

BAB II

LANDASAN TEORI

2. Kekerasan & Pengaduan Masyarakat

Kekerasan adalah tindakan yang dilakukan individu atau kelompok yang menyebabkan kerugian fisik, psikologis, atau emosional terhadap orang lain. kekerasan dapat dikategorikan menjadi kekerasan fisik, verbal, seksual, dan emosional. Di Indonesia, kasus kekerasan sering kali tidak dilaporkan secara resmi, sehingga memerlukan sistem pelaporan yang mudah diakses. (Anjelita, 2020)

Pengaduan masyarakat merupakan proses pelaporan kejadian tertentu, seperti tindak kekerasan atau pelanggaran hukum, kepada pihak yang berwenang, dalam hal ini kepolisian. Pengaduan berbasis web memberikan kemudahan bagi masyarakat untuk melaporkan kejadian tanpa harus datang langsung ke kantor polisi, sekaligus meningkatkan efisiensi penanganan laporan. (Anjelita, 2020)

2.1 Data Mining

2.1.1 Pengertian Data Mining

Data mining adalah suatu proses pengeringan atau pengumpulan informasi penting dari suatu data yang besar. Proses data mining seringkali menggunakan metode statistika, matematika, hingga memanfaatkan teknologi artificial intelligence.(Anjelita, 2020)

Nama alternatifnya yaitu Knowledge discovery (mining) in databases (KDD), knowledge extraction, data/pattern analysis, data archeology, data dredging, information harvesting, business intelligence, dan lain-lain. (Anjelita, 2020)

Jika di lihat di lihat pada gambar dalam proses KDD tersebut, Banyak konsep dan teknik yang di gunakan dalam proses data mining. Proses tersebut membutuhkan beberapa langkah untuk mendapatkan sebuah data yang di inginkan. (Anjelita 2020)

Dalam proses KDD tersebut termasuk melakukan pembersihan data, integrasi data, seleksi data, transformasi, penambangan data, evaluasi pola, dan presentasi pengetahuan. (Anjelita , 2020)

2.1.2 Fungsi Data Mining

Data mining memiliki banyak sekali fungsi, Untuk fungsi utamanya sendiri yaitu ada dua; Yaitu fungsi descriptive dan fungsi predictive. (Anjelita, 2020)

Untuk fungsi lainnya akan di bahas di bawah

1. *Descriptive*

fungsi deskripsi dalam data mining adalah sebuah fungsi untuk memahami lebih jauh tentang data yang di amati. Dengan melakukan sebuah proses di harap bisa mengetahui perilaku dari sebuah data tersebut. Data tersebut itulah yang nantinya dapat di gunakan untuk mengetahui karakteristik dari data yang dimaksud.

Dengan menggunakan Fungsi descriptive Data mining, Maka nantinya bisa menemukan pola tertentu yang tersembunyi dalam sebuah data. Dengan kata lain jika pola yang berulang dan bernilai itulah karakteristik sebuah data bisa diketahui.

2. *Predictive*

Fungsi prediksi merupakan sebuah fungsi bagaimana sebuah proses nantinya akan menemukan pola tertentu dari suatu data. Pola-pola tersebut dapat di ketahui dari berbagai variabel-variabel yang ada pada data.

Ketika sudah menemukan pola, Maka pola yang didapat tersebut bisa digunakan untuk memprediksi variabel lain yang belum diketahui nilai ataupun jenisnya.

Karena itulah fungsi satu ini dikatakan sebagai fungsi prediksi sama halnya dengan melakukan predictive analisis. Fungsi ini juga bisa digunakan untuk memprediksi sebuah variabel tertentu yang tidak ada dalam suatu data.

Sehingga fungsi ini memudahkan dan menguntungkan bagi siapapun yang memerlukan prediksi yang akurat untuk membuat hal penting tersebut menjadi lebih baik.

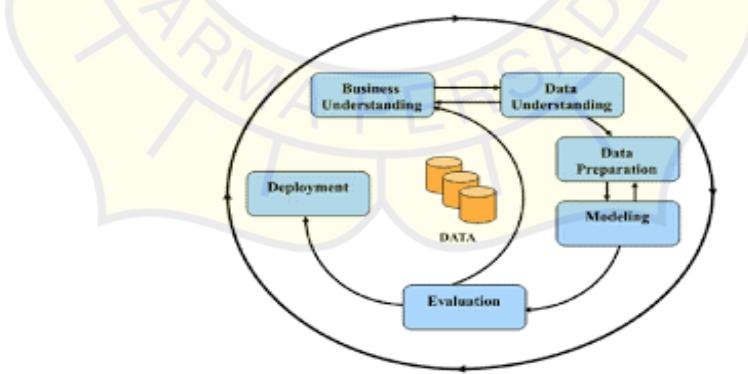
Fungsi Data mining yang lainnya yaitu : characterization, discrimination, association, classification, clustering, outlier and trend analysis, dll.

1. *Multidimensional concept description*, Karakterisasi dan diskriminasi, Atau berfungsi untuk Menggeneralisasikan, meringkas, dan membedakan karakteristik data, dll.
2. *Frequent patterns, association, correlation*
3. *Classification and prediction*, Membangun model (fungsi) yang menggambarkan dan membedakan kelas atau konsep untuk prediksi masa depan. Misalnya, Mengklasifikasikan negara berdasarkan (iklim), atau mengklasifikasikan mobil berdasarkan (jarak tempuh gas)

4. ***Cluster analysis***, Membuat data grup untuk membentuk kelas baru. Misalnya, Memaksimalkan kesamaan intra-kelas & meminimalkan kesamaan antar kelas
5. ***Outlier analysis***, Objek data yang tidak sesuai dengan perilaku umum dari data, Berguna dalam deteksi penipuan, analisis peristiwa langka.
6. ***Trend and evolution analysis***, Trend dan penyimpangan: misalnya Analisis regresi atau Mining Penambangan pola berurutan: misalnya, Kamera digital, atau Analisis periodisitas dan Analisis berbasis kesamaan.

2.1.3 Metode Data Mining

Dalam melakukan pengumpulan informasi tentunya terdapat metode, Metode tersebut akan membantu dalam proses menemukan data. (Fadilla Putri & Ekadiansyah, 2022) Data mining akan menyediakan perencanaan dari ide hingga implementasi akhir.



2.1.3 Gambar Crisp-dm 1

Langkah-langkah dalam menggunakan metode Data Mining CRISP-DM:

1. Business Understanding :

Tahap ini bertujuan untuk memahami kebutuhan pengguna secara menyeluruh. Aktivitas yang dilakukan meliputi penetapan tujuan bisnis, analisis situasi terkait ketersediaan sumber daya, perumusan tujuan pengumpulan data, serta penyusunan rencana proyek yang sistematis.

2. Data Understanding :

Tahap ini berfokus pada pemahaman dan eksplorasi data yang akan digunakan dalam proyek. Langkah-langkah yang dilakukan meliputi pengumpulan data awal, deskripsi karakteristik data, eksplorasi lebih lanjut terhadap pola atau anomali dalam data, serta verifikasi kualitas data guna memastikan kesesuaian dengan tujuan analisis.

3. Data Preparation :

Sering kali disebut sebagai tahap "pemrosesan data", fase ini bertujuan untuk menyiapkan data yang akan digunakan dalam pemodelan. Aktivitas yang dilakukan mencakup pembersihan data, transformasi variabel, serta penyusunan dataset akhir yang siap digunakan dalam proses pemodelan berikutnya..

4. Modeling :

Pada tahap ini, berbagai metode pemodelan diterapkan dan dievaluasi untuk menentukan model yang paling sesuai dengan

tujuan proyek. Aktivitas utama meliputi pemilihan teknik pemodelan, perancangan strategi pengujian, pembangunan model, serta evaluasi terhadap performa model yang dihasilkan.

5. *Evaluation* :

Tahap evaluasi bertujuan untuk meninjau hasil model secara lebih luas guna memastikan kesesuaian dengan kebutuhan bisnis. Proses ini mencakup analisis hasil model, peninjauan ulang terhadap keseluruhan proses, serta penentuan langkah strategis yang perlu diambil untuk tahap selanjutnya.

6. *Deployment* :

Sebagai tahap akhir, deployment bertujuan untuk mengimplementasikan model ke dalam sistem atau proses bisnis yang sesungguhnya. Perencanaan tahap ini telah dimulai sejak fase pemahaman bisnis dan tidak hanya berfokus pada penerapan model, tetapi juga memastikan bahwa hasil analisis dapat diterjemahkan menjadi keputusan yang dapat diintegrasikan dalam operasional bisnis.

2.1.4 Triple Exponential Smoothing

Algoritma Triple Exponential Smoothing adalah metode peramalan yang sering digunakan dalam analisis data deret waktu untuk menangkap pola musiman dan tren jangka panjang. (Fadilla Putri & Ekadiansyah, 2022)

Cara kerja algoritma Triple Exponential Smoothing adalah sebagai berikut:

- a. **Inisialisasi nilai parameter:** Tentukan parameter smoothing untuk level (*alpha*), tren (*beta*), dan musiman (*gamma*).
- b. **Hitung komponen level:** Menghitung level saat ini dengan mempertimbangkan nilai aktual dan nilai level sebelumnya.
- c. **Hitung komponen tren:** Menghitung tren dengan memperbarui tren sebelumnya berdasarkan perubahan dari nilai level.
- d. **Hitung komponen musiman:** Menghitung musiman pada periode tertentu dengan memperbarui komponen musiman sebelumnya sesuai dengan pola musiman data.
- e. **Peramalan nilai berikutnya:** Gabungkan ketiga komponen (level, tren, musiman) untuk mendapatkan perkiraan nilai masa depan.
- f. **Ulangi langkah-langkah di atas:** Perbarui setiap komponen pada setiap periode baru untuk mengembangkan peramalan yang lebih akurat.

Algoritma Triple Exponential Smoothing efektif untuk data dengan pola musiman dan tren yang jelas. Namun, algoritma ini memerlukan parameter smoothing yang harus dioptimalkan, dan kurang efektif bila data tidak memiliki pola yang konsisten. Fadilla Putri & Ekadiansyah (2022)

Triple Exponential Smoothing

1. Level (L_t):

$$L_t = a_t \cdot \frac{Y_t}{S_{t-p}} + (1 - a_t) \cdot (L_{t-1} + T_{t-1})$$

- L_t : Nilai level pada periode t .
- Y_t : Data aktual pada waktu t .
- S_{t-p} : Komponen musiman dari periode sebelumnya.
- L_{t-1} : Nilai level sebelumnya.
- T_{t-1} : Nilai tren sebelumnya.
- a_t : Konstanta pemulusan untuk level (nilai antara 0 dan 1).

2.1.4 Gambar rumus TES 1

2.1.5 Transformer Time Series

Algoritma Transformer Time Series adalah penerapan arsitektur Transformer yang dirancang khusus untuk menganalisis data deret waktu. Model ini telah membawa kemajuan besar dalam pemodelan dan prediksi data sekuensial, menggantikan pendekatan tradisional seperti ARIMA, LSTM, dan RNN Muftie (2024). Berikut adalah kelebihan dan kekurangan dari Transformer Time Series.

Kelebihan:

1. Kemampuan Memproses Secara Paralel: Tidak seperti model RNN yang memproses data secara berurutan, Transformer Time Series mampu menangani seluruh sekuens input secara bersamaan, meningkatkan efisiensi pelatihan dan inferensi.
2. Mekanisme Self-Attention: Model ini menggunakan self-attention untuk memahami hubungan jangka panjang antar titik data dalam suatu sekuens, memungkinkan analisis yang lebih akurat.
3. Fleksibilitas dan Adaptabilitas: Transformer Time Series dapat disesuaikan untuk berbagai jenis data deret waktu, mulai dari prediksi keuangan hingga analisis cuaca Muftie (2024).
4. Kemampuan Menangani Long-Term Dependencies: Model ini lebih unggul dalam menangkap pola jangka panjang dibandingkan dengan LSTM dan RNN yang mengalami keterbatasan dalam mempertahankan informasi dalam sekuens yang panjang.

5. Transfer Learning: Model yang telah dilatih sebelumnya dapat digunakan kembali untuk berbagai tugas prediksi deret waktu, menghemat sumber daya dan meningkatkan performa dengan data yang lebih sedikit.

Kelemahan:

1. Kebutuhan Komputasi yang Tinggi: Model ini memerlukan daya komputasi yang besar, terutama untuk dataset dengan jumlah titik waktu yang besar.
2. Kurangnya Interpretabilitas: Meskipun menghasilkan prediksi yang akurat, sulit untuk memahami bagaimana model mengambil keputusan, yang menjadi tantangan dalam aplikasi kritis seperti keuangan dan kesehatan.
3. Ketergantungan pada Data Besar: Transformer Time Series memerlukan banyak data historis untuk melatih model dengan baik. Jika data tidak cukup, model dapat mengalami kesulitan dalam memberikan prediksi yang akurat.
4. Kesulitan dalam Menangani Anomali: Model ini mungkin kurang efektif dalam mendekripsi atau menangani anomali pada data deret waktu, terutama jika anomali jarang muncul dalam data pelatihan.
5. Biaya Pelatihan yang Tinggi: Selain memerlukan perangkat keras yang kuat, pelatihan model Transformer Time Series memakan waktu dan sumber daya yang signifikan, membuatnya kurang ekonomis untuk penggunaan skala kecil. Dengan keunggulan dan tantangan ini, Transformer Time Series menjadi pilihan yang kuat dalam analisis deret waktu, terutama untuk tugas-tugas yang memerlukan pemahaman mendalam terhadap pola jangka panjang.

2.4 WEB

Web pertama kali diperkenalkan oleh Tim Berners-Lee pada tahun 1989 sebagai cara untuk memudahkan ilmuwan berbagi informasi. Ia menciptakan tiga teknologi utama: HTML, URL (Uniform Resource Locator), dan HTTP, yang masih digunakan hingga saat ini. Sejak itu, web telah berkembang pesat, mulai dari halaman-halaman sederhana hingga situs web dinamis yang kompleks. Widyati & Alda (2024)

2.5 Python

Python adalah bahasa pemrograman tingkat tinggi yang dirancang untuk kemudahan penggunaan dan keterbacaan kode. Python terkenal karena sintaksisnya yang sederhana dan jelas, membuatnya sangat populer di kalangan pemula maupun pengembang berpengalaman. Bahasa ini mendukung paradigma pemrograman berorientasi objek, prosedural, dan fungsional.

Beberapa fitur utama Python adalah:

1. **Mudah dipelajari:** Python dirancang agar mudah dipahami dan digunakan. Hal ini menjadikannya pilihan utama untuk pemula dalam belajar pemrograman.
2. **Bersifat dinamis:** Python tidak memerlukan deklarasi tipe data secara eksplisit, sehingga kode bisa lebih ringkas dan cepat ditulis.
3. **Pustaka standar yang kaya:** Python dilengkapi dengan pustaka standar yang sangat besar, yang menyediakan berbagai alat untuk melakukan berbagai tugas, seperti manipulasi string, pengolahan data, dan komunikasi jaringan.
4. **Portabilitas:** Python dapat dijalankan di berbagai platform, termasuk Windows, macOS, dan Linux, tanpa perlu modifikasi kode.
5. **Komunitas besar:** Python memiliki komunitas yang aktif, yang membuatnya memiliki banyak sumber daya, tutorial, dan pustaka pihak ketiga yang memperkaya fungsionalitasnya.
6. **Digunakan dalam berbagai bidang:** Python banyak digunakan dalam berbagai disiplin ilmu, termasuk pengembangan web, analisis data, kecerdasan buatan (AI), pembelajaran mesin (machine learning), otomasi, dan pemrograman aplikasi desktop

2.6 FIREBASE

Firebase adalah platform pengembangan aplikasi yang dikembangkan oleh Google, menyediakan berbagai layanan untuk membantu pengembang dalam membangun, mengelola, serta mengoptimalkan aplikasi secara lebih cepat dan efisien. Berikut adalah beberapa komponen utama dari Firebase:

- 1. *Realtime Database*:** Firebase Realtime Database adalah layanan database NoSQL cloud-hosted yang menyimpan dan mensinkronisasi data antara pengguna secara waktu nyata. Ini memungkinkan aplikasi untuk berkomunikasi secara instan dengan basis data tanpa perlu menulis kode server.
- 2. *Authentication*:** Firebase Authentication menyediakan layanan otentikasi pengguna yang aman dan mudah digunakan. Ini mendukung otentikasi melalui email/password, Google, Facebook, Twitter, dan banyak lagi, serta menyediakan mekanisme otentikasi kustom.
- 3. *Cloud Firestore*:** Cloud Firestore adalah database cloud NoSQL yang menawarkan dukungan untuk menyimpan dan menyinkronkan data antar aplikasi secara waktu nyata. Firestore menyediakan skema data yang fleksibel, serta dukungan untuk kueri yang kuat dan real-time updates.
- 4. *Storage*:** Firebase Storage menyediakan infrastruktur penyimpanan berbasis cloud untuk menyimpan, mengunggah, dan men-download

berkas, seperti gambar, video, dan file lainnya. Ini memungkinkan aplikasi untuk menyimpan data pengguna dengan mudah dan aman.

5. *Cloud Functions*: Firebase Cloud Functions adalah layanan yang memungkinkan pengembang menulis dan menyebarkan fungsi server tanpa harus mengelola infrastruktur server. Dengan Cloud Functions, Anda dapat membuat fungsi yang merespons peristiwa dalam aplikasi Anda, seperti penambahan data ke database atau autentikasi pengguna.
6. *Hosting*: Firebase Hosting memungkinkan pengembang untuk menyebarkan aplikasi web dan konten statis dengan mudah dan cepat. Ini menyediakan infrastruktur hosting global yang aman dan cepat dengan dukungan HTTP
7. *Analytics*: Firebase Analytics menyediakan informasi tentang bagaimana pengguna berinteraksi dengan aplikasi Anda, termasuk data tentang penggunaan, retensi, dan konversi. Ini membantu pengembang memahami kinerja aplikasi mereka dan membuat keputusan yang lebih baik tentang pengembangan produk.

8. *Cloud Messaging*: Firebase Cloud Messaging (FCM) adalah layanan yang memungkinkan pengembang mengirim pesan push ke pengguna aplikasi mereka di semua platform (Android, iOS, dan web). Ini memungkinkan komunikasi real-time antara aplikasi dan pengguna

9. *App Check* :

Firebase App Check adalah fitur keamanan yang dirancang untuk melindungi aplikasi Anda dari penyalahgunaan dengan memastikan bahwa lalu lintas yang masuk berasal dari aplikasi asli yang berjalan di perangkat yang tidak dimodifikasi. Dengan mengaktifkan App Check, Anda dapat mencegah akses tidak sah ke sumber daya backend Anda, termasuk layanan Firebase dan Google Cloud

2.7 Streamlit

Streamlit adalah framework open-source yang digunakan untuk membuat aplikasi web interaktif dengan Python secara cepat dan mudah, tanpa memerlukan pengetahuan mendalam tentang pengembangan web.

Streamlit memungkinkan pengembang untuk membuat antarmuka pengguna (UI) untuk proyek analisis data, pembelajaran mesin, dan visualisasi data hanya dengan menulis kode Python. Ini sangat berguna untuk prototyping dan pembuatan dashboard interaktif.

Beberapa fitur utama Streamlit adalah:

1. *Kemudahan Penggunaan:* Streamlit dirancang agar mudah digunakan, bahkan bagi mereka yang tidak memiliki pengalaman dalam pengembangan web. Cukup dengan menulis kode Python, pengembang dapat membuat aplikasi web yang dapat langsung dijalankan.
2. *Interaktivitas:* Pengguna dapat menambahkan elemen interaktif seperti slider, input teks, dan tombol untuk memungkinkan pengguna berinteraksi dengan data dan model secara langsung.
3. *Visualisasi Data:* Streamlit mendukung integrasi dengan berbagai pustaka visualisasi data seperti Matplotlib, Plotly, dan Altair. Pengguna dapat dengan mudah menampilkan grafik dan visualisasi dalam aplikasi web.
4. *Tanpa Konfigurasi yang Rumit:* Tidak perlu mengkonfigurasi server atau menulis HTML/CSS/Javascript. Semua UI dibuat menggunakan fungsi

Python yang sederhana, dan aplikasi dapat dijalankan dengan satu perintah (streamlit run).

5. *Pembaruan Otomatis:* Aplikasi Streamlit akan secara otomatis diperbarui setiap kali ada perubahan pada kode Python, memungkinkan pengembangan yang sangat cepat dan iteratif.
6. *Penggunaan dalam Data Science dan Machine Learning:* Streamlit sangat populer di kalangan data scientist dan pengembang machine learning karena memudahkan pembuatan aplikasi berbasis data yang dapat digunakan untuk berbagi hasil analisis dan model.
7. *Berfungsi di Browser:* Aplikasi Streamlit dijalankan di browser, memungkinkan pengguna untuk mengakses aplikasi dari mana saja dengan menggunakan URL.

2.9 Kajian Penelitian Terdahulu

JUDUL	Tahun	Hasil Penelitian
SISTEM PREDIKSI TINGKAT KRIMINALITAS MENGGUNAKAN METODE TRIPLE EXPONENTIAL SMOOTHING: STUDI KASUS PADA POLRES KABUPATEN PROBOLINGGO 1. (Arwin Datumaya Wahyudi Sumari 2. Ricky Yulian Adi Pratama Odhitya Desta Triswidrananta	2021	<p>Penelitian ini mengembangkan sistem prediksi tingkat kriminalitas di Kabupaten Probolinggo menggunakan metode Triple Exponential Smoothing (TES). Dengan menggunakan data kriminalitas dari 2014 hingga 2018, sistem ini memprediksi berbagai jenis tindak kejahatan, seperti pencurian, penipuan, dan kekerasan, serta mengevaluasi akurasi prediksinya dengan Mean Absolute Percentage Error (MAPE). Hasilnya menunjukkan bahwa prediksi memiliki akurasi yang masuk akal (MAPE sekitar 28,3%), meskipun beberapa jenis kriminalitas seperti pencurian hewan memiliki Akurasi rendah. Sistem ini diharapkan membantu kepolisian dalam mengantisipasi peningkatan kriminalitas dimasa mendatang.</p>
Attention Is All You Need	2023	Transformer adalah arsitektur jaringan saraf yang menggunakan mekanisme

<p>1. Noam Shazeer 2. Niki Parmar, 3. Jakob Uszkoreit, 4. Llion Jones, 5. Aidan N. Gomez, 6. Lukasz Kaiser, Illia Polosukhin.</p>		<p>self-attention untuk memproses urutan data tanpa bergantung pada urutan secara langsung. Dibandingkan dengan RNN yang lambat dalam menangani urutan panjang, Transformer lebih efisien dan mudah dilatih. Dengan konsep multi-head attention, model ini dapat menangkap berbagai hubungan dalam data secara simultan, menjadikannya unggul dalam tugas NLP seperti penerjemahan dan klasifikasi teks. Keunggulannya dalam kecepatan pelatihan dan akurasi membuat Transformer menjadi dasar bagi model canggih seperti BERT dan GPT.</p>
<p>Implementasi Algoritma Transformers BART dan Penggunaan Metode Optimasi Adam Untuk Klasifikasi Judul Berita Palsu</p>	<p>2024</p>	<p>Penelitian ini bertujuan untuk mengklasifikasikan judul berita palsu menggunakan model Transformers BART (Bidirectional and Auto-Regressive Transformers) yang dioptimalkan dengan metode Adam. Pendekatan ini diharapkan dapat menghasilkan akurasi tinggi dalam mengidentifikasi berita palsu yang marak di platform daring.</p>

<p>1. Ageng Ramdhana</p> <p>2. Subagyo Theopilus</p> <p>3. Bayu Sasongko</p>		
<p>"Perbandingan Performa Deteksi Cyberbullying dengan Transformer" –</p> <p>1. Fuad Muftie</p> <p>2. Kamal Muftie Yafi</p> <p>3. Qinthara Muftie Addina</p>	2024	<p>1. Penelitian ini menunjukkan bahwa model Transformer mencapai akurasi tertinggi 98.49% dalam klasifikasi teks cyberbullying pada dataset YouTube, mengungguli metode lain seperti SVM yang mencapai 97.82%. Dibandingkan dengan RNN, LSTM, GRU, serta metode machine learning seperti Naïve Bayes dan Logistic Regression, Transformer terbukti lebih akurat. Hasil ini menegaskan bahwa Transformer adalah model yang lebih canggih dan efisien dalam deteksi cyberbullying, berkontribusi pada upaya mitigasi masalah di media sosial dengan</p>

		solusi berbasis teknologi yang lebih efektif.
"Sistem Prediksi Tingkat Kriminalitas Menggunakan Metode Triple Exponential Smoothing: Studi Kasus pada Polres Kabupaten Probolinggo" 1. Arwin Datumaya 2. Wahyudi Sumari 3. Ricky Yulian 4. Adi Pratama 5. Odhitya Desta 6. Triswidrananta	2020	Penelitian ini bertujuan mengembangkan sistem prediksi tingkat kriminalitas berdasarkan jenis tindakan kriminal menggunakan metode Triple Exponential Smoothing (TES). Data yang digunakan berasal dari Kabupaten Probolinggo, dengan akurasi prediksi diukur menggunakan Mean Absolute Percentage Error (MAPE). Hasil menunjukkan bahwa model TES memiliki akurasi yang baik dalam memprediksi tingkat kriminalitas. Penelitian merekomendasikan penerapan sistem ini oleh kepolisian untuk perencanaan tindakan pencegahan kejahatan. Secara keseluruhan

2.9 Kajian Penelitian Terdahulu 1