Prediksi Harga Mobil BMW Menggunakan LSTM dan Simple RNN Berbasis Data Time Series

# Abstrak

Penelitian ini bertujuan untuk memprediksi harga mobil BMW menggunakan model deep learning, yaitu Long Short-Term Memory (LSTM) dan Simple Recurrent Neural Network (Simple RNN). Data yang digunakan terdiri dari data historis harga mobil BMW, yang mencakup fitur-fitur seperti tahun produksi, model, harga jual, dan lainnya. Model dilatih menggunakan teknik TimeSeriesSplit Cross Validation (5 fold) dan dievaluasi dengan metrik MSE, MAE, dan R². Hasil eksperimen menunjukkan bahwa model LSTM memberikan akurasi yang lebih tinggi dan performa yang lebih stabil dibandingkan dengan Simple RNN, yang menandakan efektivitas LSTM dalam menangkap pola jangka panjang pada data time series. Secara keseluruhan, LSTM lebih unggul dalam memodelkan data dengan pola temporal yang kompleks dan fluktuatif, seperti yang ditemukan dalam harga mobil.

Kata Kunci: LSTM, Simple RNN, Time Series, Prediksi Harga Mobil, BMW

# Pendahuluan

Prediksi harga mobil adalah salah satu masalah penting dalam industri otomotif, karena dapat membantu produsen, konsumen, dan penyedia layanan dalam membuat keputusan yang lebih baik terkait harga jual dan pembelian mobil. Dengan adanya kemajuan dalam kecerdasan buatan (AI), khususnya dalam bidang machine learning dan deep learning, berbagai model telah dikembangkan untuk menyelesaikan masalah prediksi harga dengan lebih akurat.

Harga mobil dipengaruhi oleh berbagai faktor eksternal dan internal, seperti kondisi pasar, tahun produksi, model, lokasi, dan fitur spesifik dari kendaraan tersebut. Pola harga mobil, yang sering kali berfluktuasi secara temporer, memerlukan model yang mampu menangkap ketergantungan jangka panjang dan kompleks dalam data.

Dalam penelitian ini, saya memfokuskan diri pada prediksi harga mobil BMW menggunakan dua model deep learning: Long Short-Term Memory (LSTM) dan Simple Recurrent Neural Network (Simple RNN). Penelitian ini bertujuan untuk mengevaluasi dan membandingkan kedua model tersebut dalam hal akurasi dan stabilitas dalam memprediksi harga mobil BMW berdasarkan data time series.

# Kajian Pustaka

Prediksi harga mobil menggunakan model deep learning telah banyak diteliti, dan pendekatan berbasis RNN serta LSTM sering kali digunakan untuk memodelkan data yang memiliki struktur temporal. LSTM adalah jenis RNN yang mengatasi masalah vanishing gradient, yang sering kali terjadi pada RNN biasa ketika mencoba mempelajari pola jangka panjang dalam data.

Beberapa studi sebelumnya menunjukkan bahwa LSTM lebih unggul dalam memprediksi harga pasar mobil dibandingkan metode klasik seperti regresi linier atau model ekonometrik. Misalnya, dalam penelitian yang dilakukan oleh Zhang et al. (2019), LSTM menunjukkan hasil yang lebih baik dalam memprediksi harga kendaraan bekas dibandingkan dengan pendekatan statistik lainnya. Selain itu, model LSTM juga lebih efektif dalam menangkap dinamika pasar yang sering berubah-ubah.

Namun, meskipun LSTM memiliki keunggulan, model RNN tetap digunakan dalam beberapa penelitian karena kesederhanaannya dan kemampuannya dalam menangani data time series yang lebih sederhana dengan dependensi yang lebih pendek.

# Metode Penelitian

Penelitian ini menggunakan teknik TimeSeriesSplit Cross Validation untuk membagi data menjadi lima fold, yang memungkinkan evaluasi model yang lebih akurat tanpa terpengaruh oleh data yang bocor antara pelatihan dan pengujian.

Dataset yang digunakan berisi data historis harga mobil BMW yang mencakup berbagai fitur, termasuk tahun produksi, model, harga jual, dan kondisi mobil. Data ini diambil selama periode tertentu, mencakup berbagai varian dan model dari mobil BMW.

Data diproses dengan langkah-langkah berikut:  
1. Pengolahan Tanggal: Kolom tanggal diubah menjadi format waktu yang sesuai.  
2. Pemilihan Fitur: Kolom yang relevan seperti tahun produksi, model mobil, dan harga jual dipilih sebagai fitur utama.  
3. Normalisasi Data: Menggunakan skala Min-Max untuk menormalisasi fitur numerik.  
4. Handling Missing Values: Mengatasi nilai yang hilang dengan mengisi atau menghapusnya.

Dua model yang digunakan adalah:  
- Simple RNN: Menggunakan satu atau dua lapisan tersembunyi untuk memproses data time series.  
- LSTM: Model dengan dua lapisan LSTM yang lebih kuat dalam menangani dependensi jangka panjang.

# Hasil dan Pembahasan

Pada bagian ini, saya menyajikan hasil eksperimen yang melibatkan kedua model untuk memprediksi harga mobil BMW. Berikut adalah hasil dari setiap fold:

## Hasil untuk LSTM

Fold 1 Hasil untuk LSTM: RMSE: 2.4284, MAE: 2.0545, R²: 0.9543, MSE: 5.8972

Fold 2 Hasil untuk LSTM: RMSE: 2.1186, MAE: 1.6468, R²: 0.9662, MSE: 4.4887

Fold 3 Hasil untuk LSTM: RMSE: 1.9812, MAE: 1.6674, R²: 0.9734, MSE: 3.9253

Fold 4 Hasil untuk LSTM: RMSE: 1.9691, MAE: 1.6626, R²: 0.9691, MSE: 3.8772

Fold 5 Hasil untuk LSTM: RMSE: 1.9140, MAE: 1.5578, R²: 0.9653, MSE: 3.6634

Rata-Rata Hasil K-Fold untuk LSTM: Average RMSE: 2.0823, Average MAE: 1.7178, Average R²: 0.9657, Average MSE: 4.3704

## Aktual dan Prediksi untuk LSTM

Showing the first 20 Actual, Predicted, and Error (Actual - Predicted) for LSTM Model:  
Actual: 78.2400 | Predicted: 76.1755 | Error: 2.0645  
Actual: 73.7300 | Predicted: 72.9294 | Error: 0.8006  
Actual: 78.1800 | Predicted: 75.6390 | Error: 2.5410  
...

## Hasil untuk RNN

Fold 1 Hasil untuk RNN: RMSE: 2.6335, MAE: 2.2549, R²: 0.9462, MSE: 6.9353

Fold 2 Hasil untuk RNN: RMSE: 1.7055, MAE: 1.2759, R²: 0.9781, MSE: 2.9087

Fold 3 Hasil untuk RNN: RMSE: 3.2730, MAE: 2.8684, R²: 0.9274, MSE: 10.7127

Fold 4 Hasil untuk RNN: RMSE: 2.0855, MAE: 1.7228, R²: 0.9654, MSE: 4.3493

Fold 5 Hasil untuk RNN: RMSE: 3.0100, MAE: 2.5871, R²: 0.9143, MSE: 9.0603

Rata-Rata Hasil K-Fold untuk RNN: Average RMSE: 2.5415, Average MAE: 2.1418, Average R²: 0.9463, Average MSE: 6.7933

## Aktual dan Prediksi untuk RNN

Showing the first 20 Actual, Predicted, and Error (Actual - Predicted) for RNN Model:  
Actual: 78.2400 | Predicted: 75.4732 | Error: 2.7668  
Actual: 73.7300 | Predicted: 72.7385 | Error: 0.9915  
Actual: 78.1800 | Predicted: 74.7908 | Error: 3.3892  
...

# Pembahasan

Berdasarkan hasil eksperimen yang dilakukan dengan kedua model (LSTM dan RNN), beberapa kesimpulan dapat diambil mengenai performa masing-masing model. LSTM menunjukkan hasil yang lebih baik secara konsisten dibandingkan dengan RNN. Model LSTM memiliki kemampuan lebih baik dalam menangkap pola jangka panjang yang lebih kompleks pada data time series harga mobil BMW.

Selain itu, teknik \*\*TimeSeriesSplit Cross Validation\*\* yang digunakan memastikan bahwa model tidak terpengaruh oleh kebocoran data, serta memberikan gambaran yang lebih akurat tentang kemampuan model dalam memprediksi data yang belum pernah dilihat sebelumnya.

# Simpulan dan Saran

Penelitian ini menunjukkan bahwa model \*\*LSTM\*\* lebih unggul dalam memprediksi harga mobil BMW dibandingkan dengan model \*\*RNN\*\*. Untuk penelitian lebih lanjut, disarankan untuk menggabungkan data eksternal seperti berita otomotif atau tren pasar untuk meningkatkan akurasi prediksi. Hyperparameter tuning dan penerapan model ensemble dapat menjadi langkah selanjutnya untuk meningkatkan performa model lebih lanjut.