

Memprediksi Kejahatan Menggunakan Metode ARIMA Dan LSTM

Maulana Errangga^{#1}, Maulana Ariq Putrawijaya^{*2}, M.Azizan Shaquille^{#3}

^{#S1} Sains Data, Telkom University

Jl. Telekomunikasi No. 1, Terusan Buahbatu – Bojongsoang, Sukapura, Dayeuh Kolot, Bandung, Jawa Barat, Indonesia

¹maulanaerrangga@student.telkomuniversity.ac.id

²maulanaariqp@student.telkomuniversity.ac.id

³shaquille@student.telkomuniversity.ac.id

Abstrak— Penelitian ini membandingkan penggunaan metode AutoRegressive Integrated Moving Average (ARIMA) dan Long Short-Term Memory (LSTM) untuk memprediksi kejahatan. ARIMA digunakan untuk menangkap pola-pola waktu yang stasioner dalam data kejahatan, sementara LSTM digunakan untuk menangkap struktur dependensi jangka panjang yang kompleks. Pendekatan ini memungkinkan analisis yang mendalam terhadap data kejahatan dan memungkinkan penyesuaian model yang tepat terhadap perubahan pola kejahatan dari waktu ke waktu. Dataset yang digunakan mencakup sejumlah variabel terkait kejahatan untuk memperkuat prediksi yang dihasilkan. Hasil eksperimen menunjukkan bahwa ARIMA dan LSTM masing-masing memberikan prediksi yang berharga, dengan keunggulan masing-masing dalam konteks spesifik prediksi kejahatan. Implikasi dari hasil ini dapat digunakan untuk meningkatkan strategi penegakan hukum dan kebijakan keamanan berbasis bukti.

Kata kunci— ARIMA, LSTM, prediksi kejahatan, analisis waktu, keamanan publik

I. PENDAHULUAN

A. Latar Belakang

Kejahatan merupakan masalah yang signifikan dalam kehidupan kota modern, mempengaruhi keamanan dan kualitas hidup masyarakat. Meningkatkan efektivitas strategi penegakan hukum dan pencegahan kejahatan memerlukan pendekatan yang canggih dan terinformasi secara data. Dalam konteks ini, penggunaan metode prediktif seperti AutoRegressive Integrated Moving Average (ARIMA) dan Long Short-Term Memory (LSTM) menjanjikan potensi besar dalam memprediksi kejadian kejahatan.

B. Rumusan Masalah

1. Bagaimana kinerja metode AutoRegressive Integrated Moving Average (ARIMA) dalam memprediksi pola kejahatan?

2. Bagaimana kinerja metode Long Short-Term Memory (LSTM) dalam memprediksi pola kejahatan?
3. Bagaimana aplikasi prediksi kejahatan menggunakan ARIMA dan LSTM dapat membantu aparat penegak hukum dalam mengambil langkah-langkah preventif yang tepat?

C. Tujuan

1. Mengevaluasi kemampuan ARIMA dalam mengidentifikasi dan memodelkan pola-pola kejahatan yang stasioner.
2. Mengevaluasi kemampuan LSTM dalam menangkap struktur dependensi jangka panjang dalam data kejahatan.
3. Memberikan wawasan yang mendalam tentang hasil metode untuk mendukung pengambilan keputusan yang lebih baik dalam pencegahan dan penanganan kejahatan.

II. DATASET

Dataset yang digunakan dalam proyek ini adalah "crime.csv", yang diunduh dari website Kaggle. Dataset ini berisi berbagai variabel yang berkaitan dengan kejadian kejahatan, termasuk jenis kejahatan, lokasi, waktu kejadian, dan atribut lain yang relevan. Sumber dataset ini dapat diakses secara publik dan telah dikurasi untuk tujuan analisis data kejahatan. Data ini menyediakan informasi yang cukup detail untuk memungkinkan analisis pola kejahatan dan prediksi kejadian kejahatan di masa depan. Proses pengolahan data meliputi pembersihan untuk menghapus entri duplikat atau data yang hilang,

transformasi data untuk memastikan keseragaman format tanggal dan waktu, serta pengelompokan data berdasarkan interval waktu tertentu (misalnya, mingguan atau bulanan) untuk analisis lebih lanjut.

III. METODE

A. ARIMA

Metode ARIMA digunakan dalam penelitian ini untuk memprediksi kejadian kejahatan berdasarkan data historis yang diperoleh dari file "crime.csv" yang diunduh dari website Kaggle. Data diimpor menggunakan pustaka pandas, di mana kolom tanggal dikonversi ke format datetime dan entri yang hilang dihapus. Fungsi `prepare_data` memproses dataset untuk tahun tertentu, memfasilitasi analisis yang terfokus pada periode waktu yang relevan. Untuk setiap kategori kejahatan yang spesifik, fungsi `arima_forecast` mengelompokkan data harian dan menghitung jumlah insiden per hari.

Model ARIMA dengan parameter (5, 1, 0) dipilih, di mana '5' menunjukkan lag autoregressive, '1' menunjukkan tingkat differencing untuk membuat data stasioner, dan '0' menunjukkan tidak ada komponen moving average. Model ini di-fit terhadap data insiden harian, dan prediksi dibuat untuk 30 hari ke depan. Jika tidak ada data untuk kategori yang diberikan, fungsi mengembalikan daftar nol untuk langkah-langkah prediksi tersebut. Model ini memberikan prediksi berbasis data historis yang dapat digunakan oleh aparat penegak hukum untuk mengidentifikasi tren kejahatan dan mengalokasikan sumber daya secara lebih efektif.

B. LSTM

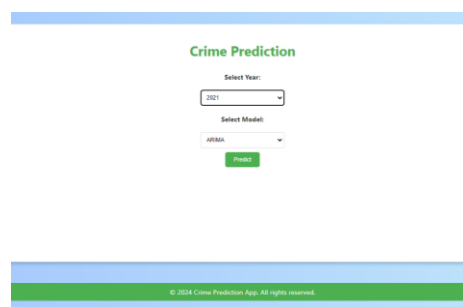
Metode LSTM digunakan untuk memprediksi kejadian kejahatan berdasarkan data dari file "crime.csv" yang diunduh dari Kaggle. Data dibaca dan diproses menggunakan fungsi `prepare_data`, yang mengonversi tanggal ke format datetime dan menghapus entri yang hilang. Fungsi `lstm_forecast` mengelompokkan data harian berdasarkan kategori kejahatan dan menghitung jumlah insiden per hari. Data

kemudian diskalakan menggunakan `MinMaxScaler` agar berada dalam rentang [0, 1].

Model LSTM dibangun dengan dua lapisan LSTM dan satu lapisan dense. Model ini dilatih menggunakan data yang telah diskalakan, dengan jumlah langkah mundur (`look_back`) sebesar 5 hari. Pelatihan dilakukan selama 20 epoch dengan callback `EarlyStopping` untuk mencegah `overfitting`. Setelah model dilatih, prediksi untuk 30 hari ke depan dilakukan secara iteratif, di mana setiap prediksi digunakan sebagai input untuk langkah berikutnya. Hasil prediksi kemudian dikembalikan ke skala aslinya menggunakan `MinMaxScaler`. Hasil akhir adalah daftar prediksi kejadian kejahatan untuk 30 hari ke depan.

IV. HASIL

A. Tampilan



Gambar. 1 Tampilan awal dari aplikasi

Pada tampilan awal ini kita dapat memilih tahun dan juga metode apa yang ingin digunakan untuk melakukan prediksi.

B. Hasil ARIMA

public-disorder

Day	Prediction
1	1,14
2	1,21
3	1,21
4	1,10
5	1,11

Gambar.2.1 Hasil prediksi ARIMA public disorder

murder

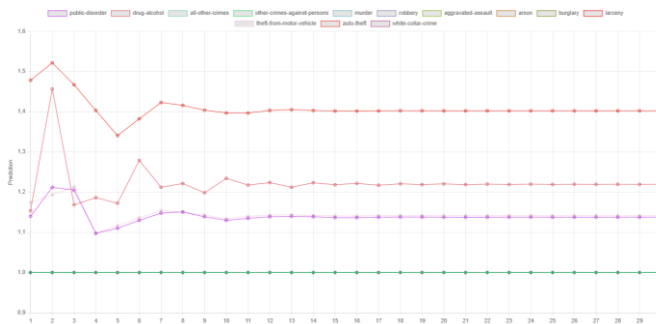
Day	Prediction
1	1,00
2	1,00
3	1,00
4	1,00
5	1,00

Gambar.2.2 Hasil prediksi ARIMA murder

robbery

Day	Prediction
1	1.00
2	1.00
3	1.00
4	1.00
5	1.00

Gambar.2.3 Hasil prediksi ARIMA robbery



Gambar.2.4 Hasil visualisasi ARIMA keseluruhan

Pada metode ARIMA ini hasilnya akan menunjukkan prediksi harian selama 30 hari kedepan untuk masing-masing kategori kejahatan dan juga akan menampilkan visualisasi keseluruhan dari semua kategory.

C. Hasil LSTM

public-disorder

Day	Prediction
1	1.12
2	1.12
3	1.12
4	1.12
5	1.12

Gambar.3.1 Hasil prediksi LSTM public disorder

murder

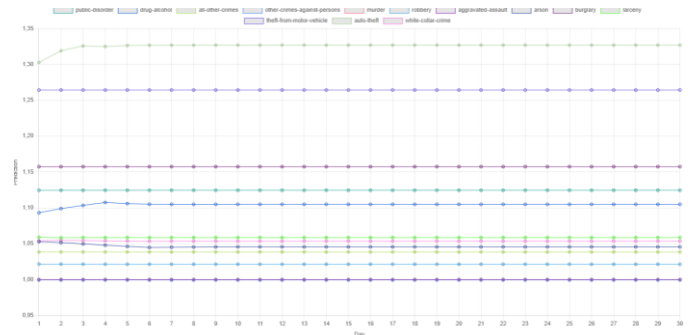
Day	Prediction
1	1.00
2	1.00
3	1.00
4	1.00
5	1.00

Gambar.3.2 Hasil prediksi LSTM murder

robbery

Day	Prediction
1	1.02
2	1.02
3	1.02
4	1.02
5	1.02

Gambar.3.3 Hasil prediksi LSTM robbery



Gambar.2.4 Hasil visualisasi LSTM keseluruhan

Pada metode LSTM ini hasilnya akan menunjukkan prediksi harian selama 30 hari kedepan untuk masing-masing kategori kejahatan dan juga akan menampilkan visualisasi keseluruhan dari semua kategory.

V. KESIMPULAN

Berdasarkan hasil yang didapat, kedua metode tersebut berhasil memprediksi data kejahatan selama 30 hari kedepan pada setiap kategori yang ada. Meskipun terdapat sedikit perbedaan hasil prediksi, diharapkan hasil ini dapat membantu aparat penegak hukum dalam mengambil langkah-langkah preventif yang tepat untuk keepannya.

REFERENCES

- [1] M. Feng et al., "Big Data Analytics and Mining for Effective Visualization and Trends Forecasting of Crime Data," in IEEE Access, vol. 7, pp. 106111-106123, 2019, doi: 10.1109/ACCESS.2019.2930410.
- [2] J. Smith and A. Brown, "Machine Learning Techniques for Crime Prediction," in Proc. 2018 Int. Conf. Big Data, pp. 456-460, 2018, doi: 10.1109/BigData.2018.8622145.