ISSN (ONLINE): 2454-9762
ISSN (PRINT): 2454-9762
Available online at www.ijarmate.com

International Journal of Advanced Research in Management, Architecture, Technology and Engineering (IJARMATE)
Vol. 7, Issue 4, April 2021

FUTURE PREDICTION OF AIR POLLUTION

T.Anto Theepak[1] , A.Inba Jothi[2] , A.Arthy[3] ,

1(Assistant Professor, Department of IT, Francis Xavier Engineering College, Tirunelveli, India)
2(Department of IT, UG Scholar, Francis Xavier Engineering College, Tirunelveli, India)
3(Department of IT, UG Scholar, Francis Xavier Engineering College, Tirunelveli, India)

ABSTRACT: Air pollution refers to the release of pollutants into the air that are detrimental to human health and the planet as a whole. It can be described as one of the most dangerous threats that the humanity ever faced. It causes damage to animals, crops, forests, and water-bodies. It also contributes to the depletion of the ozone layer, which protects the Earth from the sun's UV rays. Some of the other environmental effects of air pollution are haze, eutrophication, and global climate changes.Air pollution can be determined by air using air quality index. Some of the major air pollutants which increases air quality index are No2, So2, and O3 etc. In our Project we are going to predict future AQI value from the existing AQI in the dataset by using a deep learning based Long Short Term Memory algorithm (LSTM).

I.INTRODUCTION

To manage and control our environment from adverse effects of air pollution there is a need of effective prediction and analysis of the same. There are multiple reasons that contribute towards the increase in air pollution like vehicular emissions, construction and demolition sites, factories mining activities to name a few. We have numerous pollutants that contribute to air pollution are nitrogen dioxide, sulphur-dioxide, polycyclic-aromatic hydrocarbons [PAHs] and suspended particulate matter [PM] covering dust, soot, and smoke. Several methods including soft- computing techniques

ISSN (ONLINE): 2454-9762
ISSN (PRINT): 2454-9762
Available online at www.ijarmate.com

IJARMATE
Your ult-MATE Research Paper !!!

International Journal of Advanced Research in Management, Architecture, Technology and Engineering (IJARMATE)
Vol. 7, Issue 4, April 2021

like genetic-

algorithms, fuzzy-logic, artificial neural networks, artificial intelligence, back propagation algorithms, etc,. To name a few can be used in the analysis and prediction of air pollutants.

Air pollution occurs when harmful or excessive quantities of substances are introduced into Earth's atmosphere. Sources of air pollution include gases (such as ammonia, carbon

monoxide, sulphur dioxide, nitrous oxides, methane and chlorofluorocar bons), particulates (both organic and inorganic), and biological molecules. It may cause diseases, allergies and even death to humans; it may also cause harm to other living organisms such as animals and food crops, and may damage the natural or built environment. Both human activity and natural processes can generate air pollution.

Air pollution is a significant risk factor for a number of pollution-related diseases, including respiratory infections, heart disease, and COPD, stroke and lung cancer. The human health effects of poor air quality are far reaching, but principally affect the body's respiratory system and the cardiovascular system. Individual reactions to air pollutants depend on the type of pollutant a person is exposed to, the degree of exposure, and the individual's health status and genetics. Indoor air pollution and poor urban air quality are listed as two of the world’s worst toxic pollution problems in the 2008 Blacksmith Institute World's Worst Polluted Places report.Outdoor air pollution alone causes 2.1 to 4.21 million deaths annually.Overall, air pollution causes the deaths of around 7 million people worldwide each year, and is the world's largest single environmental health risk. Productivity losses and degraded quality of life caused by air pollution are estimated to cost the world economy $5 trillion per year. Various pollution control technologies and strategies are available to reduce air pollution.

ISSN (ONLINE): 2454-9762
ISSN (PRINT): 2454-9762
Available online at www.ijarmate.com

International Journal of Advanced Research in Management, Architecture, Technology and Engineering (IJARMATE)
Vol. 7, Issue 4, April 2021

II.EXISTING SYSTEM

In the existing System they analyzed the proportionality of pollutant in the air based on the- time-of the day and day-of-the- week. Also estimate the effect of various environmental parameters like temperature, wind speed, and NO, NO2, CO, PM10, and SO2.This

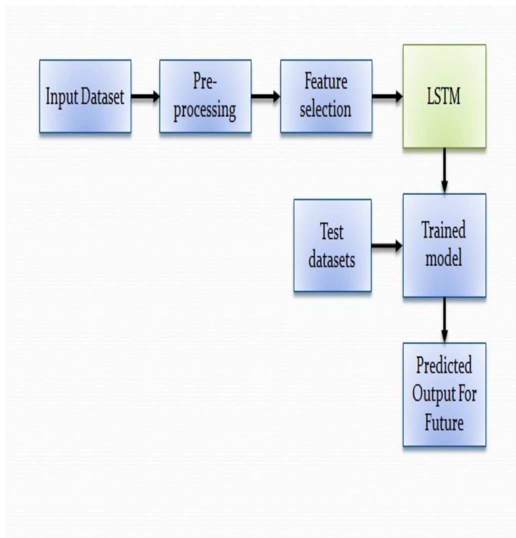
is estimated using the WEKA tool to analyze the air pollution data sets collected from the pollution control board. The problem which is occurred here is they cant able to predict the air quality index of future.

III. PROPOSED SYSTEM

In our project, we detects the future AQI value by using the existing AQI from the air pollution data sets collected from the pollution control board.In our project for future prediction we used a machine learning algorithm which is called as Long short term algorithm. By using long short term algorithm the cost is highly reduced, so that we got high accuracy

ISSN (ONLINE): 2454-9762
ISSN (PRINT): 2454-9762
Available online at www.ijarmate.com

International Journal of Advanced Research in Management, Architecture, Technology and Engineering (IJARMATE)
Vol. 7, Issue 4, April 2021



System Architecture

ISSN (ONLINE): 2454-9762
ISSN (PRINT): 2454-9762
Available online at www.ijarmate.com

IJARMATE
Your uh-MATE Research Paper !!!

International Journal of Advanced Research in Management, Architecture, Technology and Engineering (IJARMATE)
Vol. 7, Issue 4, April 2021

The input data’s are pre-processed to extract the needed data’s from the data set. At first we drop the null rows from the dataset after that we change the data type of date column from object to date time data format.

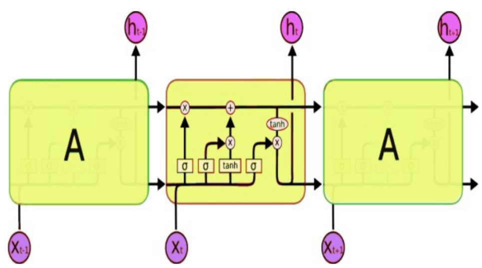
After that the Input features are selected from the pre-processed data. Here the date column and the AQI values are considered as the Input Features Then the input features are feeded into the LSTM Model for training process. The model is trained continuously for 100 times to get the accurate result. After training the test data's are feeded into the model to get the predicted future AQI values.

ISSN (ONLINE): 2454-9762
ISSN (PRINT): 2454-9762
Available online at www.ijarmate.com

International Journal of Advanced Research in Management, Architecture, Technology and Engineering (IJARMATE)
Vol. 7, Issue 4, April 2021

IV. LSTM ALGORITHM AND IT'S ARCHITECTURE

Long short-term memory (LSTM) is an artificial recurrent neural network (RNN) architecture used in the field of deep learning. LSTM networks are well-suited to classifying, processing and making predictions based on time series data, since there can be lags of unknown duration between important events in a time series. 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.



Architecture of LSTM

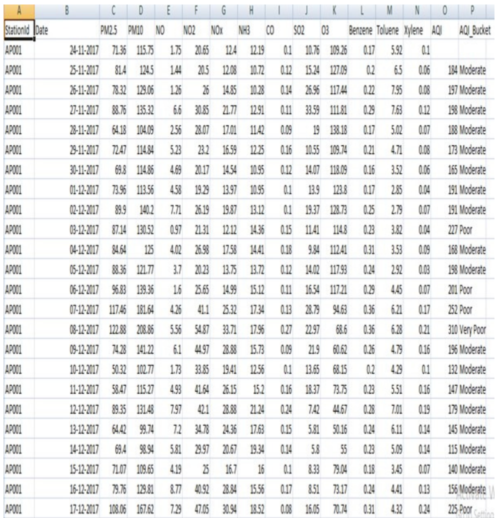
Long short-term memory (**LSTM**) is an artificial recurrent neural network (RNN) architecture used in the field of deep learning. Unlike standard feedforward neural networks, LSTM has feedback connections. It can not only process single data points (such as images), but also entire sequences of data (such as speech or video).

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

ISSN (ONLINE): 2454-9762
ISSN (PRINT): 2454-9762
Available online at www.ijarmate.com

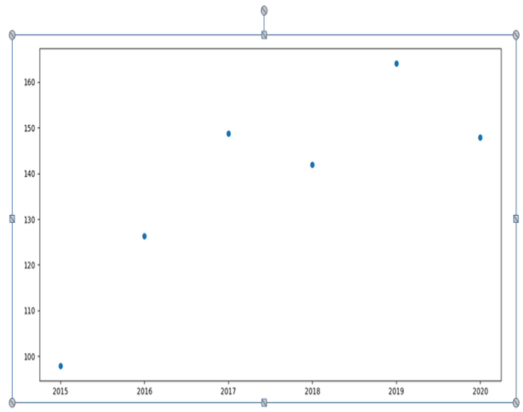
International Journal of Advanced Research in Management, Architecture, Technology and Engineering (IJARMATE)
Vol. 7, Issue 4, April 2021

Air pollution dataset



StationId	Date	PM2.5	PM10	NO	NO2	NOx	NH3	CO	SO2	O3	Benzene	Toluene	Xylene	AQI	AQI_Bucket
AP001	24-11-2017	71.36	115.75	1.75	20.65	12.4	12.19	0.1	10.76	109.26	0.17	5.92	0.1		
AP001	25-11-2017	81.4	124.5	1.44	20.5	12.08	10.72	0.12	15.24	127.09	0.2	6.5	0.06	184	Moderate
AP001	26-11-2017	78.32	129.06	1.26	26	14.85	10.28	0.14	26.96	117.44	0.22	7.95	0.08	197	Moderate
AP001	27-11-2017	88.76	135.32	6.6	30.85	21.77	12.91	0.11	33.59	111.81	0.29	7.63	0.12	198	Moderate
AP001	28-11-2017	64.18	104.09	2.56	28.07	17.01	11.42	0.09	19	138.18	0.17	5.02	0.07	188	Moderate
AP001	29-11-2017	72.47	114.84	5.23	23.2	16.59	12.25	0.16	10.55	109.74	0.21	4.71	0.08	173	Moderate
AP001	30-11-2017	69.8	114.86	4.69	20.17	14.54	10.95	0.12	14.07	118.09	0.16	3.52	0.06	165	Moderate
AP001	01-12-2017	73.96	113.56	4.58	19.29	13.97	10.95	0.1	13.9	123.8	0.17	2.85	0.04	191	Moderate
AP001	02-12-2017	89.9	140.2	7.71	26.19	19.87	13.12	0.1	19.37	128.73	0.25	2.79	0.07	191	Moderate
AP001	03-12-2017	87.14	130.52	0.97	21.31	12.12	14.36	0.15	11.41	114.8	0.23	3.82	0.04	227	Poor
AP001	04-12-2017	84.64	125	4.02	26.98	17.58	14.41	0.18	9.84	112.41	0.31	3.53	0.09	168	Moderate
AP001	05-12-2017	88.36	121.77	3.7	20.23	13.75	13.72	0.12	14.02	117.93	0.24	2.92	0.03	198	Moderate
AP001	06-12-2017	96.83	139.36	1.6	25.65	14.99	15.12	0.11	16.54	117.21	0.29	4.45	0.07	201	Poor
AP001	07-12-2017	117.46	181.64	4.26	41.1	25.32	17.34	0.13	28.79	94.63	0.36	6.21	0.17	252	Poor
AP001	08-12-2017	122.88	208.86	5.56	54.87	33.71	17.96	0.27	22.97	68.6	0.36	6.28	0.21	310	Very Poor
AP001	09-12-2017	74.28	141.22	6.1	44.97	28.88	15.73	0.09	21.9	60.62	0.26	4.79	0.16	196	Moderate
AP001	10-12-2017	50.32	102.77	1.73	33.85	19.41	12.56	0.1	13.65	68.15	0.2	4.29	0.1	132	Moderate
AP001	11-12-2017	58.47	115.27	4.93	41.64	26.15	15.2	0.16	18.37	73.75	0.23	5.51	0.16	147	Moderate
AP001	12-12-2017	89.35	131.48	7.97	42.1	28.88	21.24	0.24	7.42	44.67	0.28	7.01	0.19	179	Moderate
AP001	13-12-2017	64.42	99.74	7.2	34.78	24.36	17.63	0.15	5.81	50.16	0.24	6.11	0.14	145	Moderate
AP001	14-12-2017	69.4	98.94	5.81	29.97	20.67	19.34	0.14	5.8	55	0.23	5.09	0.14	115	Moderate
AP001	15-12-2017	71.07	109.65	4.19	25	16.7	16	0.1	8.33	79.04	0.18	3.45	0.07	140	Moderate
AP001	16-12-2017	79.76	129.81	8.77	40.92	28.84	15.56	0.17	8.51	73.17	0.24	4.41	0.13	156	Moderate
AP001	17-12-2017	108.06	167.62	7.29	47.05	30.94	18.52	0.08	16.05	70.74	0.31	4.32	0.24	225	Poor

Actual AQI value

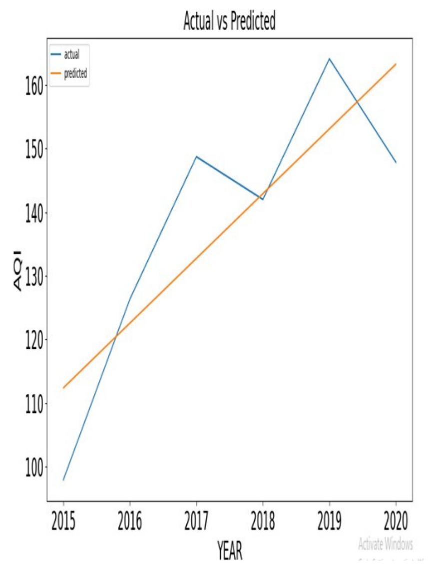


Actual AQI value

ISSN (ONLINE): 2454-9762
ISSN (PRINT): 2454-9762
Available online at www.ijarmate.com

International Journal of Advanced Research in Management, Architecture, Technology and Engineering (IJARMATE)
Vol. 7, Issue 4, April 2021

Actual vs Predicted value for existing years



Predicted values for future

```
Predicted AQI of 2021 is 262.9427141407758
Predicted AQI of 2022 is 237.60897496745898
Predicted AQI of 2023 is 288.27645331408394
Predicted AQI of 2024 is 262.9427141407758

Predicted AQI_Bucket of 2021 is poor
Predicted AQI_Bucket of 2022 is poor
Predicted AQI_Bucket of 2023 is poor
Predicted AQI_Bucket of 2024 is poor
```

CONCLUSION

The dataset collected from pollution control board is analyzed using the deep learning based Long Short Term Memory algorithm (LSTM). From the analysis it is clear thatpollutants have a tendency to increase during workingdays and especially peak hours of the day andsubsides during the week-ends and general holidays.In our Project we are predict future AQI value from the existing AQI in the dataset by using a machine learning algorithm.The results show the dependencies and their relationships. As a futurescope dependencies from one pollutant to the otherwith respect to the distance and future predictions canbe estimated using the various soft computing algorithms.

REFERENCE

[1] Mouhammd Alkasassbeh, Alaa F. Sheta, Hossam Faris and Hamza Turabieh, "Prediction of PM10 and TSP air pollution parameters using artificial neural network autoregressive, external input models: a case study in Salt, Jordan", Middle-East Journal of Scientific Research14 (7): pp. 999-1009, 2013.

[2] Ahmed Boubrima, WalidBechkit, and HerveRivano,"Optimal WSN Deployment Models for Air Pollution Monitoring", EEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, VOL. 16, NO. 5, MAY 2017, 2723-2735

[3]Shweta Taneja, Dr. Nidhi Sharma, KettunOberoi,YashNavoria, "Predicting Trends in Air Pollution in Delhi using Data Mining", Conference: 2016 1st India International Conference on Information Processing (IICIP).

[4] R. G. Derwent, D. R. Middleton, R. A. Field and R. Perry,"Analysis and Interpretation of Air Quality Data from anurban roadside location in Central London over the periodfrom July 1991 to July 1992," In Atmospheric Environment29(8),923- 946, April 1995.

[5] Edward Ming-Yang Wu, and Shu-Lung Kuo, "Air Quality Time Series Based GARCH Model Analyses of Air QualityInformation for a Total Quantity Control District," Aerosoland Air Quality Research, Vol. 12, pp. 331–343, 2012.

[6] R. Sindhwani , "Assessment of gaseous and respirable suspended Particulate matter (PM10) emission estimates overmegacity Delhi: past trends and future scenario (2000- 2020)", Center for atmospheric sciences, November 2012, pp. 123- 140.

[7] A.S. Nag pure, K. Sharma, B.R. Gur jar, "Traffic induced emission estimates and trends (2000–2005) in megacityDelhi. ", Urban Climate, vol. 4, 2013, pp. 61–73.

[8] Sarath Guttikunda, and Puja Jawahar, "Urban Air Pollution Analysis in India," Urban Emissions. Info, NewDelhi, India, 2011.