Advanced Time Series Forecasting with Deep Learning and Attention Mechanisms

1. Introduction

This project focuses on multivariate electricity consumption forecasting using deep learning models. The goal is to predict future power consumption using historical data and multiple correlated variables. Traditional sequence models are extended with attention mechanisms to improve performance and interpretability.

2. Dataset Description

A multivariate electricity consumption dataset was used, consisting of more than 1000 time steps and 7 features. The dataset exhibits clear trends, seasonality, and correlations between variables. Missing values were handled by removing incomplete rows due to the large dataset size.

3. Data Preprocessing

The data was scaled using MinMaxScaler. Sliding window technique was applied with a 24-hour input window to predict the next 6 time steps. The dataset was split into training, validation, and test sets.

4. Baseline Models

A Dense Neural Network and a standard LSTM model were implemented as baselines. The Dense model flattened temporal information, while the LSTM captured sequential dependencies. The LSTM significantly outperformed the Dense model.

5. LSTM with Attention Model

An attention mechanism was added on top of LSTM outputs to allow the model to focus on the most relevant historical time steps. This improves both forecasting accuracy and interpretability. The attention layer learns importance weights across the input sequence.

6. Model Evaluation

Performance was evaluated using Root Mean Squared Error (RMSE). Results:

- Dense Model RMSE: 1.106 kW
- LSTM RMSE: 0.562 kW
- LSTM + Attention RMSE (scaled): 0.115

The attention-based model achieved the best performance.

7. Attention Interpretation

Attention weight visualization showed that recent time steps contribute more strongly to predictions, while earlier time steps have lower influence. This aligns with real-world electricity consumption behavior.

8. Conclusion

The project demonstrates that incorporating attention mechanisms into LSTM models improves forecast accuracy and model explainability for multivariate time series forecasting.

9. Future Work

Future extensions include Transformer-based architectures, hyperparameter optimization, probabilistic forecasting using quantile loss, and feature-wise attention analysis.