Assignment 3
Time-Series Data
Report

1. Introduction

Time-series forecasting is integral to weather prediction, where RNNs, especially LSTMs (Long Short-Term Memory) and GRUs (Gated Recurrent Units), excel due to their ability to capture temporal dependencies. This report explores and compares LSTM and GRU layers, as well as a hybrid 1D CNN + RNN model, which combines CNN's spatial-temporal pattern recognition with CNN's long-term sequence learning. The goal was to develop a model capable of accurately predicting critical weather metrics.

2. Data Preparation

The dataset used, **`jena_climate_2009_2016.csv`**, contains hourly climate measurements from 2009 to 2016, covering variables such as temperature, humidity, atmospheric pressure, and wind speed.

Preprocessing Steps:

- **Data Cleaning**: Missing values were handled to ensure consistency.
- **Feature Scaling**: All features were normalised to accelerate training convergence.
- **Splitting the Dataset**: The data was divided into training, validation, and test sets to enable fair evaluation of model performance.

3. Exploratory Data Analysis (EDA)

EDA provided insights into the patterns within the dataset, such as:

- **Seasonal Temperature Cycles**: Clear seasonal trends were observed in temperature, which helped set model parameters.
- **Correlation Among Variables**: Temperature and humidity displayed an inverse relationship, providing direction for feature engineering.
- **Wind and Humidity Cycles**: Periodic variations were noted in wind and humidity, indicating seasonal influences on these factors.

Visualisations helped identify the temporal dependencies and patterns relevant for model training.

4. Methodology

Three models were tested, each with distinct configurations:

Model 1: Stacked LSTM

A stacked LSTM architecture served as a baseline. By adjusting the number of units across layers, we aimed to enhance the model's ability to capture long-term dependencies in weather data.

Model 2: GRU-Based RNN

A GRU-based RNN was explored to compare against the LSTM. GRUs are computationally less intensive and can train faster, providing a more efficient alternative for time-series tasks with moderate temporal complexity.

Model 3: Hybrid 1D CNN + LSTM

This model combined a 1D CNN layer with an LSTM to leverage CNN's ability to capture local patterns within the time sequence. This hybrid architecture allowed the model to recognize both short-term and long-term dependencies more effectively, resulting in improved forecasting accuracy.

5. Model Optimization

Key optimization strategies included:

1. **Tuning Layer Units**: Different unit configurations were tested in each recurrent layer to better capture the dataset's complexity.
2. **Layer Selection (LSTM vs. GRU)**: While LSTM generally outperformed GRU in terms of validation accuracy, GRU showed faster convergence.
3. **CNN-RNN Combination**: The hybrid model provided the most accurate predictions by capturing both localised and extended patterns within the data.

6. Results and Discussion

Each model's performance was evaluated based on **Mean Absolute Error (MAE)** and **Mean Squared Error (MSE)**. The following observations were made:

- The stacked LSTM provided strong baseline results, but struggled with some long-term dependencies.
- The GRU-based model was efficient in terms of training time, though it slightly underperformed compared to the LSTM model.
- The hybrid 1D CNN + LSTM achieved the highest accuracy, indicating its superior ability to capture both localised and long-term weather patterns. This model's success can be attributed to the complementary strengths of CNNs and LSTMs in identifying and retaining complex temporal relationships.

Evaluation of Models :

Results Summary :

- Summarise the performance of each model, referencing the table of MAE scores :
  - Naive Method : 2.62
  - Densely Connected Network : 2.63
  - 1D Convolutional Model: 3.21
  - Simple LSTM Model : 3.17
  - RNN Model : 9.93
  - Stacked RNN Model : 9.91
  - Dropout LSTM Model : 2.59
  - Simple LSTM with 32 Units : 2.63
  - Stacked LSTM with 64 Units : 2.57
  - Stacked LSTM with 8 Units : 2.78
  - 1D Convolutional with RNN : 3.7

7. Conclusion

This assignment demonstrated the effectiveness of advanced RNN architectures for time-series forecasting, particularly for weather data. The 1D CNN + LSTM model emerged as the most accurate, benefiting from CNN's capability to capture local temporal structures and LSTM's strength in retaining long-term dependencies. Future work could explore additional hyperparameter tuning or incorporate external climate data to further enhance the model's predictive robustness.