15	20	134
Total		173	155	208	167	149	153	1005

3.2.2 Data Hasil Responden

Selain menggunakan data laporan korban tindak kejahatan curanmor di Poresta Malang Kota tahun 2018 -2023, peneliti juga melakukan survei dengan kuesioner yang dibagikan kepada 10 responden penegak hukum di wilayah hukum Polresta Malang Kota. Identitas responden yang mengisi survei tentang tindak kejahatan pencurian kendaraan bermotor ditunjukkan pada Tabel.

Tabel 3.7 Identitas responden

Nama	Pangkat	Jabatan	Masa Dinas
Aji Yulembariono	Bripka	Anggota Lesrolub	14 tahun
Ardi Guruh Wijayanto	Aipda	Anggota Satreskrim	-
Harsono	Aiptu	P.S Kasubnit 2 unit 6	27 tahun
Heri Nurcahyo	Aiptu	P.S Kasubnit 2 unit 5 Satreskrim	25 tahun
Imam Subekti	Aiptu	Kasubnit 6 unit 1	29 tahun
Jajang Eri Bowo	Ipda	P.S Kanit Resmob	22 tahun
Lisa Merisa	Brigadir	BA Satreskrim	9 tahun
Nindi Asatullah	Brigpol	BA Satreskrim	10 tahun
Purwanto	Aiptu	Anggota Reskrim	24 tahun
Sutomo	Bripka	Anggota Satreskrim	11 tahun

Hasil survei mendapatkan data dan atribut yang cukup beragam. Namun dengan mengikuti atribut yang akan digunakan pada penelitian yang ditunjukkan pada Tabel 3.6, penulis hanya mengambil sebagian data yang relevan untuk penelitian. Berikut karakteristik data yang dihasilkan dari survei menggunakan kuesioner.

a. Karakteristik data responden jenis kelamin korban

Dari data responden, korban tindak pidana pencurian kendaraan bermotor mayoritas berjenis kelamin laki – laki dengan presentase 75%. Menurut mereka laki – laki lebih banyak menggunakan kendaraan bermotor dikarenakan untuk keperluan pekerjaan dikarenakan laki – laki merupakan tulang punggung keluarga.

b. Karakteristik data responden usia korban

Korban tindak pidana pencurian kendaraan bermotor menurut para responden mayoritas memiliki usia dewasa awal yaitu rentang 26 – 35 tahun dengan presentase 59% disusul dengan remaja awal yaitu rentang 17 - 25 tahun. Hal tersebut dikarenakan untuk usia tersebut korban sudah dapat melakukan pembuatan surat ijin mengemudi dan pada rentang usia tersebut mayoritas korban sedang dalam masa produktif. Sehingga penggunaan kendaraan bermotor pada rentang usia tersebut cenderung lebih banyak karena membutuhkan alat transportasi yang efektif.

- c. Karakteristik data responden kecamatan TKP
Kecamatan yang sering terjadi tindak pidana pencurian kendaraan bermotor menurut para responden adalah kecamatan Lowokwaru. Menurut mereka kecamatan tersebut merupakan pusat kegiatan perekonomian dan pendidikan. Berbanding lurus dengan usia korban pada rentang 17 - 35 tahun yang mayoritas berprofesi sebagai pelajara/mahasiswa dan pekerja.
- d. Karakteristik data responden waktu kejadian
Waktu yang sering terjadi tindak pidana pencurian kendaraan bermotor menurut para responden jika durutkan dari yang terbanyak adalah malam dan dini hari dengan presentase secara berturut – turut yaitu 77% dan 15%. Sama seperti alasan pada karakteristik yang ada pada data laporan korban tindak kejahatan curanmor di Poresta Malang Kota tahun 2018 -2023 mereka juga menyatakan pada rentang waktu tersebut mayoritas penduduk dan korban sedang dalam kondisi beristirahat. Sehingga pelaku dapat dengan mudah melakukan pencurian kendaraan bermotor dikarenakan kemungkinana besar pada rentang waktu tersebut sedikit atau bahkan tidak ada orang yang melakukan aktivitas yang dapat mengganggu pelaku melakukan aksi pencurian kendaraan bermotor.
- e. Karakterisitik data responden hari kejadian
Hari yang sering terjadi tindak pidana pencurian kendaraan bermotor menurut para responden jika durutkan dari yang terbanyak adalah hari jum'at, sabtu, dan minggu atau hari sekitaran weekend dengan presentase 24%. Hal tersebut kemungkinan pada hari – hari tersebut mayoritas korban sedang tidak melakukan pekerjaan atau beristirahat. Sehingga korban kemungkinan sedang dalam kondisi tidak waspada terhadap tindak pidana curanmor.
- f. Karakteristik data responden jenis tindak pidana
Survei yang dilakukan hanya seputar tidak pidana penucian kendaraan pencurian kendaraan bermotor. Sehingga untuk jenis tindak pidana yang terjadi hanya tindak pidana pencurian kendaraan bermotor.

3 HASIL DAN PEMBAHASAN

3.1 Prediksi Tindak Pidana Pencurian Kendaraan Bermotor Menggunakan *Machine learning* di Polresta Malang Kota

Prediksi tindak pidana pencruian kendaraan bermotor menggunakan *machine learning* di Polresta Malang Kota menggunakan metode *Artificial Neural network* dan *K-Nearest Neighbours* akan dilakukan meggunakan aplikasi microsoft Excel dan Rapidminer. Penggunaan *microsoft excel* dilakukan untuk pengolahan data sebelum dilakukan pembuatan model menggunakan Rapidmier.

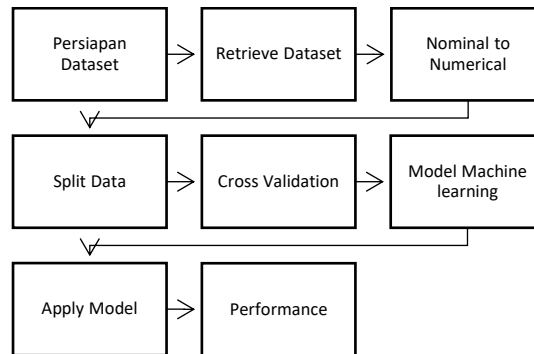
3.1.1 *Preprocessing*

Preprocessing merupakan tahap pembersihan data agar data yang digunakan untuk pelatihan model dalam kondisi optimal. *Dataset* yang digunakan merupakan *dataset* laporan korban tindak pidana pencurian kendaraan bermotor di Poresta Malang Kota dengan total 1005 data yang kemudian akan dijadikan data primer pada penelitian. Selanjutnya, dilakukan pemelihan atribut yang sesuai dengan objek penelitian. Atribut yang dipilih ditunjukkan pada Tabel 3.1. Kolom yang tidak ada pada Tabel 3.1 akan dihapus karena tidak memiliki korelasi dengan penelitian ini. *Dataset* setelah dilakukan pemilihan atribut ini kemudian dilakukan pembersihan dan normalisasi data seperti merubah *value* data yang *typo* (salah) dan melakukan normalisasi usia korban kedalam bentuk range usia. Kemudian menggunakan *dataset* yang ada dilakukan penghapusan duplikasi data. Sehingga data memiliki kombinasi 606 data. Kemudian dilakukan perhitungan setiap kombinasi yang berada pada data awal atau 1005 data untuk

membuat kolom target atau label. Jika kombinasi sama dengan 1 maka diberi *value* rendah. Sedangkan jika kombinasi diatas 2 diberi *value* tinggi.

Tabel 4.1 Data tingkat kejahatan

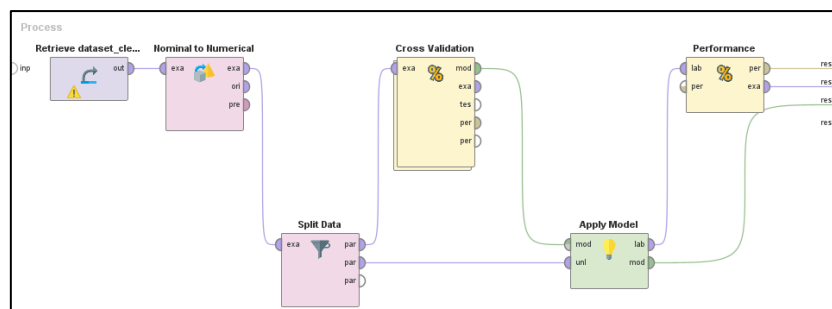
Tingkat Kejahatan	Jumlah
Rendah	358
Tinggi	248



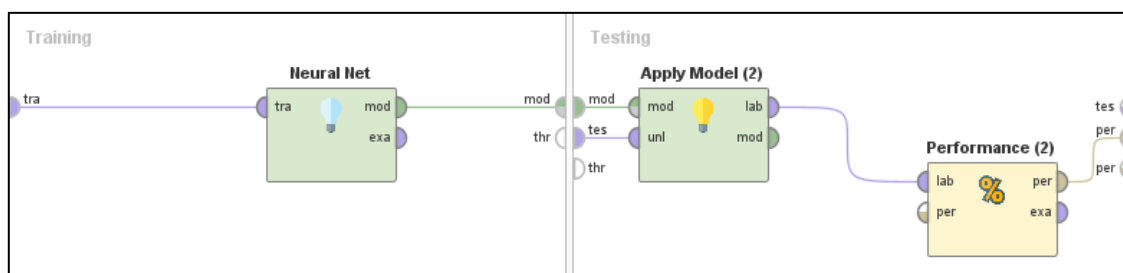
Gambar 4.1 Desain Processing

3.2 Metode Artificial Neural network

Desain metode *Artificial Neural network* seperti yang ditunjukkan pada Gambar 4.2 dimulai dari pengambilan *dataset* menggunakan *retrieve* dari *local repository*, perubahan *dataset* nominal to numerical, *split* data dengan skenario rasio yang ditentukan, *cross validation* menggunakan data *training*, *apply* model menggunakan hasil *cross validation* dan data *testing*, dan evaluasi *performance* dari hasil *apply* model. Dalam desain metode ANN tersebut, menggunakan operasi *cross validation* untuk evaluasi kinerja model yang dibuat. Desain *cross validation* yang digunakan ditunjukkan pada Gambar 4.3.



Gambar 4.2 Desain metode ANN



Gambar 4.3 Desain *cross validation*

Berdasarkan desain *cross validation* yang ditunjukkan pada Gambar 4.2, operator yang digunakan yaitu *neural net*, *apply model* dan *performance*. *Neural net* atau model ANN yang akan digunakan merupakan model default dari Rapidminer yang memiliki parameter *confidence layers* dengan jumlah 2, *training cycles* dengan nilai 200, *learning rate* dengan nilai 0.01, *momentum* dengan nilai 0.9, dan *error epsilon* 1.0E-4. Berikut hasil pengujian metode ANN menggunakan beberapa skenario yang sudah ditetapkan

- a. Rasio data *training* 0.6 dan data *testing* 0.4
Pengujian pertama dilakukan menggunakan skenario 0.6 untuk data *training* dan 0.4 untuk data *testing* dengan rincian data sebesar 364 data *training* dan 242 data *testing*. Akurasi metode ANN dengan rasio pembagian data 06:04 untuk data *training* dan *testing* yaitu 93.8%. Menggunakan data *testing* 242 data, total data yang diprediksi benar yaitu 227 data. Hasil *recall*, *precision*, dan *f1-score* untuk masing – masing kelas yaitu rendah *recall* 94.97%, *precision* 95.57%, *f1-score* 95.27%; dan tinggi *recall* 91.57%, *precision* 90.48% dan *f1-score* 91.02%.
- b. Rasio data *training* 0.7 dan data *testing* 0.3
Pengujian kedua dilakukan menggunakan skenario 0.7 untuk data *training* dan 0.3 untuk data *testing* dengan rincian data sebesar 424 data *training* dan 182 data *testing*. Akurasi metode ANN dengan rasio pembagian data 07:03 untuk data *training* dan *testing* yaitu 97.25%. Menggunakan data *testing* 182 data, total data yang diprediksi benar yaitu 179 data. Hasil *recall*, *precision*, dan *f1-score* untuk masing – masing kelas yaitu rendah *recall* 99.15%, *precision* 96.67%, *f1-score* 97.89%; dan tinggi *recall* 93.85%, *precision* 98.39% dan *f1-score* 96.07%.
- c. Rasio data *training* 0.8 dan data *testing* 0.2
Pengujian ketiga dilakukan menggunakan skenario 0.8 untuk data *training* dan 0.2 untuk data *testing* dengan rincian data sebesar 485 data *training* dan 121 data *testing*. Akurasi metode ANN dengan rasio pembagian data 08:02 untuk data *training* dan *testing* yaitu 95.04%. Menggunakan data *testing* 121 data, total data yang diprediksi benar yaitu 115 data. Hasil *recall*, *precision*, dan *f1-score* untuk masing – masing kelas yaitu rendah *recall* 94.81%, *precision* 97.33%, *f1-score* 96.05%; dan tinggi *recall* 95.45%, *precision* 91.30% dan *f1-score* 93.33%.
- d. Rasio data *training* 0.9 dan data *testing* 0.1
Pengujian keempat dilakukan menggunakan skenario 0.9 untuk data *training* dan 0.1 untuk data *testing* dengan rincian data sebesar 454 data *training* dan 61 data *testing*. Hasil *recall*, *precision*, dan *f1-score* untuk masing – masing kelas yaitu rendah *recall* 95%, *precision* 100%, *f1-score* 97.44%; dan tinggi *recall* 100%, *precision* 91.30% dan *f1-score* 95.45%.

Hasil pengujian model ANN tidak hanya menghasilkan confusion matrix, tetapi juga menampilkan struktur model ANN yang terdiri dari 1 confidence layer dengan 11 node, 17 node input, dan 3 output. Model dengan rasio pembagian data training dan testing 0.6:0.4 memiliki akurasi terendah, tetapi meningkat secara signifikan pada rasio 0.7:0.3. Namun, terjadi penurunan kembali pada rasio 0.8:0.2, mungkin menandakan overfitting karena model terlalu cocok dengan data training. Peningkatan signifikan pada pembagian 0.7:0.3 menunjukkan bahwa model awalnya underfitting dan peningkatan jumlah data training membantu model untuk belajar dengan lebih baik.

Evaluasi model ANN menunjukkan bahwa untuk kelas rendah, performa model kurang stabil saat rasio pembagian data training dan testing di atas 0.6:0.4, dengan penurunan performa terutama pada recall dan f1-score pada pembagian 0.8:0.2, menandakan adanya overfitting. Namun, precision tetap tinggi karena fitur-fitur yang digunakan mampu membedakan kelas rendah secara akurat, meskipun model cenderung overfitting. Untuk kelas tinggi, hasilnya

data 07:03 untuk data *training* dan *testing* yaitu 90.11%. Menggunakan data *testing* 182 data, total data yang diprediksi benar yaitu 164 data. Hasil *recall*, *precision*, dan *f1-score* untuk masing – masing kelas yaitu rendah *recall* 92.31%, *precision* 82.31%, *f1-score* 87.02%; tinggi *recall* 86.15%, *precision* 86.15% dan *f1-score* 86.15%.

C. Rasio data *training* 0.8 dan data *testing* 0.2

Pengujian pertama dilakukan menggunakan skenario nilai k 7, rasio pembagian data 0.8 untuk data *training* dan 0.2 untuk data *testing* dengan rincian data sebesar 485 data *training* dan 121 data *testing*. Akurasi metode KNN dengan nilai k 7, rasio pembagian data 07:03 untuk data *training* dan *testing* yaitu 92.56%. Menggunakan data *testing* 121 data, total data yang diprediksi benar yaitu 112 data. Hasil *recall*, *precision*, dan *f1-score* untuk masing – masing kelas yaitu rendah *recall* 92.21%, *precision* 95.95%, *f1-score* 94.04%; dan tinggi *recall* 93.18%, *precision* 87.23% dan *f1-score* 90.11%.

D. Rasio data *training* 0.9 dan data *testing* 0.1

Pengujian pertama dilakukan menggunakan skenario nilai k 7, rasio pembagian data 0.9 untuk data *training* dan 0.1 untuk data *testing* dengan rincian data sebesar 545 data *training* dan 61 data *testing*. Akurasi metode KNN dengan nilai k 7, rasio pembagian data 09:01 untuk data *training* dan *testing* yaitu 90.16%. Menggunakan data *testing* 61 data, total data yang diprediksi benar yaitu 55 data. Hasil *recall*, *precision*, dan *f1-score* untuk masing – masing kelas yaitu rendah *recall* 90%, *precision* 94.74%, *f1-score* 92.31%; dan tinggi *recall* 90.48%, *precision* 82.61% dan *f1-score* 86.37%.

Hasil pengujian menunjukkan bahwa implementasi KNN dengan nilai k=7 pada rasio 0.8:0.2 menghasilkan peningkatan signifikan dalam performa akurasi model, sementara model dengan rasio 0.7:0.3 memiliki akurasi terendah. Kemungkinan model pada rasio 0.6:0.4 dan 0.7:0.3 mengalami underfitting karena penggunaan data training yang terlalu sedikit, sedangkan rasio di atas 0.8:0.2 cenderung mengalami overfitting, mendekati akurasi rasio 0.7:0.3. Hasil evaluasi untuk kelas rendah menunjukkan ketidakstabilan model terutama pada rasio 0.7:0.3, sedangkan untuk kelas tinggi model cenderung stabil dengan peningkatan performa pada rasio 0.8:0.2, namun mengalami penurunan pada rasio 0.9:0.1, kemungkinan karena overfitting. Dengan demikian, rasio 0.8:0.2 dianggap optimal untuk implementasi dataset pada penelitian ini.

Hasil evaluasi *recall*, *precision* dan *f1-score* KNN dengan implementasi nilai k=7 menunjukkan ketidakstabilan model untuk kelas rendah, terutama pada pembagian data training dan testing dengan rasio 0.7:0.3. Implementasi nilai k=7 mungkin kurang sesuai untuk kasus tersebut, di mana jumlah tetangga terdekat yang digunakan untuk prediksi tidak cocok dengan struktur data, terutama untuk kelas rendah. Kenaikan performa tidak signifikan dibandingkan dengan penurunan yang cukup jauh. Namun, untuk kelas tinggi, model cenderung stabil pada berbagai rasio pembagian data training dan testing, meskipun mengalami penurunan pada rasio 0.9:0.1, kemungkinan karena overfitting.

Tabel 4.3 Summary pengujian KNN nilai k=7

Rasio	accuracy %	Class	recall %	precision %	f1-score %
0.6:0.4	90.91	Rendah	93.08	93.08	93.08
		Tinggi	86.75	86.75	86.75
0.7:0.3	90.11	Rendah	92.31	82.31	87.02
		Tinggi	86.15	86.15	86.15
0.8:0.2	92.56	Rendah	92.21	95.95	94.04
		Tinggi	93.18	87.23	90.11
0.9:0.1	90.16	Rendah	90	94.74	92.31

Tinggi	90.48	82.61	86.37
--------	-------	-------	-------

3.3.2 Metode *K-Nearest Neighbours* Nilai $K=9$

1. Rasio data *training* 0.6 dan data *testing* 0.4
 Pengujian pertama dilakukan menggunakan skenario nilai $k=9$, rasio pembagian data 0.6 untuk data *training* dan 0.4 untuk data *testing* dengan rincian data sebesar 364 data *training* dan 242 data *testing*. Akurasi metode KNN dengan nilai $k=9$, rasio pembagian data 06:04 untuk data *training* dan *testing* yaitu 90.91%. Menggunakan data *testing* 242 data, total data yang diprediksi benar yaitu 220 data. Hasil *recall*, *precision*, dan *f1-score* untuk masing – masing kelas yaitu rendah *recall* 93.08%, *precision* 93.08%, *f1-score* 93.08%; dan tinggi *recall* 86.75%, *precision* 86.75% dan *f1-score* 86.75%.
2. Rasio data *training* 0.7 dan data *testing* 0.3
 Pengujian pertama dilakukan menggunakan skenario nilai $k=9$, rasio pembagian data 0.7 untuk data *training* dan 0.3 untuk data *testing* dengan rincian data sebesar 424 data *training* dan 182 data *testing*. Akurasi metode KNN dengan nilai $k=9$, rasio pembagian data 07:03 untuk data *training* dan *testing* yaitu 92.86%. Menggunakan data *testing* 182 data, total data yang diprediksi benar yaitu 169 data. Hasil *recall*, *precision*, dan *f1-score* untuk masing – masing kelas yaitu rendah *recall* 94.87%, *precision* 94.07%, *f1-score* 94.47%; dan tinggi *recall* 89.23%, *precision* 90.62% dan *f1-score* 89.92%.
3. Rasio data *training* 0.8 dan data *testing* 0.2
 Pengujian pertama dilakukan menggunakan skenario nilai $k=9$, rasio pembagian data 0.8 untuk data *training* dan 0.2 untuk data *testing* dengan rincian data sebesar 485 data *training* dan 121 data *testing*. Akurasi metode KNN dengan nilai $k=9$, rasio pembagian data 07:03 untuk data *training* dan *testing* yaitu 94.21%. Menggunakan data *testing* 121 data, total data yang diprediksi benar yaitu 114 data. Hasil *recall*, *precision*, dan *f1-score* untuk masing – masing kelas yaitu rendah *recall* 92.21%, *precision* 98.61%, *f1-score* 95.30%; dan tinggi *recall* 97.73%, *precision* 87.76% dan *f1-score* 92.48%.
4. Rasio data *training* 0.9 dan data *testing* 0.1
 Pengujian pertama dilakukan menggunakan skenario nilai $k=9$, rasio pembagian data 0.9 untuk data *training* dan 0.1 untuk data *testing* dengan rincian data sebesar 545 data *training* dan 61 data *testing*. Akurasi metode KNN dengan nilai $k=9$, rasio pembagian data 09:01 untuk data *training* dan *testing* yaitu 93.44%. Menggunakan data *testing* 61 data, total data yang diprediksi benar yaitu 57 data. Hasil *recall*, *precision*, dan *f1-score* untuk masing – masing kelas yaitu rendah *recall* 90%, *precision* 100%, *f1-score* 94.74%; dan tinggi *recall* 100%, *precision* 84% dan *f1-score* 91.30%.

Hasil pengujian model KNN dengan nilai $k=9$, menunjukkan peningkatan secara bertahap dalam performa model dari rasio 0.6:0.4 hingga 0.8:0.2. Namun, terjadi penurunan pada rasio 0.9:0.1. Peningkatan bertahap ini mungkin menandakan bahwa model awalnya underfitting dan penambahan data training membantu meningkatkan performa. Namun, penurunan akurasi pada rasio 0.9:0.1 mungkin mengindikasikan overfitting karena model menjadi terlalu rumit dan terlalu sesuai dengan data training.

Hasil evaluasi *recall*, *precision* dan *f1-score* KNN dengan implementasi nilai $k=9$ untuk kelas rendah menunjukkan stabilitas model dalam pembagian data training dan testing di atas rasio 0.6:0.4 hingga 0.7:0.3, Namun, pada rasio 0.8:0.2, model mulai tidak stabil dengan *recall* yang menurun dan *precision* yang meningkat, menunjukkan tanda-tanda overfitting karena model terlalu cocok dengan data training. Meskipun demikian, nilai *precision* tetap tinggi karena fitur-fitur yang digunakan mampu membedakan kelas rendah dengan akurat, sehingga meskipun ada kecenderungan overfitting, kemampuan model dalam membedakan kelas rendah tetap tinggi.

Hasil evaluasi *recall*, *precision* dan *f1-score* KNN dengan implementasi nilai $k=9$ untuk kelas tinggi juga menunjukkan stabilitas model dalam pembagian data training dan testing dengan rasio 0.6:0.4 hingga 0.7:0.3. Namun, pada rasio 0.8:0.2, model mengalami penurunan nilai *precision* namun *recall* meningkat, menunjukkan tanda-tanda overfitting. Meskipun model cenderung overfitting, kemampuannya dalam menangkap informasi penting dari fitur-fitur tetap menghasilkan *recall* yang tinggi untuk kelas tinggi.

Tabel 4.4 Summary pengujian KNN nilai $k=9$

Rasio	<i>accuracy</i> %	Class	<i>recall</i> %	<i>precision</i> %	<i>f1-score</i> %
0.6:0.4	90.91	Rendah	93.08	93.08	93.08
		Tinggi	86.75	86.75	86.75
0.7:0.3	92.86	Rendah	94.87	94.07	94.47
		Tinggi	89.23	90.62	89.92
0.8:0.2	94.21	Rendah	92.21	98.61	95.30
		Tinggi	97.73	87.76	92.48
0.9:0.1	93.44	Rendah	90	100	94.74
		Tinggi	100	84	91.30

3.4 Evaluasi Model *Machine learning*

Evaluasi model dilakukan menggunakan nilai *accuracy*, *recall*, *precision*, dan *f1-score* dari setiap skenario pada model *Artificial Neural network* dan *K-Nearest Neighbours*. Namun dikarenakan tidak semua kelas pada setiap skenario menghasilkan nilai *recall*, *precision*, dan *f1-score* maka akan dilakukan perhitungan *macro-average* pada nilai *recall*, *precision*, dan *f1-score* setiap kelas dengan menghitung rata – rata dari nilai tersebut. Sehingga menggunakan hasil perhitungan tersebut dapat ditentukan model terbaik yang dapat diimplementasikan. Berikut untuk hasil *macro-average* setiap skenario pada model ANN dan KNN.

Berdasarkan hasil pengujian yang ditunjukkan pada Tabel 5.5 metode ANN menggunakan pembagian data 0.9:0.1 merupakan metode terbaik yang dapat diterapkan untuk memprediksi tindak pidana pencurian kendaraan bermotor di Polresta Malang Kota. Dengan mempertimbangkan nilai *recall* dikarenakan nilai tersebut lebih valid karena ketika tingkat kejahatan rendah di prediksi tinggi lebih aman dibandingkan sebaliknya. Metode ANN dengan skenario tersebut menghasilkan nilai *accuracy* 96.72%, *recall* 97.50%, *precision* 95.65%, dan *f1-score* 96.44% untuk nilai *macro-average*.

Tabel 4.5 Summary pengujian model *machine learning*

Metode	Rasio	<i>accuracy</i> %	<i>recall</i> %	<i>precision</i> %	<i>f1-score</i> %
ANN	0.6:0.4	93.80	93.27	93.03	93.15
	0.7:0.3	97.25	96.50	97.53	96.98
	0.8:0.2	95.04	95.13	94.32	94.69
	0.9:0.1	96.72	97.50	95.65	96.44
KNN (k = 7)	0.6:0.4	90.91	89.92	89.92	89.92
	0.7:0.3	90.11	89.23	84.23	86.59
	0.8:0.2	92.56	92.70	91.59	92.07
	0.9:0.1	90.16	90.24	88.68	89.34
KNN (k = 9)	0.6:0.4	90.91	89.92	89.92	89.92
	0.7:0.3	92.86	92.05	92.35	92.19
	0.8:0.2	94.21	94.97	93.19	93.89
	0.9:0.1	93.44	95.00	92.00	93.02

Penurunan performa saat implementasi model KNN terutama untuk nilai $k=7$. Pada implementasi $k=7$ dapat dilihat jika menghasilkan model menghasilkan performa yang kurang stabil dengan grafis yang fluktuatif. Sementara untuk implementasi $k=9$ cenderung lebih stabil dengan peningkatan performa secara bertahap pada setiap rasio pembagian data. Hal tersebut bisa terjadi karena perbedaan kompleksitas model jika dibandingkan dengan $k=9$. Sehingga model KNN dengan nilai $k=7$ dapat dikatakan terlalu sederhana untuk kasus prediksi menggunakan data kejahatan tindak pidana pencurian kendaraan bermotor di Kota Malang. Namun jika dibandingkan dengan model ANN, model KNN memiliki performa yang lebih rendah. Itu dimungkinkan karena pendekatan model ANN menggunakan jaringan saraf tiruan yang lebih superior daripada pendekatan jarak antar tetangga pada model KNN.

3.5 Analisis Pencegahan Tindak Pidana Pencurian Kendaraan Bermotor

Menggunakan data hari kejadian, waktu kejadian, kecamatan TKP dan tingkat kejahatan dapat ditentukan strategi untuk pencegahan tindak pencurian kendaraan bermotor. Berikut beberapa pendekatan kriminologi untuk tindakan pencegahan yang dapat dilakukan berdasarkan data yang diperoleh.

3.6 Pendekatan Deskriptif

Strategi pencegahan yang diusulkan didasarkan pada pengamatan dan pengumpulan data terkait kejahatan, seperti hari kejadian, waktu kejadian, dan lokasi (kecamatan TKP). Ini mencerminkan pendekatan deskriptif dalam kriminologi yang melibatkan pemahaman tentang karakteristik kejahatan dan pelaku, serta cara kejahatan dilakukan. Dengan mengetahui pola-pola ini, tindakan pencegahan yang spesifik dapat diambil, seperti larangan parkir pada waktu dan lokasi tertentu.

3.7 Pendekatan Kausalitas

Strategi pencegahan juga mencoba memahami hubungan sebab-akibat dari kejahatan. Misalnya, meningkatkan kehadiran pihak berwenang pada malam dan dini hari untuk mengurangi kesempatan bagi pelaku kejahatan. Ini mencerminkan pendekatan kausalitas dalam kriminologi yang mempelajari hubungan antara faktor-faktor tertentu, seperti kehadiran pihak berwenang, dengan tingkat kejahatan.

3.8 Pendekatan Normatif

Melakukan pendidikan terhadap masyarakat dan mendorong partisipasi mereka dalam pos ronda merupakan contoh strategi yang mencerminkan pendekatan normatif. Ini berfokus pada fakta-fakta dan cara-cara untuk mengatasi masalah kejahatan dengan melibatkan masyarakat secara aktif dalam upaya pencegahan.

3.9 Pendekatan Geografi Kejahatan

Pendekatan Geografi Kejahatan dalam penelitian memberikan pemahaman yang mendalam tentang pola kejahatan secara spasial dan temporal, khususnya dalam konteks pencegahan pencurian kendaraan bermotor. Metode analisis seperti analisis kluster digunakan untuk mengidentifikasi area rentan dengan kecenderungan kejahatan yang tinggi, memungkinkan untuk menentukan target pencegahan yang lebih efektif. Penggunaan Sistem Informasi Geografis (SIG) memungkinkan pemetaan data spasial untuk menyoroti daerah-daerah dengan tingkat pencurian kendaraan bermotor yang tinggi, memungkinkan perancangan strategi pencegahan yang lebih efektif seperti peningkatan patroli atau penempatan kamera pengawas. Pendekatan ini konsisten dengan pendekatan deskriptif dan kausalitas dalam menentukan strategi pencegahan kejahatan yang lebih efektif.

3.10 Pendekatan Analisis Data Temporal dan Spasial

Pendekatan ini menggabungkan informasi waktu dan lokasi kejahatan untuk mengidentifikasi pola-pola yang signifikan dalam data. Dalam konteks pencurian kendaraan bermotor, analisis data temporal dapat membantu dalam mengetahui pola-pola waktu kejadian yang rentan terhadap tindak pencurian tersebut, seperti pada malam hari atau saat keramaian rendah.

Selain itu, dengan mempertimbangkan data spasial, seperti kecamatan TKP, dapat diidentifikasi daerah-daerah yang sering menjadi target pencurian kendaraan bermotor. Dengan pemahaman yang lebih mendalam tentang pola temporal dan spasial kejahatan ini, strategi pencegahan yang spesifik dapat dirancang, seperti peningkatan kehadiran pihak berwenang pada waktu-waktu dan lokasi-lokasi yang rentan.

4 SIMPULAN

Hasil penelitian menunjukkan bahwa algoritma Artificial Neural Network (ANN) dan K-Nearest Neighbours (KNN) **dapat digunakan** untuk memprediksi berdasarkan klasifikasi kejahatan tindak pidana pencurian kendaraan bermotor di Kota Malang. Hasil uji performance terbaik dengan algoritma ANN menggunakan rasio dataset training : testing sebesar 70%:30% didapat akurasi 97,25 %. Untuk uji performansi algoritma NB terbaik, dengan K=9 dan rasio dataset training : testing sebesar 80%:20% didapat akurasi 94,21 %. Hasil evaluasi kedua algoritma tersebut peneliti menentukan untuk persoalan prediksi dengan klasifikasi tingkat kejahatan pencurian kendaraan bermotor di Kota Malang, peneliti memilih algoritma ANN sebagai solusinya. Hal ini mengingat performance algoritma yang dihasilkannya lebih tinggi disbanding dengan algoritma KNN. Pencegahan tindak pidana pencurian kendaraan bermotor di Kota Malang dapat dilakukan dengan memanfaatkan hasil prediksi tingkat kejahatan, data waktu dan lokasi kejadian, serta data kecamatan TKP untuk mendukung upaya pencegahan yang lebih efektif oleh pihak berwajib. Model *machine learning* dengan algoritma ANN dapat dijadikan sebagai alat bantu dalam pengambilan keputusan pimpinan di Polresta Malang Kota. Adapun saran pada penelitian ini adalah: 1) Polresta Malang Kota dapat menggunakan hasil prediksi tingkat kejahatan pencurian kendaraan bermotor yang dihasilkan dari model terbaik yaitu Artificial Neural Network (ANN) sebagai alat bantu dalam perencanaan dan pelaksanaan tindakan pencegahan kejahatan curanmor di Kota Malang. Dengan memanfaatkan data hasil prediksi dan faktor-faktor yang berkontribusi pada tingkat kejahatan, pihak berwenang dapat mengalokasikan sumber daya secara lebih efektif dan efisien untuk mengurangi kejadian pencurian kendaraan bermotor; 2) Untuk masa depan, penting untuk terus meningkatkan kualitas dan akurasi prediksi dengan mempertimbangkan penggunaan teknik-teknik pemrosesan data yang lebih canggih dan model-model machine learning yang lebih kompleks. Selain itu, mengintegrasikan data dari sumber-sumber yang lebih luas dan mendalam seperti data cuaca, pola lalu lintas, dan faktor-faktor lingkungan sosial juga dapat meningkatkan kemampuan prediksi untuk masa depan; 3) Pada penelitian mendatang diharapkan dapat menggunakan dataset yang lebih besar dan mempertimbangkan data lokasi dan waktu yang lebih rinci agar dapat dilakukan prediksi tingkat kejahatan tindak pidana pencurian kendaraan bermotor di Kota Malang secara tepat dan akurat.

DAFTAR PUSTAKA

Aldhira, Erlitsya Maharani (2023) Efektivitas Penanganan Tindak Pidana Pencurian Kendaraan Bermotor Roda Dua (Studi di Wilayah Hukum Polres Metro, Lampung). Fakultas Hukum, Universitas Lampung.

- Adi Firmansyah dan Jarot Prianggono. Prediksi Tindak Pidana UU Ite Dengan Metode K-Nn Oleh Personel Polres Metro Bekasi Era Police 4.0. *Jurnal of Comprehensive Science-Vol 2 no 5* (2023). <https://doi.org/10.59188/jcs.v2i5.341>
- Febry Hermawan dan Jarot Prianggono.. Crime of theft prediction using Machine Learning K-Nearest Neighbour Algorithm at Polresta Bandar Lampung. *Jurnal Teknik Informatika-Sinkron*. Vol 7 no 3 (2023). DOI: 10.33395/sinkron.v8i3.12422.
- Aryanti, W., & Rifai, N. A. K. (2023). Penerapan *Artificial Neural network* dengan Algoritma Backpropagation untuk Memprediksi Harga Saham. *Jurnal Riset Statistika*, 107-118.
- Bagus, K. T. A. G. I., & Subawa, G. (2021). Upaya Pencegahan Dan Pemberantasan Tindak Pidana Pencurian Kendaraan Bermotor Dalam Perspektif KUHP di Ditreskrim Polda Bali. *Jurnal Hukum Mahasiswa*, 1(1).
- Bhardwaj, A. S., dkk, (2019). Deep learning architectures for crime occurrence detection and prediction. *Int. J. Advance Res., Ideas Innov. Technol*, 5(2), 822-824.
- BPS Kota Malang (2023). Kota Malang Dalam Angka, Malang Municipality in Figures 2023. ©BPS KOTA MALANG/BPS-Statistics of Malang Municipality
- Etriyanti, E. (2021). Perbandingan Tingkat Akurasi Metode Knn Dan Decision Tree Dalam Memprediksi Lama Studi Mahasiswa. *Jurnal Ilmiah Binary STMIK Bina Nusantara Jaya Lubuklinggau*, 3(1), 6-14.
- Fitriani, E., Sampara, S., & Agis, A. (2020). Pelaksanaan Diversi Dalam Sistem Peradilan Pidana Anak: Studi Kota Makassar. *Journal of Lex Generalis (JLG)*, 1(2), 268-280.
- Hakim, L. N. (2020). Urgensi Revisi Undang-Undang tentang Kesejahteraan Lanjut Usia. *Aspirasi: Jurnal Masalah-Masalah Sosial*, 11 (1), 43–55.
- He, J., & Zheng, H. (2021). Prediction of crime rate in urban neighborhoods based on *machine learning*. *Engineering Applications of Artificial Intelligence*, 106, 104460.
- Hidayatullah, A. (2020). Perbedaan kecenderungan perilaku nomophobia ditinjau berdasarkan jenis kelamin pada mahasiswa uin ar-raniry banda aceh (Doctoral dissertation, UIN AR-RANIRY).
- I Gusti Ngurah Parwata. (2017). *Kriminologi*. Denpasar: Universitas Udayana.
- Ilham, M., & Wibisono, B. H. (2023). Pola Spasial Kejahatan Pencurian Berdasarkan Aspek Temporal di Kecamatan Kadia: Spatial Pattern of Theft Crime Based on Temporal Aspects in Kadia District. *SPECTA Journal of Technology*, 7(3), 711-722.
- Indonesia, S. (2014). Klasifikasi Baku Jabatan Indonesia. Kementerian Ketenagakerjaan dan Badan Pusat Statistik.
- Ingre, B., & Yadav, A. (2015). Performance analysis of NSL-KDD dataset using ANN. In 2015 international conference on signal processing and communication engineering systems (pp. 92-96). IEEE.
- Jesia Quader Yuki, dkk. (2019). Predicting crime using time and location data. In Proceedings of the 7th International Conference on Computer and Communications Management (pp. 124-128).
- Karabo Jenga, Cagatay Catal dan Gorkem Kar. (2022). *Machine learning* in Crime Prediction. *Journal of Ambient Intelligence and Humanized Computing* (2023) 14:2887–2913.
- Khairi, A., Ghozali, A. F., & Hidayah, A. D. N. (2021). Implementasi K-Nearest Neighbor (KNN) Untuk Klasifikasi Masyarakat Pra Sejahtera Desa Sapikerep Kecamatan Sukapura. vol, 2, 319-323.
- Khatun Sameya, dkk. (2023). *Machine learning* based Advanced Crime Prediction and Analysis. International Conference on Sustainable Computing and Data Communication Systems (ICSCDS), Erode, India, 90-96. doi: 10.1109/ICSCDS56580.2023.10104655.
- Kurniawan, Y. I., & Barokah, T. I. (2020). Klasifikasi Penentuan Pengajuan Kartu Kredit Menggunakan K-Nearest Neighbor. *Jurnal Ilmiah Matrik*, 22(1), 73-82.

- Kusuma, P. D. (2020). *Machine learning* Teori, Program, dan Studi Kasus. Deepublish.
- Lestary, R. M. I. (2021). Perilaku Konsumtif Ditinjau dari Jenis Kelamin dan Tingkat Ekonomi Orang Tua pada Mahasiswa Universitas Islam Negeri Ar-Raniry Banda Aceh (Doctoral dissertation, UIN Ar-Raniry).
- Maifriandi, A. (2021). Analisis pelaksanaan pembangunan desa di Desa pulau komang sentajo kecamatan Sentajoraya kabupaten kuantan singingi Tahun2018. *Juhanperak*, 2(1), 55-73.
- Maulidah, dkk. (2020). Algoritma Klasifikasi Decision Tree Untuk Rekomendasi Buku Berdasarkan Kategori Buku. *E-Bisnis: Jurnal Ilmiah Ekonomi dan Bisnis*, 13(2), 89-96.
- Mufidah, A. U. (2021). Tinjauan kriminologis tentang pencurian kendaraan bermotor selama pandemi di Polres Malang. *Dinamika*, 27(10), 1405-1423.
- Mulyana, D. I., & Sumarsono, A. S. (2023). Penerapan Metode *Neural Network* Dengan Struktur Backpropagation Untuk Memprediksi Kebutuhan Stok Pada Toko Umkm Perlengkapan Bayi Babyqu. *Jurnal Indonesia: Manajemen Informatika dan Komunikasi*, 4(1), 121-128.
- Musid, F. N., & Dikananda, A. R. (2023). Implementasi Algoritma K-Means Clustering Dalam Pengelompokkan Data Jumlah Kerusakan Rumah Berdasarkan Kondisi Di Jawa Barat. *Journal of Student Research*, 1(3), 101-114.
- Muzammil Khan, Azmat Ali, and Yasser Alharbi. (2022). Predicting and Preventing Crime: A Crime Prediction Model Using San Francisco Crime Data by Classification Techniques. *Hindawi Complexity*, Volume 2022, Article ID 4830411, 13 pages.
- Nadir, R. A., & Sukmana, R. N. (2023). Sistem Prediksi Harga Emas Berdasarkan Data Time Series Menggunakan Metode *Artificial Neural network* (ANN). *Digital Transformation Technology*, 3(2), 426-437. <https://doi.org/10.47709/digitech.v3i2.2877>
- Nasution, H. M., & Cipta, H. (2023). Analisis Spasial dan Temporal Data Kejadian Bencana Banjir dengan Model Generalized Space Time Autoregressive Integrated Moving Average (GSTARIMA). *Jurnal Absis: Jurnal Pendidikan Matematika dan Matematika*, 6(1), 810-825.
- Nuzula, F. F., dkk. (2021). Perbedaan fear of failure pada mahasiswa bidikmisi ditinjau dari jenis kelamin. *Proyeksi: Jurnal Psikologi*, 16(2), 174-182.
- Peraturan Kepala Kepolisian Negara Republik Indonesia Nomor 6 Tahun 2010 Tentang Manajemen Penyidikan Oleh Penyidik Pegawai Negeri Sipil.
- Peraturan Menteri Negara Perumahan Rakyat Nomor: 11/PERMEN/M/2008 Tentang Pedoman Keresasian Kawasan Perumahan dan Permukiman
- Pratama, B. P., Sembiring, S., & Febrina, Y. (2023). Wild Race by Teenagers in Padang City in the Aspect of Criminology. *Greenation International Journal of Law and Social Sciences*, 1(1), 35-43.
- Puspitorini, dkk. (2023). Analisis Klasifikasi Kasus Tindak Pidana Pencurian Dengan Pohon Keputusan Menggunakan Algoritma C4. 5 (Studi Kasus Polsek Telanaipura). *FORTECH (Journal of Information Technology)*, 7(1), 19-25.
- Sinaga, L. V., & Simatupang, M. Y. M. (2020). Fungsi Olah Tempat Kejadian Perkara (Tkp) Guna Mengungkapkan Kasus Penganiayaan Berat Ditinjau Dari Sudut Hukum Acara Pidana. *JURNAL RECTUM: Tinjauan Yuridis Penanganan Tindak Pidana*, 2(2), 129-136.
- Soesilo, R. (1995). *Kitab Undang-Undang Hukum Pidana (KUHP): Serta Komentar-Komentarnya Lengkap Pasal Demi Pasal*.
- Somayeh Shojaee, Aida Mustapha, Fatimah Sidi, Marzanah A. Jabar. 2013. A Study on Classification Learning Algorithms to Predict Crime Status. *International Journal of Digital Content Technology and its Applications (JDCTA)*, 7(9).

- Subhan, S. (2021). Klasifikasi Konten Web Radikal di Indonesia Menggunakan Web Content Mining dan Algoritma K-Nearest Neighbor. *Jurnal Informasi, Sains dan Teknologi*, 4(2), 70-77.
- Supriyadi, A. (2022). Telaah Terhadap Norma Gratifikasi Dengan Modus Operandi Pemberian Kenaikan Jabatan. *Adil: Jurnal Hukum*, 13(1), 61-77.
- Sya'ban, D. R., Hamzah, A., & Susanti, E. (2022). Klasifikasi Buah Segar dan Busuk Menggunakan Algoritma Convolutional *Neural network* dengan TFLite sebagai Media Penerapan Model *Machine learning*. PROSIDING SNAST, F7-16.
- Syafie, I. K. (2022). Ilmu pemerintahan. Bumi Aksara.
- Tamir, A., dkk, (2021). Crime Prediction and Forecasting using *Machine learning* Algorithms. *International Journal of Computer Science and Information Technologies*, 12(2), 26-33.
- Tisnawan, R., & Marsal, M. (2020). Sistem Informasi Geografis Indikasi Tindak Kriminal di Kecamatan Ujungbatu Rokan Hulu. *Rabit: Jurnal Teknologi dan Sistem Informasi Univrab*, 5(2), 117-129.
- Vaddang, W. (2023). Tinjauan Kriminologis Terhadap Opzet Pemasangan Pagar Listrik Oleh Petani Yang Mengakibatkan Matinya Orang (Studi Kasus di Kabupaten Luwu Tahun 2018-2020) (Doctoral dissertation, Universitas Hasanuddin).
- Varun Mandalapu, dkk. (2023). Crime Prediction Using *Machine learning* and Deep Learning: A Systematic Review and Future Directions *IEEE Access*, Volume 11.
- Wu Shaobing, dkk. (2020). Crime prediction using data mining and *machine learning*. The 8th International Conference on Computer Engineering and Networks (CENet2018). Springer International Publishing, 360-375.
- Zhang, X., dkk, (2020). Comparison of *machine learning* algorithms for predicting crime hotspots. *IEEE access*, 8, 181302-181310.
- “Data BPS” dalam <https://malangkota.bps.go.id/>. 3 April 2024
- ”Keadaan Geografi Kota Malang” dalam <https://malangkota.go.id/sekilas-malang/geografis/>. 29 Januari 2024
- “*Job Discription* Satuan Reserse Kriminal” dalam <https://tribratanews.malangkota.jatim.polri.go.id/satreskrim/>. 31 Januari 2024
- “Memahami Confusion Matrix: Accuracy, Precision, Recall, Specificity, dan F1-Score untuk Evaluasi Model Klasifikasi” dalam <https://esairina.medium.com/memahami-confusion-matrix-accuracy-precision-recall-specificity-dan-f1-score-610d4f0db7cf>. 5 Februari 2024