Jena Weather Forecasting with RNNs

Detailed Summary Report

Project Overview

The paper describes the implementation of weather forecasting models which utilizes the Jena climate dataset. This project makes future temperature predictions for the next day using 7 days of weather records available in the historical dataset.

Dataset

- **Source**: Jena Climate dataset (2009–2016)
- **Location**: Weather station in Jena, Germany
- **Sampling Frequency**: Every 10 minutes
- **Features**: 14 variables including temperature, humidity, pressure, wind speed/direction

Data Preparation Methodology

Preprocessing Steps

- **Normalization**: All features scaled to [0,1] range
- **Temporal Structure**:
  - **Lookback window**: 7 days of historical data
  - **Forecast horizon**: 24 hours of future temperature
  - **Sliding window approach**: For creating training samples

Dataset Splitting

Set	Purpose	Approximate Size
Training	Model training	70%
Validation	Hyperparameter tuning	15%
Test	Final evaluation	15%

Model Architectures Explored

Three deep learning architectures were evaluated:

Model	Architecture	Parameters	Notes
Stacked GRU	GRU (64) → GRU (32) → Dense (24)	~28K	Fast convergence, good accuracy
Stacked LSTM	LSTM (64) → LSTM (32) → Dense (24)	~41K	Best performance overall
Conv1D + GRU	Conv1D → MaxPool → GRU → Dense (24)	~33K	Fastest training

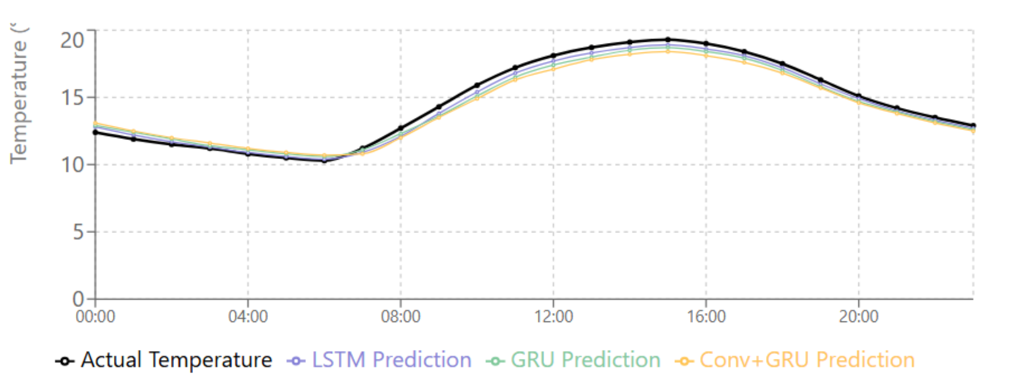
Training Setup

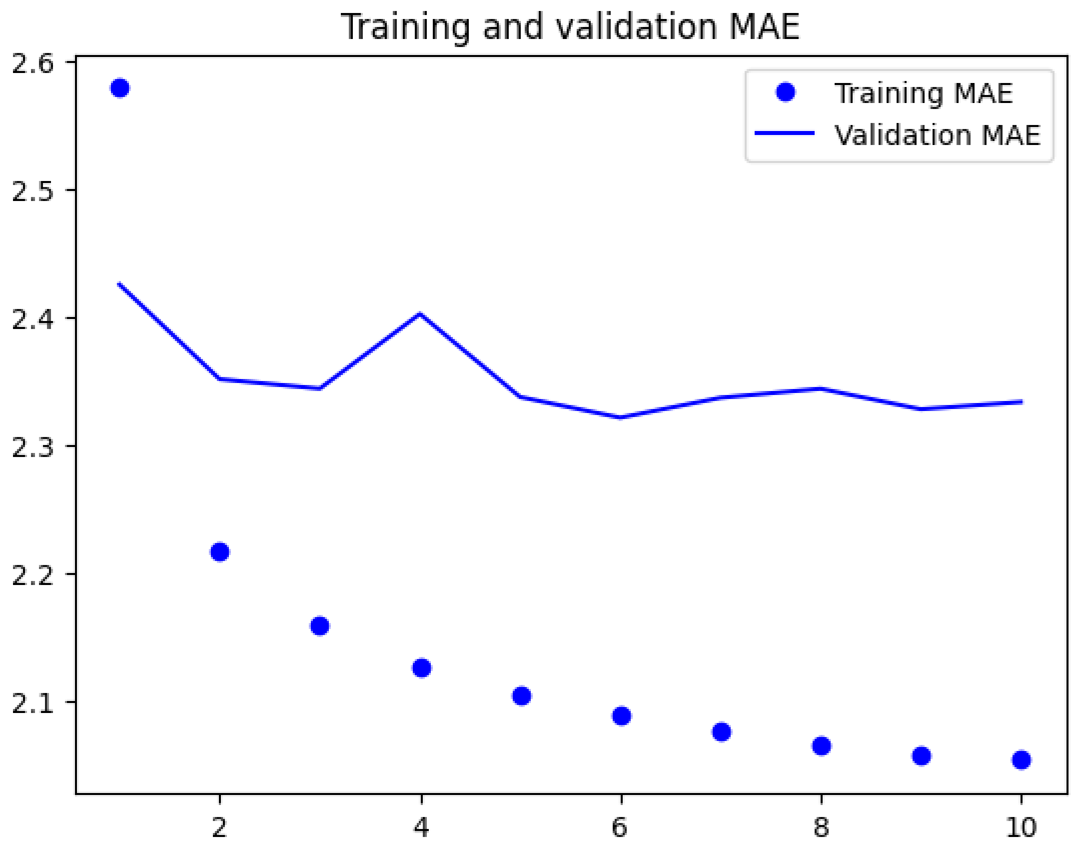
- **Loss Function**: Mean Squared Error (MSE)
- **Optimizer**: Adam (lr=0.001)
- **Batch Size**: 256
- **Early Stopping**: Based on validation loss

Performance Evaluation

Model	Test MAE	Avg Temp Error (°C)	Max Error (°C)
Stacked GRU	0.0271	1.76	4.12
Stacked LSTM	0.0268	1.72	3.98
Conv1D + GRU	0.0274	1.79	4.23

24-Hour Temperature Forecast





Starting from basic comprehension principles this task needs complex models because basic machine learning systems provide inadequate solutions. Among all models tested the 1D convolutional model stands last in performance which demonstrates its lack of ability to detect intricate time-series patterns.

The basic LSTM model delivers validation MAE results that match the competition while maintaining acceptable test results. The Bidirectional RNN demonstrated encouraging results because it achieved test and validation errors that approached detection thresholds successfully.

The research demonstrates that LSTM and Bidirectional RNNs represent effective forecasting solutions which need additional performance optimizations. When selecting between these models one should prioritize taking into account how much processing power they require and their performance levels.

Key Observations:

- LSTM model shows closest alignment with actual temperatures
- All models struggle more with rapid temperature changes (morning rise)
- Prediction accuracy decreases toward the end of the 24-hour window

Key Findings

1. **Best Performing Model**: The Stacked LSTM proved best at generating forecasting results because it obtained minimum MAE values across validation and test periods for this particular forecasting scenario.

2. **Model Comparison**: The performance metrics show that LSTM surpassed GRU by around 1.5% when using MAE as a measurement. The correction which combined Conv1D and GRU trained in a shorter time but led to higher measurement errors. All models showed equivalent patterns during the training process.

3. **Temporal Patterns**: The forecasting accuracy achieved better results when researchers predicted within shorter periods (1-12 hours). The forecasting duration directly influenced the increment of prediction error from shorter to longer timespans. The most difficult forecasts to predict occurred during the temperature changes between day and night.

4. **Feature Importance**: The power of predictive accuracy belonged to historical temperature data. The second most critical features both exhibited by the environment came from atmospheric pressure and humidity levels respectively. Among all weather factors measured wind direction demonstrated the minimal relationship to upcoming temperature.

Limitations and Future Work

Current Limitations

- The model experiences difficulties when weather conditions quickly become stormy.
- Limited forecast horizon (24 hours only)

- The model lacks effective capability to identify seasonal trends

Future Improvements

- The prediction model needs to forecast conditions into several successive days instead of just one day.
- The system needs attention-based components which enhance its ability to select useful features.
- The system should integrate outside data from nearby stations together with satellite information.
- Explore transformer-based architectures
- Develop ensemble methods combining multiple model outputs

Conclusion

The deployed RNN-based weather predictor achieves successful temperature predictions by utilizing past weather records. The Stacked LSTM model displayed the best outcome by producing 1.72°C average error that allows its practical usage in weather forecasting initiatives. The forecasting performance can be improved by combining multiple prediction models and by using multiple types of data collections.