**Temperature Prediction**

**Goal:**

* To predict daily temperature based on historical data.
* To compare a traditional algorithm that deals with the analysis of time series data with AI based Deep Learning models and evaluate the performance of both these algorithms based on their predictions on Weather data.

**Dataset:**

The Dataset consists of two parts, training set and test set. These are stored in separate excel files. The training set consists of 1462 rows with attributes of Date, mean Temperature, Humidity, Wind speed, and mean Pressure. The test set consists of 114 rows.

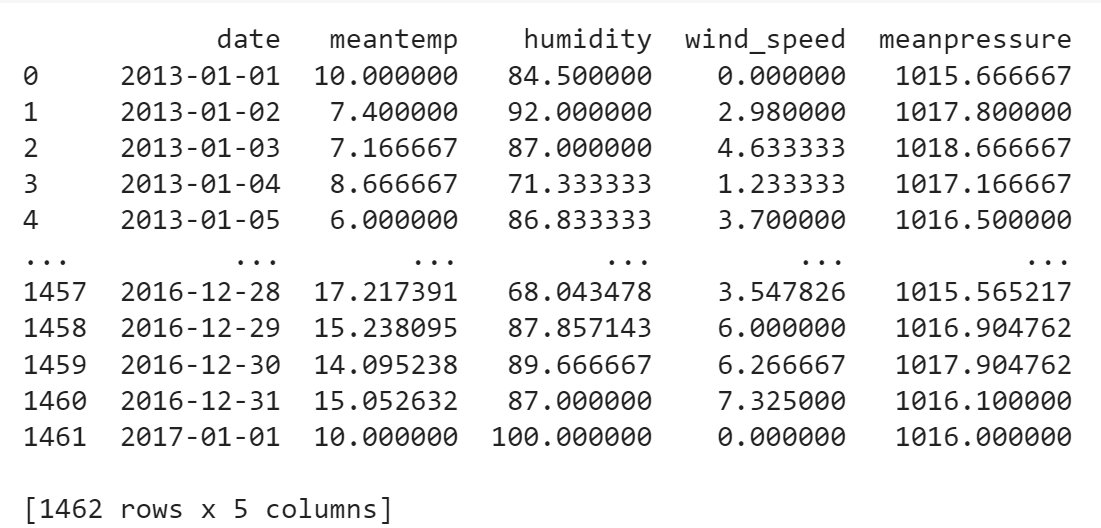


Figure 1: Training Dataset

The training data starts from 01.01.2013 till 01.01.2017. And the test data starts from 01.01.2017 till 24.04.2017. The figure 2 explains about the trends and seasonality in the training data.

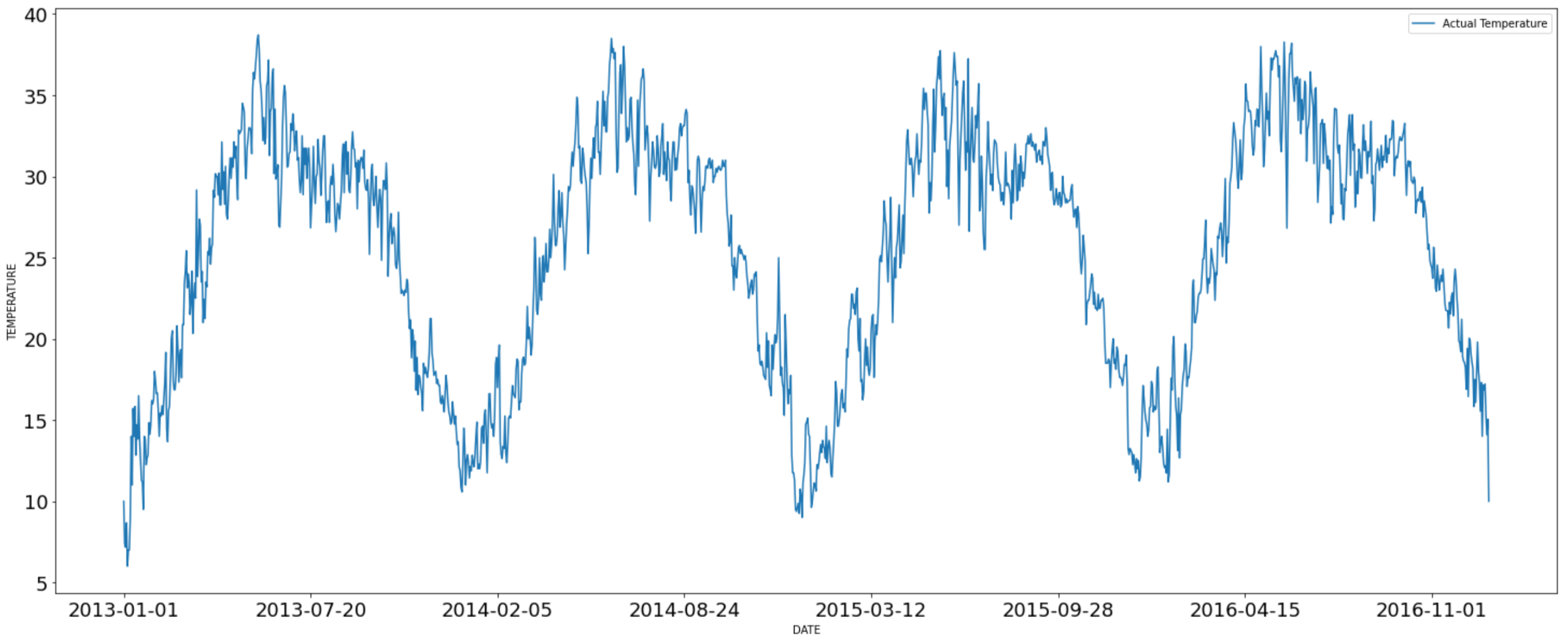


Figure 2: Visualization of training data

**Algorithms Description:**

1. **Trailing Moving Average:** This is a traditional algorithm that can be used in time series forecasting. Here we select a window of size n, which is used to calculate the temperature of the current day by calculating the mean of the temperatures of previous n days. The window passes through all instances of the dataset until it predicts the temperature of the last day in the dataset.
2. **Long short-term memory** (**LSTM**)**:** LSTM’s are special type of RNN (Recurrent Neural Networks) which tackle the problem of short-term memory. LSTM’s works well with the sequential data (e.g., Time-series data). They contain a cell state which is responsible for smooth flow of information for longer sequences (e.g., window). They take sequence of length n (window size) as input and predicts the output for that sequence.

Initially we train the network by passing the training data. So that the network will find the underlying features (e.g., trends, seasonality) in the data by learning the weights. Then we pass the test data and the network will try to predict the output based on the previously learned weights. Next, we compare the predicted values to the original values to determine the performance of the model on the unseen test set.

**Experiments and Results:**

By experimenting with different windows sizes. It is clear that the error increases with the increase in the window size.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Window length** | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| **MSE** | 3.87 | 4.58 | 5.15 | 5.63 | 6.04 | 6.39 | 6.72 | 7.08 | 7.44 |

Table 1: MSE for different window sizes

Training Moving average is a naïve algorithm. It doesn’t learn any patterns in the data. It just calculates the average of the previous days. For datasets which contains lot of patterns this algorithm struggles. For example, let’s say we have data which contains a pattern like, after every 10 days there is a huge drop of temperature in 11th day. This drop in temperature cannot be predicted by the trailing window approach. Because this algorithm doesn’t learn that there is drop after every 10 days. But this pattern will be learned by the AI algorithms.

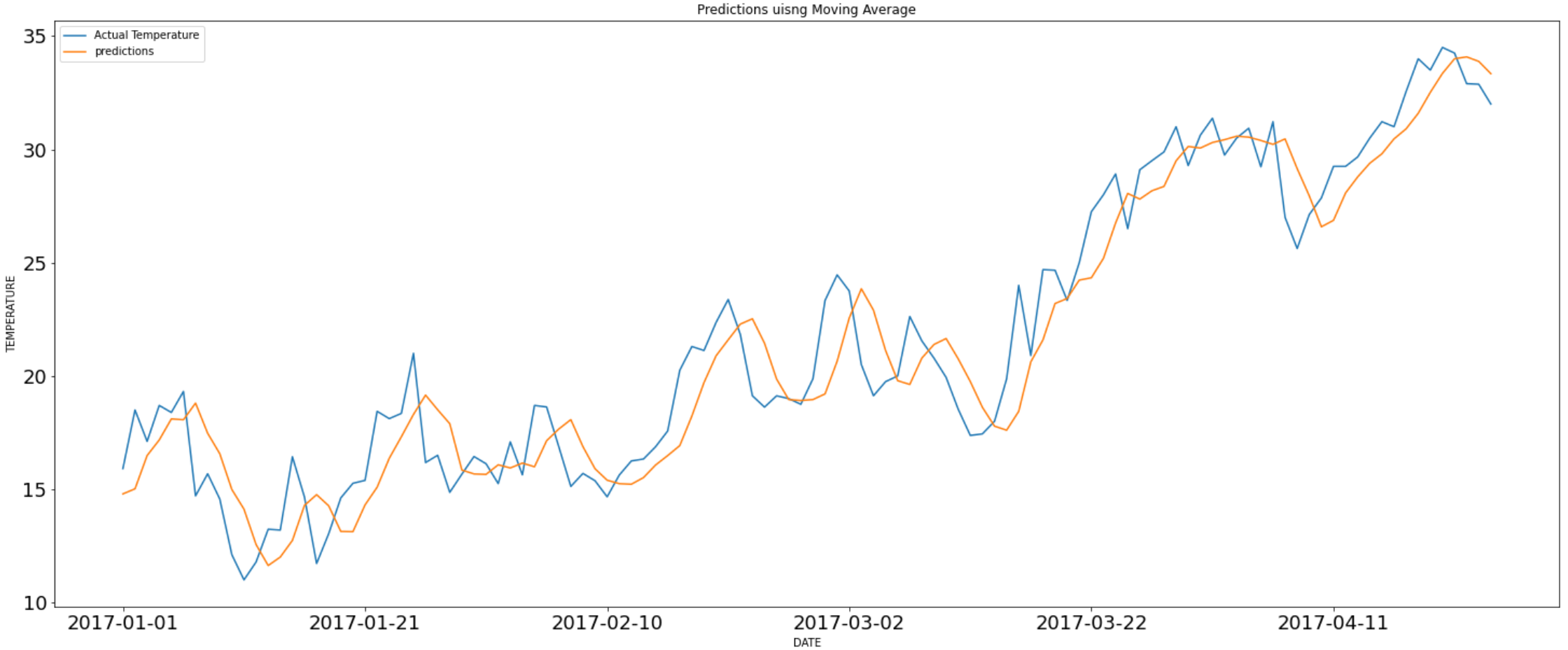
By experimenting with different LSTM architectural choices, it is found that choosing the right batch size is crucial. By changing the other hyperparameters there is only slight differences observed in the MSE values.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Batch Size** | 8 | 16 | 32 | 64 |
| **MSE** | 3.54 | 2.79 | 3.06 | 3.33 |

Table 2: MSE values for different batch sizes using LSTM

Though significant difference in performance is not observed between both the models. The AI model is more reliable because it learns the patterns and trends in the data pretty well.

**Evaluation:**



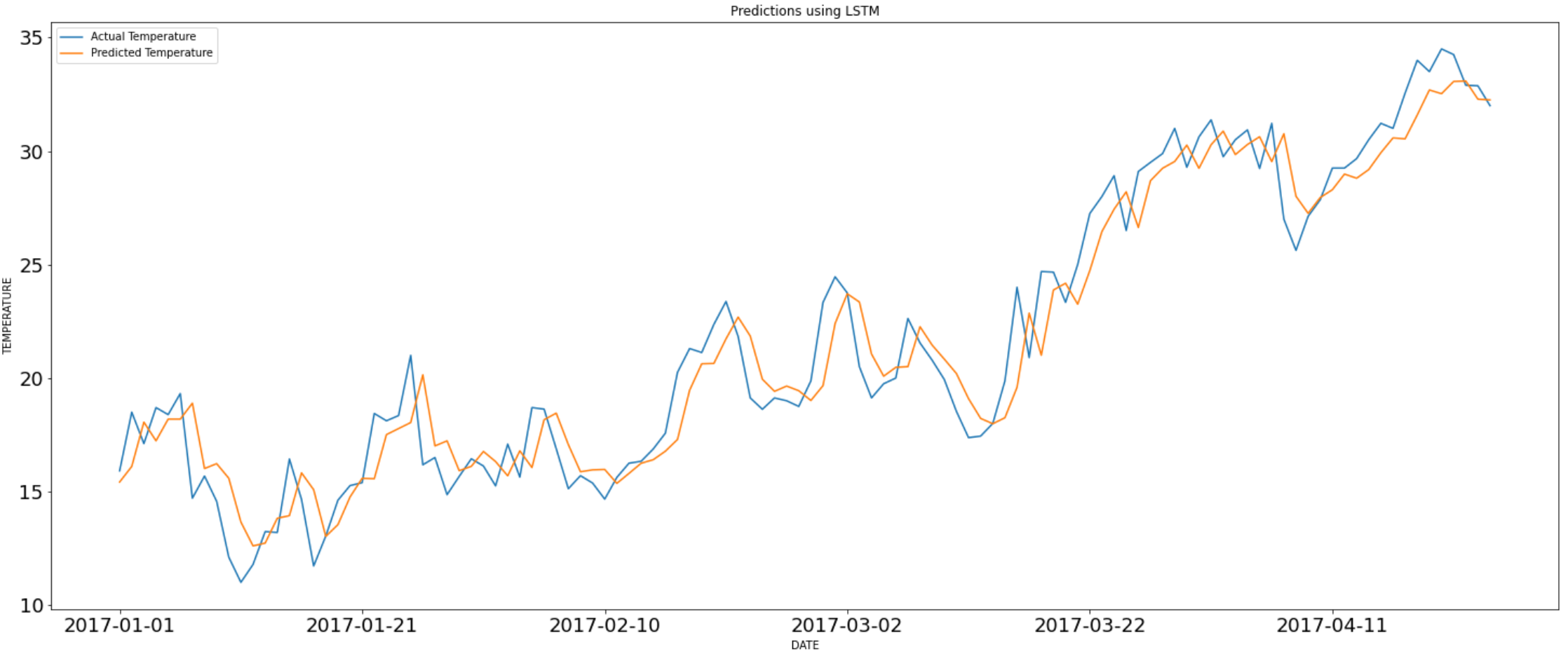


Figure 3 is the predictions of Moving Average and figure 4 is the predictions of LSTM. By looking at the graph it can be observed that the predictions of LSTM are lot better compared to Moving Average.

There should be a single evaluation metric to quantitatively measure the performance of the models. Mean squared error is chosen to be the metric, it is well suited for this problem. The below table shows the MSE for the two algorithms.

|  |  |
| --- | --- |
| Algorithms | MSE |
| Mean Average | 3.87 |
| LSTM | 2.79 |

Figure 7: MSE of the two algorithms

From the table it can be observed that LSTM have lower MSE and it outperforms Moving Average algorithm.

**Summary:**

In this experiment there is a comparison made between traditional Moving Average model and AI based LSTM model on weather dataset. There are several drawbacks mentioned about Moving Average but there are some drawbacks for the LSTM as well. Firstly, the training time of LSTM’s are very huge and they consume lot of memory. Despite its drawbacks it performs very efficiently on sequential data. Moreover, for regression problem it is highly unlikely to predict the exact value. Getting as close to the original value is good enough. Based on the analysis AI based LSTM algorithm have lower MSE and better predictions compared to the traditional Moving Average algorithm. Various architectures like GRU’s can also be experimented on this problem.