Temperature Prediction using Deep Learning

Maitreya Sameer Ganu , IISER Thiruvananthapuram

10 th December 2024

Abstract

This project involves using a Long Short-Term Memory (LSTM) model for predicting daily maximum and minimum temperatures. The dataset used contains daily weather data from January 1 , 2012 to December 31 , 2015. The task is to predict future temperatures based on historical data, using a 7-day moving average to improve model performance.

1 Introduction:

Temperature prediction is an important problem in various fields, including meteorology and climate science. In this project, we aim to predict the maximum and minimum temperatures for the next day based on historical data using an LSTM model, a type of recurrent neural network (RNN) that is well-suited for sequential data.

2 Overview of the data set:

I have used the seattle weather data set and we shall only consider the maximum and minimum temperature columns. Now , we shall see the variation of maximum and minimum temperature with respect to time and then apply a 7 - day moving average to smoothen the data , enhance the accuracy and capture weekly trends .

	precipitation	temp_max	temp_min	wind
count	1461.000000	1461.000000	1461.000000	1461.000000
mean	3.029432	16.439083	8.234771	3.241136
std	6.680194	7.349758	5.023004	1.437825
min	0.000000	-1.600000	-7.100000	0.400000
25%	0.000000	10.600000	4.400000	2.200000
50%	0.000000	15.600000	8.300000	3.000000
75%	2.800000	22.200000	12.200000	4.000000
max	55.900000	35.600000	18.300000	9.500000

Figure 1: Data Description.

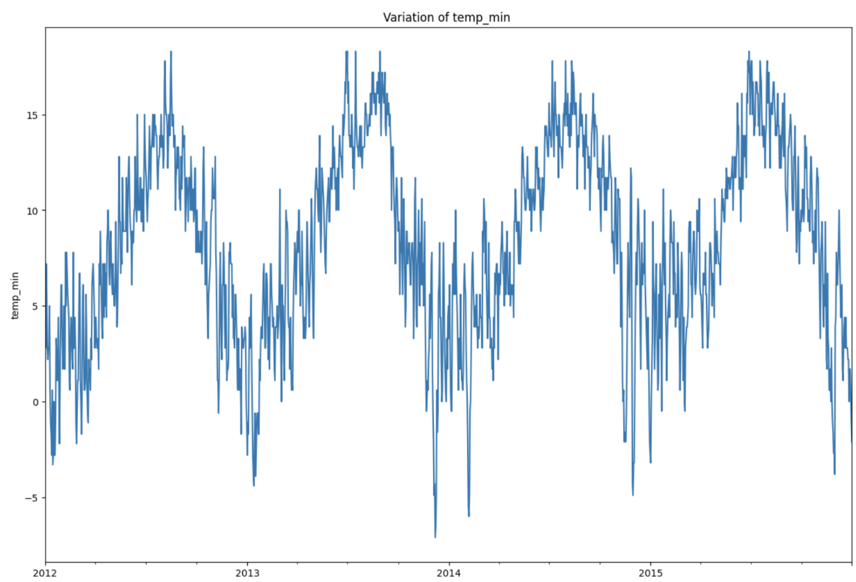


Figure 2: Variation of Minimum Temperature. The data illustrates the variation in the minimum temperature over the period of interest. This information is crucial for predicting temperature fluctuations, especially when combined with other weather variables. By observing the trends, we can make informed predictions for future temperature behavior.

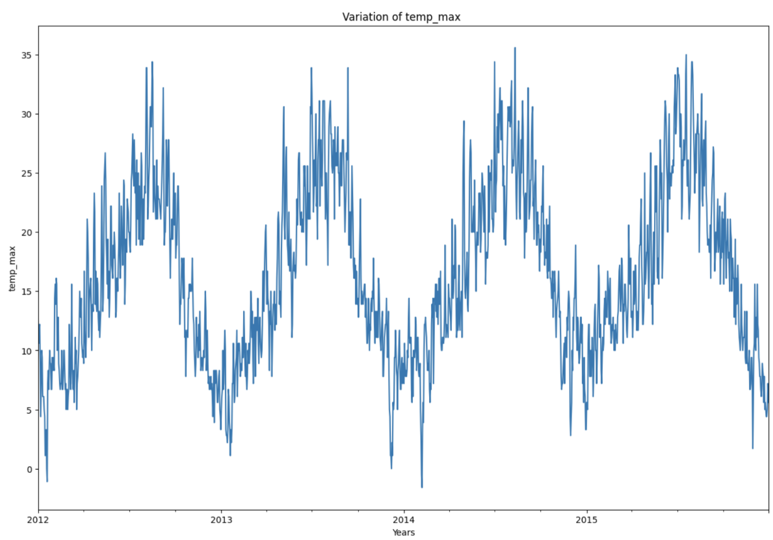


Figure 3: Variation of Maximum Temperature. This figure presents the variation of maximum temperature over the same period. The maximum temperature is a critical factor in weather forecasting models, as it influences the overall climate conditions, energy usage, and agricultural practices. The data shows the peaks and troughs in maximum temperature for the studied years.

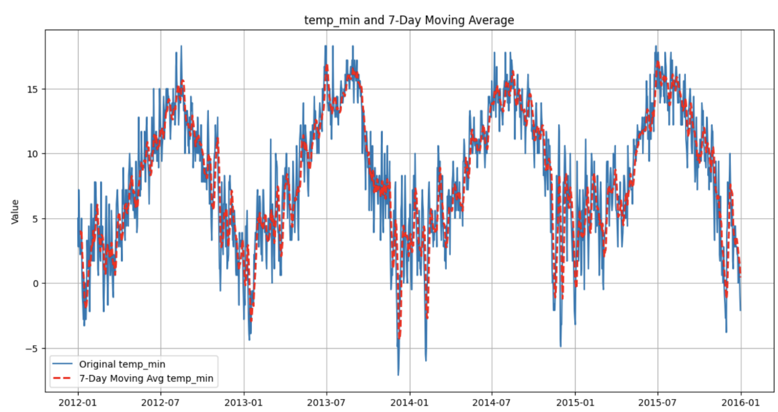


Figure 4: 7-Day Moving Average on Minimum Temperature Data. The 7-day moving average smooths out the daily fluctuations in minimum temperature. By averaging the temperatures over a week, this graph helps to observe underlying trends and patterns, filtering out short-term noise and allowing for more accurate temperature predictions.

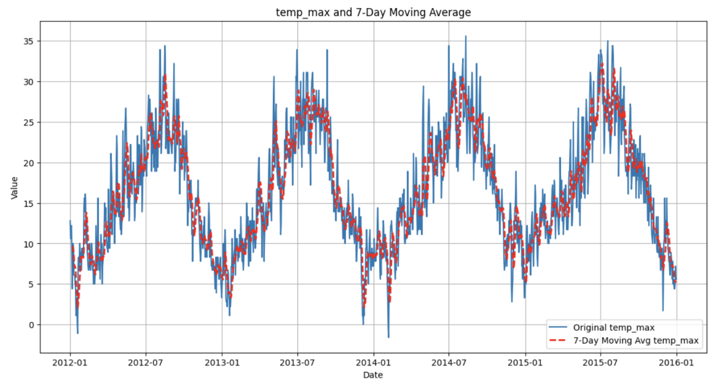


Figure 5: 7-Day Moving Average on Maximum Temperature Data. The 7-day moving average for maximum temperature data helps identify longer-term trends and variations. By smoothing the data, it becomes easier to observe the peak and valley cycles that typically occur on a weekly basis, assisting in more precise predictions in forecasting models.

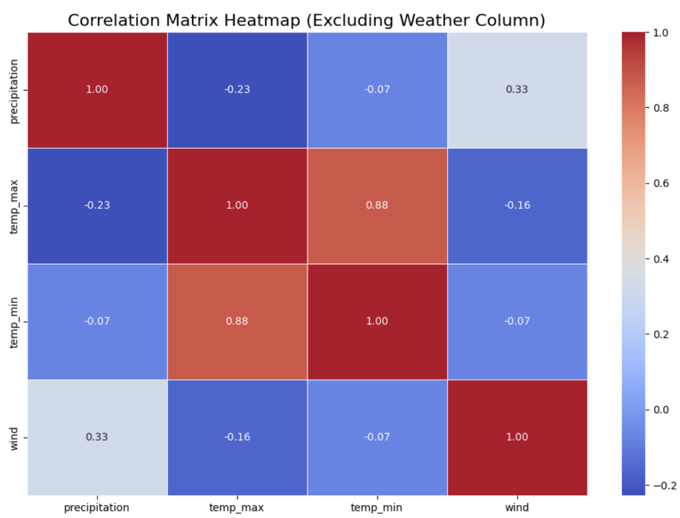


Figure 6: This heatmap illustrates the correlations between the various features in the dataset, excluding the 'weather' column. The values represent the Pearson correlation coefficients, ranging from -1 to 1. Positive values indicate a direct relationship, while negative values suggest an inverse relationship between the features. The heatmap uses the 'coolwarm' color map to visually highlight the strength of these correlations, with warmer colors indicating higher correlations and cooler colors indicating lower correlations.

3 Data Preprocessing:

The dataset used in this project contains daily temperature records from the years 2012 to 2015. The dataset includes the following columns:

- `date`: The date of the observation.
- `temp_max`: The maximum temperature recorded on that day.
- `temp_min`: The minimum temperature recorded on that day.

The data is collected from daily temperature measurements, spanning from January 1, 2012, to December 31, 2015. As this dataset contains time-series data, preprocessing is a critical step to prepare it for use with machine learning models, especially LSTM (Long Short-Term Memory) networks, which are used for sequential data. The preprocessing steps undertaken for preparing the data are as follows:

1. **Calculate the 7-day moving average for the `temp_max` and `temp_min` columns:**
   - A moving average is a statistical technique used to smooth out fluctuations in data and highlight long-term trends. For this project, a 7-day moving average is calculated for both `temp_max` and `temp_min` columns. The purpose of this is to reduce the noise in the temperature data caused by daily fluctuations and capture the overall trend of the temperatures over a week.
2. **Apply MinMaxScaler scaling to normalize the data:**
   - Machine learning models, including LSTMs, generally perform better when the input features are on a similar scale. This helps the model converge faster and makes it easier to train.
   - MinMax scaling normalizes the data by rescaling it to a fixed range, usually between 0 and 1. The formula for MinMax scaling is given by:

$$\text{X\_scaled} = \frac{X - \min(X)}{\max(X) - \min(X)}$$

4 Model Architecture:

Model: "sequential"

Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 100)	40,800
dense (Dense)	(None, 1)	101

Total params: 40,901 (159.77 KB)
Trainable params: 40,901 (159.77 KB)
Non-trainable params: 0 (0.00 B)

Model: "sequential_1"

Layer (type)	Output Shape	Param #
lstm_1 (LSTM)	(None, 100)	40,800
dense_1 (Dense)	(None, 1)	101

Total params: 40,901 (159.77 KB)
Trainable params: 40,901 (159.77 KB)
Non-trainable params: 0 (0.00 B)

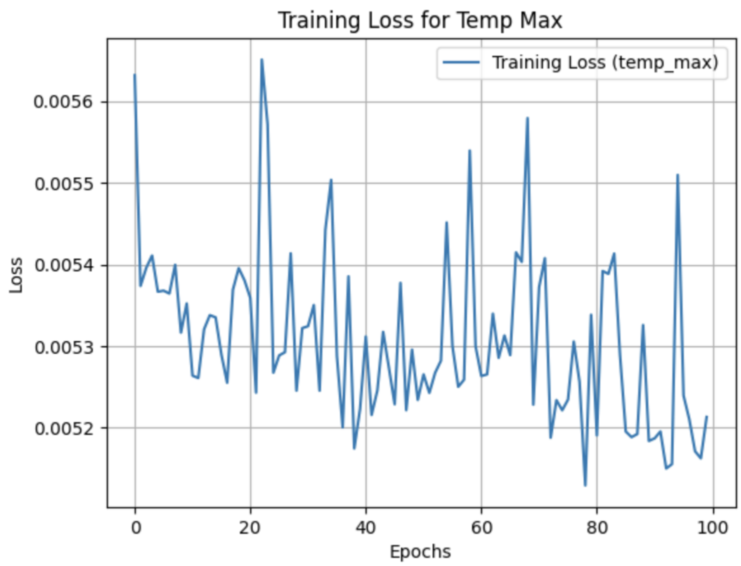
We use a Long Short-Term Memory (LSTM) network for this task, which is a type of Recurrent Neural Network (RNN) capable of learning long-term dependencies in sequential data. The architecture of the LSTM model is as follows:

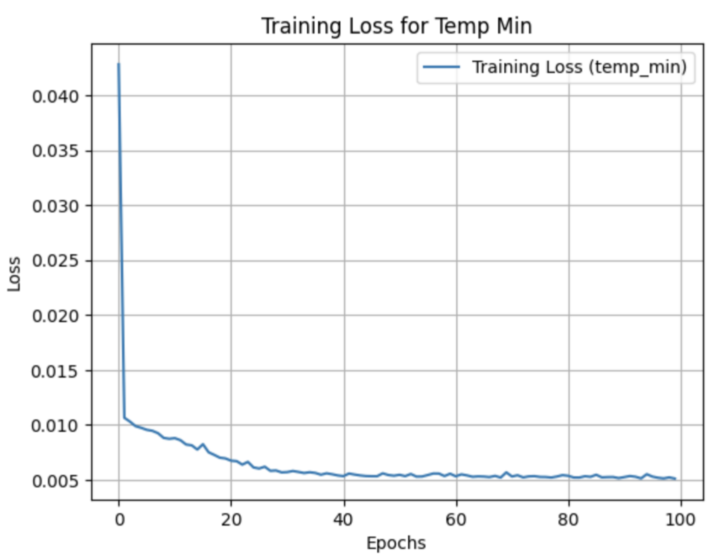
- The input to the model is a sequence of 7 days of historical temperature data.
- An LSTM layer with 100 units is used to learn the patterns in the temperature data.
- A dense layer is used to output the predicted temperature for the next day.

The model is compiled with the Adam optimizer and mean squared error (MSE) loss function.

5 Training and Evaluation:

The model is trained using the 7-day moving average data with 100 epochs and a batch size of 32. The training process includes monitoring the loss during the training. The model's performance is evaluated using Mean Squared Error (MSE) and Root Mean Squared Error (RMSE).





6 Results and Discussion:

In this section, we evaluate the performance of our trained models using two common metrics: Mean Squared Error (MSE) and Root Mean Squared Error (RMSE). These metrics help in understanding the accuracy of the model predictions compared to the actual temperature values.

The evaluation metrics for both the maximum temperature (`temp_max`) and minimum temperature (`temp_min`) are as follows:

- **For Temp Max:**
  - MSE = 3.285
  - RMSE = 1.813
- **For Temp Min:**
  - MSE = 3.421
  - RMSE = 1.850

The low values of MSE and RMSE indicate that the model is performing well in predicting both the maximum and minimum temperatures.

The results indicate that the model is able to predict the temperature values with a reasonable degree of accuracy, but further improvements could be made by using more features or tuning the model hyperparameters.

7 Conclusion:

In this project, we successfully used an LSTM model to predict daily maximum and minimum temperatures. By using a 7-day moving average and normalizing the data, the model was able to learn the temporal patterns in the temperature data. The resulting errors were low , thus indicating the high accuracy of the model.

8 Data and code availability:

- Access the data set used in this project :
  Seattle weather prediction dataset
- Access the code used in this project :
  Temperature prediction using Seattle weather data