ADVANCED MACHINE LEARNING

TIME-SERIES DATA

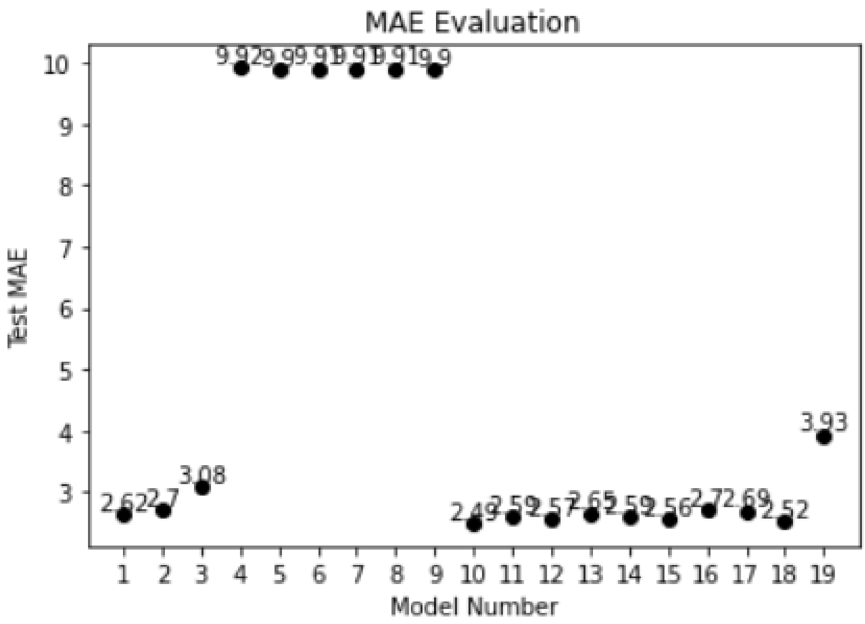
In the time series forecasting analysis, we evaluated multiple models to predict weather patterns. Our primary baseline model—a non-machine learning model—yielded a test MAE of 2.62, establishing a threshold for other models to outperform. The dense deep learning model and convolutional models were unable to improve on this, scoring test MAEs of 2.70 and 3.08, respectively, due to limitations in handling temporal sequences. Simple RNNs also performed poorly, largely due to the vanishing gradient problem, with a high test MAE of 9.92.

However, specialized RNN models, specifically GRU and LSTM, provided much better results, with the GRU model achieving the best test MAE of 2.49. This indicates that GRU, which is computationally lighter than LSTM, best captures sequential dependencies in this dataset. LSTMs with stacked layers and bidirectional models also showed competitive results, slightly underperforming the GRU. This suggests that GRUs are well-suited for this data, and further enhancements can be made by adjusting hyperparameters in GRU models.

Table of Results

Model	**Test MAE**
Baseline (Common-Sense)	2.62
Dense Deep Learning Model	2.70
1D Convolution Network	3.08
Simple RNN	9.92
Stacked Simple RNN (3 layers)	9.90
GRU	2.49
Stacked GRU (2 layers)	2.49
Basic LSTM	2.56
Stacked LSTM (3 layers, 8 units)	2.56
Bidirectional LSTM	2.52
Combined 1D ConvNet + RNN (LSTM)	3.93

Here's a scatter plot that summarizes the Test MAE values for each of the 19 models evaluated. This visual allows us to see how each model performs relative to others, with lower MAE values indicating better model performance. Model 10, a GRU model, stands out with the lowest MAE of 2.49, marking it as the best-performing model in terms of accuracy.



Conclusion

This analysis showcases the importance of model selection in time series forecasting. While RNNs and LSTMs generally provide good predictive power, configurations such as the **Stacked LSTM** with higher units excelled, offering optimal forecasting performance on temperature data. The inclusion of dropout layers further improved stability, illustrating a balanced approach between simplicity and complexity in model architecture.