**Weather Forecasting Using Long Short-Term Memory Model (LSTM) and Autoregressive Integrated**

**Moving Average (ARIMA) Model**

**Satkar jain**

ABSTRACT :

The aim of this paper is to present a deep neural network architecture and use it in time series weather prediction. It uses multi stacked LSTMs to map sequences of weather values of the same length. The final goal is to produce a models which can forecast the weather data (Visibility) per day of Kota. Approximately 8 years (2011-2018) of daily meteorological data was used to train the model.The results show that LSTM based neural networks are competitive with the traditional methods and can be considered a better alternative to forecast general weather conditions.

1) INTRODUCTION:

Weather is the state of the atmosphere. Most weather occurs in the troposphere, or the lowest layer of the atmosphere. Weather is made up of multiple parameters, including air Temperature, atmospheric pressure, humidity, precipitation, solar radiation and wind. This factors can be measured to define typical weather patterns and to determine the quality of local atmospheric

Conditions. High Temperature can increase the heat transfer to local bodies of water in addition to heating the air. Lack of precipitation affects not only the weather conditions, but also the soil moisture and water levels due to evaporation. Wind speed and direction can be indicative of a front moving into the area, or it can create waves and encourage a stratified water column to mix. Hence, Weather Forecasting is important. In Machine Learning, Artificial Neural Networks (ANNs) are classes of models inspired by biological neural networks, which comprise interconnected adaptive processing nodes or units. What makes ANNs important is their adaptive nature, this feature makes ANNs a well suited tool to approximate highly nonlinear and multivariate functions. Both Feed-Forward Neural Networks and Recurrent Neural Networks can be used for Accurate Prediction. In the actual complex system, there are multiple variables evolving together and influencing each other, therefore multivariate prediction is more important.

2) Study Area:

Kota is a city located in the southeast of northern Indian state of Rajasthan.

The cartographic coordinates are 25.18°N 75.83°E.

3) Models:

3.1 ARIMA

The acronym ARIMA stands for Auto-Regressive Integrated Moving Average. The stationarized series in the forecasting equation are called "autoregressive" terms. The forecast errors are called "moving average" terms.

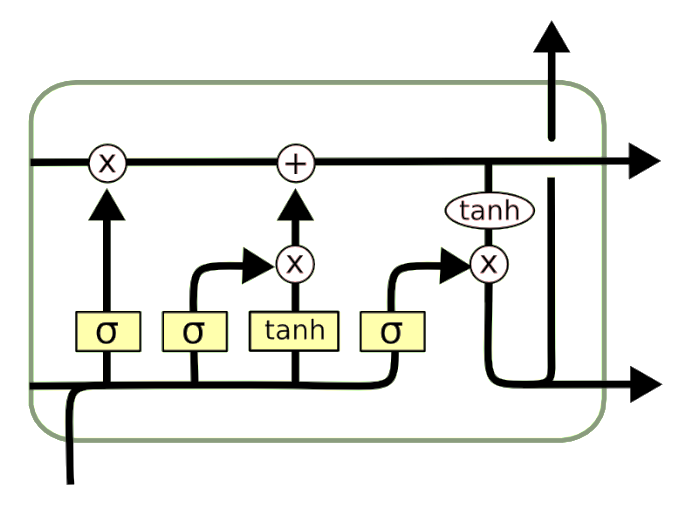
An Autoregressive Integrated Moving Average (ARIMA) is statistical properties and the well-known Box–Jenkins methodology (Box and Jenkins, 1976).

A non-seasonal ARIMA model is classified as an "ARIMA(*p*,*d*,*q*)" model, where:

* *p* is the number of autoregressive terms
* *d* is the number of non-seasonal differences needed for stationary
* *q* is the number of lagged forecast errors in the prediction equation

3.2 LSTM

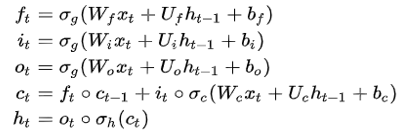
LSTMs are a type of Recurrent Neural Networks capable of learning long-term dependencies. They were introduced by Hochreiter and Schmidhuber . LSTMs remember information for long periods of time thanks to their inner cells which can carry information unchanged at will. The network have the complete control over the cell state, it can add,edit or remove information in the cell using special structures called gates.



A common LSTM unit is composed of a **cell**, an **input gate**, an **output gate** and a **forget gate**. The cell remembers values over arbitrary time intervals and the three *gates* regulate the flow of information into and out of the cell.

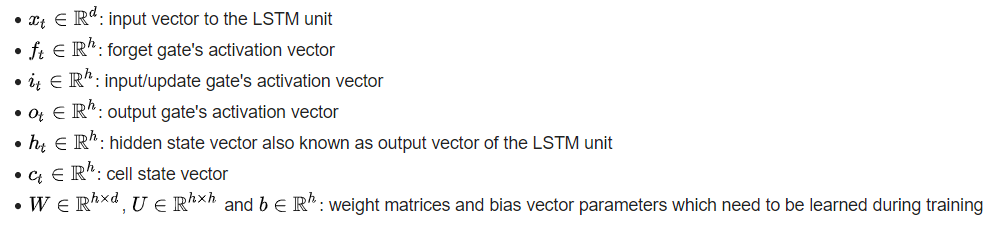
LSTMs were developed to deal with the exploding and [**vanishing gradient problems**](https://en.wikipedia.org/wiki/Vanishing_gradient_problem) that can be encountered when training traditional RNNs.

Equation involved:



where the initial values are {\displaystyle c\_{0}=0}and {\displaystyle h\_{0}=0}and the operator ({\displaystyle \circ }ο) denotes the Hadamard product (element-wise product). The subscript t {\displaystyle t}jjindexes the time step.

Variables:



4) EXPERIMENTS

Around 8 year data was trained. With the following variables:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Variable | Pressure | Humidity | Min Temp | Max Temp | Windspeed | Visibility |

4.1 Dataset

This dataset is obtained from google big data and timeanddate (https://www.timeanddate.com/)(via Web scraping).The range of data for this study was from year 2011 to year 2019 that consist of 7 series data.

4.2 Data Pre-processing

All of the values were normalized to points in [0,1] to avoid training problems (example: local optima problem) and also weight decay:

4.3 Validation

The model and trained on 8 years of data and tested on 1 year of data

The Mean Squared Error (MSE) was used as a measure of error made by the neural network model.

4.4 Models used

In this study two LSTM models and ARIMA model were explored:

4.4.1 SARIMA:

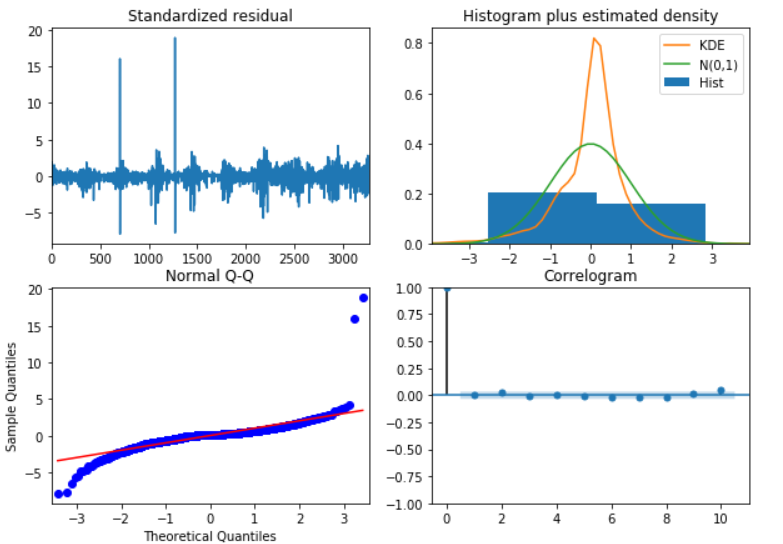
In this study SARIMA model is used because is used because Visibility have the seasonality.

The model is trained using the auto arima\_model of **pmdarima** keeping the seasonal=True ,by doing this we got the parameters of SARIMA for which **AIC**(Akaike's Information Criterion) score is minimum.

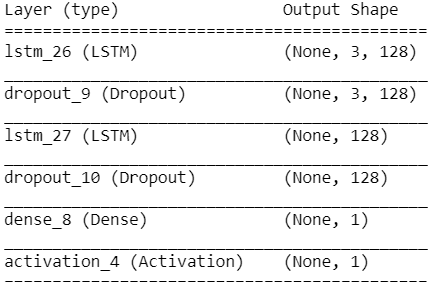
Model:

**SARIMAX(2, 0, 1)x(0, 1, 1, 12)**

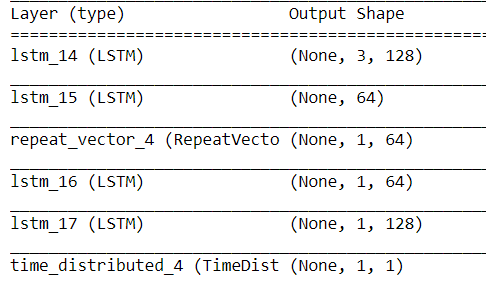
Residual plots in ARIMA model:



4.4.2 Stacked LSTM:

Stacking LSTM hidden layers makes the model deeper, more accurately earning the description as a deep learning technique.

4.4.3 Encoder- Decoder(LSTM):



**Repeat Vector** : this layer repeat the final output vector from

the encoding layer as a constant input to each timestep of the

decoder.

**Time Distributed Dense** : Applies a same Dense

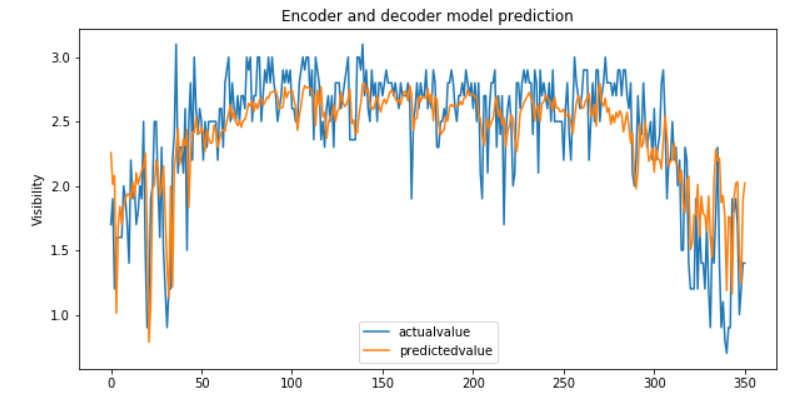
(fully-connected) operation to every timestep of a 3D tensor.

4.5 Result

The Mean Squared Error was used as a score function to evaluate

the quality of the predictions :

|  |  |  |  |
| --- | --- | --- | --- |
| Model | SARIMAX | Stacked LSTM | LSTM(Encoder-Decoder) |
| MSE | 0.2711 | 0.1158 | 0.1090 |



5 Conclusion

LSTM and ARIMA model have been used in time series forecasting. This study compares the LSTM Model and ARIMA Model in the weather forecasting. The visibility variable of the weather is used in this study

because it is the most important weather variable that has big impact on all phases of flight, especially when the aircraft is manoeuvering on or close to the ground LSTM model of this research set visibility as the predictor variable.

Encoder and decoder LSTM shown the best result as compared to other 2 model