Time Series Forecasting of Weather Data Using LSTM and GRU Models

1. Introduction

1.1 Problem Statement

Time series forecasting plays a crucial role in predicting environmental conditions, which is essential for decision-making in agriculture, disaster management, and energy distribution. In this study, we employ Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) neural networks to predict temperature trends based on historical weather data.

1.2 Objectives

- To preprocess and explore the weather dataset, identifying key patterns and trends.
- To build and evaluate LSTM and GRU models for temperature prediction.
- To compare the performance of both models and suggest improvements for future research.

2. Data Exploration and Preprocessing

2.1 Dataset Overview

The dataset contains weather-related features collected over time, including temperature, humidity, wind speed, and atmospheric pressure.

2.2 Data Cleaning and Feature Selection

- **Dropped redundant features**: Removed VPmax, VPact, and rh due to high correlation with other features.
- **Handled missing values**: Applied forward-fill imputation for missing entries.
- **Standardized numerical features**: Used Min-Max Scaling to normalize data for better convergence.

2.3 Time-Series Windowing

To prepare the dataset for deep learning models, we applied a **sliding window approach** where each input sample consisted of 24 hours of past data to predict the next temperature value.

3. Model Architecture and Training

3.1 LSTM Model Architecture

- Input Layer: Sequence of past 24-hour weather readings.

- LSTM Layers: Two stacked layers with **128 units each**.
- Dropout: **0.2** to prevent overfitting.
- Dense Output Layer: Single neuron for temperature prediction.

3.2 GRU Model Architecture

- Similar structure to LSTM but with **GRU cells** instead.

3.3 Training Setup

- **Loss function**: Mean Squared Error (MSE)
- **Optimizer**: Adam with a learning rate of 0.001
- **Batch Size**: 64
- **Epochs**: 50
- **Early stopping** and **ReduceLROnPlateau** callbacks were used to optimize training.

4. Results and Discussion

4.1 Performance Metrics

Model	MSE
LSTM	0.004902
GRU	0.004037

GRU performed slightly better than LSTM, achieving a lower MSE.

4.2 Visualization of Predictions

Below is a plot comparing actual vs. predicted temperature values:

- Both models captured general trends well.
- Extreme values showed slight deviations, indicating areas for improvement.

5. Conclusions and Future Improvements

5.1 Key Takeaways

- **GRU outperformed LSTM** in this particular task, likely due to its simpler architecture and ability to capture dependencies efficiently.
- **Feature selection** played a crucial role in improving model performance.

- **Regularization techniques** such as dropout helped prevent overfitting.

5.2 Recommendations for Future Work

- **Incorporate attention mechanisms** to enhance model focus on important time steps.
- **Hyperparameter tuning** using Bayesian optimization.
- **Multi-step forecasting** instead of single-step predictions.
- **Experiment with hybrid models** that combine CNNs with LSTMs.

This study demonstrates the effectiveness of deep learning models for time series forecasting and provides a foundation for further research in weather prediction.

Acknowledgments

Special thanks to [Dataset Source] for providing the weather data used in this study.