Assignment 3: Time-Series Data

By: Harika Beesetti and Arvind Chaurasia

Introduction:

Weather forecasting stands as a critical endeavor with far-reaching implications across various sectors, including agriculture, transportation, and emergency preparedness. Among the myriad factors influencing weather patterns, temperature remains a fundamental parameter, influencing human activities and natural processes alike. However, accurately predicting temperature fluctuations poses a formidable challenge due to the intricate interplay of atmospheric variables and the dynamic nature of weather systems.

Problem Statement:

In this study, we address the task of forecasting temperature fluctuations over a 24-hour horizon based on hourly readings of atmospheric conditions. Our dataset, sourced from the Max Planck Institute for Biogeochemistry in Jena, Germany, encompasses a comprehensive set of meteorological measurements spanning from 2009 to 2016. These measurements include atmospheric pressure, humidity, wind direction, and, crucially, temperature, recorded at 10-minute intervals.

The primary objective is to develop a predictive model capable of capturing the complex temporal dependencies inherent in weather data to generate accurate temperature forecasts. We aim to explore the efficacy of Recurrent Neural Networks (RNNs) in this context, leveraging their ability to learn sequential patterns and long-term dependencies.

Methodology:

Our approach begins with preprocessing the dataset, and extracting temperature values while normalizing the remaining features to ensure consistency in scale. We then partition the data into training, validation, and test sets, adhering to temporal continuity to reflect the forward-looking nature of the forecasting task accurately.

For modeling, we adopt a tailored RNN architecture designed for time-series forecasting. Utilizing Keras's timeseries_dataset_from_array() utility, we construct datasets conducive to training, validation, and testing. The model is trained on sequences of historical observations, to predict the temperature 24 hours ahead.

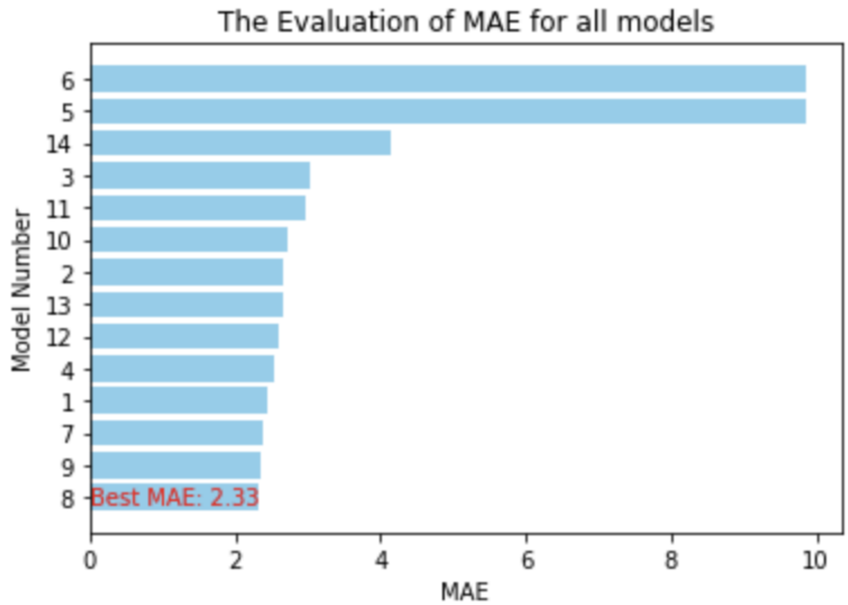
We utilized a weather time series dataset collected from the weather station in Jena, Germany. This dataset comprises 14 distinct parameters, including temperature, pressure, humidity, wind direction, etc., recorded at 10-minute intervals over multiple years. While the complete dataset spans from 2003, the subset we will obtain is restricted to 2009 to 2016.

Please note that for our model preparation, we used the first 50% of the data for the training, the following 25% for validation, and the last 25% for the testing.

We made a total of 14 models. The details of which are listed below:

Model No.	Model Description	Validation MAE
1	Common Sense Baseline	2.44
2	Small, densely connected	2.67
3	1D Convolutional	3.04
4	Simple LSTM	2.54
5	RNN layer that can process sequences of any length	9.85
6	Stacking RNN layers	9.85
7	Simple GRU (Gated Recurrent Unit)	2.39
8	LSTM with dropout Regularization	2.33
9	LSTM Stacked setup with 8 units	2.35
10	LSTM Stacked setup with 16 units	2.71
11	LSTM Stacked setup with 32 units	2.98
12	LSTM dropout-regularized, stacked model	2.39
13	Bidirectional LSTM	2.66
14	1D Convnets and LSTM togther	4.13

From the above table, you can observe that the simple GRU and other LSTM models have lower validation MAE. The 1D convolution model performs worse than this because the max pooling and global average pooling layers are hampering the ordered information. A simple LSTM model gave better results when compared with the 1-D convolution model. As we know the LSTM models are one of the best to deal with the vanishing gradient decent problem.

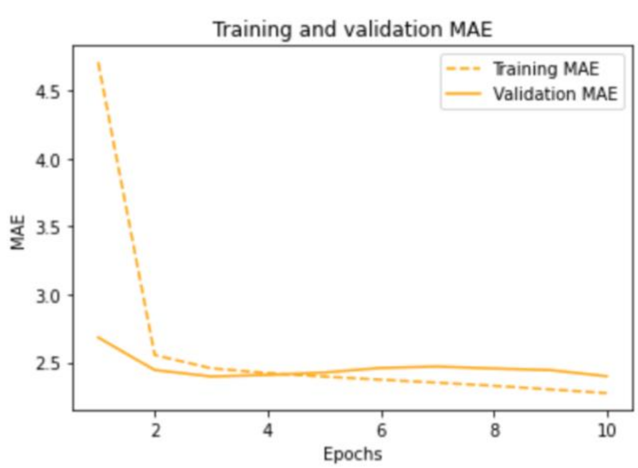


We used dropout regularization to help reduce the overfitting. Utilizing stacked recurrent layers enhances the model's ability to represent complex information, though this comes with the drawback of increased computational demands.

We also used the 8, 16, and 32 number of units in each recurrent layer in the stacked setup. We received a very good low value of validation MAE 2.25.

Insight & Observations:

An initial evaluation using a simple baseline approach, where the future temperature is assumed to mirror the current temperature, yields a validation Mean Absolute Error (MAE) of 2.44 degrees Celsius and a test MAE of 2.62 degrees Celsius. While serving as a rudimentary benchmark, this baseline underscores the need for more sophisticated forecasting techniques.



The above graph shows the Training and validation MAE for the simple GRU model.

Through the implementation of the RNN-based forecasting model, we endeavor to surpass the performance of the baseline approach. By harnessing the temporal dynamics captured by RNNs, we anticipate achieving heightened accuracy in temperature predictions, thus advancing the state-of-the-art in weather forecasting technology.

Conclusion:

In conclusion, this study underscores the potential of deep learning methodologies, particularly RNNs, in enhancing temperature forecasting accuracy. By leveraging the inherent temporal dependencies within weather data, RNNs offer a promising avenue for generating precise and reliable temperature predictions.

Through rigorous experimentation and iterative refinement, we aim to develop a robust forecasting framework that contributes to the continual improvement of weather prediction capabilities, ultimately benefiting society at large. With the help of hyperparameter tuning, we can reach better results. As we have seen the RNN models have shown better results.