Journal of Inventive Engineering and Technology (JIET)
ISSN: 2705-3865
Volume-2, Issue-3, pp-15-29
www.jiengtech.com
Research Paper Open Access

PM10.0 AIR POLLUTION MODELING AND ESTIMATION USING ARTIFICIAL NEURAL NETWORK (ANN)

[1](**Asogwa, Samuel Chibuzor,** Department of Computer Science, Micheal Okpara University of Agriculture, Umudike, Abia State, Nigeria) sasogwa@gmail.com

[2](**Obodoeze, Fidelis Chukwujekwu,** Department of Computer Engineering, Akanu Ibiam Federal Polytechnic Unwana, Ebonyi State, Nigeria) fcobodoeze@gmail.com

[3](**Ugwuja N.E.,** Department of Computer Science, Micheal Okpara University of Agriculture, Umudike, Abia State, Nigeria)

[4](**Etuk Enefiok A.,** Department of Computer Science, Micheal Okpara University of Agriculture, Umudike, Abia State, Nigeria)

Corresponding Author: fcobodoeze@gmail.com

ABSTRACT : *Air pollution is a huge challenge to the residents of highly populated cities and their municipal managers over the years because of the serious threats it poses to human health and environment. Several smart Artificial Intelligence based techniques such as the use of Machine learning and deep learning algorithms methods such as Artificial Neural Network (ANN) can be deployed to predict or estimate the levels of emissions of pollutants in the ambient environment within a particular locality or city provided past historical dataset of other air and meteorological parameters are available to use to identify patterns in an occurrence in the dataset which can be used to predict or forecast future occurrences of air pollution in that particular location. This paper used historical dataset of air pollutants and meteorological parameters such as PM2.5, PM10.0, PM1.0, and other parameters obtained in Awka Metropolis from October 14th 2021 to December 4, 2021 to model the proposed models. In this work, the particulate matter PM10.0 prediction modeling were carried out using Artificial Neural Network (ANN) and eight other machine learning models using the Train-Test Data Split method. The prediction performances of these machine learning models were evaluated using statistical performance metrics such as MAE, RMSE and R^2. After more than 100 experimental runs, Artificial Neural Network (ANN) algorithm using Multi Layer Perceptron (MLP) with 200 hidden neurons, we obtained an RMSE value of 1.2637 µg/m^3, MAE of 0.5885 µg/m^3 and R^2 of 0.9828 or 98.28%, which is a high accurate prediction of the measured value. Other machine learning algorithms tried on the training and testing data gave different results of RMSE, MAE and R^2. This shows that apart from ANN, other machine learning algorithms can accurately predict or forecast in advance time the levels of dispersion of air pollutants such as PM10.0 in the ambient environment of a particular city or location.*

KEYWORDS: *PM10.0, PM2.5, particulate matters, Artificial Neural Network (ANN), smart city*

Date of Submission: 15-03-2023 Date of acceptance: 24-03-2023

I. INTRODUCTION

Air pollution is a very big challenge all over the world, especially in big or mega cities where there is population explosion. Air pollution deteriorates the quality of air in the ambient environment and creates several health problems. Due to this reasons, there is an urgent need to provide smart-city air pollution monitoring and management solutions that can be deployed to manage and mitigate the challenges arising from air pollution in highly populated cities or metropolis; such smart solutions may include the deployment of intelligent machine learning models such as Artificial Neural Network (ANN) to predict or forecast beforehand before the dispersion of dangerous air pollutants to the ambient environment. Population explosion is the main reason that leads to increased air pollution issues in big cities; this is due to increased human activities, use of industrial machines, motor vehicle exhausts, and also emissions from small-scale businesses and domestic activities. The main air pollutants include particulate matters such as PM1.0, PM2.5, PM10 and gaseous effluents such as SO2, NO_2, NO, CO, CO_2, Ozone (O_3), Volatile Organic Compounds (VOCs), Formaldehydes (HCHO) as well as noise pollution. Nigerian is ranked 18th as the countries with the highest air pollution index in the world (IQAir, 2023); see also Table 1. Also according to UNICEF (2021), Nigeria has highest number of air pollution-related child pneumonia deaths in the world. Air pollution, especially in the home, is the biggest driver of child deaths from pneumonia in Nigeria

Some of the various pollutants found in the air in Nigeria could come from a wide range of the different polluting sources. One of such sources include cars and other vehicles and these can release large amounts of nitrogen dioxide (NO_2) and sulfur dioxide (SO_2), as well as carbon monoxide (CO) and black carbon, the main component in soot; this can be very poisonous when inhaled. This can also lead to poor visibility to motor vehicle drivers which can be very dangerous to road users as this can cause auto accidents on highways. Black carbon can also have a profound effect on the environment due to its ability to absorb solar radiation from the sun and convert it directly into heat.

Table 1: World's most polluted countries & regions (historical data 2018-2021) (IQAir, 2023)

Rank	Country/Region	2021	2020	2019	2018	Population
1	Bangladesh	76.9	77.1	83.3	97.1	164,689,383
2	Chad	75.9	-	-	-	16,425,859
3	Pakistan	66.8	59	65.8	74.3	220,892,331
4	Tajikistan	59.4	30.9	-	-	9,537,642
5	India	58.1	51.9	58.1	72.5	1,380,004,385
6	Oman	53.9	44.4	-	-	5,106,622
7	Kyrgyzstan	50.8	43.5	33.2	-	6,524,191
8	Bahrain	49.8	39.7	46.8	59.8	1,701,583
9	Iraq	49.7	-	-	-	40,222,503
10	Nepal	46	39.2	44.5	54.1	29,136,808
11	Sudan	44.1	-	-	-	43,849,269
12	Uzbekistan	42.8	29.9	41.2	34.3	33,469,199
13	Qatar	38.2	44.3	-	-	2,881,060
14	Afghanistan	37.5	46.5	58.8	61.8	38,928,341
15	United Arab Emirates	36	29.2	38.9	49.9	9,890,400
16	Montenegro	35.2	26.1	-	-	628,062
17	Indonesia	34.3	40.7	51.7	42	273,523,621
18	Nigeria	34	-	-	-	206,139,587

Others would include ones from factories and open burn sites, where fossil fuels, organic matter and synthetic materials are all burnt. They include pollutants such as volatile organic compounds (VOC's), polychlorinated biphenyls, dioxins, furans and even heavy metals such as lead, mercury and cadmium. Some examples of VOC's include formaldehyde, benzene, toluene, xylene and methylene chloride. They can also find their release from household items of products, with varnishes and other similar materials emitting these chemicals. This is further compounded if such materials are burnt in an open fire, releasing large amounts of these VOC's and other chemical compounds and particulate matter into the air. Also, Nigerian cities are polluted with coarse particles sometimes called particulate matters (PM2.5 and PM10.0) which can cause very severe health complications and even death.

Coarse (bigger) particles, called PM10.0, can irritate your eyes, nose, and throat. Dust from roads, farms, dry riverbeds, construction sites, and mines are types of PM10. Fine (smaller) particles, called PM2.5, are more dangerous because they can get into the deep parts of your lungs — or even into your blood. On long term basis, patients with respiratory diseases such as asthma and lung diseases can suffer more severe consequences such as death if affected by air pollutants such as PM2.5.

Some of the health issues that arise from the inhalation of these tiny particles and various chemicals include short term acute ones such as increased bouts of coughing, chest infections as well as aggravation of preexisting conditions such as asthma. Irritation to the mucous membranes can also occur, with the eyes, nose, mouth and ears all susceptible to aggravation or breakouts, with allergies being triggered off in vulnerable groups that include young children and those with a predisposition towards chemical sensitivities.

Machine learning algorithms such as Artificial Neural Network (ANN) can be used to estimate or predict the present or future levels or concentrations of PM10.0 air pollutants or any other pollutants in the atmosphere within a particular city or locality if historical dataset of such air pollutants and some weather variables are easily available. This can help city managers and residents to take proactive safety measures to protect the health of residents especially the health sensitive groups such as people with respiratory diseases.

a. Artificial Neural Networks (ANN)

ANN has shown that it is very powerful in pattern recognition applications with strong classification and recognition capabilities by following the human logic system which is the brain. Artificial Neural Networks (ANNs) are able to learn from experience, so it can process very large and intelligent tasks. ANN is a parallel system which fulfils the most complicated operations or tasks of realization in different fields of business industry and science. It predicts and detects without increasing the complexity of the problem.
ANN has hidden layers in the middle of one input and one output that will process the information data to the next layer and each layer to the next layer by forwarding the result until it arrive at the final layer which is output layer. ANN is the most used algorithm of recent in several human disciplines. It is the most popular machine learning (ML) algorithm in artificial intelligence (AI). It uses particularly for different processing as FBP feed forward back propagation. These effective machines solve complex problems every second and can used to make people’s life much easier. (Yegnanarayana, 2009).

Neurons:

Human brain includes set of biological neurons connected as network structures. Hence, human brains are interconnected set of neural networks realizing our thinking, reading, breathing, and motion. Some of these neural structures were given at birth and other parts have been learned through experience and intuition. Artificial Neural Network (ANN) processes these huge and complex tasks that feed into its hidden layers just like the neurons that operate in human brains for processing tasks. ANN neurons, in order to work properly, need to be trained with past dataset or data samples in order to forecast or predict future data or trend. Fig.1 shows the structure of a neuron.

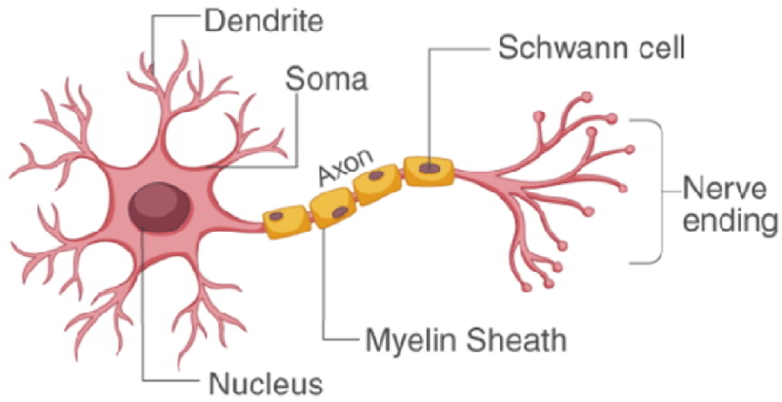

Fig.1: The Structure of a Neuron (Byju's, 2022)

Next, the testing and training data are holding out to test the result with other data to gain the difference by feeding the network system with numbers of neural network neurons, which we can say the number of neurons is changeable and depend on processing complexity and to the data that we are going to feed it in the network system, subsequently, depending on the output and input complication and the layers on the network system. Therefore, the architecture of ANN may vary from one to another (Demuth, Beale, Jess, & Hagan, 2014).

b. Structure of Artificial Neural Network (ANN)

Neural network includes a large number of units arranged in a series of layer which is made up of several artificial intelligent neurons. The architecture or structure of Artificial Neural Network (ANN) is shown in Fig.2.

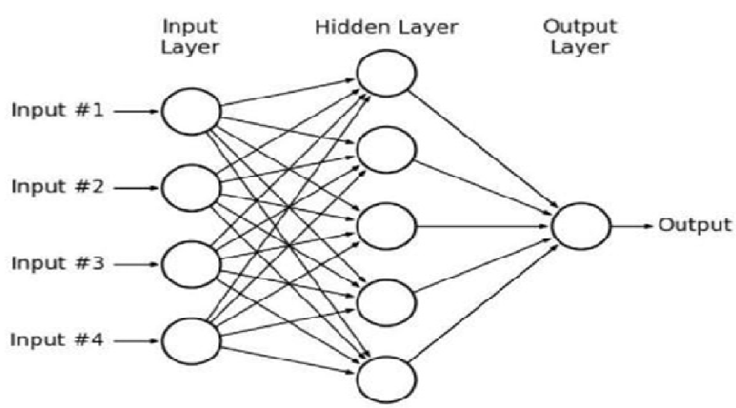

Fig. 2: Structure of ANN (Beyza Ecem Erce, 2023)

INPUT LAYER: it is the first layer that contains artificial neurons which receive the input data from outside in order to learn to recognize or processes.

HIDDEN LAYER: The hidden consists of the neurons at the middle between the input layer and the output layer; that is the reason it is referred to as the middle layers. The job of these layers is to process data and transform the input data through the network neurons to the output. For fineness and validity, the weights are continuously updated to the output of the hidden layer.

OUTPUT LAYER: the final layer in the structure of ANN contains units that respond to the data through learning to obtain the final result. Most neural networks are fully linked and connected that means the hidden layer fully linked between each neuron in the next output layer and to the previous layer or input layer at first

.
Weights:

When a neural network takes the large dataset through its input layer, it split the dataset into tiny fragments, then it transmits these fragments through all the neurons. The neurons take the data, process them using its stored weight, then send the results to the output layer. So, in ANN architecture the information and data are stored in memory storages. The weight is also modified at each step during the training, testing and validation. So, the output accuracy is carried out and the data is saved for any future operation.

Feed -forward Neural Network:

A Feed-forward neural network is a classification algorithm that is modeled after biological cells or neurons of the brain. It is made up of several simple neuron-like processing units, organized in layers and every unit in a layer is connected with all the units in the previous layer.

The simplest type of Feed-forward Neural network is the Perceptron; it is a type of Feed-forward neural network that has no hidden units or layers. This implies that a Perceptron has only an input layer and an output layer. There are some kinds of more complex Feed-forward networks that have input layer, hidden layers and output layer and they are called Multi Layer Perceptron (MLP). The architecture of Feed-Forward Neural Network is shown in Fig.3.

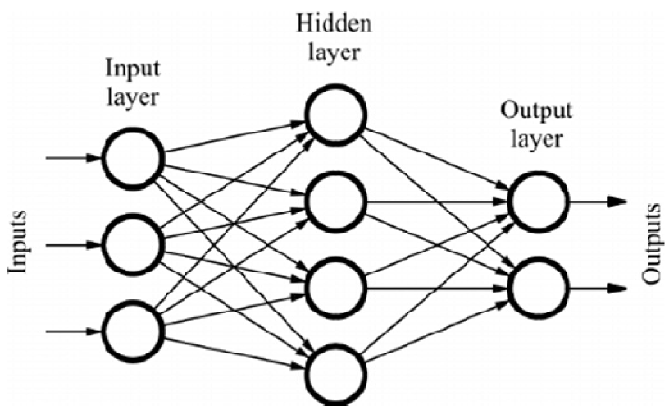

Fig.3: The architecture of a Feed-Forward Multi Layer Perceptron ANN (Beyza Ecem Erce, 2023)

It is called Feed-Forward (FF) because the data flow through the function that evaluated from the input. Feed-Forward neural network is a type of classification algorithm that is impacted biologically. It depends on a large number of normal neurons like in a process the units and orderly in a layer with all previous layers. The input layer processes the data that it receives and sends the obtained result to the next layer. Each of the linked layers has a different weight or strength.. The result can gain through the processing of each layer. Lastly, it can be gained from the output layer. Any layer that is not an output layer or input layer is a hidden layer. The artificial neurons work like a human brain that processes the input data in the neurons. Neurons in the middle layers send the data or information from the input layer through a channel called connection and the layer only connect to the previous layers.

c. Applications of Artificial Neural Network

Artificial Neural Network can be applied in different domains of industry to accomplish the following results:-

1. *Data Clustering*
2. *Data Classification*
3. *Data Regression, and*
4. *Time Series prediction*

From the results obtained in the various experiments carried out by the authors in the papers, it is evident that ANN is very useful in the estimation of the concentrations of the various air pollutants in the atmosphere in a locality. In this paper, data-driven models such as the supervised machine learning algorithms like the Artificial

Neural Network (ANN) will be used to model PM10.0 pollutant concentrations in Awka Metropolis of Anambra State, Nigeria. Awka is the most populated Metropolitan City in Anambra State of Nigeria after the commercial city of Onitsha. Awka was created as the administrative capital of Anambra State since 1991. Awka has quite a lot of government buildings, civilian residential buildings, universities, polytechnics and several factories and thousands of vehicles plying the city 24/7.

Several works have been carried out in Literature on the use of smart city air pollution prediction systems using machine learning algorithms such as Artificial Neural Network (ANN) and other statistical deterministic approaches especially for big cities or metropolis. For instance, Goulier et al (2020) presented a research paper on the application of Multi Layer Perceptron (MLP) Artificial Neural Network (ANN) to model and predict air pollutant concentrations in Canyon Street of the city of Münste, Germany. The inputs to the model are air pollutant concentrations of CO2, NH3, NO, NO2, NOx, O3, PM1, PM2.5, PM10 and PN10 as well as equivalent sound pressure level, the total number of vehicles, time of the day, hour of the day, day of the week, and motor traffic intensity. Meteorological data such as wind speed, wind direction, relative humidity and air pressure were also added as inputs. The experimental results showed that there is a clear positive correlation between modelled and observed values.

Aljanabi et al. (2020) presented a research study for the prediction of ground-level ozone (O3) in the city of Amman, Jordan using ANN Multi Layer Perceptron (MLP). It was concluded in the study that MLP outperformed other considered models using during pre-processing filter experiments. Allam (2019) presented a paper on a theoretical model of using Artificial Neural Network (ANN) to analyse air pollution index from an IoT sensors using several artificial neurons. The paper justified the use of ANN to work with IoT sensors in order to analyse pollution data accurately and on a timely fashion. Cortina-Januchs et al (2015) presented a research paper on the development of PM10 concentrations forecasting model for Salamanca Mexico. The proposed model combined Multi-Layer Perceptron (MLP) Artificial Neural Network and Clustering algorithm to predict next day's maximum hourly daily PM10 concentrations within the city. The source dataset used historical time series of meteorological variables such as wind direction, wind speed, air temperature, and relative humidity together with PM10 concentrations. Experimental results showed that the combination of Artificial Neural Network (MLP) and clustering algorithm improved the accuracy of PM10 forecasting compare to using either conventional analytical modeling tool or Artificial Neural Network (ANN) alone. Khoshand et al (2017) presented a study on the prediction of ground-level air pollution using Multi-Layer Perceptron (MLP) Artificial Neural Network (ANN) using Back Propagation (BP) algorithm with MATLAB program for training the model. The developed model was used to manage and predict daily concentrations of various air pollutants such as Ozone, PM10, NO2, CO and PM2.5 in Tehran city of Iran within a four-year period from 2012 to 2015. The experimental results showed appropriate agreement between the observed and predicted concentrations. This shows that ANN can be used to improve prediction accuracy of air pollution.

II. MATERIALS AND METHODS

This section presents the materials and methodology used in the experiments of this research paper

A. MATERIALS

About 12,958 historical datasets of air and noise pollution, as well as meteorological parameters captured in Awka Metropolis of Anambra State, Nigeria from October 25 to December 4 of 2021 using real-time air pollutants and weather wireless sensors, were used as input predictors for the proposed Artificial Neural Network (ANN) model. These datasets included PM10.0, PM2.5, VOCs, carbon dioxide, noise etc and also meteorological parameters such as air temperature, relative humidity, air pressure and light intensity. The historical dataset for the experiments can be found at http://www.myrasoft.ng/awka-pollution-monitor/AWKA-POLLUTION-2022NEW.csv

B. METHODS

The methodology employed in this research paper followed the seven steps of machine learning modeling as depicted in Fig.4.

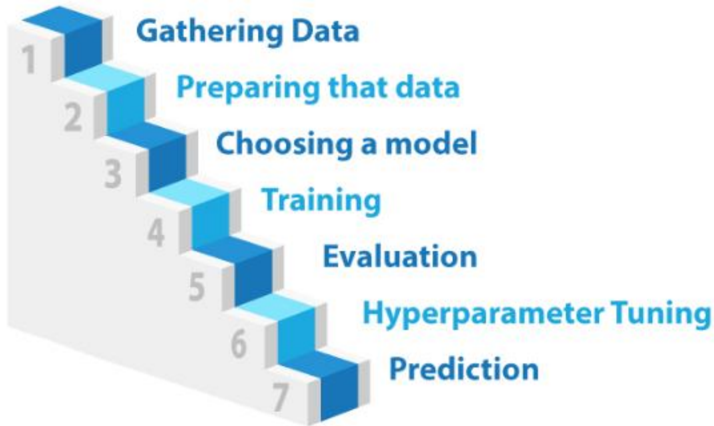

Fig.4: The steps of the Methodology used in the experiments

A typical machine learning process must follow the methodology as shown in Fig.5 which includes training, validation and testing in order to obtain an accurate and reliable machine learning model for deployment in solving a particular problem.

Fig5: A typical machine learning process

a. Train-Test Data Split Method

The train-test split was used to estimate the performance of each of the eight machine learning algorithms in predicting the PM2.5 air pollutant levels. This method is a fast and easy procedure to compute prediction accuracy (R^2) and prediction errors (RMSE and MAE) for each pollutant's concentration of each machine learning model results so as to compare them to one another.

By default Test set is split into 30% of actual data and the Training set is split into 70% of the actual data as shown in Fig.6.

The dataset is split into training and testing sets to evaluate how well the machine learning models are performing. The *train set* is used to train or *fit the model*, the statistics of the train set are known. The second set is called the *test or testing dataset,* this set is solely used to *perform predictions* only and evaluate the model.

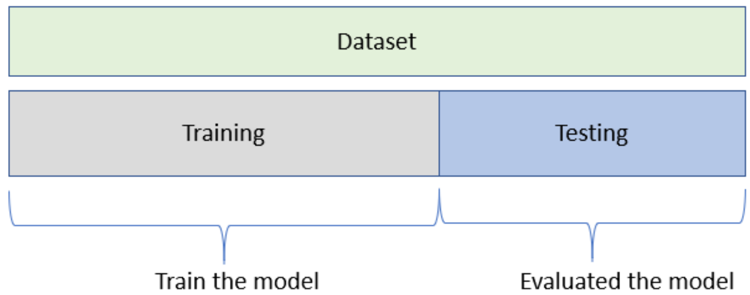

Fig. 6: Dataset splitting into training and testing datasets for Train-Test Data Split method

Fig. 7 shows the following steps taken to actualize Experiment 4 using Train-Test Split Method:

1. *Load the Awka Metropolis Pollution historical Dataset into Python Pandas or memory storage*
2. *Split the loaded dataset into training samples (70%) and the remaining 30% as testing dataset*
3. *Fit the models on the training dataset*
4. *Perform pollutant's concentration prediction using the test dataset*
5. *Plot the model (measured data versus predicted data)*
6. *Print the accuracy and prediction errors scores*
7. *Compare the scores of accuracy and prediction errors of each model*

Fig.7: Train-Test Data Split method used in the experiments

b. Artificial Neural Network (ANN) MLP Design and Hyperparameters

Hyperparameters are tuning of various features of the machine learning model in order to obtain the best predictive power of the dataset during training and testing. For the design and implementation of the MLP Artificial Neural Network (MLP ANN) in the experimental runs, the following network and hyperparameters settings as shown in Table 2 were used to achieve network convergence and best optimal performance results in the experimental runs. Python's *MLP Regressor () function or method* was used to instantiate or implement MLP ANN network in this paper. Fig. 8 depicts the structure of the MLP ANN used in this research thesis.

Table 2: Hyperparameters used in the experiments for Artificial Neural Network (ANN) MLP

Parameter	Value
No of input layers	9
No of output layer	1
No of hidden layers(hidden_layer_size)	100
Learning_rate_init	0.001
Learning_rate	constant
Activation_function	Relu
Max_iter	200
Alpha	0.0001
Batch_size	auto=max_ter=200
Power_t	0.5
Random_state	none
Solver	adam

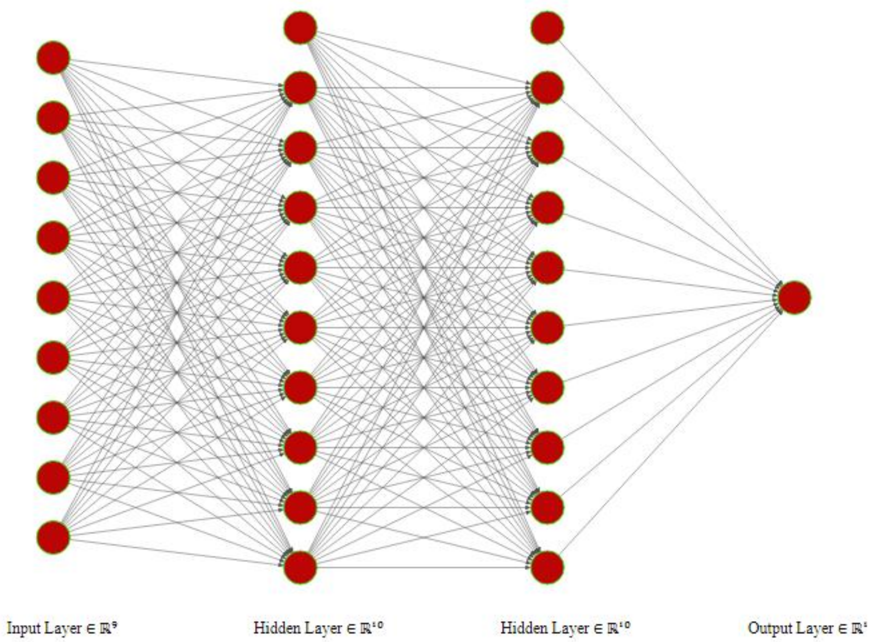

Fig. 8: The Artificial Neural Network (ANN MLP) structure diagram used in the research experiment

The Input LAYER:

The input layer includes nine (9) input predictors such as air and noise pollutant variables such as PM2.5, PM_1(PM1.0), PM10, TVOC, ecarbondioxide, noise and meteorological variables such as temperature, pressure, humidity, light.

The Output LAYER:

The output layer includes any output or dependent variable which can be any of the air and noise pollutant variable such as PM2.5, PM2.5, PM_1, PM10, TVOC, ecarbondioxide, noise.

The Hidden LAYER:
The hidden layer contains about 100 neurons interconnected together in the network and to the input and output layers.

A. PERFORMANCE EVALUATION METRICS

In order to determine or evaluate the best machine learning Air pollution prediction models quantitatively in terms of error bands or the prediction accuracy, the following statistical performance metrics- Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Coefficient of Determination or Variance (R^2) were employed and calculated as shown in Eqs. (1)-(3).

A. Mean Absolute Error (MAE)

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|P_i - M_i| \quad (1)$$

B. Root Mean Square Error (RMSE)

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}|P_i - M_i|^2} \quad (2)$$

C. Coefficient of Determination (R^2)

$$R^2 = 1 - \frac{\sum(M_i - P_i)^2}{\sum(M_i - \overline{M_i})^2} \quad (3)$$

where n is the number of data in the test dataset, P_i and M_i are the predicted and measure value for the ith hour and $\overline{M_i}$ is the mean of all the measured values for the ith hour. The higher the value of R^2, the more accurate and better the prediction result while the lower the values of RMSE and MAE, the higher the accuracy of the prediction model or algorithm.

III. RESULTS AND DISCUSSION

This section describes the experimental results and discussion of the results obtained from the experiments.

A. RESULTS

The sensor readings obtained from the outdoor deployment of the air and noise pollution monitoring system is in consonance with pollution and weather datasets from some online forecasting websites for Awka Metropolis. Fig.9 shows the PM10.0 concentration or levels (in µg/m^3) in Awka Metropolis on November 4, 2021. This shows that the maximum PM10.0 distribution is at 90 µg/m^3.at around 8.45 AM in the morning and the lowest PM10.0 sensor reading of 10 µg/m^3 at around 10.10 AM.

Fig.9 Plot of PM10.0 Concentration against Time on November 4, 2021

Fig. 10 shows the PM10.0 statistical distribution in Awka Metropolis. Fig. 11 depicts the statistical distribution of PM10.0 in Awka Metropolis using Histogram plot. Fig.12 shows the density plot for the PM10.0 distribution within Awka Metropolis.

```
In [126]: dataset['PM10'].value_counts() #generate counts

Out[126]: 0.0      231
          26.0      85
          25.0      76
          27.0      74
          21.0      69
                  ...
          107.0      1
          91.0       1
          96.0       1
          106.0      1
          94.0       1
          Name: PM10, Length: 96, dtype: int64
```

Fig.10: PM10.0: Statistical distribution of Awka Metropolis by Python value-counts() method

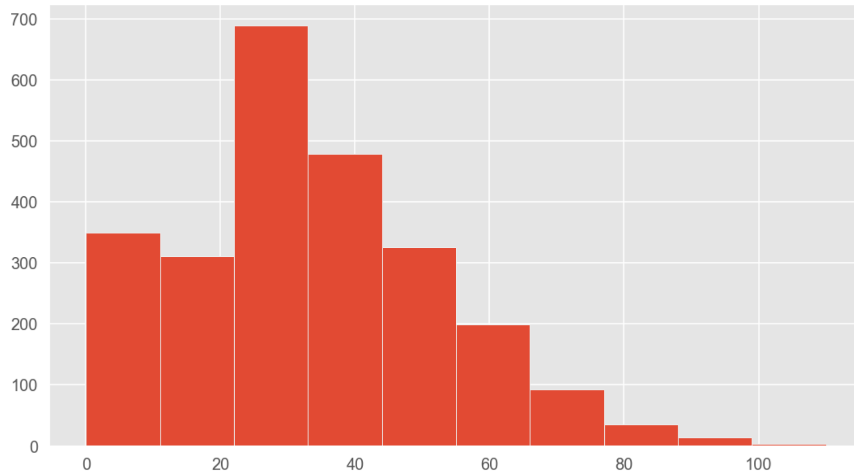

Fig.11: Histogram plot for PM10.0 Distribution in Awka Metropolis

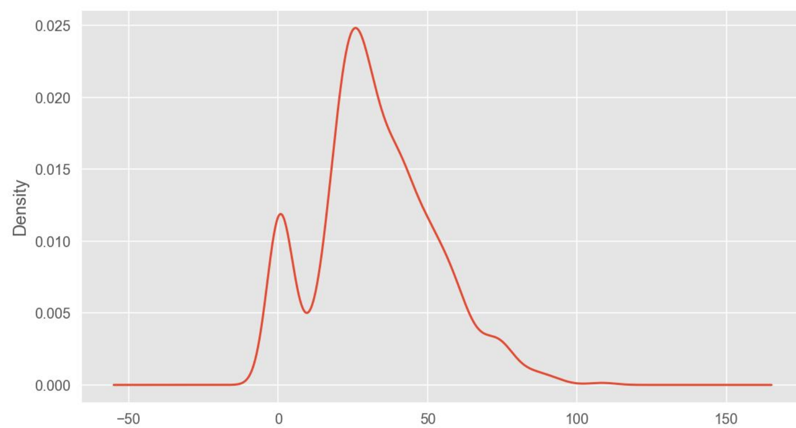

Fig.12: Density plot distribution for PM10.0 concentrations in Awka Metropolis

Table 3 presents the performance evaluation results obtained for PM10.0 Pollution prediction results using the Artificial Neural Network (ANN). Table 4 shows the performance evaluation results from modeling the PM10.0 emission in Awka Metropolis using seven (7) other machine learning Table 5 shows the printout of the actual values versus the predicted values for PM10.0 prediction using MLP ANN algorithm.

Table 3: PM10.0 pollution prediction result using MLP ANN

RMSE	MAE	R^2
1.2637	0.5885	0.9828

Table 4: PM10.0 pollution prediction modeling results with other Machine Learning algorithms

ML Algorithm	RMSE	MAE	R^2
MLR	2.3396	1.6277	0.8037
SVR	2.2782	0.9386	0.8139
Decision Tree	1.1350	0.2996	0.9538
Random Forest	0.8952	0.2996	0.9713
AdaBoost	0.9677	0.3068	0.9664
XGBoost	0.9326	0.33181	0.9688
Extra Trees	0.9087	0.3084	0.9704

Table 5: MLP ANN PM10.0 Prediction: Actual versus predicted values

	Actual-Values PM10 (ug/m3)	Predicted-Values PM10 (ug/m3)
0	25.0	25.021405
1	25.0	24.944466
2	25.0	25.051864
3	25.0	24.908995
4	25.0	25.056245
5	26.0	26.669581
6	25.0	25.069622
7	25.0	24.973060
8	25.0	25.114676
9	25.0	25.020091
10	25.0	25.028941
11	25.0	24.962883
12	25.0	26.925041
13	21.0	22.073756
14	25.0	24.977969
15	25.0	24.988212
16	25.0	24.991902
17	25.0	25.018644
18	25.0	25.029647
19	16.0	16.945953

Fig. 13 shows the regression plot result of $PM_{10.0}$ concentrations prediction using MLP Artificial Neural Network (ANN) algorithm and the following results were obtained: RMSE= 1.2637 $\mu g/m^3$, MAE= 0.5885 $\mu g/m^3$ and R^2= 0.9828 or 98.28% prediction accuracy.

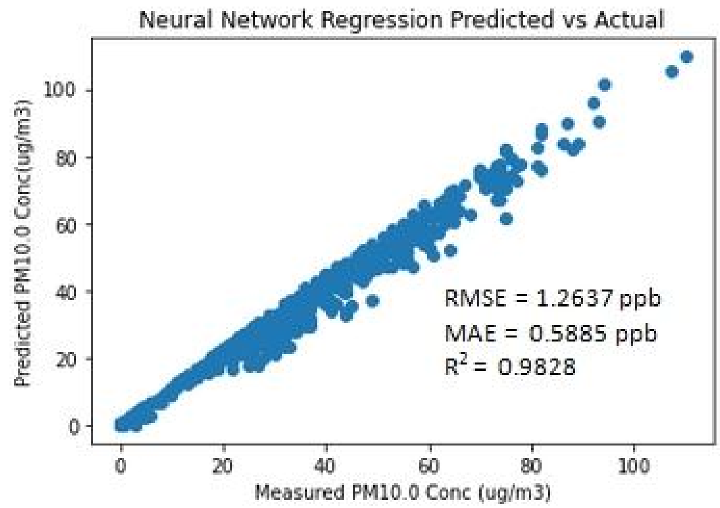

Fig.13: Regression scatterplot of PM10.0 pollution using MLP ANN algorithm

IV. CONCLUSIONS

Air pollution or air quality deterioration is a major challenge in the management of Mega cities or Metropolis. Smart city solutions such as artificial intelligence (AI) based machine learning algorithms can be deployed in the mega cities to estimate the future concentrations of air pollutant parameters present in a city or metropolis in advance before it occurs. This is important so as to alert city administrators and the general public to know the implications of the environment in order to protect their health. This paper has demonstrated that it is possible to use machine learning algorithms to model and predict air quality of a city or metropolis in advance with time .

In this work, the particulate matter PM10.0 prediction modeling were carried out using Artificial Neural Network (ANN) and eight other machine learning models using the Train-Test Data Split method. The prediction performances of these machine learning models were evaluated using statistical performance metrics such as MAE, RMSE and R^2. After more than 100 experimental runs, Artificial Neural Network (ANN) algorithm using Multi Layer Perceptron (MLP) with 200 hidden neurons obtained an RMSE value of 1.2637 $\mu g/m^3$, MAE of 0.5885 $\mu g/m^3$ and R^2 of 0.9828 or 98.28%, which is a high accurate prediction of the measured value. Other machine learning algorithms we tried on the training and testing data gave different results of RMSE, MAE and $R^{2.}$ This shows that apart from ANN, other machine learning algorithms can accurately predict or forecast in advance time the levels of dispersion of air pollutants such as PM10.0 in the ambient environment of a particular city or location.

From the experimental results obtained, it is very clear the Artificial Neural Network (ANN) is very useful in the estimation and prediction of future air pollutant concentrations within a particular city or location. Other methods can be used also to improve the prediction modeling results of the air pollutants concentrations within a particular location or city.

Conflict of interest

The authors declare that there is no conflicting interest in the publication of this paper.

REFERENCES

Aljanabi, M., Shkoukani, M., Hijjawi, M. (2020). “Ground-level Ozone Prediction Using Machine Learning Techniques: A Case Study in Amman, Jordan”. International Journal of Automation and Computing, 1-11. https://doi.org/10.1007/s11633-020-1233-4

Allam, Z. (2019). Achieving Neuroplascticity in Artificial Neural Networks through Smart Cities. Smart Cities 2019, 2, 118-134; doi:10.3390/smartcities2020009

Aceves-Fernández, M.A., Domínguez-Guevara, R., Pedraza-Ortega, J.C. and Vargas- Soto, J.E. (2020). *Evaluation of Key Parameters Using Deep Convolutional Neural Networks for Airborne Pollution (PM10) Prediction*. Hindawi. Discrete Dynamics in Nature and Society, Volume 2020, February 2020. PP 1-14. https://doi.org/10.1155/2020/2792481

Beyza Ecem Erce(2023)."Basic Structure of Artificial Neural Networks". Data Science Earth March 10, 2023. Accessed online at https://medium.com/datasciencearth/basic-structure-of-artificial-neural-networks-9aef29df9d

Byju's (2022). "Neurons". Accessed online at https://byjus.com/biology/important-diagrams-for-cbse-class-8/

Cortina-Januchs, M.G., Quintanilla-Dominguez, G., Vega-Corona, A., and Andina, O. (2015). Development of a model for forecasting of PM10 concentrations in Salamanca Mexico. Atmospheric Pollution Research 6(2015).pp.626-634.

IQAir (2023)..”World's most polluted countries & regions (historical data 2018-2021”. Retrieved online at https://www.iqair.com/world-most-polluted-countries”.

Goulier Laura, Bastian Paas, Laura Ehrnsperger and Otto Klemm (2020). "Modelling of Urban Air Pollutant Concentrations with Artificial Neural Networks Using Novel Input Variables". International Journal of Research on Public Health 2020,17, 2025; doi:10.3390/ijerph17062025

Khoshand, A., Sehrani, M.S., Kamalan, H. and Bodaghpour, S. (2017). *Prediction of Ground-Level Air Pollution Using Artificial Neural Network in Tehran.* Anthropogenic Pollution, Vol. 1(1), 2017, pp.61- 67. DOI: 10.22034/apj.2017.1.1.6167

UNICEF (2021). “UNICEF for every child, Nigeria” Retrieved online at https://www.unicef.org/nigeria/press-releases/nigeria-has-highest-number-air-pollution-related-child-pneumonia-deaths-world