ISSN No:-2456-2165

Air Quality Index Forecasting Using RNN and LSTM

Rahul Singh
Department of Electronics and Communication
Galgotias College of Engineering and Technology
Gr. Noida, India

Satyam Sagar
Department of Electronics and Communication
Galgotias College of Engineering and Technology
Gr. Noida, India

Hitesh Kumar
Department of Electronics and Communication
Galgotias College of Engineering and Technology
Gr. Noida, India

Vaibhav Rai
Department of Electronics and Communication
Galgotias College of Engineering and Technology
Gr. Noida, India

Abstract:- Nowadays AIR Pollution is a great concern and totally dominant topic in this world of industrialization. This needs to be controlled by predicting some future predicting techniques like AQI. AQI decides whether whether the air is pure or having hazardous air quality. AQI follows the regular pattern which is prominent in predicting the future air pollution quality. AQI follows the regular pattern which is prominent in predicting the future air pollution quality. LSTM is a relevant RNN technique which plays a crucial role in these deep learning methods. LSTM is used in time-series forecasting problems in which the inputs are taken in the form of regression values and can be predicted by using graphical techniques . This paper consist of deep information regarding RNN and LSTM which proved to be helpful is predicting the air pollution quality. This paper gives a serious impact of particulate matter (as pm2.5) in elaborating the data of particular city for some span of time.

Keywords:- *LSTM Cell , RNN, , Air Quality Index(AQI), Regression Analysis , PM 2.5.*

I. INTRODUCTION

Air plays a vital role in our daily lives and is one of the most prominent source for living our life. So it is our prime responsibility to keep it perfect for humans and other living beings. In this world full of various urban techniques around the globe we are getting in a disastrous situation such as Air Pollution. Air Pollution consist of various harmful particulate such as PM2.5, PM10 and other hazardous pollutants as CO, SO2, NO2,03 etc. AQI(Air Quality Index) is used for tracking the purity of air in any region around the globe. AQI Forecasting is basically the prediction of purity of air in the near future. The ancient technologies like Deep Learning, Forecast by analogy, DELPHI Method played a crucial role in enhancing some new techniques for forecasting and predicting of data such that using various regressive analysis and graphical representation of data. Within the past years the advancing behaviour of numerous tools that came into existence which are thoroughly used for forecasting of data.

In this paper we present RNN based model with the help of LSTM cell which is prominently used for time series forecasting. In this paper we will briefly discuss about Regression analysis[1] and graphically represents the data to be predicted in near future of a particular city.

II. LITERATURE REVIEW

Air Quality Index is an exhaustive index used to describe the freshness or purity of air on daily basis. AQI[12]is computed on the basis of the mean concentration of various pollutants or impurities present in air measured over a particular span of time. This paper comprises of descriptive data for AQI values of Kanpur(INDIA)which was corroborated from various government specified agencies that indicates the AQI values for pm2.5. AQI absolutely reviews the air contamination craze which can proved to be constructive for various Public-Interest Departments .

This paper explains us about the outcomes of forecasting[13]which came after correlating and examining the input data (training data and testing data) through RNN with additional support of MATLAB9.6.

Neural Networks has played a noteworthy development in the field of forecasting which made it the most favoured tool in machine learning. This paper demonstrates the advancement of RNN[6] in time series forecasting and predicting the futuremost value of AQI across some polluted cities which made it an intimating tool behind the bars.

The RNN Method has been skillfully applied in the the field of machine learning through various waveforms[5] by using various techniques like training , testing and comparing. This paper includes 2 sets of data that includes (a)training and target (b)testing and target. This paper also includes the comparison of the past and future values of AQI.

LSTM is one of the most substantial part of time-series forecasting model[9] that is favorably used to predict the expected futuremost AQI values of new data which is required to predict the AQI.

Forecasting is one of the most prominent and reliable method in predicting the futuremost value thus can be clearly suggested for improving the quality of AQI. Forecasting method constitutes Time Series Analysis[10] in which historical data are contributed and analysed to do the prediction. This time series anaylization technique proved to be time consuming and more accurate process to determine the future predicted values and analyzing it with the historical values shown in the dataset amd representing it graphically. It is primarily part of RNN that provides some sort of perfectionism in predicting the upcoming AQI values by

ISSN No:-2456-2165

inspecting the historical or previous data shown in the model.

This model gingerly represents the accurate pattern in AQI state and deliberately accede with the observed states. This paper deals with various random time-varying prediction[7] processes that includes past AQI values, various climatological processes. AQI state forecasting is barely with the help of RNN and LSTM technique.

The air pollutant which is seecking much awareness in the atmospheric region is due to contamination of fine particulate matter (PM2.5).PM2.5[4] conducts excess morbidity and also increases the mortality rate and toxicness in the environment. Variation of PM2.5 concentration is gradually determined by series of indefinitine climatological factors. For instance the numerous critically examined paper reveals that Aerosol Optical Thickness (AOT) is is just closely corresponded to PM2.5 concentration. This paper demonstrates that the valus can be earlier predictedby using machine learning with the gradual need of RNN and LSTM cell and can prove to be a long-term investment in the field of Health and Human-Management.

However the composition or percentage of error gradually describes the accurateness of the model.The batch gradient[11] or steepest gradient refurbish the framework of this model by considering the mean error of all the data samples available. The gradient conquer the obstacles by considering single data sample at a particular time and updates the parameter through the training set involved in the data given.As the gradient value of the analyzed output of the model decreases or step down , the accurateness of the system increases with it. This problem is decently examined by the graphical analysis of the realized data and it proves to be a accurately prediction technique in this complete model.

The most competitive thing in Forecasting[13] approach is the determination of proper sequence in which the data needs to be predicted in future. To overcome and challenge these obstacles this paper briefly explains the advancement of RNN[6] by the proper functioning of machine learning using LSTM. These pattern and sequences needs to analyze time series pattern and analyze long sequence data. RNN[6] is gradually used for learning and description of historical data. These neural networks are well proficient to understand the concept of long and short term memory differences as well as other transient differences.

In this paper the prediction[4] is properly described with the help of various graphical analysis and waveform[5] and waveform comparision of training and testing sets of the historical data. This suitably elaborates the maximum error rate of the forecasted[13] data and perfectly intricates the the accurateness of the experiment performed.

III. METHODOLOGY

A. *Data Interpretation:*

This dataset used is of Kanpur pm2.5 data which has values from 1st January 2018 to 31st January 2019. The factors that affect these values of pm 2.5 includes effects of black carbon , sulphate aerosols etc.

January	*February*	*March*	*April*	*May*	*June*	*July*
278	165	159	141	115	116	138
277	195	212	158	127	134	126
296	212	155	144	88	148	109
384	169	175	157	111	152	0

Table 1: Training Data

August	*September*	*October*	*November*	*December*
85	101	0	319	234
98	92	0	355	244
129	86	143	271	265
143	97	0	280	283

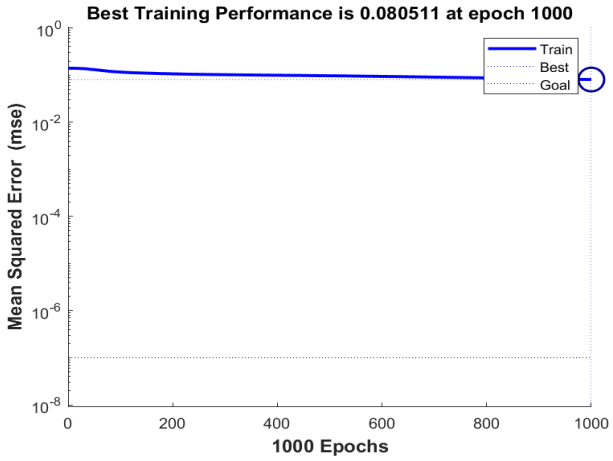
Table 2 : Testing Data

A. *Training RNN Model*

The model was trained for 62 days per iteration for 1000 iterations. The learning rate (lr) was 0.05 and the performance parameter was 10 till it attains the target value.

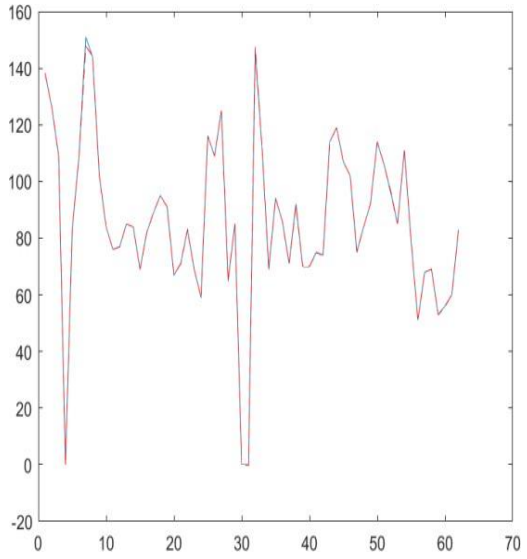
B. Testing Model

The historical data tested was from August,2018 to November 2018, against the LSTM[15] cell and forecasted it with the historical data set and the waveform has been created between the testing and training set of data involved in the model.

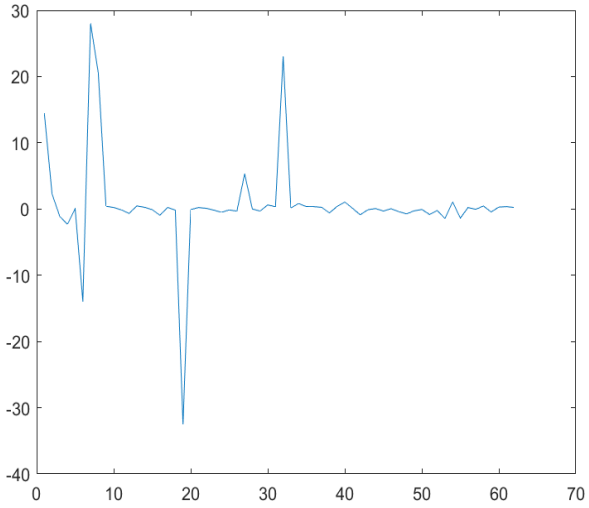


Graph 1: Best training performance .

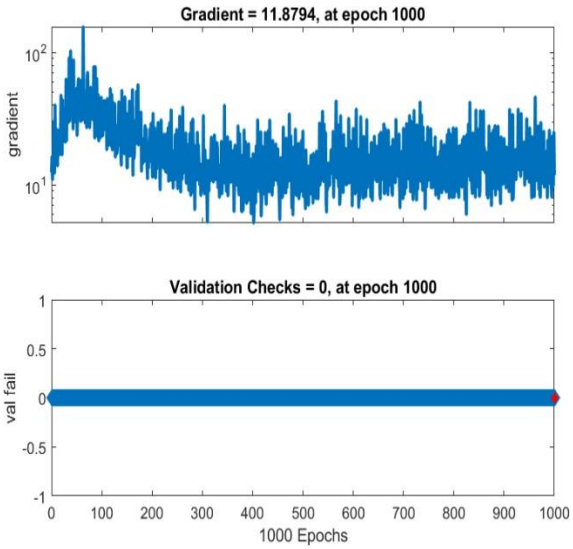
ISSN No:-2456-2165



Graph 2: The prediction vs actual graph of network



Graph 3: Error plot of training data



Graph 4: Gradient of training data

IV. CONCLUSION

From the various processes performed , we can identify that various factors are involved in determining AQI values We found that a blended composition of RNN and LSTM performs more accurately in forecasting the historical data. The mixture of RNN and LSTM performs much better than traditional LSTM network and works more efficiently in varying time series. We can clearly suggest that the best training performance increases with decrease in the value of epoch or error rate involved.

REFERENCES

[1]. Forecasting Air Quality Index using Regression Models: A Case Study on Delhi and Houston. S. S. Ganesh, S. H. Modali, S.

[2]. R. Palreddy and P.Arulmozhivarman, *2017 International Conference on Trends in Electronics and Informatics (ICEI)*, Tirunelveli, 2017, pp. 248-254.

[3]. Air Quality Forecast Monitoring and Its Impact on Brain Health Based on Big Data and the Internet of ThinY. Huang, Q. Zhao, Q. Zhou and W. Jiang, in IEEE Access, vol. 6, pp. 78678- 78688, 2018

[4]. Meteorology-based Forecasting of Air Quality Index Using Neural Network. M. Sharma, S. Aggarwal, P.

[5]. Bose and A. Deshpande, ,IEEE International Conference on Industrial Informatics, 2003. INDIN 2003.

[6]. Recurrent Air Quality Predictor Based on Meteorology- and Pollution-Related Factors. K. Gu, J. Qiao and W. Lin, in IEEE Transactions on Industrial Informatics, vol. 14, no. 9, pp. 3946- 3955, Sept. 2018

[7]. An Adaptive Kalman Filtering Approach to Sensing and Predicting Air Quality I. J. Chen, K. Chen, C. Ding, G. Wang, Q.

[8]. Liu and X. Liu, in IEEE Access, vol. 8, pp. 4265-4272, 2020

[9]. Forecasting of Air Quality Index in Delhi Using Neural Network Based on Principal Component Analysis . Kumar, A., Goyal, P. *Pure Appl. Geophys.* **170,** 711–722 (2013).

[10]. Ordinal time series model for forecasting air quality index for ozone in Southern California. Kim, S.E. Ordinal time series model for forecasting air quality index for ozone in Southern California. *Environ Model Assess* **22,** 175–182 (2017).

[11]. Air pollution forecasting in Ankara, Turkey using air pollution index and its relation to assimilative capacity of the atmosphere. Genc, D.D., Yesilyurt, C. & Tuncel, G. *Environ Monit Assess* **166,** 11–27 (2010).

[12]. Air quality forecasting based on cloud model granulation. Lin, Y., Zhao, L., Li, H. *et al. J Wireless Com Network* **2018,** 106 (2018).

[13]. Artificial neural networks and fuzzy time series forecasting: an application to air quality. Rahman, N.H.A., Lee, M.H., Suhartono *et al. Qual Quant* **49,** 2633–2647 (2015).

[14]. A rolling forecast approach for next six-hour air quality index track.Chen-Jui Liang Jeng-Jong Liang Chiao-Wun Jheng Ming Chang Tsai, Volume60, November2020,101153
[15]. The air quality index trend forecasting based on improved error correction model and data preprocessing for 17 port cities in China Suling Zhu Jianan Sun Yafei Liu Mingming Lu Xingrong Liu Volume 252, August 2020, 126474
[16]. A novel dynamic ensemble air quality index forecasting. HongminLiJianzhouWangHufangYang, Volume 11, 8, August 2020, Pages 1258-1270,