

# Penerapan Metode *Hybrid LSTM* dan *Seasonal Decomposition* untuk Peramalan Permintaan Tiket Harian PT KAI

Rahmanizar Maksun Yuniyanto<sup>1,\*</sup>, Ahmad Mukhlason<sup>2</sup>

<sup>1</sup>Departemen Magister Manajemen Teknologi, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

<sup>2</sup>Departemen Sistem Informasi, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

<sup>1</sup>rahmanizar.maksum@gmail.com; <sup>2</sup>mukhlason@is.its.ac.id

\*penulis korespondensi

## INFO ARTIKEL

### Sejarah Artikel

Diterima: 21 April 2025

Direvisi: 18 Juni 2025

Diterbitkan: 28 Agustus 2025

### Kata Kunci

LSTM

Machine Learning

Peramalan

Permintaan Tiket Harian

Seasonal Decompose

## ABSTRAK

Peramalan permintaan tiket harian berperan penting dalam mendukung efisiensi operasional dan peningkatan kualitas layanan transportasi, termasuk pada kereta api. Penelitian ini membangun model peramalan permintaan tiket Kereta Api 1 Argo Bromo Angrek dengan teknik *machine learning* yang menggabungkan pendekatan hibrida metode dekomposisi deret waktu dan algoritma *Long Short-Term Memory* (LSTM). Metode dekomposisi deret waktu memisahkan data menjadi komponen *trend*, *seasonal*, dan *residual* yang dianalisis secara terpisah guna memahami pola permintaan secara lebih mendalam. Hasil menunjukkan bahwa komponen trend memiliki korelasi kuat terhadap permintaan dan berperan penting dalam model peramalan. Model LSTM yang dikembangkan menunjukkan performa baik dengan nilai *Mean Absolute Percentage Error* (MAPE) sebesar 7,53% pada data uji. Model yang dihasilkan dapat dimanfaatkan untuk mendukung perencanaan operasional PT Kereta Api Indonesia (PT KAI). Pengembangan selanjutnya disarankan untuk mempertimbangkan variabel eksternal seperti hari libur, cuaca, dan kebijakan harga, serta diarahkan untuk mendukung sistem manajemen tarif dinamis (*dynamic pricing*) guna mengoptimalkan pendapatan dan efisiensi layanan.

## PENDAHULUAN

Dalam industri transportasi modern, kemampuan untuk meramalkan permintaan tiket memiliki peran penting dalam menentukan efisiensi operasional, kualitas pelayanan, dan keberhasilan strategi pemasaran. Hal ini disebabkan oleh sifat industri transportasi yang sangat dinamis, sensitif terhadap waktu, dan bergantung pada permintaan yang fluktuatif akibat faktor musiman, tren sosial, kondisi ekonomi, hingga peristiwa tak terduga seperti bencana atau pandemi [1]. Peramalan jumlah permintaan tiket merupakan bagian penting dalam pengambilan keputusan manajerial, seperti penjadwalan, perencanaan armada, dan pengaturan harga tiket [2]. Pesawat terbang merupakan salah satu moda transportasi yang telah menerapkan sistem peramalan permintaan tiket untuk mendukung strategi pemasaran. Pola musiman dan tren dalam data penjualan tiket terbukti dapat digunakan untuk memprediksi permintaan dengan akurasi tinggi [3]. Peramalan permintaan tiket pada sektor ini juga telah banyak mengadopsi teknik *machine learning*, yang mampu mengenali pola data pelatihan dan memberikan prediksi yang andal terhadap data baru.

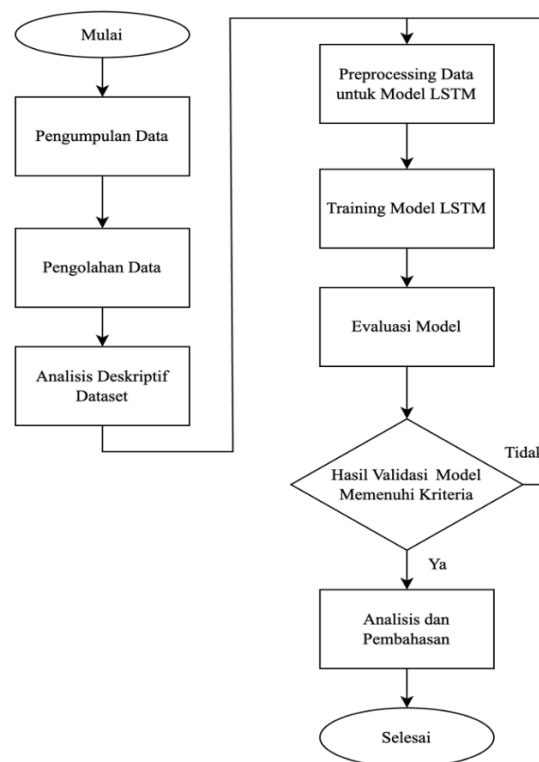
Kebutuhan akan sistem peramalan tiket tidak hanya terbatas pada transportasi udara. Di sektor transportasi darat, PT Kereta Api Indonesia (PT KAI) sebagai penyedia layanan kereta api nasional juga terus berupaya meningkatkan strategi pemasaran, salah satunya dengan menerapkan sistem peramalan permintaan tiket harian [4]. Upaya ini menjadi penting karena permintaan tiket penumpang bersifat fluktuatif dan sering tidak sesuai dengan kapasitas yang

tersedia. Oleh karena itu, dibutuhkan sistem peramalan yang akurat dan adaptif untuk mendukung pengambilan keputusan operasional [5]. Penelitian ini merespons tantangan tersebut dengan mengusulkan model peramalan hibrida yang menggabungkan metode dekomposisi deret waktu dan *machine learning* untuk menangkap dimensi tren, musiman, dan noise dalam data permintaan harian [6], [7].

Salah satu algoritma *machine learning* yang menjadi solusi penting dalam proses peramalan yaitu *Long Short-Term Memory* (LSTM [8]. LSTM sebagai bagian dari *Recurrent Neural Network* (RNN) efektif dalam menangani pola data yang bersifat deret waktu yang kompleks seperti permintaan tiket [9]. Beberapa penelitian membuktikan bahwa LSTM memberikan performa akurasi peramalan data *time series* lebih baik dibanding algoritma lain, seperti algoritma ARIMA, SR, dan SVM [10], [11][12]. LSTM mampu menurunkan nilai RMSE lebih dari 83% dibanding ARIMA dalam peramalan *time series* dengan pola fluktuatif [12]. Sedangkan untuk dekomposisi deret waktu bekerja dengan memisahkan deret waktu menjadi komponen tren, musiman, dan residual, sehingga memiliki keunggulan dapat mengungkap pola-pola mendasar yang mungkin tersembunyi. Pendekatan multi-faset ini bertujuan menghasilkan peramalan permintaan yang sangat akurat dengan mempertimbangkan baik tren jangka panjang maupun fluktuasi jangka pendek [13].

Penelitian ini menerapkan metode *Jump Upon Spectrum and Trend* (JUST) sebagai teknik dekomposisi deret waktu, dan mengintegrasikannya dengan LSTM dalam model hibrida. JUST memisahkan data menjadi komponen tren, musiman, dan residual, sedangkan LSTM menangkap ketergantungan jangka panjang dalam data sekuensial yang kompleks. Integrasi ini dirancang untuk menangani dinamika struktural dan temporal secara simultan. Meskipun JUST dan LSTM telah digunakan dalam berbagai studi [14], [15], penerapannya secara eksplisit dalam konteks perkeretaapian di Indonesia masih jarang ditemukan. Oleh karena itu, pendekatan ini menawarkan kontribusi metodologis baru dalam peramalan berbasis *machine learning* untuk sektor transportasi nasional.

## METODE



Gambar 1. Diagram Metode Penelitian

Pendekatan metodologis yang sistematis dan berbasis data dilakukan untuk menjawab permasalahan penelitian dalam peramalan permintaan tiket harian oleh PT KAI. Oleh karena itu, penelitian dirancang menggunakan alur kerja berbasis analisis deret waktu dan *machine learning*, khususnya pada penerapan model LSTM untuk memodelkan permintaan tiket harian. Metode penelitian ditunjukkan seperti pada Gambar 1. Gambar 1 menyajikan alur proses penelitian ini, yang dimulai dengan pengumpulan dan pembersihan data permintaan tiket harian, yang kemudian dianalisis secara deskriptif untuk memahami pola awal. Selanjutnya, dilakukan pemodelan berbasis LSTM dengan mempertimbangkan fitur hasil dekomposisi tren, musiman, dan residual. Pada proses pengumpulan dan pengolahan data, data input yang digunakan merupakan data per transaksi permintaan tiket periode tahun 2019 hingga 2024 untuk kereta api KA 1 Argo Bromo Anggrek. Proses pengolahan mencakup pembersihan data, perhitungan sisa waktu, serta transformasi data menjadi format harian. Data kemudian dianalisis secara deskriptif untuk mengidentifikasi pola umum permintaan, sebelum diproses lebih lanjut melalui dekomposisi deret waktu. Hasil dekomposisi digunakan sebagai input model LSTM. Model kemudian dievaluasi menggunakan data pengujian dengan mengacu pada grafik *loss* untuk menilai performa peramalan.

### Metode Dekomposisi Deret Waktu

Salah satu metode dekomposisi deret waktu dan deteksi titik perubahan (*change point detection*) yaitu *Jump Upon Spectrum and Trend* (JUST) [16]. JUST menggunakan pendekatan *anti-leakage least-squares spectral analysis* (ALLSSA), yang unggul dalam mempertahankan informasi komponen musiman serta mengurangi kebocoran informasi akibat dekomposisi. JUST memisahkan data deret waktu menjadi tiga komponen utama yaitu Trend ( $T_v$ ), Seasonal ( $S_v$ ), dan Residual ( $R_v$ ). Deret waktu ( $F$ ) dengan panjang  $n$  dapat dinyatakan seperti pada Persamaan (1).

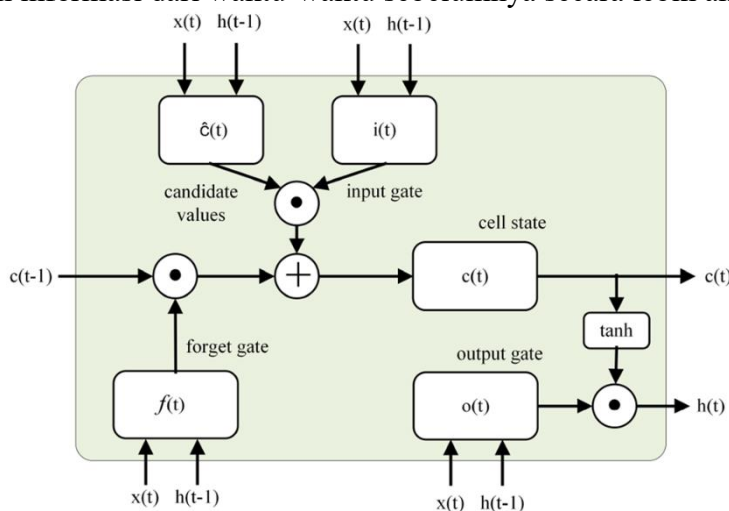
$$F = \{f_{t1}, f_{t2}, \dots, f_{tn}\} \quad (1)$$

Di mana  $f_{ti}$  merepresentasikan nilai pengamatan (observasi) pada waktu ke- $i$ , dan  $n$  adalah total panjang data deret waktu yang akan didekomposisi. Notasi ini menjadi dasar bagi penerapan algoritma JUST dalam memproses potongan data secara berurutan, untuk mengidentifikasi struktur musiman dan tren tersembunyi yang relevan terhadap peramalan permintaan tiket. Untuk melakukan dekomposisi terhadap deret waktu  $F$  seperti didefinisikan pada persamaan (1), algoritma JUST mengikuti serangkaian langkah sistematis. Proses ini dimulai dengan segmentasi data ke dalam jendela-jendela waktu, kemudian memanfaatkan ALLSSA untuk mengidentifikasi pola musiman dan tren lokal secara adaptif. Berikut langkah-langkah operasional dari algoritma JUST:

1. Menentukan ukuran jendela ( $R$ ) dan ukuran langkah ( $\delta$ ), lalu menginisiasi parameter awal dengan  $d = 0$ .
2. Mengambil potongan deret waktu  $f_{t1+d\delta}, f_{t2+d\delta}, \dots, f_{tR+d\delta}$ .
3. Inisialisasi titik perubahan pertama  $l = 3 + d\delta$ .
4. Untuk setiap potongan data, terapkan algoritma ALLSSA untuk mendekomposisi  $T_w, S_w$ , dan  $R_w$  lalu perbarui  $l = l + 1$ .
5. Jika potongan masih dalam batas jendela ( $l < R + d\delta$ ), kembali ke langkah 4 untuk melanjutkan dekomposisi
6. Mencatat jumlah kuadrat dari komponen tersisa untuk nilai  $l$  optimal sebagai  $l_d$ , lalu atur  $d = d + 1$
7. Ulangi proses hingga seluruh data deret waktu selesai didekomposisi.

### Model Long Short-Term Memory (LSTM)

*Long Short-Term Memory* (LSTM) merupakan salah satu jaringan saraf tiruan (*Recurrent Neural Network*) atau RNN yang dirancang untuk menangani masalah *vanishing gradient* yang sering dihadapi oleh RNN tradisional. Masalah ini muncul ketika informasi yang relevan dari input sebelumnya perlahan-lahan menghilang ketika diumpan kembali melalui lapisan tersembunyi (*hidden layers*) dari RNN [17]. Dalam proses peramalan, LSTM menangkap pola jangka panjang dalam data historis, sehingga peramalan dapat mempertimbangkan informasi dari waktu-waktu sebelumnya secara lebih akurat.



Gambar 2. Arsitektur Blok LSTM [18]

Gambar 2 menunjukkan arsitektur dari LSTM yang terdiri dari tiga gerbang utama yaitu *forget gate*, *input gate*, dan *output gate*. *Forget gate* menerima input dari data  $x(t)$  dan *hidden state* sebelumnya  $h(t-1)$ , kemudian memutuskan informasi mana dari *cell state* sebelumnya  $c(t-1)$  yang harus dilupakan melalui fungsi aktivasi sigmoid yang menghasilkan nilai antara 0 dan 1. Selanjutnya, *input gate* juga menerima  $x(t)$  dan  $h(t-1)$  untuk menentukan informasi baru mana yang perlu disimpan dalam memori. Hasil  $i(t)$  dan  $\hat{c}(t)$  digabungkan untuk memperbarui *cell state*  $c(t)$ . *Output gate* kemudian menentukan bagian dari *cell state* yang akan diteruskan sebagai output  $h(t)$  atau *hidden state* baru. Output ini diperoleh dengan terlebih dahulu menerapkan fungsi *tanh* pada  $c(t)$  untuk mengatur skala nilainya, lalu dikalikan dengan hasil dari *output gate*  $o(t)$  yang juga menggunakan fungsi sigmoid.

LSTM sebagai salah satu teknik dalam *machine learning* memerlukan data dalam format input-output yang spesifik. Dalam konteks pemodelan deret waktu, data disusun menggunakan pendekatan *sliding window* dengan membentuk subsekuensi dengan panjang  $L + 1$  dibuat dari deret waktu. Titik pertama hingga ke- $L$  dari subsekuensi tersebut berfungsi sebagai input, sementara titik ke- $L + 1$  dijadikan sebagai target atau output peramalan [18].

### Evaluasi Model

Untuk menilai performa model kombinasi *machine learning* dengan dekomposisi waktu untuk peramalan tiket kereta harian, dilakukan evaluasi menggunakan dua metrik utama, yaitu *Mean Squared Error* (MSE) dan *Mean Absolute Percentage Error* (MAPE) [19]. Selain itu, validasi dilakukan secara visual dengan membandingkan hasil peramalan terhadap pola tren data historis, guna memastikan kesesuaian tren dan pola peramalan terhadap kenyataan [20].

*Mean Squared Error* (MSE) merupakan metrik evaluasi yang digunakan untuk mengukur rata-rata dari kuadrat selisih antara nilai aktual dan nilai hasil peramalan [19]. MSE bersifat sensitif terhadap outlier, karena setiap kesalahan dipangkatkan dua, sehingga kesalahan besar memiliki dampak yang signifikan terhadap nilai akhir. Fungsi utama MSE adalah untuk menunjukkan seberapa jauh hasil peramalan menyimpang dari nilai aktual dalam skala kuadrat. Rumus perhitungan MSE ditunjukkan pada Persamaan (2).

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y})^2 \quad (2)$$

Dimana  $y_i$  adalah nilai aktual pada waktu ke- $i$ ,  $\hat{y}$  adalah nilai hasil peramalan pada waktu ke- $i$ , dan  $n$  adalah jumlah total observasi. Selain itu, digunakan juga metrik *Mean Absolute Percentage Error* (MAPE) yang mengukur tingkat kesalahan dalam bentuk persentase rata-rata dari kesalahan absolut terhadap nilai aktual. MAPE sering digunakan dalam studi peramalan karena memberikan hasil yang lebih intuitif dan mudah dipahami dalam konteks bisnis dan operasional. Namun, kelemahan MAPE muncul ketika nilai aktual mendekati nol, karena dapat menyebabkan nilai error yang sangat besar [20]. Rumus dari MAPE ditunjukkan pada Persamaan (3).

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{\hat{y}_t - y_t}{Y_t} \right| \times 100 \quad (3)$$

Dengan  $y_t$  sebagai nilai aktual,  $\hat{y}_t$  sebagai nilai peramalan dan  $n$  adalah jumlah observasi. Sebagai pelengkap dari evaluasi kuantitatif, dilakukan juga validasi visual dengan membandingkan grafik antara hasil peramalan dan data aktual dalam bentuk deret waktu. Visualisasi ini bertujuan untuk menilai sejauh mana model mampu mengikuti pola tren, musiman, atau fluktuasi yang terdapat pada data historis [16].

## HASIL DAN PEMBAHASAN

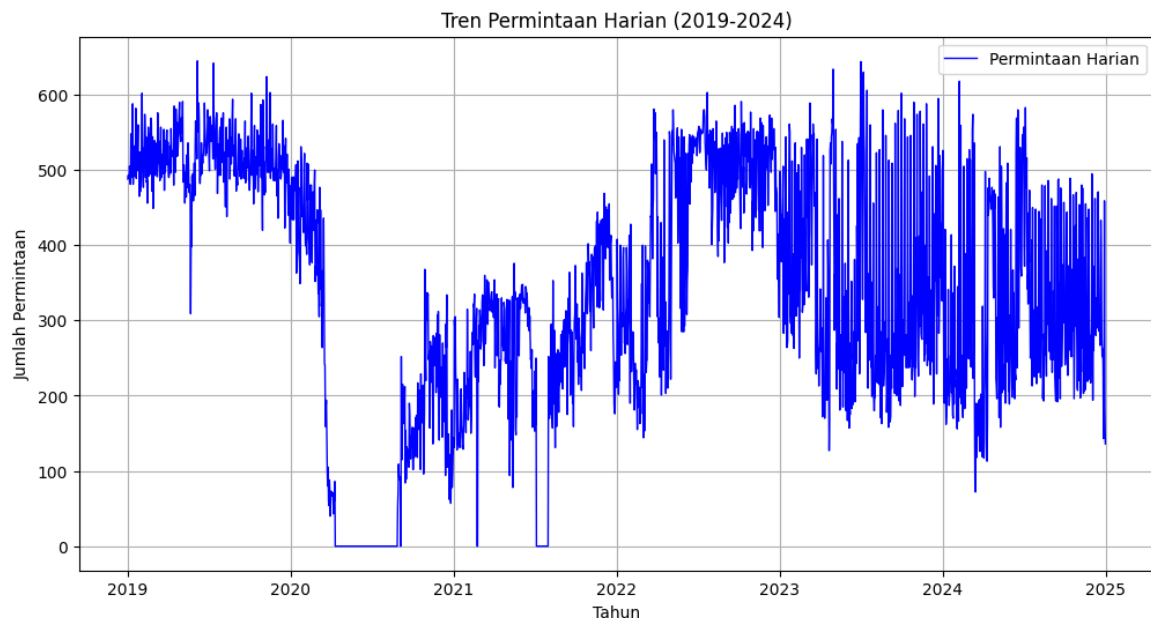
### Analisis Dataset

Dataset penjualan harian KA 1 Argo Bromo Anggrek diperoleh dari data sekunder PT KAI untuk periode 2019–2024, meliputi fase sebelum, saat, dan setelah pandemi COVID-19. Tabel 1 menyajikan hasil data *cleaning*, yang memastikan bahwa seluruh data telah divalidasi tanpa adanya nilai kosong atau duplikasi.

Tabel 1. Jumlah Data Penjualan Harian

Tahun	Data Awal	Missing Values	Data Duplicate	Data Akhir
2019	365	0	0	365
2020	366	0	0	366
2021	365	0	0	365
2022	365	0	0	365
2023	365	0	0	365
2024	366	0	0	366
Total	<b>2.192</b>	<b>0</b>	<b>0</b>	<b>2.192</b>

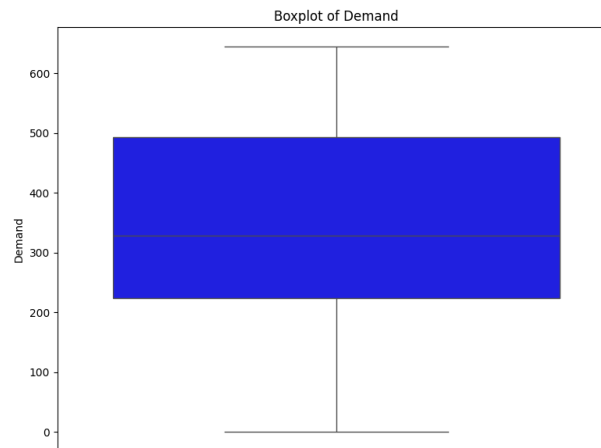
Analisis tren permintaan harian selama enam tahun terakhir dilakukan untuk mengidentifikasi pola fluktuasi jumlah penumpang dalam periode tertentu. Gambar 3 menunjukkan penurunan signifikan pada tahun 2020 akibat kebijakan pembatasan perjalanan selama pandemi. Permintaan mulai meningkat kembali pada tahun 2021 dan mencapai puncaknya di tahun 2022 dengan tingkat okupansi tertinggi sebesar 101%. Namun, tren ini mengalami penurunan pada tahun 2023 dan 2024, dipengaruhi oleh kenaikan tarif serta perubahan stamformasi kereta, yang mengurangi kapasitas tempat duduk dari 450 menjadi 400.



Gambar 3. Trend Permintaan KA 1 Argo Bromo Anggrek (2019-2024)

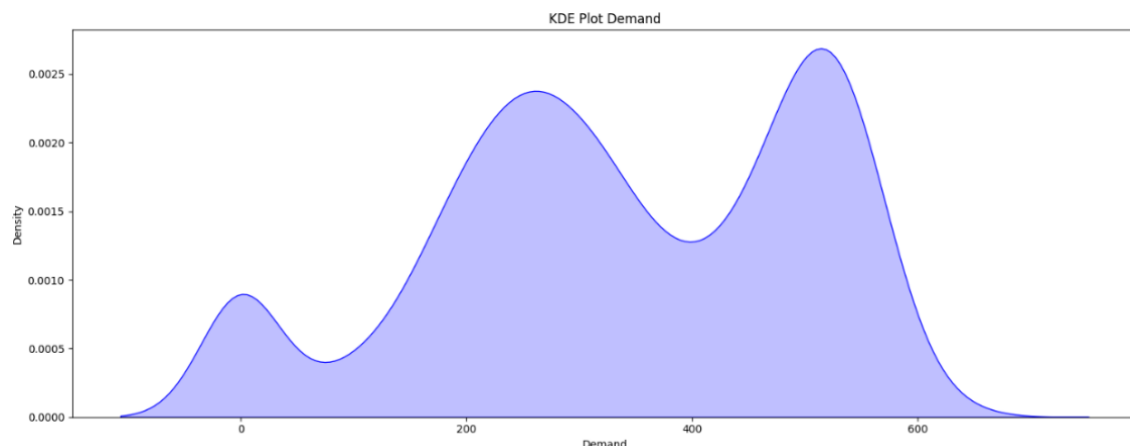
Gambar 4 menunjukkan bahwa nilai median permintaan berada di sekitar 300, dengan 50% data terkonsentrasi dalam rentang 200 hingga 500. Nilai minimum mendekati 0, sedangkan maksimum sekitar 650, menunjukkan variasi permintaan yang cukup besar. Distribusi terlihat relatif simetris karena median berada di tengah antara kuartil pertama (Q1

$\approx 200$ ) dan kuartil ketiga ( $Q3 \approx 500$ ). Tidak terdapat outlier ekstrem yang terlihat, sehingga sebaran data cenderung merata dalam batas IQR.



Gambar 4. Distribusi Data Permintaan

Gambar 5 memperlihatkan distribusi bimodal berdasarkan KDE, dengan dua puncak dominan yang mencerminkan pola permintaan berbeda. Mayoritas data terkonsentrasi dalam rentang 0 hingga 650, dengan kepadatan tertinggi di sekitar 500, menunjukkan bahwa banyak hari dengan permintaan tinggi. Sebaran kecil pada nilai 0–100 menunjukkan adanya beberapa periode dengan permintaan sangat rendah.

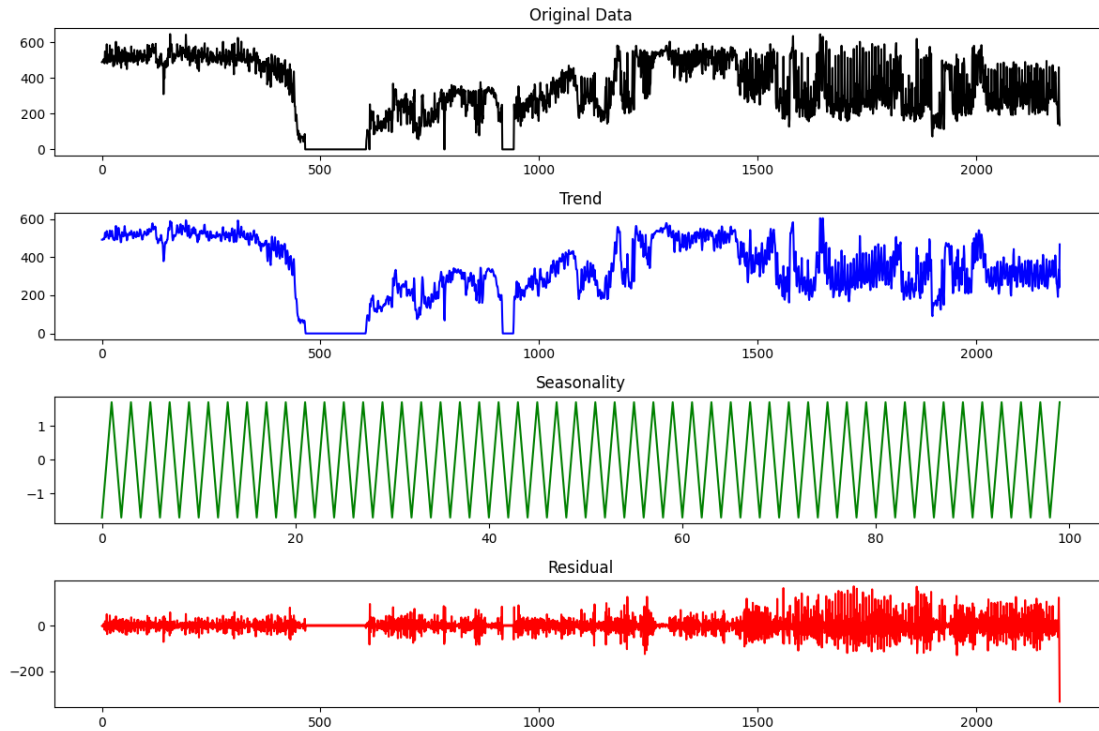


Gambar 5. Kernel Density Estimation Data Permintaan

Analisis dataset menunjukkan bahwa pola permintaan perjalanan KA 1 Argo Bromo Angrek mengalami variasi yang dipengaruhi oleh faktor eksternal, seperti pandemi COVID-19, kebijakan tarif, dan perubahan stamformasi kereta. Distribusi permintaan harian menunjukkan adanya fluktuasi signifikan pada periode tertentu, dengan tren pemulihan setelah tahun 2021. Perbedaan pola distribusi yang terlihat dari *boxplot* dan KDE plot mengindikasikan adanya perubahan dalam kebiasaan perjalanan penumpang dari tahun ke tahun. Variabel-variabel dalam dataset ini memiliki korelasi yang dapat dimanfaatkan untuk membangun model peramalan berbasis *deep learning*. Oleh karena itu, tahap berikutnya akan difokuskan pada pemrosesan lebih lanjut serta pengembangan model LSTM untuk memahami pola peramalan permintaan perjalanan KA secara lebih mendalam.

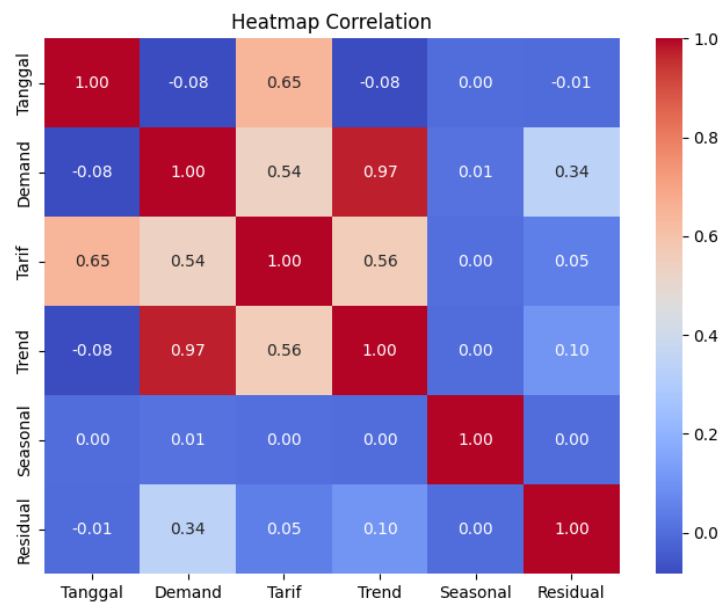
## Preprocessing Data

Dataset yang digunakan dalam pemodelan berasal dari penjualan harian KA 1 Argo Bromo Anggrek selama periode 2019–2024. Seluruh data telah divalidasi dan dibersihkan pada tahap pengolahan sebelumnya. Pada tahap preprocessing, dilakukan serangkaian proses yang mencakup ekstraksi fitur menggunakan *seasonal decomposition*, pembagian dataset, normalisasi data, dan penerapan *sliding window*.



Gambar 6. Seasonal Decomposition Permintaan Harian

Gambar 6 menunjukkan hasil dekomposisi deret waktu musiman menggunakan model aditif dengan periode dua hari, yang memisahkan data permintaan harian menjadi tiga komponen utama: *Trend*, *Seasonality*, dan *Residual*. Komponen *Trend* berhasil mengekstraksi pola jangka panjang, komponen *Seasonality* merepresentasikan pola berulang dalam siklus harian, dan komponen *Residual* mencerminkan variasi acak (*noise*). Ketiga fitur hasil ekstraksi tersebut kemudian dianalisis korelasinya terhadap data permintaan aktual (*Demand*) untuk memastikan relevansinya dalam pemodelan LSTM.



Gambar 7. Uji Korelasi Fitur

Gambar 7 menunjukkan hasil analisis korelasi antar fitur dalam bentuk *heatmap*, yang merepresentasikan kekuatan hubungan linear antar variabel. Komponen “*Trend*” memiliki korelasi sangat kuat terhadap “*Demand*” sebesar 0,97, sedangkan “*Seasonal*” dan “*Residual*” menunjukkan korelasi rendah. Berdasarkan hasil ini, keempat fitur tetap digunakan sebagai input dalam pemodelan.

Setelah pemilihan fitur, dataset dibagi menjadi tiga bagian dengan rasio 70% untuk *training set*, 15% untuk *validation set*, dan 15% untuk *testing set*. Ukuran data masing-masing subset dapat dilihat pada Tabel 2.

Tabel 2. Ukuran Pembagian Dataset

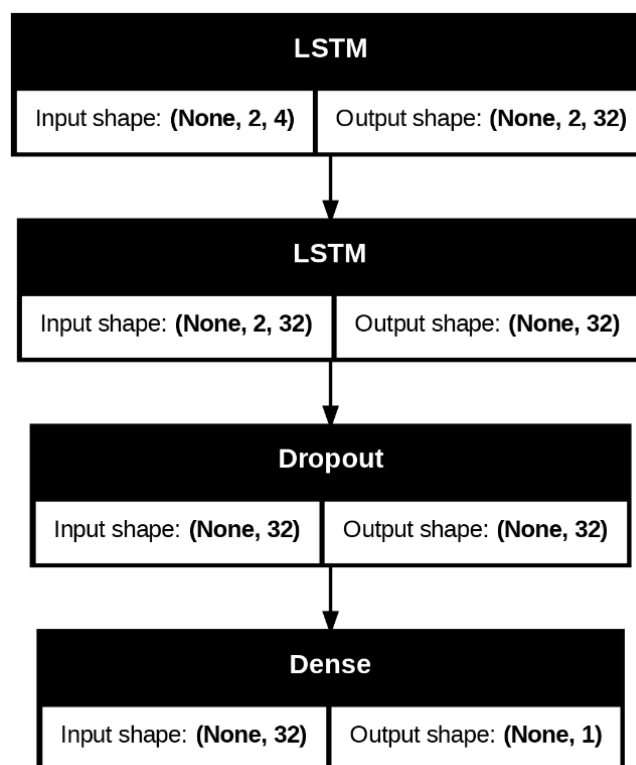
Dataset	Jumlah Sampel
<i>Training Set</i>	1531 Data
<i>Validation Set</i>	328 Data
<i>Test Set</i>	327 Data

Proses normalisasi menggunakan *standard scaler* dilakukan setelah pembagian data, dengan fitting hanya pada data pelatihan untuk mencegah data leakage. Normalisasi tersebut diterapkan pada fitur hasil *seasonal decomposition* (*Trend*, *Seasonal*, *Residual*) serta variabel target setelah dilakukan pembagian dataset. Normalisasi dilakukan dengan menerapkan *fitting* skala hanya pada data *training*, kemudian skala yang sama digunakan untuk mentransformasikan data *validation* dan *testing*. Pendekatan ini memastikan bahwa tidak terjadi data *leakage* antara tahap pelatihan dan evaluasi model.

Tahap akhir preprocessing adalah penerapan *sliding window*, di mana fitur “*Trend*”, “*Seasonal*”, “*Residual*”, dan “*Demand*” digunakan untuk menyusun data *time-series* dalam bentuk sekuensial yang dapat diproses oleh LSTM. Setiap sampel input terdiri dari dua timestep sebelumnya untuk setiap fitur, sehingga membentuk input dengan dimensi (jumlah sampel, 2, 4). Proses *windowing* menghasilkan  $(N - \text{timestamp} - 1)$  sampel, dengan output model berupa nilai “*Demand*” pada waktu setelah *timestamp* yang jumlahnya selalu sesuai dengan jumlah sampel hasil *windowing*.

## Pembuatan Model LSTM

Setelah preprocessing selesai, dataset yang telah diproses digunakan untuk melatih model LSTM yang dirancang secara khusus dalam penelitian ini. Gambar 8 menampilkan struktur arsitektur model yang dibangun dan divisualisasikan menggunakan *library deep learning* berdasarkan hasil implementasi dalam penelitian ini. Model terdiri dari dua lapisan LSTM bertingkat, masing-masing dengan 32-unit neuron dan fungsi aktivasi *tanh*. Lapisan pertama berperan dalam mempertahankan urutan data dari waktu ke waktu, sedangkan lapisan kedua merangkum pola yang telah dipelajari untuk mendukung proses prediksi. Selanjutnya, model menerapkan lapisan *Dropout* sebesar 0.2 untuk mengurangi risiko *overfitting*, dan diakhiri dengan lapisan *Dense* tunggal dengan aktivasi linear untuk menghasilkan nilai peramalan. Input model terdiri dari dua langkah waktu berturut-turut (*timestep*), masing-masing mencakup empat fitur utama: “*Trend*”, “*Seasonal*”, “*Residual*”, dan “*Demand*”. Arsitektur ini dirancang untuk menangkap ketergantungan jangka panjang serta variabilitas permintaan harian secara efektif.

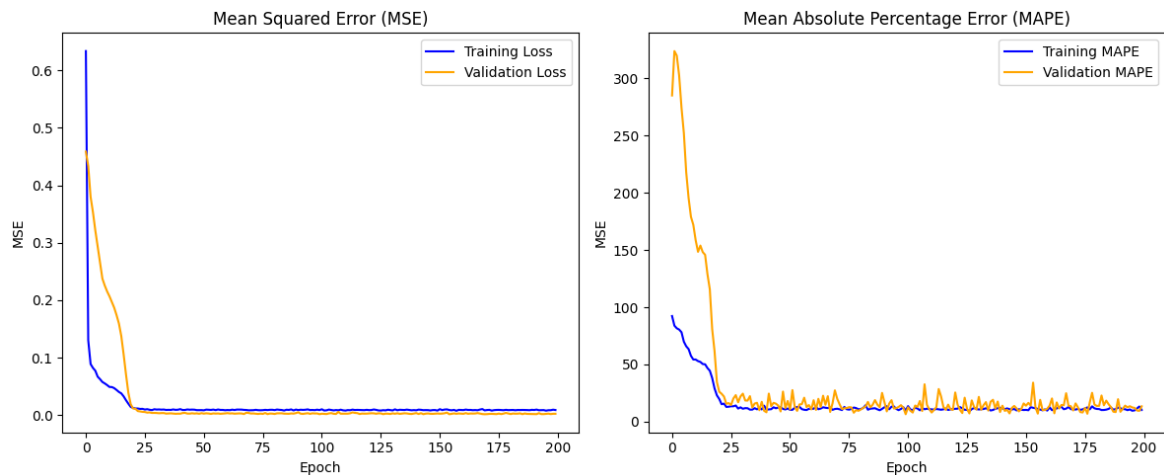


Gambar 8. Model Arsitektur LSTM dalam Studi Ini

Model memiliki total 13.089 parameter yang seluruhnya bersifat *trainable*, sehingga dapat diperbarui selama proses pelatihan untuk meningkatkan performa peramalan. Proses pelatihan dilakukan selama 200 *epoch* dengan *batch size* 32 dan learning rate sebesar 0.001, yang dioptimasi menggunakan optimizer Adam untuk menyesuaikan bobot secara adaptif dan mempercepat konvergensi. Pada akhir pelatihan, bobot model yang diperoleh pada epoch ke-200 digunakan sebagai bobot akhir untuk melakukan inferensi. Evaluasi model dilakukan menggunakan *Mean Squared Error* (MSE) sebagai *loss function* untuk meminimalkan selisih kuadrat antara nilai peramalan dan aktual, serta *Mean Absolute Percentage Error* (MAPE) sebagai metrik tambahan untuk mengukur tingkat kesalahan relatif dalam bentuk persentase.

## Analisis dan Evaluasi Model LSTM

Dalam bagian ini, dilakukan analisis terhadap performa model LSTM yang telah dibangun untuk meramalkan permintaan perjalanan KA 1 Argo Bromo Anggrek. Model ini dirancang untuk menangkap pola permintaan berdasarkan data historis yang telah diproses sebelumnya. Evaluasi dilakukan untuk memastikan bahwa model tidak hanya mampu memberikan peramalan yang akurat tetapi juga memiliki tingkat generalisasi yang baik terhadap data baru. Analisis melibatkan beberapa metrik utama, seperti MAPE dan RMSE, guna menilai tingkat akurasi model dalam memperkirakan permintaan penumpang. Selain itu, dilakukan perbandingan antara peramalan model dan data aktual untuk mengidentifikasi pola kesalahan serta area yang memerlukan perbaikan.



Gambar 9. Grafik MSE (Loss) dan MAPE selama pelatihan model LSTM berdasarkan data KA 1 Argo Bromo Anggrek.

Gambar 9 merupakan hasil pelatihan model LSTM yang telah diimplementasikan, dengan menampilkan nilai error terhadap data training dan validasi selama 200 epoch. Grafik menunjukkan bahwa error menurun drastis pada 25 epoch pertama sebelum akhirnya stabil pada nilai yang rendah. Pola konvergensi ini menunjukkan bahwa model cepat belajar dari data *training* tanpa mengalami masalah *overfitting*, ditunjukkan dengan error validasi yang mengikuti *error training* dalam rentang yang wajar. Nilai error validasi yang lebih tinggi di awal pelatihan ( $MAPE > 300$ ) menurun drastis dan stabil yang menunjukkan peningkatan performa model seiring berjalannya *epoch*. Selain itu, perbedaan kecil antara *error training* dan validasi memberikan indikasi bahwa model tidak hanya baik dalam meramalkan data yang sudah dipelajari, tetapi juga cukup kuat dalam memproyeksikan data baru yang tidak pernah dilihat sebelumnya. Dengan mempertahankan error yang rendah, model menunjukkan bahwa kombinasi hyperparameter dan arsitektur yang digunakan mampu memberikan performa yang optimal dalam meramalkan permintaan harian KA 1.

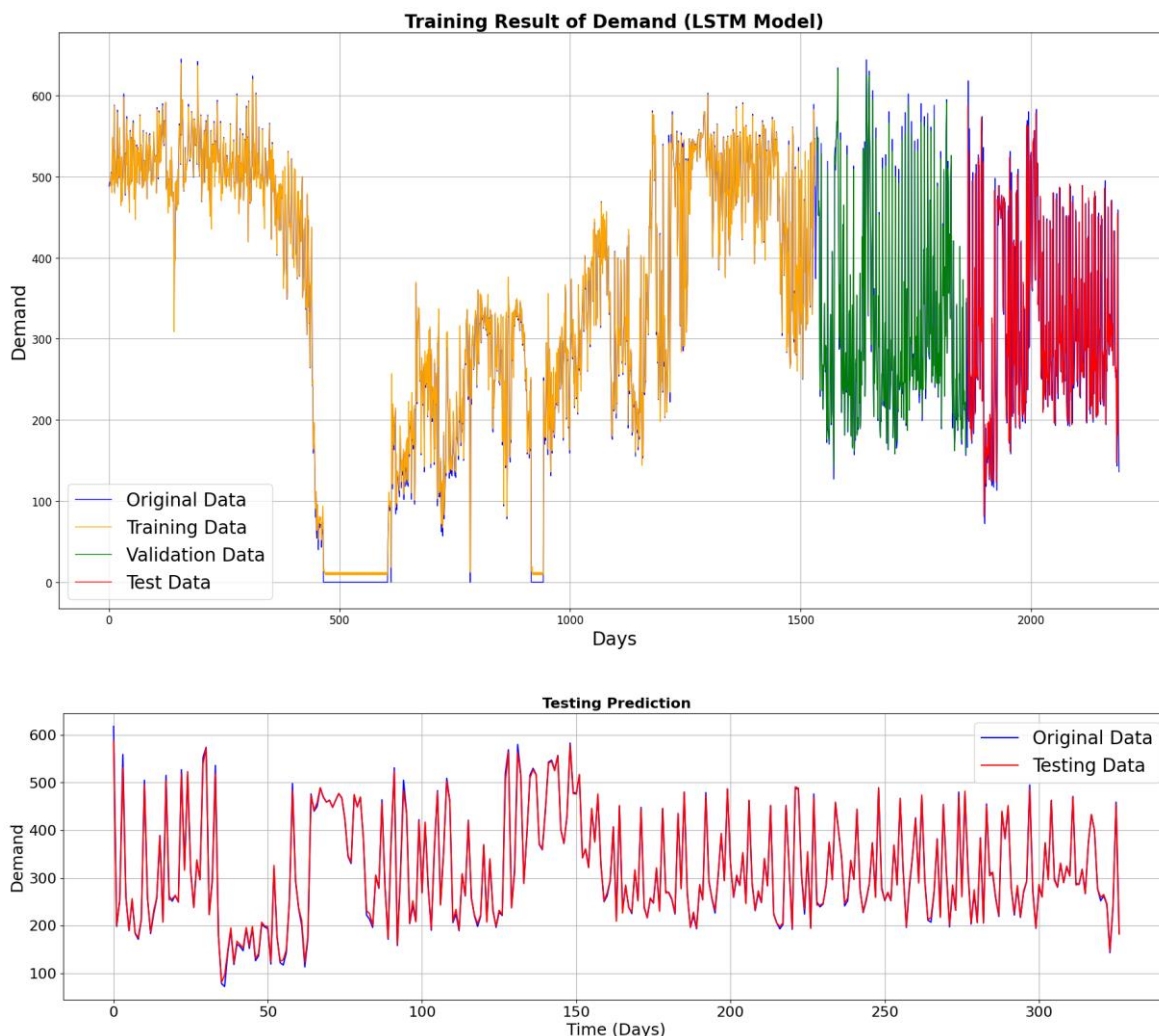
Hasil evaluasi model pada data *training*, *validation*, dan *testing* disajikan dalam Tabel 3. Dari hasil tersebut, dapat dilihat sejauh mana model mengenali pola data dan menghasilkan peramalan dengan nilai MSE dan MAPE yang optimal.

Tabel 3. Evaluasi Kinerja Model LSTM

Dataset	MSE	MAPE
Training	0,0015	7,61%
Validation	0,0021	13,45%
Testing	0,0013	7,53%

Hasil evaluasi tersebut menunjukkan bahwa model memiliki performa yang stabil di semua set data, dengan nilai MAPE terendah pada data testing sebesar 7.53%. Hal ini

menandakan bahwa model tidak hanya efektif pada data yang sudah dipelajari (*training* dan *validasi*), tetapi juga mampu mempertahankan akurasi peramalan pada data baru.



Gambar 10. Visualisasi hasil peramalan model LSTM pada data training, validation, dan testing.

Gambar 10 menampilkan hasil peramalan model LSTM terhadap data permintaan harian KA 1 Argo Bromo Angrek, yang diperoleh dari proses pelatihan dan pengujian model. Grafik menunjukkan kesesuaian pola antara data aktual dan hasil peramalan pada fase training, validation, dan testing. Garis biru menunjukkan data aktual (*original data*), sedangkan garis oranye, hijau, dan merah masing-masing merepresentasikan hasil peramalan pada fase *training*, *validation*, dan *testing*. Seluruh garis peramalan mengikuti tren data aktual dengan baik, terutama pada data training dan validation yang hampir seluruhnya tumpang tindih, menandakan akurasi tinggi. Pada fase pengujian, peramalan tetap konsisten meskipun terdapat sedikit deviasi pada titik-titik dengan fluktuasi tajam.

## KESIMPULAN

Penelitian ini mengembangkan model peramalan permintaan tiket berbasis arsitektur hybrid yang mengintegrasikan metode dekomposisi spektral JUST ke dalam model Long Short-Term Memory (LSTM). Integrasi dilakukan dengan menempatkan hasil dekomposisi trend, seasonal, dan residual sebagai fitur input terpisah yang merepresentasikan dinamika jangka pendek dan panjang dari permintaan tiket kereta api harian. Hasil evaluasi

menunjukkan bahwa model hybrid JUST–LSTM secara konsisten menghasilkan performa peramalan yang baik, dengan nilai MAPE sebesar 7,53% pada data uji. Pencapaian ini menunjukkan bahwa pendekatan dekomposisi mampu memperjelas struktur data sebelum dipelajari oleh LSTM, serta menegaskan bahwa kombinasi JUST dan LSTM dapat mengatasi keterbatasan metode konvensional dalam memodelkan fluktuasi permintaan berbasis data aktual. Temuan ini memberikan kontribusi praktis bagi sektor perkeretaapian nasional dan menawarkan solusi yang kompetitif dalam pengembangan sistem peramalan tiket untuk mendukung manajemen kapasitas dan strategi penetapan tarif dinamis.

## REFERENSI

- [1] W. A. Degife and B. S. Lin, “Deep-Learning-Powered GRU Model for Flight Ticket Fare Forecasting”, *Applied Sciences (Switzerland)*, vol. 13, no. 10, May 2023, doi: 10.3390/app13106032.
- [2] P. B. Floranica and V. Sari, “Peramalan Jumlah Penumpang Bandara Internasional Ahmad Yani Semarang Menggunakan Metode Neural Network Time Series Berbasis Genetic Algorithm”, in *The 9th University Research Colloquium*, Purworejo: Universitas Purworejo, Mar. 2019, pp. 244–250.
- [3] V. Ch. Jwala, K. Jahnavi, K. Mukthamu, K. Durga Madhavi, and K. Jaya Lakshmi, “An Ensemble Learning Method to Predict Airline Ticket Price Using Machine Learning”, *International Journal of Advanced Research in Science, Communication and Technology*, pp. 294–298, Mar. 2024, doi: 10.48175/ijarsct-16664.
- [4] D. A. Lestari and A. A. Mustafidah, “Strategi Perusahaan PT Kereta Api Indonesia (PERSERO) Dalam Meningkatkan Pelayanan Jasa Transportasi Kereta”, *Gudang Jurnal Multidisiplin Ilmu*, vol. 2, pp. 152–157, Feb. 2024, doi: 10.59435/gjmi.v2i2.315.
- [5] B. S. Adicahya, S. Wulandari, and D. Avianto, “Metode Neural Network Dalam Prediksi Jumlah Penumpang Kereta Api Berbasis Web”, *Journal of Information System Research (JOSH)*, vol. 6, no. 1, pp. 291–303, Oct. 2024, doi: 10.47065/josh.v6i1.6001.
- [6] X. Liu and Q. Zhang, “Combining Seasonal and Trend Decomposition Using LOESS With a Gated Recurrent Unit for Climate Time Series Forecasting”, *IEEE Access*, vol. 12, pp. 85275–85290, 2024, doi: 10.1109/ACCESS.2024.3415349.
- [7] S. Balaji and D. Manikavelan, “Hybrid Demand Forecasting Model for Big Marts using Seasonal Decomposition, ARIMA, and Gradient Boosting Regressor”, in *6th International Conference on Mobile Computing and Sustainable Informatics, ICMCSI 2025 - Proceedings*, Institute of Electrical and Electronics Engineers Inc., 2025, pp. 1687–1693. doi: 10.1109/ICMCSI64620.2025.10883480.
- [8] I. Varshavskiy, E. Stavinova, and P. Chunaev, “Forecasting Railway Ticket Demand with Search Query Open Data”, in *Procedia Computer Science*, Elsevier B.V., 2022, pp. 132–141. doi: 10.1016/j.procs.2022.10.215.
- [9] J. Wang and A. Duggasani, “Forecasting Hotel Reservations with Long Short-Term Memory-based Recurrent Neural Networks”, *Int J Data Sci Anal*, vol. 9, no. 1, pp. 77–94, Feb. 2020, doi: 10.1007/s41060-018-0162-6.
- [10] A. Kumar Dubey, A. Kumar, V. García-Díaz, A. Kumar Sharma, and K. Kanhaiya, “Study and Analysis of SARIMA and LSTM in Forecasting Time Series Data”, *Sustainable Energy Technologies and Assessments*, vol. 47, Oct. 2021, doi: 10.1016/j.seta.2021.101474.
- [11] M. Rahimzad, A. Moghaddam Nia, H. Zolfonoon, J. Soltani, A. Danandeh Mehr, and H. H. Kwon, “Performance Comparison of an LSTM-based Deep Learning Model versus Conventional Machine Learning Algorithms for Streamflow Forecasting”, *Water Resources Management*, vol. 35, no. 12, pp. 4167–4187, Sep. 2021, doi: 10.1007/s11269-021-02937-w.
- [12] S. Siami-Namini, N. Tavakoli, and A. Siami Namin, “A Comparison of ARIMA and LSTM in Forecasting Time Series”, in *Proceedings - 17th IEEE International Conference on Machine Learning and Applications, ICMLA 2018*, Institute of Electrical and Electronics Engineers Inc., Jul. 2018, pp. 1394–1401. doi: 10.1109/ICMLA.2018.00227.
- [13] P. Das and S. Barman, “Perspective Chapter: An Overview of Time Series Decomposition and Its Applications”, in *Applied and Theoretical Econometrics and Financial Crises [Working Title]*, IntechOpen, 2025. doi: 10.5772/intechopen.1009268.
- [14] D. A. Kristiyanti and Y. Sumarno, “Penerapan Metode Multiplicative Decomposition (Seasonal) Untuk Peramalan Persediaan Barang Pada PT. Agrinusa Jaya Santosa”, *Jurnal Sistem Komputer dan Kecerdasan Buatan*, vol. 3, no. 2, pp. 45–51, Mar. 2020, doi: 10.47970/siskom-kb.v3i2.145.
- [15] M. Davi and E. Winarko, “Rancang Bangun Aplikasi Peramalan Jumlah Penumpang Menggunakan Long Short-Term Memory (LSTM)”, *Infotekmesin*, vol. 14, no. 2, pp. 303–310, Jul. 2023, doi: 10.35970/infotekmesin.v14i2.1911.

- 
- [16] K. Zhang, X. Huo, and K. Shao, “Temperature Time Series Prediction Model Based on Time Series Decomposition and Bi-LSTM Network”, *Mathematics*, vol. 11, no. 9, p. 1, May 2023, doi: 10.3390/math11092060.
- [17] S. M. Al-Selwi *et al.*, “RNN-LSTM: From Applications to Modeling Techniques and Beyond—Systematic Review”, *Journal of King Saud University - Computer and Information Sciences*, vol. 36, no. 5, Jun. 2024, doi: 10.1016/j.jksuci.2024.102068.
- [18] H. Abbasimehr and R. Paki, “Improving Time Series Forecasting Using LSTM and Attention Models”, *J Ambient Intell Humaniz Comput*, vol. 13, no. 1, pp. 673–691, Jan. 2022, doi: 10.1007/s12652-020-02761-x.
- [19] R. S. Amanu, F. A. Ramadhan, and A. H. Saputra, “Perbandingan Model Prediksi Data Mining dalam Memprediksi Konsentrasi Polutan Karbon Monoksida (CO) di Jakarta”, *Jurnal Teknologi Informasi*, vol. 18, no. 1, p. 7, Jan. 2024, doi: 10.47111/JTI.
- [20] Q. H. Do, S. K. Lo, J. F. Chen, C. L. Le, and L. H. Anh, “Forecasting Air Passenger Demand: A comparison of LSTM and SARIMA”, *Journal of Computer Science*, vol. 16, no. 7, pp. 1063–1084, 2020, doi: 10.3844/JCSSP.2020.1063.1084.
-