

Non-Linear Smog Prediction Model using Recurrent Long Short Term Memory

S. Geetha

Assistant Professor (Sr. Grade), Department of Computer Applications, Mepco Schlenk Engineering College, Sivakasi, Tamil Nadu, India

ABSTRACT: An enriched non-linear smog prediction model is developed using Artificial Neural Network and Recurrent Neural Networks' Long Short Term Memory. The air pollution and other meteorological data was collected from Central Pollution Control Board, Delhi. The collected data had lot of missing values. To overcome with missing values, imputation techniques was applied on the data before used for modelling. The collected data was split into two sets as training set and testing set. The non-linear prediction models are developed using ANN's Multiple Layer Perceptron and RNN's Long Short Term Memory networks. These models are evaluated using the Mean Absolute Error and Root Mean Square Error. The RNN-LSTM model provides better result for PM_{10} , $PM_{2.5}$, NO_2 , SO_2 when compared to ANN-MLP.

KEYWORDS: Air Pollution; Non-Linear Prediction; SO_2 ; CO; PM_{10} ; $PM_{2.5}$; NO_2 ; Smog; ANN; MLP; LSTM; Recurrent Neural Network

I. INTRODUCTION

Prediction become inevitable in all the domains. Especially, in the atmospheric environment, the pollutants concentration should be predicted well in advance to control it. Smog is one of the biggest problem in Delhi. Smog is formed when air is contaminated with fog and other atmospheric conditions.

The main air pollutants [1] are carbon monoxide (CO), carbon dioxide (CO_2), nitrogen monoxide (NO), nitrogen dioxide (NO_2) and nitrogen oxides (NO_x), Particulate Matters 10 (PM_{10}), and Particulate Matters 2.5 ($PM_{2.5}$), sulfur dioxide (SO_2), ozone (O_3) and other meteorological factors.

As per World Health Organization records, smog dangling over countries are causing approximately seven million mortalities each year from stroke, lung cancer, cardiovascular disease, and acute respiratory infections [2]. Smog causes visibility problem which is generating more percentage of road accidents. Smog has also affected the investment in industries [3]. Hence, the Prediction of smog becomes essential to enrich the quality of human life and to save living beings. All pollutants (NO_2 , CO, NO_x , SO_2 , O_3 , $PM_{2.5}$, PM_{10}) are keep on spreading and changing. There are many traditional prediction models such as Logistic Regression, Time series analysis, Decision Trees, and Machine Learning Algorithms are available for predicting air pollutions. However, as the air pollution data is so complex, it requires non-linear approach for prediction. Deep learning is the latest technology applied in various domains to bring automation and accuracy in work. A surfeit of studies has been done using Deep learning to create prediction model especially with time series data. This paper derived to bring Artificial Neural Network's Multiple Layer Perceptron along with deep learning technique RNN-LSTM to deal with air pollution dataset and to create prediction model for smog.

II. LITERATURE REVIEW

Artificial Neural Networks (ANN) models have been shown better for deterministic models for complex non-linear data [4]. significant progress has been made in air-quality dispersion models [5]. However, being highly non-linear, they require large amounts of accurate input data and are computationally expensive [6]. The main purpose of exploration of real data PM_{10} concentration temporal data, it seems crucial that non-linear models; for instance, Artificial Neural Networks (ANN) were developed for more accurate and precise PM_{10} concentration forecasting [7]. Furthermore, ANN quickly processes the information with no particular presumptions of the way of the non-linearity [8].

The special type of deep learning network under the recurrent neural network is LSTM which supports to generate model with time series data. As the air quality dataset consists of time series data, the LSTM can be used for prediction.

Although, lot of models were already developed using LSTM [9] and provides better results, the hyper parameter selection to be optimized manually in LSTM network. Prediction of solar irradiance is done based on weather forecasts and long short-term memory(LSTM) networks [10]. Thus, the ANN-MLP and RNN-LSTM were experimented and discussed in this paper.

III. MATERIALS AND METHODS

Artificial Neural Networks

Artificial neural networks are biologically inspired modeling tools for complex processes and phenomena. They can solve noisy problems and their computational power is given by their ability to estimate missing values. An ANN can model complex nonlinear phenomenon like air pollution [11], stock market prediction [12]. The artificial neural networks were first proposed for the adaptive control of nonlinear dynamical systems in [13]. They are well suited for modeling nonlinear problems such as the air quality forecasting[14], stock price prediction [15], etc. Moreover, the adaptive nature of the ANNs enhances their ability for time series forecasting. Out of many ANN types, here we have chosen Multiple Layer Perceptron (MLP). It was developed with an input layer, one or more hidden layers and an output layer. The generic architecture of MLP is shown in Fig. 1.

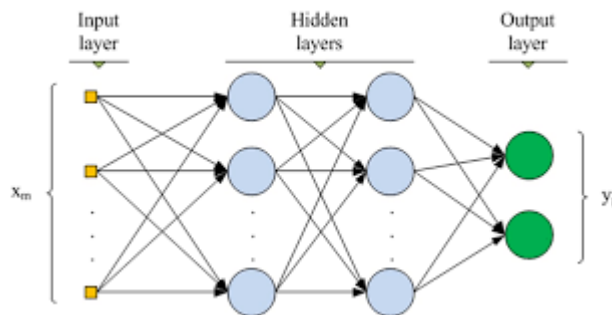


Fig. 1. Architecture of MLP

Long-Short Term Memory Networks (LSTM)

LSTM Networks is one kind of Recurrent Neural Network (RNN). To overcome the problem of failing to store previous data in memory cell, LSTM over RNN is emerged. In LSTM Network, memory cells are repeatedly connected, with three vital gate units: Input Gate, Output Gate, and Forget Gate. The input gate controls the maximum flow of values into the cell, the forget gate controls how long the values can be in the cell and the output gate controls the cell output by activation. The Fig. 2 shows the traditional architecture of LSTM.

LSTM is using activation function to determine whether the memory cell is containing the value or forgotten [12], with the help of the following formula.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{1}$$

Input gate determines which values used for updating and tanh is used to compute \tilde{C}_t . The formulas are

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{2}$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{3}$$

To update the old cell value, the following formula is used.

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{4}$$

Finally, to get the output of the model, the following formulas are used.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{5}$$

$$h_t = o_t * \tanh(C_t) \tag{6}$$

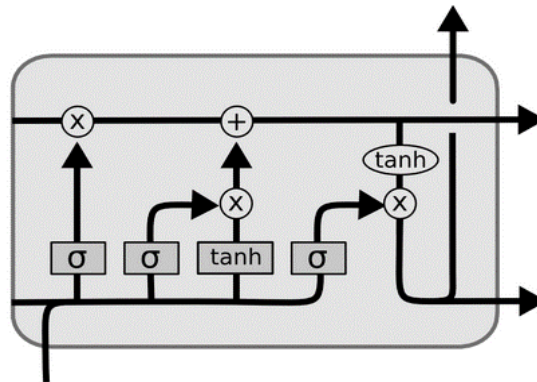


Fig. 2: LSTM Architecture

Here, x is the input, t is the time, h is the hidden layer output, c is the cell output state, f_t , i_t and o_t are the outputs of the forget gate, input gate and output gate. Sigmoid activation function is used in three gates.

As a result of the advantages of LSTM, we use LSTM as the basic part of our model to predict the smog and the performance through the abstract features generated with the help of LSTM layers. In this study, the computations are done using Python 3.6.

IV. EXPERIMENT RESULTS

Artificial Neural Networks – Multiple Layer Perception Model

Initially, many experiment were performed to fine the best topology and that was used for model creation. Different values of the nodes' number on the input and hidden layers, of the number from the hidden layers, for the learning rate and momentum, for the training epochs were tested. Fig.3 presents the best prediction that had 9 nodes on the input layer corresponding to the ones determined by the PCA and 1 unit on the output layer for the air quality index, and 2 hidden layers each one with 10 units. The validation method that was used is a 10-fold cross validation approach known as repeated random sub-sampling validation. The cross validation accuracy is given by the root mean squared error.

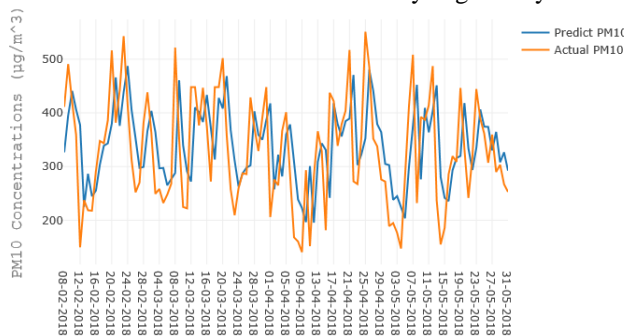


Fig.3.ANN-MLP Prediction for PM₁₀

Long-Short Term Memory Networks (LSTM) Model

The forecasting of air pollutants SO₂, CO, PM_{2.5}, PM₁₀ and NO₂ with actual and predicted comparison using LSTM. LSTM model created with fitted hyper parameters (epochs, neurons). No. of neurons lies between 50 - 200, No. of epochs lies between 50 – 1000. Batch size varies from 40 – 100. The model created using ‘non-linear’ activation function and ‘adam’ optimizer. The Fig. 4 shows only the NO₂ predictions as sample.

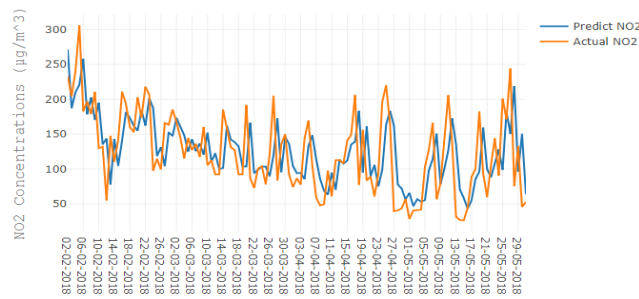


Fig.4. LSTM Prediction for NO2

The models are created using two approaches and evaluated with RMSE and MAE. The LSTM-RMSE and LSTM-MAE values for PM₁₀, PM_{2.5}, NO₂ and SO₂ are very lower than the RMSE and MAE value of ANN-MLP. But for CO the MAE and RMSE values are lower than the RNN-LSTM model.

Among two models, RNN-LSTM is predicting the values more accurate than other for many meteorological factors.

| | ANN – MLP | | LSTM | |
|-------------------|--------------|--------------|--------------|--------------|
| | MAE | RMSE | MAE | RMSE |
| PM ₁₀ | 42.64 | 85.10 | 39.01 | 49.04 |
| PM _{2.5} | 41.23 | 62.37 | 05.16 | 50.06 |
| CO | 00.66 | 00.88 | 00.90 | 96.14 |
| NO ₂ | 16.28 | 24.33 | 11.66 | 14.19 |
| SO ₂ | 01.79 | 03.64 | 00.91 | 01.01 |

Table 1: MAE & RMSE values for ANN-MLP and LSTM

All the developed models have been trained and validated with the daily average data of air pollutants. The RMSE and MAE computations between observed and predicted SO₂, CO, PM_{2.5}, PM₁₀ and NO₂ are given in Table 2. Overall, the performance of the RNN-LSTM model is found satisfactory for most of the air pollutant. Thus, the developed model can be used for predicting smog over urban areas.

V. CONCLUSION

Usually linear regression models were created for time series data. But it grows more complex, the result arrived with linear regression were not satisfactory. Then so many researches were carried out to prove non-linear regression for complex time series data which contains multiple variables. Out of so many non-linear regression methods, the recent two techniques were taken to experimentation. Finally, based on the evaluation result, the RNN-LSTM works better than ANN-MLP. The other powerful non-linear regression models to be developed in future.

REFERENCES

- [1]. Banerjee T, Singh SB, Srivastava RK., “Development and performance evaluation of statistical models correlating air pollutants and meteorological variables at Pantnagar India,” Atmosphere Research, 99:505–17, 2011.
- [2]. Sundeep Mishra, “Is smog innocuous? Air pollution and cardiovascular disease”, Indian Heart Journal, 2017.
- [3]. Jingbo Luo, “How does smog affect firms’ investment behavior? A natural experiment based on a sudden surge in the PM2.5 index”, China Journal of Accounting Research, 2017.
- [4]. M.A. Elangasinghe, N. Singhal, K.N. Dirks, J.A. Salmond, “Development of an ANN-based air pollution forecasting system with explicit knowledge through sensitivity analysis”, Atmospheric Pollution Research, 5, pp. 696-708, 2014.
- [5]. R. Arasa, M.R. Soler, S. Ortega, M. Olid, M. Merino, “A performance evaluation of MM5/MNEQA/CMAQ air quality modelling system to forecast ozone concentrations in Catalonia”, Tethys, 7, pp. 11-23, 2010.



- [6] A.L. Dutot, J. Rynkiewicz, F.E. Steiner, J. Rude, “A 24-h forecast of ozone peaks and exceedance levels using neural classifiers and weather predictions”, *Environmental Modelling & Software*, 22, pp. 1261-1269, 2007.
- [7] Paschalidou, A.K., Karakitsios, S., Kleanthous, S., Kassomenos, P.A., “Forecasting hourly PM10 concentration in Cyprus through artificial neural networks and multiple regression models: Implications to local environmental management”, *Environmental Science Pollution Research*, 18, 316–327, 2011.
- [8]. Al-Alawi, S., Abdul-Wahab, S., Bakheit, C., “Combining principal component regression and artificial neural networks for more accurate forecasting of ground-level ozone”, *Environmental Modelling Software*, 23, 396–403, 2008.
- [9] S. Geetha, L. Prasika, “Smog Prediction Model using Time Series with Long-Short Term Memory”, *International Journal of Mechanical Engineering and Technology (IJMET)*, Volume 10, Issue 01, pp. 1026–1032, 2019.
- [10] AfanGalih Salman et al., “Weather Forecasting Using Merged Long Short-Term Memory Model (LSTM) and Autoregressive Integrated Moving Average (ARIMA) Model”, *Journal of Computer Science*, 14 (7): 930-938, 2018.
- [11]. Chelani, A.B., Gajghate, D.G., Hasan, M.Z., “Prediction of ambient PM10 and toxic metals using artificial neural networks”, *Journal of Air and Waste Management Assessment*, 805–810, 2002.
- [12]. Moghaddam A. H., Moghaddam M. H., and Esfandyari M., “Stock market index prediction using artificial neural network”, *Journal of Economics, Finance and Administrative Science*, 21 (41): 89-93, 2016.
- [13]. Narendra, K.S., Parthasarathy, K., “Identification and control of dynamical systems using neural networks”, *IEEE Transactions on Neural Networks*, 4–27, 1990.
- [14]. Haiming, Z., Xiaoxiao, S., “Study on prediction of atmospheric PM2.5 based on RBF neural networks”, *International Proceedings of International Conference on Digital Manufacturing and Automation*, pp. 1287–1289, 2013.
- [15]. Guresen E., Kayakutlu G., and Daim T. U., “Using artificial neural network models in stock market index prediction”, *Expert Systems with Applications*, 38 (8): 10389-10397, 2011.