

# Prediksi Jumlah Penonton Program Televisi “Stasiun Dangdut” Menggunakan Model *Long Short-Term Memory* (LSTM)

Alyssa Amorita Azzah<sup>1</sup>, Aviolla Terza Damaliana<sup>2</sup>, Wahyu Syaifullah Jauharis Saputra<sup>3</sup>

<sup>1,2,3</sup>Program Studi Sains Data, Universitas Pembangunan Nasional Veteran Jawa Timur

Corresponding author email: [21083010057@student.upnjatim.ac.id](mailto:21083010057@student.upnjatim.ac.id)

## ABSTRAK

Abstrak— Perubahan perilaku audiens di era digital telah menurunkan jumlah penonton televisi konvensional dan menuntut stasiun televisi untuk memanfaatkan analisis data dalam pengambilan keputusan strategis. Penelitian ini bertujuan untuk memprediksi jumlah penonton harian program “Stasiun Dangdut” pada JTV menggunakan model Long Short-Term Memory (LSTM), yang mampu mengenali pola non-linear serta hubungan jangka panjang pada data deret waktu. Dataset yang digunakan terdiri atas data jumlah penonton harian periode 2020–2024, yang melalui tahapan pra-pemrosesan, pembagian data, normalisasi, serta pelatihan model. Hasil evaluasi menunjukkan bahwa model LSTM memberikan performa prediksi yang cukup baik dengan nilai Mean Absolute Error (MAE) sebesar 25.208,12; Root Mean Squared Error (RMSE) sebesar 30.639,20; dan Mean Absolute Percentage Error (MAPE) sebesar 29.76%. Nilai tersebut menunjukkan bahwa dalam penelitian ini, LSTM cukup efektif digunakan untuk peramalan jumlah penonton televisi dengan pola data kompleks serta berpotensi menjadi alat bantu strategis bagi pengelola media dalam merancang kebijakan konten berbasis analitik prediktif.

Keywords: peramalan deret waktu, Long Short-Term Memory, jumlah penonton televisi, prediksi

## I. PENDAHULUAN

Perkembangan teknologi digital dalam satu dekade terakhir telah mengubah lanskap industri media secara fundamental. Laporan Digital 2024: Indonesia menunjukkan bahwa penetrasi internet dan penggunaan media sosial terus meningkat, yang berimplikasi pada terfragmentasinya atensi audiens dari media konvensional ke platform digital [1]. Fenomena ini menuntut stasiun televisi lokal, seperti JTV, untuk lebih adaptif dalam mempertahankan basis pemirsanya. Dalam konteks ini, data jumlah penonton (viewership) bukan sekadar statistik, melainkan aset strategis untuk pengambilan keputusan.

Permasalahan utama yang dihadapi adalah karakteristik data penonton harian yang sangat fluktuatif dan memiliki pola non-linear. Penelitian terdahulu mencatat bahwa metode statistik klasik seperti ARIMA (Auto-Regressive Integrated Moving Average) sering kali memiliki keterbatasan akurasi saat dihadapkan pada data dengan volatilitas tinggi atau tidak memiliki pola musiman yang konsisten [2, 3]. Metode konvensional cenderung bekerja optimal pada data stasioner, sedangkan data penonton televisi sering kali mengandung noise akibat faktor eksternal yang acak.

Untuk mengatasi keterbatasan tersebut, pendekatan berbasis Deep Learning menjadi alternatif yang menjanjikan. Salah satu arsitektur yang terbukti andal adalah Long Short-Term Memory (LSTM). Berbagai literatur memvalidasi keunggulan LSTM dalam menangkap dependensi jangka panjang (long-term dependencies). Sebagai contoh, Julian dan Pribadi [4] serta Rizkillah dan Widiyanesti [5] membuktikan bahwa LSTM mampu memprediksi data time series yang fluktuatif (seperti saham dan kripto) dengan

akurasi yang baik. Senada dengan itu, survei komprehensif oleh Lim dan Zohren [6] menempatkan arsitektur Recurrent Neural Network (RNN) seperti LSTM sebagai metode yang unggul untuk time series forecasting modern.

Penelitian ini bertujuan untuk memprediksi jumlah penonton harian program “Stasiun Dangdut” menggunakan metode LSTM. Urgensi penelitian ini terletak pada kebutuhan industri penyiaran lokal akan alat bantu analitik yang presisi untuk optimalisasi strategi konten di tengah persaingan digital.

## II. METODOLOGI PENELITIAN

Metodologi penelitian ini dirancang secara sistematis untuk memastikan validitas model prediksi yang dibangun. Tahapan penelitian meliputi pengumpulan dataset, pra-pemrosesan data, pembentukan sequence, perancangan arsitektur model, pengaturan parameter pelatihan, dan evaluasi kinerja model.

### 2.1 Dataset

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari stasiun televisi JTV. Dataset mencakup data harian jumlah penonton (viewership) program "Stasiun Dangdut" dengan rentang waktu dari 1 Januari 2020 hingga 31 Desember 2024. Format data asli berupa deret waktu (time series) dengan atribut utama berupa Tanggal dan Jumlah Penonton. Total data mencakup 1.827 baris data harian.

Tabel 1. Data Jumlah Penonton

Tanggal	Jumlah Penonton
2020-01-01	76.340
2020-01-02	82.628
2020-01-03	151.077
2020-01-04	79.328
...	...
2024-12-28	53.332
2024-12-29	103.292
2024-12-30	22143
2024-12-31	72.079

### 2.2 Preprocessing Data

Tahap pra-pemrosesan bertujuan untuk meningkatkan kualitas data sebelum dimasukkan ke dalam model. Langkah-langkah yang dilakukan meliputi:

#### a. Pembersihan Data

Dilakukan pengecekan terhadap nilai yang hilang (missing values). Jika ditemukan data kosong, dilakukan imputasi menggunakan metode interpolasi linier. Metode ini dipilih untuk mengisi kekosongan data dengan nilai rata-rata di antara dua titik data yang ada, sehingga kontinuitas tren data tetap terjaga.

#### b. Konversi Tipe Data

Kolom Tanggal dikonversi menjadi tipe datetime dan ditetapkan sebagai indeks data untuk memastikan urutan kronologis yang tepat dalam analisis deret waktu.

#### c. Normalisasi Data

Algoritma LSTM menggunakan fungsi aktivasi (seperti sigmoid dan tanh) yang sensitif terhadap skala data input. Data dengan rentang nilai yang sangat besar (puluhan hingga ratusan ribu) dapat memperlambat proses konvergensi gradien. Oleh karena itu, dilakukan normalisasi menggunakan teknik MinMax Scaler untuk mengubah rentang nilai data menjadi [0, 1].

### 2.3 Arsitektur Model LSTM

Model prediksi dibangun menggunakan kerangka kerja (framework) TensorFlow/Keras. Arsitektur jaringan saraf yang diusulkan terdiri dari lapisan-lapisan berikut:

- a. Input Layer  
Menerima sequence data hasil pra-pemrosesan.
- b. LSTM Layer  
Menggunakan 64 unit neuron. Lapisan ini bertugas menangkap pola temporal dan dependensi jangka panjang dari data historis melalui mekanisme gerbang (gates) internalnya.
- c. Dense Layer  
Satu neuron pada lapisan output (fully connected layer) untuk menghasilkan nilai prediksi tunggal (regresi).
- d. Optimizer  
Menggunakan algoritma Adam (Adaptive Moment Estimation) karena efisiensinya dalam menangani data non-stationary dan kemampuan adaptasi learning rate secara otomatis [7].
- e. Loss Function  
Menggunakan Mean Squared Error (MSE) sebagai fungsi objektif yang diminimalkan selama pelatihan.

### 2.4 Parameter Pelatihan

Proses pelatihan model dilakukan dengan konfigurasi parameter yang disesuaikan berdasarkan hasil eksperimen untuk mencapai konvergensi optimal:

- a. Pembagian Data  
Dataset dibagi menjadi 80% data latih (training set) dan 20% data uji (testing set) secara sekuensial (tanpa pengacakan) untuk menjaga integritas urutan waktu.
- b. Epochs  
Epochs dilakukan sebanyak 50 iterasi. Jumlah ini dipilih untuk memastikan model belajar cukup lama hingga mencapai loss yang minimum tanpa terjebak dalam overfitting.
- c. Batch Size  
Batch size yang dipilih adalah 16, yang berarti bobot model diperbarui setiap 16 sampel data. Ukuran batch yang lebih kecil membantu model memperbarui gradien lebih sering dan menangkap detail fluktuasi data dengan lebih baik.

### 2.5 Evaluasi Model

Untuk mengukur akurasi model secara kuantitatif, penelitian ini menggunakan tiga metrik evaluasi standar dalam peramalan deret waktu:

- a. Mean Absolute Error (MAE)  
Mean Absolute Error atau MAE digunakan untuk mengukur rata-rata besaran kesalahan absolut tanpa memperhatikan arah kesalahan.
- b. Root Mean Squared Error (RMSE)  
Root Mean Squared Error atau RMSE digunakan untuk memberikan

penalti lebih besar pada kesalahan prediksi yang ekstrem (besar) karena proses pengkuadratan selisih error.

c. Mean Absolute Percentage Error (MAPE)

Mean Absolute Percentage Error atau MAPE digunakan untuk mengukur rata-rata persentase kesalahan relatif terhadap nilai aktual, yang memudahkan interpretasi kinerja model dalam konteks bisnis. Secara umum, interpretasinya adalah sebagai berikut :

Tabel 2. Interpretasi MAPE

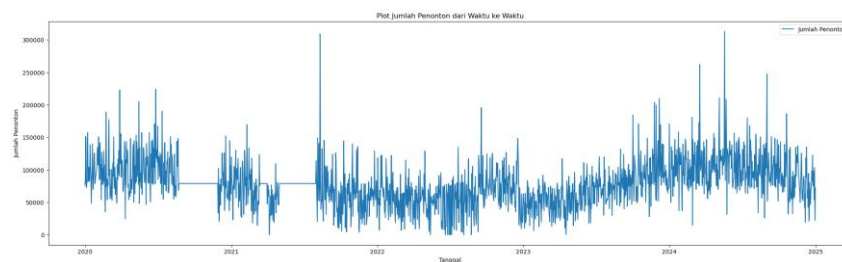
No	Akurasi	Error
1	< 10%	Akurasi sangat baik
2	10–20%	Akurasi baik
3	20–50%	Akurasi cukup / wajar
4	> 50%	Akurasi buruk

### III. HASIL DAN ANALISIS

Bagian ini memaparkan hasil eksperimen yang telah dilakukan, mulai dari analisis karakteristik data asli, proses dekomposisi untuk memahami komponen penyusun data, hingga evaluasi kinerja model *Long Short-Term Memory* (LSTM) yang dibangun. Analisis mendalam dilakukan untuk memvalidasi apakah model mampu menangkap pola kompleks dari data jumlah penonton program "Stasiun Dangdut" yang bersifat dinamis.

#### 3.1 Analisis Data Asli

Langkah awal dalam memahami karakteristik dataset adalah melalui visualisasi data deret waktu secara keseluruhan. Dataset mencakup data harian dari tahun 2020 hingga 2024.



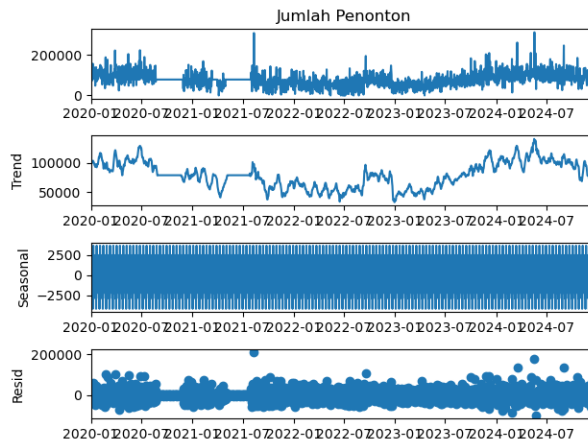
Gambar 3.1 Visualisasi data harian

Berdasarkan Gambar 3.1, visualisasi data harian memperlihatkan pola distribusi yang sangat fluktuatif dengan variabilitas yang tinggi antar-waktu. Grafik menunjukkan adanya lonjakan (*spikes*) jumlah penonton yang signifikan pada periode tertentu, diikuti dengan penurunan tajam yang berulang. Tidak terlihat adanya tren kenaikan atau penurunan yang monoton dan mulus secara jangka panjang. Ketidakberaturan pola ini mengindikasikan bahwa data memiliki tingkat volatilitas yang tinggi (*high volatility*), yang sulit dimodelkan menggunakan metode statistik linear biasa. Hal ini mengonfirmasi perlunya penerapan metode non-linear berbasis *Deep Learning* seperti LSTM yang dirancang khusus untuk menangkap variasi dinamis dan dependensi jangka panjang tersebut.

#### 3.2 Dekomposisi Deret Waktu

Untuk menggali lebih dalam struktur data yang tersembunyi, dilakukan proses dekomposisi menggunakan metode `seasonal_decompose`. Teknik ini

memisahkan data deret waktu asli menjadi tiga komponen penyusun utama: *Trend* (kecenderungan jangka panjang), *Seasonal* (pola musiman yang berulang), dan *Residual* (sisa atau *noise*).

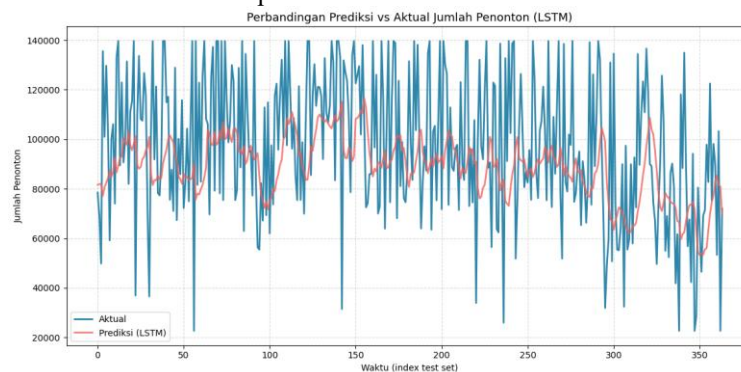


Gambar 3.2 Plot Dekomposisi

Sebagaimana terlihat pada Gambar 3.2, hasil dekomposisi memperlihatkan karakteristik masing-masing komponen. Komponen *Trend* menunjukkan pergerakan data yang relatif lebih halus namun tetap dinamis. Sementara itu, komponen *Seasonal* memperlihatkan pola repetitif yang sangat rapat, mengindikasikan adanya siklus mingguan atau bulanan yang konsisten. Yang paling menarik perhatian adalah komponen *Residual*. Terlihat bahwa komponen residual memiliki variansi yang cukup tinggi dan pola yang acak (*random*). Besarnya nilai residual ini menandakan adanya faktor eksternal atau *noise* dominan yang tidak dapat dijelaskan hanya oleh tren maupun musiman. Tingginya komponen acak ini menegaskan kembali bahwa permasalahan prediksi ini memiliki kompleksitas tinggi yang sulit diselesaikan oleh metode peramalan sederhana.

### 3.3 Prediksi vs Aktual

Setelah proses pelatihan selesai, performa model diuji pada data uji (test set) untuk melihat seberapa baik model memprediksi data masa depan yang belum pernah dilihat sebelumnya. Gambar 3.3 memvisualisasikan perbandingan antara nilai aktual dan nilai prediksi model LSTM.



Gambar 3.3 Plot Prediksi vs Aktual

Visualisasi hasil prediksi pada data uji memperlihatkan bahwa model LSTM mampu mengikuti pola pergerakan data aktual secara umum. Model berhasil menangkap momen-momen kenaikan dan penurunan jumlah penonton, yang ditandai dengan himpitan antara garis prediksi dan garis aktual. Hal ini membuktikan bahwa model telah berhasil mempelajari pola dasar dari data historis.

Meskipun demikian, jika diamati lebih teliti, terdapat sedikit keterlambatan (*lag*) dalam respon model terhadap perubahan tren yang sangat ekstrem atau mendadak. Fenomena ini merupakan karakteristik umum pada prediksi deret waktu satu langkah ke depan (*one-step ahead prediction*), di mana model cenderung sangat dipengaruhi oleh nilai data terakhir ( $t-1$ ). Kendati demikian, arah pergerakan prediksi tetap konsisten dengan data aktual, menunjukkan bahwa model cukup reliabel untuk melihat tren jangka pendek.

### 3.4 Evaluasi Model

Untuk mengukur kinerja model secara kuantitatif dan objektif, dilakukan perhitungan *error* menggunakan tiga metrik evaluasi standar, yaitu Persamaan (2), (3), dan (4). Hasil evaluasi disajikan dalam Tabel 3 berikut:

Tabel 3. Hasil Evaluasi Model

No	Metrik	Nilai
1	MAE	25.208,12
2	RMSE	30.639,20
3	MAPE	29,76%

Nilai MAPE sebesar 29,76% menunjukkan bahwa rata-rata penyimpangan prediksi model terhadap data aktual adalah sekitar 29%. Angka ini mengindikasikan bahwa model memiliki kemampuan prediksi yang cukup akurat, mengingat karakteristik data yang sangat fluktuatif dan penuh *noise*. Hasil ini sejalan dengan temuan Puteri [8] pada penelitian dengan domain *time series* serupa, di mana nilai MAPE di bawah 50% sering kali dianggap dapat diterima (*acceptable*) untuk data dengan volatilitas tinggi yang dipengaruhi oleh banyak faktor eksternal tak terduga.

## IV. KESIMPULAN

Penelitian ini berhasil menerapkan metode Long Short-Term Memory (LSTM) untuk memprediksi jumlah penonton harian program televisi lokal "Stasiun Dangdut". Berdasarkan hasil analisis, dapat disimpulkan bahwa LSTM memiliki kemampuan yang efektif dalam menangkap pola data yang non-linear dan sangat fluktuatif, yang dibuktikan dengan nilai MAPE sebesar 29,76%. Meskipun model mampu mengikuti tren pergerakan data, keterbatasan utama penelitian ini adalah penggunaan data univariat (hanya data historis penonton). Performa prediksi pada titik-titik ekstrem masih mengalami sedikit keterlambatan (*lag*). Oleh karena itu, untuk penelitian selanjutnya, sangat disarankan untuk menambahkan variabel eksogen (*multivariate analysis*) seperti data hari libur nasional, data interaksi media sosial, atau jadwal bintang tamu acara [10] guna meningkatkan sensitivitas dan akurasi model terhadap lonjakan penonton.

## REFERENSI

- [1] We Are Social, "Digital 2024: Indonesia," *DataReportal*, 2024.
- [2] Y. S. Nugroho and A. H. Setyawan, "Analisis Prediksi Harga Saham PT. Telkom Indonesia Menggunakan Metode Arima Dan LSTM," *Jurnal Ilmiah Komputasi*, vol. 20, no. 3, pp. 385-392, 2021.
- [3] A. S. Aribowo, "Penerapan Deep Learning Menggunakan Long Short-Term Memory untuk Prediksi Harga Saham Gabungan," *Jurnal Media Informatika Budidarma*, vol. 6, no. 1, pp. 313-319, 2022.

- [4] R. Julian and M. R. Pribadi, "Peramalan Harga Saham Pertambangan Pada Bursa Efek Indonesia (BEI) Menggunakan Long Short Term Memory (LSTM)," *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, vol. 8, no. 3, 2021.
- [5] M. Rizkilloh and F. Widiyanesti, "Prediksi Harga Cryptocurrency Menggunakan Algoritma Long Short Term Memory (LSTM)," *Jurnal RESTI*, vol. 6, no. 1, pp. 25–31, 2022.
- [6] B. Lim and S. Zohren, "Time-series forecasting with deep learning: a survey," *Philosophical Transactions of the Royal Society A*, vol. 379, no. 2194, 2021.
- [7] J. Brownlee, *Deep Learning for Time Series Forecasting: Predict the Future with MLPs, CNNs and LSTMs*, Machine Learning Mastery, 2020.
- [8] D. I. Puteri, "Implementasi Long Short Term Memory (LSTM) dan Bidirectional Long Short Term Memory (BiLSTM) Dalam Prediksi Harga Saham Syariah," *Euler: Jurnal Ilmiah Matematika, Sains dan Teknologi*, vol. 11, no. 1, pp. 35-43, 2023.
- [9] Cahyani, "Implementasi Metode Long Short Term Memory (LSTM) untuk Memprediksi Harga Bahan Pokok Nasional," *JUSTIN (Jurnal Sistem dan Teknologi Informasi)*, vol. 12, no. 4, 2024.
- [10] W. Widayat, "Analisis Sentimen Movie Review Menggunakan Word2Vec dan Metode LSTM Deep Learning," *Jurnal Media Informatika Budidarma*, vol. 5, no. 3, p. 1018, 2021.