Project Report

Weather Forecasting Using Linear Regression (LSTM Model)

Submitted by

Abhi Kathiria (19BCP162D)

Priyanshi Pansara (19BCP171D)

Ravi Patel (19BCP172D)

Tanvi Modi (19BCP173D)

Vainavi Gangvekar (19BCP174D)

SCHOOL OF TECHNOLOGY
PANDIT DEENDAYAL ENERGY UNIVERSITY
GANDHINAGAR, GUJARAT, INDIA

May 2022

Abstract:

Numerical weather prediction (NWP) uses mathematical models of the atmosphere and oceans to predict the weather based on current weather conditions. NWP requires considerable computer power to solve complex mathematical equations to obtain a forecast based on current weather conditions. Here, we propose a lightweight data-driven weather forecasting model by exploring state-of-the-art deep learning techniques based on Artificial Neural Network (ANN). Weather information is captured by time-series data and thus, we explore the latest Long Short-Term Memory (LSTM) layered model, which is a specialized form of Recurrent Neural Network (RNN) for weather prediction. The aim of this project is to develop and evaluate a short-term weather forecasting model using the LSTM. Our experiment shows that the proposed lightweight model produces better results compared to the well-known and complex models, demonstrating its potential for efficient and accurate short-term weather forecasting.

Table of Contents:

List of Figures:

List of Abbreviation:

CHAPTER 1 - INTRODUCTION:

Weather forecasting refers to the scientific process of predicting the state of the atmosphere based on specific time frames and locations [1]. Numerical Weather Prediction (NWP) utilizes computer algorithms to provide a forecast based on current weather conditions by solving large systems of nonlinear mathematical equations, which are based on specific mathematical models [2]. Meteorology adopted a more quantitative approach with the advance of technology and computer science, and forecast models became more accessible to researchers, forecasters, and other stakeholders. Many NWP systems were developed in recent years, such as the Weather Research and Forecasting (WRF) model [3]. As a consequence, the WRF model became the world's most-used atmospheric NWP model due to its higher resolution rate, accuracy, open source nature, community support, and wide variety of usability within different domains[4], [5]. According to [1], data-driven computer modeling systems can be utilized to reduce the computational power of NWPs. In particular, Artificial Neural Networks (ANN) can be used for this purpose due to their adaptive nature and learning capabilities based on prior knowledge. This feature makes the ANN techniques very appealing in application domains for solving highly nonlinear phenomena. Recurrent Neural Networks (RNN) and deep learning have attracted considerable attention due to their superior performance[6]. More specifically, the proposed weather prediction model is defined on the Long Short-Term memory (LSTM), a specialized form of the RNN, for weather prediction.

CHAPTER 2 - LITERATURE REVIEW

- Deep learning based effective fine-grained weather forecasting model.
  https://link.springer.com/article/10.1007/s10044-020-00898-1

  This paper talks about weather prediction using the LSTM model which is a lightweight data driven weather forecasting model and compares its performance with the existing classical machine learning approaches. The proposed deep learning networks with LSTM and TCN layers are assessed in two different regressions, namely multi-input multi-output and multi-input single-output.
  We demonstrate that the proposed lightweight deep model can be utilized for weather forecasting up to 12 h for 10 surface weather parameters.These experiments show that we can apply the neural network approach for weather prediction. Based on the geographical appearance of location (such as the top of a mountain, land covered by several mountains, the slope of the land, etc.) the regional weather forecasting may not be accurate.

- Intelligent Methods for Weather Forecasting: A Review.
  https://ieeexplore.ieee.org/abstract/document/6136289

  This paper talks about some hybrid models for weather forecasting using different measures of assessments. Accuracy is expectable by constructing a consortium of statistical and artificial intelligent methods. For the last three decades, artificial intelligent based learning models like neural networks, genetic algorithms and neuro-fuzzy logic have shown much better results as compared to Box-Cox modeling approaches.

- Using Artificial Intelligence to improve Real-Time Decision-Making For High-Impact Weather
  https://journals.ametsoc.org/view/journals/bams/98/10/bams-d-16-0123.1.xml?tab_body=pdf

  Artificial intelligence (AI) and data science technologies, specifically machine learning and data mining, bridge the gap between numerical model prediction and real-time guidance by improving accuracy. In this work, we demonstrate that applying AI techniques along with a physical understanding of the environment can significantly improve the prediction skill for multiple types of high-impact weather.Application of modern AI techniques to high-impact weather forecasting is improving our ability to sift through the deluge of big data to extract insights and accurate, timely guidance for human weather forecasters and decision-makers. AI techniques build on traditional methods, such as MOS, by providing more flexible and powerful models capable of identifying complex relationships between a huge number of modeled and observed weather features or derived quantities.

- Weather Monitoring Using Artificial Intelligence
  https://ieeexplore.ieee.org/abstract/document/7556813

  The paper talks about the analysis and prediction is based on linear regression which predicts the next day's weather with good accuracy. An accuracy of more than 90% is obtained, based on the data set. Recent studies have reflected that machine learning techniques achieved better performance than traditional statistical methods. Machine learning, a branch of artificial intelligence, has been proved to be a robust method in predicting and analyzing a given data set. The module plays a vital role in agricultural, industrial and logistical fields where the weather forecast is an important criterion.

CHAPTER 3 - PROPOSED METHODOLOGY

3.1 Weather Research and Forecasting Model:

The Weather Research and Forecasting (WRF) model was developed by Norwegian physicist Vilhelm Bjerknes and applied various thermodynamic equations so that numerical weather-based predictions can be made mainly through different vertical levels [8]. The primary role of the WRF was to carry out analysis focusing on climate time scale via linking physics data between land, atmosphere and ocean. The WRF model became the world's most-used atmospheric model since its initial public release was in the year 2000 [5]. One of the primary challenges in the WRF is its requirement for massive computational power to solve the equations that describe the atmosphere. Furthermore, atmospheric processes are associated with highly chaotic dynamical systems, which causes a limited model accuracy. As a consequence, the model forecast capabilities are less reliable as the difference between current time and the forecast time increases [1], [12]. In addition, the WRF is a large and complex model with different versions and applications, which lead to the need for greater understanding of the model, its implementation and the different options associated with its execution [5]. As depicted in Figure 1, the proposed model is based on state-of-the-art deep learning techniques that use Artificial Neural Network (ANN) and modern LSTM layers technology.

3.2 Proposed Deep Model using Long Short-Term Memory (LSTM) Network:

The approach discussed in this project is based on temporal weather data to identify the patterns and produce weather predictions. As discussed above, we use the state-of-theart Long Short-Term Memory (LSTM), which is a specialized form of Recurrent Neural Network (RNN) and it is widely applied to handle temporal data. The key concepts of the LSTM include layers of nodes that allow the passing of data through a multistep process to enable the recognition of the right pattern [7], [14].

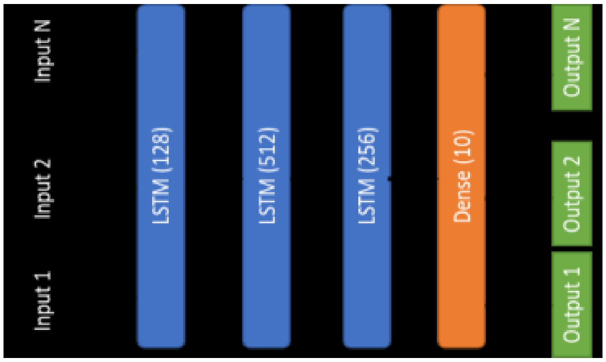

Fig. 3.2.1 Proposed Weather Prediction Model

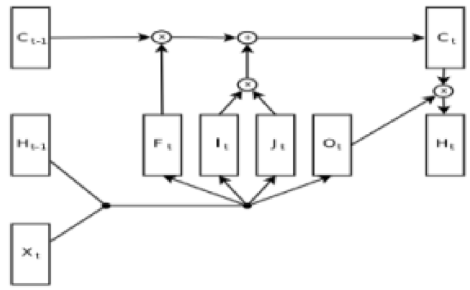

Fig. 3.2.2 LSTM memory cell [6]

Figure 3.2.1 shows the proposed lightweight model consisting of stacked LSTM layers for weather forecasting using surface weather parameters. The model provides outputs, which are the predicted weather parameters. It is well-known that the LSTM has the ability to learn long-term dependencies by incorporating memory units [6]. These memory units allow the network to learn, forget previously hidden states, and update hidden states. Figure 3.2.2 shows the LSTM memory architecture used in our model. Compared to the RNN, these additional memory cells give the ability to learn enormously complex and long-term temporal dynamics with the LSTM. We use Keras for this implementation of the proposed model using LSTM [14], [16], [17].

CHAPTER 4 – IMPLEMENTATION DETAILS

4.1 Surface Weather parameters:

For monitoring and forecasting purposes, the surface weather parameters are observed and reported [18]. The surface parameters of Temperature (C), Humidity, Pressure (millibars), Wind Speed (km/h) are calculated [4]. Further we have calculated Mean Squared Error (MSE), Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) for all 4 surface weather parameters.

4.2 Data Collection and preparation:

We have used the 'weatherHistory.csv' dataset from Kaggle having 12 columns of weather parameters for the period of 1st of January 2006 to 31st of December on hourly basis. Further we have divided the dataset into training and testing parts. This is used as the training dataset to train the proposed models. Similarly, a dataset has been created for the period of 1st of June 2016 to 30th of June to test the network. The July 2016 dataset is also extracted to evaluate the overall model as the ground truth. The training data set has been normalized to keep each value in between -1 and 1 and the same maximum and minimum variable values are used to normalize the testing data set and the evaluation data set.

4.3 Proposed model with optimal number of LSTM layers:

Several different configurations have been utilized to train and test the proposed models. Figure 3.2.1 depicts the general architecture of the proposed model. Each layer consists of a number of nodes. Optimizations are generally used in deep learning networks to minimize a given cost function by updating the model parameters such as weights and bias values. For better learning, it is often useful to reduce learning rates as the training progresses when training a deep network. This can be achieved by using adaptive learning rate methods [14]. Therefore, the fixed learning rate and adaptive learning rate methods have been explored to train the proposed deep models.

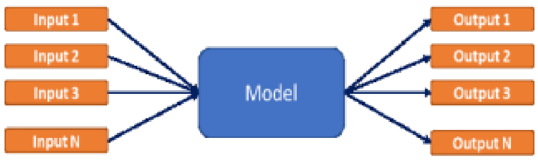

Fig. 4.3.1 MIMO layout

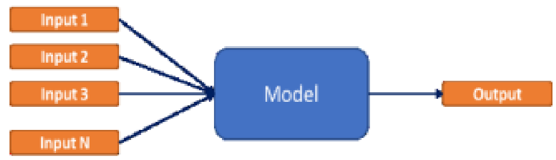

Fig. 4.3.2 MISO layout

The proposed model is evaluated using two different types, namely multi-input multi-output (MIMO) and multi-input single-output (MISO). In the MIMO, all 10 variables are fed into the network, which are expected to predict the same 10 variables as the output. In contrast, in the MISO approach, all 10 variables are fed into the network with a single variable output. In the MIMO, only one model is required for weather forecasting involving 10 different variables. Whereas, in the MISO, 10 different models are required as each of them is trained to predict a particular weather parameter with all 10 variables as input. Figure 2 depicts the basic arrangement of the MIMO and the MISO. Both MIMO and MISO classifications are utilized within these experiments. The most common metrics for evaluating regression models include Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and Explained Variance (EV) [19]. In this project, we use the MAE, RMSE and MSE to evaluate the models. As described in Section 4.3, there are different deep learning models, which are trained with different optimisers, different learning rates and regressions (MIMO and MISO). These models are evaluated using the January 2006 data to select the best model or a model with the lease MSE, which can be used as a tool for future forecasting.

CHAPTER 5 – RESULT ANALYSIS

As described in Section 4.4, deep learning models are trained with different configurations and controls. The results are subsequently evaluated via the Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and Mean Squared Error (MSE). This is used to assess the least error model and after comparing all evaluation reports. The least MSE for the MIMO is identified in the configuration with three LSTM layers, each layer has a number of nodes 128, 512, 256 respectively. We use the SGD optimizer with a fixed learning rate of 0.01 to optimize the MSE regression loss function. The model is trained for 230 epochs for optimum performance. The best values for each variable for the MISO are found in different configurations with different optimisers.

The LSTM model was trained using a Python package called 'keras' on top of Tensorflow backend. Hyper parameter searching is an important process prior to the learning process. Weather data starting from January 2016 until December 2016 is the dataset considered for the prediction process. This dataset is split into training dataset and testing dataset. Weather data is taken as the dataset for training the proposed Intensified LSTM based RNN model. The experimental variables used are Temperature, Humidity, Pressure, and Wind Speed. The network is fed all the experimental variables and its corresponding outcome variable during the training phase, in order for it to learn and remit the predicted outcome variable Weather as the output.

As a performance assessment for the reduced prediction error of the proposed methodology, Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and Mean Squared Error (MSE) is calculated. Computed MAE, RMSE and MSE is plotted against the number of epochs for the Intensified LSTM model and is shown in Figure 3.

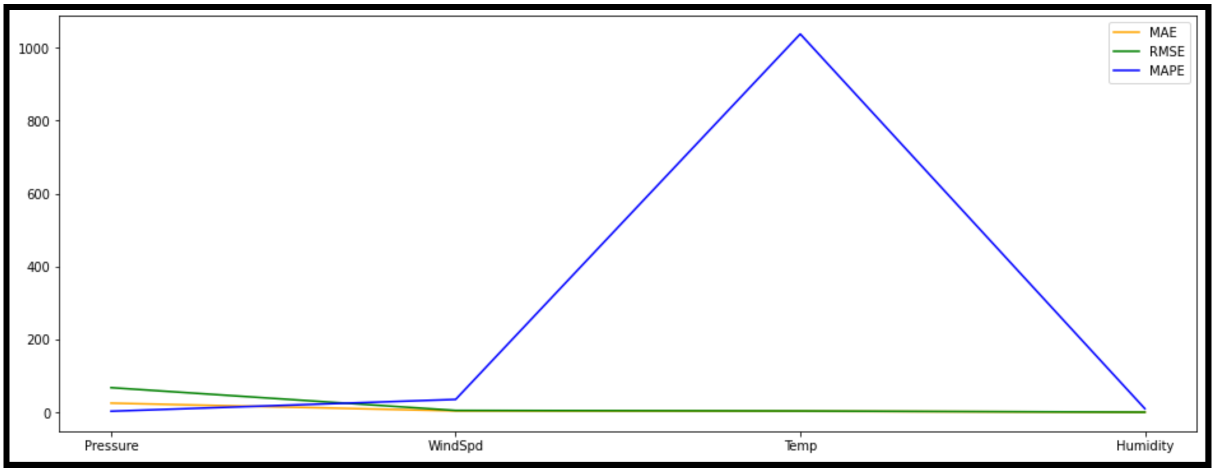

Fig. 5.1 Comparison graph of MAE, RMSE, MAPE

When the difference between actual value and predicted value is defined as error, accuracy can be stated as how close the predicted value is to the actual value. In such a way, accuracy is calculated at every epoch. It can be seen that accuracy increases as the number of epochs increases. Holding such a huge dataset is possible only because of the use of LSTM units in RNN networks.

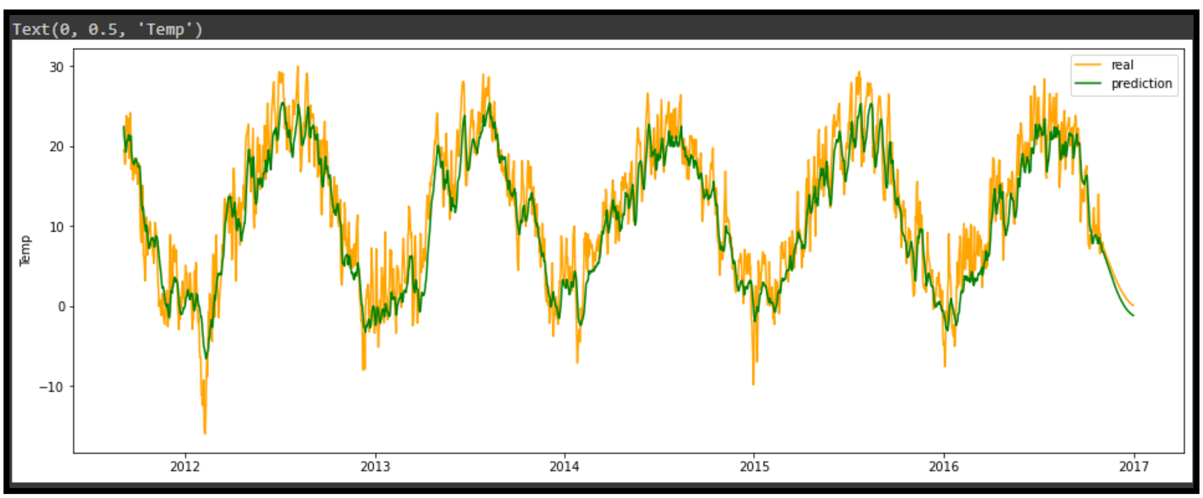

Fig. 5.2 Comparison graph of real and predicted forecast

This plot is compared with the plot of actual rainfall data of the same year to assess the performance of the model to track the actual graph. It can be seen that the predicted values almost match the actual values.

CHAPTER 6 - CONCLUSION & FUTURE WORK

In this project report, we demonstrate that the proposed lightweight deep model can be utilized for short-term weather forecasting. The model outperformed the state-of-the-art WRF model. The proposed model could run on a standalone computer unit and it could easily be deployed on a selected geographical region. Furthermore, the proposed model is able to overcome some challenges within the WRF model, such as the understanding of the model and its installation, as well as its running and portability. In particular, the deep model is portable and can be easily installed into a Python environment for effective results [9], [14]. This process is highly efficient compared to the WRF model. This project is performed using four different surface weather parameters and an increased number of inputs would probably lead to enhanced results. However, it will increase the model complexity requiring a large number of parameters to estimate. Furthermore, January to December weather data is utilized to train the deep model and the increase in the size of the training dataset could help towards an improved results in a deep learning network [14]. Besides, we used the MIMO approach within this research to predict weather data. Figure 4 shows that the MISO approach produces better MSE 8 values compared to the MIMO. Therefore, there is a huge potential that the MIMO approach will increase the accuracy of the results even if this method is less efficient compared to the MIMO.

References:

1. M. Hayati and Z. Mohebi, ‘Application of Artificial Neural Networks for Temperature Forecasting’, Int. J. Electr. Comput. Eng., vol. 1, no. 4, p. 5, 2007.
2. P. Lynch, The emergence of numerical weather prediction: Richardson’s dream. Cambridge: Cambridge Univ. Press, 2006.
3. L. Oana and A. Spataru, ‘Use of Genetic Algorithms in Numerical Weather Prediction’, in 2016 18th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC), Timisoara, Romania, 2016, pp. 456–461.
4. ‘WRF Model Users Site’, 2019. [Online]. Available: http://www2.mmm.ucar.edu/wrf/users/. [Accessed: 21-Jan-2019].
5. J. G. Powers et al., ‘The Weather Research and Forecasting Model: Overview, System Efforts, and Future Directions’, Bull. Am. Meteorol. Soc., vol. 98, pp. 1717–1737, Jan. 2017.
6. R. Jozefowicz, W. Zaremba, and I. Sutskever, ‘An Empirical Exploration of Recurrent Network Architectures’, in ICML, pp. 2342-2350, 2015.
7. G. E. Hinton, ‘Reducing the Dimensionality of Data with Neural Networks’, Science, vol. 313, no. 5786, pp. 504–507, Jul. 2006.
8. D. J. Stensrud, Parameterization schemes: keys to understanding numerical weather prediction models. 2007.
9. C. Skamarock et al., ‘A Description of the Advanced Research WRF Version 3’, 2008.
10. Noaa, ‘Reading GRIB Files’, 2017. [Online]. Available: http://www.cpc.ncep.noaa.gov/products/wesley/reading_grib.html. [Accessed: 23-Jan-2019].
11. N. C. for E. P. W. S. S. D. of Commerce, ‘NCEP GFS 0.25 Degree Global Forecast Grids Historical Archive’. 26-Jan-2015.
12. S. S. Baboo and I. K. Shereef, ‘An Efficient Weather Forecasting System using Artificial Neural Network’, Int. J. Environ. Sci. Dev., pp. 321–326, 2010.
13. A. Routray, U. C. Mohanty, K. K. Osuri, S. C. Kar, and D. Niyogi, ‘Impact of Satellite Radiance Data on Simulations of Bay of Bengal Tropical Cyclones Using the WRF-3DVAR Modeling System’, IEEE Trans. Geosci. Remote Sens., vol. 54, pp. 2285–2303, Apr. 2016.
14. A. Gulli and S. Pal, Deep Learning with Keras. Packt Publishing Ltd, 2017.
15. A. Behera, A. Keidel, and B. Debnath, ‘Context-driven Multi-stream LSTM (M-LSTM) for Recognizing Fine-Grained Activity of Drivers’, Pattern Recognit., p. 16, 2018.
16. Keras, ‘Home - Keras Documentation’, 2019. [Online]. Available: https://keras.io/. [Accessed: 28-Jan-2019].
17. A. Krizhevsky, I. Sutskever, and G. E. Hinton, ‘ImageNet Classification with Deep Convolutional Neural Networks’, in Advances in Neural Information Processing Systems 25, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 1097–1105.
18. ‘Federal Meteorological Handbook Number 1: Surface Weather Observations and Reports’, p. 98, 11 2017.