sciendo International Journal of Mathematics and Computer in Engineering 1(1) (2023) 67–78



International Journal of Mathematics and Computer in Engineering

https://sciendo.com/journal/IJMCE

Original Study

Automated neural network forecast of $PM_{2.5}$ concentration

Kerehalli Vinayaka Prasad[1], Hanumesh Vaidya[1†], Choudhari Rajashekhar[2], Kumar Swamy Karekal[3], Renuka Sali[3]

[1]Department of Studies in Mathematics, Vijayanagara Sri Krishnadevaraya University, Ballari, Karnataka, India

[2]Department of Mathematics, Manipal Institute of Technology Bengaluru, Manipal, Karnataka, India

[3]Department of Studies in Computer Science, Vijayanagara Sri Krishnadevaraya University, Ballari, Karnataka, India

Communicated by Haci Mehmet Baskonus; Received: 23.04.2023; Accepted: 18.05.2023; Online: 09.06.2023

Abstract

Particulate Matter 2.5 ($PM_{2.5}$) is a major contributor to air pollution and its exposure has substantial health consequences. As a result, precise prediction of $PM_{2.5}$ concentrations is required in order to establish emission reduction strategies for air quality management. The article presents an Artificial Neural Network (ANN) model to forecast $PM_{2.5}$ levels in a particular region. The model uses data such as air temperature, carbon monoxide, nitric oxide, nitrogen dioxide, ozone, suspended particles, rainfall, relative humidity, sulfur dioxide, wind direction and wind speed to predict $PM_{2.5}$ concentrations in the air accurately. The model's efficacy is evaluated using statistical measures such as the Coefficient of Determination, the Root Mean Squared Error and the Mean Absolute Error. The study results indicate that the ANN model outperforms more traditional statistical models, with R^2 values of 0.987, which is higher than the values achieved by the Linear Regression and Decision Tree Regressor models, which are 0.88 and 0.89 respectively. The study's findings have significant implications for public health and environmental policy, as they can provide more accurate and rapid statistics on air quality. The ability to forecast $PM_{2.5}$ concentrations can help policymakers and health professionals take proactive measures to mitigate the impact of air pollution on public health.

Keywords: Artificial neural network, particulate matter 2.5, root mean square error, mean absolute error, coefficient of determination.
AMS 2020 codes: 68T07; 62G32; 62M10; 74L10; 37M10.

1 Introduction

Particulate matter (PM) is one of the most dangerous contaminants in the world's air and a major environmental problem. PM refers to microscopic particles of solid or liquid matter suspended in the air that can enter deep into the airways and cause respiratory and cardiovascular issues; it is one of the most dangerous forms of air pollution. PM comes in two sizes: PM_{10} and $PM_{2.5}$. PM_{10} particles have a diameter of 10 micrometers or less. PM with a width of 2.5 micrometers or less, known as $PM_{2.5}$, is especially harmful because it can travel through the body and damage vital systems. Zhao et al. [1] have carried out a study to envisage the characterization of spatial and seasonal concentration variation for $PM_{2.5}$ and its chemical composition and their interrelations between different compositions. They evaluated the primary and secondary sources of $PM_{2.5}$

†Corresponding author.
Email address: hanumeshvaidya@gmail.com

sciendo
ISSN 2956-7068
doi:10.2478/ijmce-2023-0005

Open Access.© 2023 Vaidya et al., published by Sciendo.
This work is licensed under the Creative Commons Attribution alone 4.0 License.

and analyzed variations of $PM_{2.5}$ chemical compositions. The study revealed that Motor vehicle exhausts and coal combustion emissions both played important roles in $PM_{2.5}$ pollution. The effects of $PM_{2.5}$ particulates on the pulmonary system of humans were studied by Xing et al. [2] through clinical, laboratory and mechanistic investigations. A pollution-related health indicator was also proposed and the public was encouraged to reduce their exposure to polluted air. To reduce the health effects of $PM_{2.5}$, the World Health Organization (WHO) recently established initial interim standards for $PM_{2.5}$, as reported in the WHO global air quality guidelines-2021 (ISBN 978-92-4-003422-8:electronic version), with a daily average threshold of 75 g /m3 and an annual upper limit of 35 g /m3. Grantz et al. [3] describes how PM can harm ecosystems such as plants, soil and water via deposition and subsequent uptake by plants, or by deposition into water, where it can alter water quality and clarity. The impacts of atmospheric PM on human health and ecosystems have become a topic of interest in recent years. As a result of these factors, there is an increasing interest in researching the formation, evolution and potential management strategies of PM. So, a precise forecast of $PM_{2.5}$ levels protects public health and reduces pollution. Artificial neural networks (ANNs) have been widely used in $PM_{2.5}$ predictions due to their impressive ability to model complex relationships between input and output variables. $PM_{2.5}$ forecasts at any time of day in central Santiago, Chile, were proposed using multilayer neural networks, linear regression and persistence models by Perez et al. [4]. They found that neural networks did the best when compared to other types of networks. Ordieres et al. [5] developed three neural network architectures, the Multilayer Perceptron (MLP), the Radial Basis Function (RBF) and the Square Multilayer Perceptron (SMPL) to forecast $PM_{2.5}$ along the Texas and Chihuahua border with the United States. They also noted that RBF did better than the alternative models. Suppose you want to predict $PM_{2.5}$ levels in Xi'an Province, China, one day in advance. In that case, you can use the Ensemble Empirical Mode Decomposition-General Regression Neural Network (EEMD-GRNN) model proposed by Zohu et al. [6]. Experimental results show that the developed hybrid EEMD-GRNN model beats the pure GRNN model without EEMD, multiple linear regression (MLR) models, PCR models and the classic autoregressive integrated moving average (ARIMA) model. Indoor $PM_{2.5}$ and $PM_{2.5-10}$ concentrations in naturally ventilated classrooms in the Mediterranean environment were forecasted using multiple linear regression (MLR) and feedforward backpropagation (FFBP), by Elbayoumi et al. [7]. The FSR method selects the most important input factors for the best possible model structure to boost the accuracy of predictions made by FFBP neural networks. Based on the empirical data, it appears that the FFBP model is superior to the MLR model. Three models, including a recursive artificial neural network, a feedforward neural network and multiple linear regression, were created by Biancofiore et al. [8] to predict PM10 and $PM_{2.5}$ concentrations in advance in an urban region on the Adriatic shore. The recursive neural network model beats the other two methods in all tests. Furthermore, $PM_{2.5}$ can be predicted even when it is not identified by using PM_{10} concentrations as an input component in the algorithm. Chen et al. [9] discovered that the Back Propagation artificial neural network model can properly predict how the $PM_{2.5}$ mass concentration in an agricultural park will alter under various weather conditions. Ong et al. [10] suggest a deep recurrent neural network (DRNN) enhanced with a novel pre-training approach using an auto-encoder to predict $PM_{2.5}$ concentrations in Japan. For time series prediction, numerical experiments demonstrate that the proposed DRNN outperforms the canonical and state-of-the-art auto-encoder training methods. Even the suggested model outperformed the VENUS (Visual Atmospheric Environment Utility System) model in Japan. Jiang et al. [11] created a novel hybrid model called Adaptive Fuzzy Neural Network (AFNN) to assess and anticipate $PM_{2.5}$ concentrations in the Jing-Jin-Ji and Pearl River Delta areas of China for a specified number of forward time steps. The experimental findings indicated that the AFNN model outperforms state-of-the-art models. Zhang et al. [12] conducted a study to reclassify and specifically describe the most popular techniques for estimating $PM_{2.5}$ concentrations across the entire research region; to list enhancements and integrations of conventional techniques and their applications; and, finally, to compare existing $PM_{2.5}$ estimation techniques based on precision and applicability. The study's findings indicate that machine learning and artificial intelligence technology are likely to result in more accurate techniques of $PM_{2.5}$ estimate on a spatial scale that are adaptable to a wide range of research kinds. Zhang et al. [13] developed a nationwide land-use regression model (LUR) to predict the monthly average $PM_{2.5}$, PM_{10}, and NO_2 in China

from 2014 to 2016. The model demonstrated good accuracy and prediction ability, as well as high geographical resolution and low bias, and it may be employed in studies concentrating on the long-term negative impacts of air pollution. To forecast $PM_{2.5}$ concentrations that are significantly non-stationary and do not obey standard distributions, Liu et al. [14] proposed a hybrid model that includes wavelet packet decomposition (WPD), gradient boost regression tree (GBRT), linear programming boosting (LPBoost), multi-layer perceptron (MLP), and Dirichlet process mixture model (DPMM) algorithms. The results show that the suggested model performs exceptionally well in forecasting. According to Harishkumar et al. [15], a machine-learning model was proposed to forecast $PM_{2.5}$ concentrations in Taiwan and this model was shown to beat competing methods. $PM_{2.5}$ emission levels across Delhi were predicted using an Extra-Trees regression with the AdaBoost hybrid model, as detailed by Kumar et al. [16]. Their model performed better than Long Short-Term Memory (LSTM) and the eXtreme Gradient Boosting (XGBoost). Recurrent Neural Network-Community Multiscale Air Quality (RNN-CMAQ l) was proposed by Chang-Hoi et al. [17] as a statistical-dynamical hybrid forecast method for predicting $PM_{2.5}$ concentrations in the Seoul metropolitan region with up to two days' notice. Additionally, they showed that the RNN model performed better than the current CMAQ prediction model. In this article, we present our findings from an investigation into the performance of an ANN in forecasting $PM_{2.5}$ concentration and we evaluate these findings in light of those obtained by other models. Studies examining the efficacy of using ANNs for $PM_{2.5}$ predictions have been conducted repeatedly and consistently with positive results. The following section briefly reviews the methodology and data collection. Section 2 gives the methodology. Section 3 describes model development and discusses the three different models compared: Linear Regression, decision tree regressor and ANN. Section 4 gives the result of model fitting and prediction. Finally, section 5 concludes the paper.

2 Methodology

2.1 Data collection

Eight synoptic meteorological sites in various towns across Taiwan Air Quality Monitoring stations (TAQMN) gather $PM_{2.5}$ concentration and other meteorological data hourly. Lunbei, Taixi, Annan, Taiman, Douliu, Xingang, Puzi and Newport are some of these stops. From 2012 to 2017, statistics were gathered by the hour and then aggregated into daily and annual totals. Table 1 shows the maximum levels of PM recorded at meteorological sites across Taiwan.

Table 1 Average upper limit of $PM_{2.5}$ in eight different stations of Taiwan.

Cities	Lunbei	Taixi	Annan	Taiman	Douliu	Xingang	Puzi	Newport
Avg. Upper limit of $PM_{2.5}$	132 $\mu g/m^3$	100 $\mu g/m^3$	101 $\mu g/m^3$	104 $\mu g/m^3$	150 $\mu g/m^3$	126 $\mu g/m^3$	108 $\mu g/m^3$	161 $\mu g/m^3$

Table 1 demonstrates that the upper limit of $PM_{2.5}$ in all Taiwanese cities is significantly higher than the WHO intermediate criteria. With Newport having the highest upper limit of $PM_{2.5}$ among them, raising major concerns about air pollution. Therefore, this study mainly concentrates on meteorological data obtained from Newport synoptic station situated in Taiwan (23^0 $28'$ $45''$ N, 120^0 $26'$ $56''$).

The sample subset of the TAQMN meteorological data about Newport city is summarised in Table 2. The meteorological data included Date, Atmospheric temperature (AMB), carbon monoxide (CO), nitric oxide (NO), nitrogen dioxide (NO_2), nitrogen oxide (NOX), ozone (O3), suspended particles ($PM_{2.5}$), Rainfall (RAINFALL), relative humidity (RH), sulfur dioxide (SO_2), wind direction hourly value (WD-HR), Wind direction (WIND-

D), Wind Speed (WIND-S), hourly wind speed value (WS-HR) and suspended particles $PM_{2.5}$. The below figure intuitively shows the concentration of $PM_{2.5}$, PM_{10}, NO_2 and O_3 in the meteorological data gathered from TAQMN data of Newport city from the year 2012 to 2017.

Table 2 Sample TAQMN Dataset pertaining to Newport.

DATE	AMB	CO	NO	NO2	NOX	O3	PM10	RAINFALL	RH	SO3	WD-HR	WIND-D	WIND-S	WS-HR	RM2.5
01.01.2012	16.20	0.35	1.84	9.50	11.37	32.75	68.33	42.75	70.16	2.99	192.60	210.41	3.33	0.88	42.75
02.01.2012	15.45	0.51	2.22	13.98	16.23	27.60	97.79	70.58	70.95	4.10	227.04	208.25	2.25	0.95	70.58
03.01.2012	16.37	0.49	5.53	16.57	22.12	25.37	81.33	56.25	73.87	4.09	250.00	240.87	2.14	1.17	56.25
04.01.2012	12.62	0.53	2.02	13.54	15.57	20.79	49.87	36.25	70.08	4.54	198.45	173.83	4.77	0.82	36.25
05.01.2012	12.00	0.45	3.82	17.78	21.58	11.88	35.70	24.96	80.66	3.01	158.46	154.70	2.59	0.83	30.75
06-01-2012	14.95	0.38	5.50	15.57	21.04	14.48	23.33	21.30	84.66	2.43	144.64	166.87	3.18	0.62	22.03
07-01-2012	16.41	0.32	1.70	9.62	11.45	25.75	32.33	18.45	79.58	2.62	127.58	149.95	3.14	0.73	18.45
08-01-2012	17.70	0.34	2.13	9.20	11.24	25.53	39.91	26.50	77.95	2.05	179.27	184.16	3.37	1.01	26.50
09-01-2012	17.75	0.33	1.49	10.46	12.00	26.75	36.95	22.37	77.75	2.90	186.78	180.98	3.63	0.93	22.37
10-01-2012	17.29	0.42	2.21	12.77	15.00	26.57	54.16	35.70	75.04	3.52	165.22	171.70	3.42	0.85	35.70
11-01-2012	16.75	0.47	1.79	10.86	12.75	29.79	77.79	50.91	67.91	3.64	193.45	161.33	4.34	0.69	50.91
12-01-2012	16.54	0.59	1.85	12.81	14.64	28.87	67.87	54.29	72.12	4.01	206.91	190.05	2.42	0.87	54.29
13-01-2012	17.50	0.40	2.09	13.49	15.57	19.12	47.50	24.62	84.79	3.37	236.25	213.20	2.30	0.88	27.08
14-01-2012	18.25	0.47	3.74	16.70	20.50	14.50	56.00	37.45	85.41	2.01	229.57	246.58	1.81	0.98	41.12
15-01-2012	19.16	0.46	3.20	14.19	17.21	18.66	41.12	25.40	80.83	3.54	256.33	241.20	2.58	1.40	26.41
16-01-2012	15.83	0.32	2.95	10.62	13.55	24.50	20.25	11.17	82.04	2.46	178.50	156.83	3.68	1.67	15.16
17-01-2012	15.70	0.37	3.80	10.64	14.38	25.45	51.91	22.12	79.70	2.16	203.50	173.45	2.55	1.26	22.12
18-01-2012	18.20	0.45	3.96	15.48	19.51	25.17	68.12	39.12	78.45	3.02	218.27	197.62	1.80	1.18	39.12
19-01-2012	18.25	0.48	3.33	17.14	20.37	19.74	65.62	44.29	80.12	2.90	233.36	204.81	2.80	1.15	44.29
20-01-2012	17.33	0.59	2.05	14.05	16.12	27.88	71.62	44.29	80.04	2.88	188.50	172.45	2.29	1.30	44.29
21-01-2012	17.29	0.70	1.87	13.42	15.29	27.75	82.29	50.41	79.62	2.67	232.91	191.70	2.32	1.33	50.41
22-01-2012	14.62	0.45	1.01	7.78	8.77	25.75	21.12	15.70	80.66	1.57	195.62	187.70	3.99	0.80	19.83

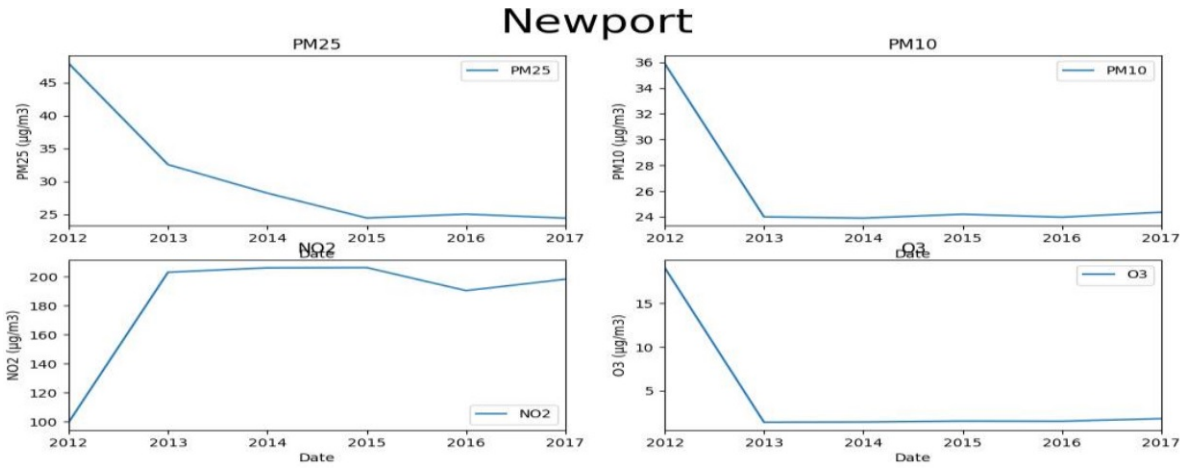

Fig. 1 Concentration of $PM_{2.5}$, PM_{10}, NO_2 and O_3 from 2012 to 2017.

3 Model development

3.1 Artificial neural network

ANNs are algorithmic models that are inspired by the structure and behavior of neurons in the human brain. Like human brains, ANNs comprise many layers of interconnected neurons.

The term "nodes" is used to describe these neurons. Multilayer neural networks (ANNs) consist of an input layer, a hidden layer(s) and an output layer. Each manufactured neuron, or node, is connected to the others and has its own relative importance and sensitivity. There is a level at which a node's output is enabled and it begins sending information to the next network layer. Not doing so prevents the information from moving to the next network tier [18]. The effectiveness of a neural network can be tweaked with various factors and hyperparameters. These elements significantly impact the results that ANNs produce. You can adjust things like weights, biases, learning rate and group size. Using Python's TensorFlow framework, a supervised ANN model is built with one input layer, two intermediate layers and one final output layer. The input layer's number of neurons equals the number of features in the data being processed. The output layer has the same number of neurons as the number of outputs associated with each input. So, fifteen neurons represented the weather factors in the input layer and one represented $PM_{2.5}$ in the output layer. Furthermore, the number of neurons and layers required for the hidden layer are determined by the number of training cases, the complexity of the data to be learned and the type of activation functions used. Less neurons will result in underfitting and significant statistical bias, whereas too many neurons will result in overfitting, high variance and increased training time. So, to balance the model an ANN structure with 15-16-8-1 architecture is used in this study and is shown in figure 2. In the concealed layers, the ReLU activation function is employed; in the output layer, the linear activation function, Sigmoid is employed. The Adam optimizer algorithm is utilized during model training by setting the learning rate to 0.001. Mean squared error (MSE) was used as the loss function. Table 3 summarizes the network's parameters.

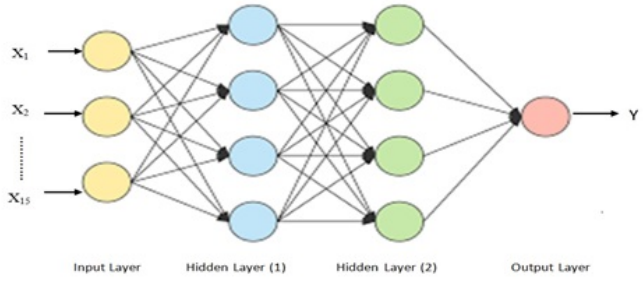

Fig. 2 ANN architecture used in this study.

Table 3 ANN Parameters.

Parameter Name	Parameter
Hidden Layer number	2
Hidden Layer Activation function	Relu
Output Layer Activation function	Sigmoid
Learning Rate	0.001
Maximum Number of Epochs to Train	150
Optimizer	Adam
Loss Function	MSE

3.2 Decision tree regressor

A well-known example of a machine learning technique that falls under the category of supervised learning is the Decision Tree Regressor (DTR) algorithm. In addition to that, it is useful for tasks involving categorization [19] and prediction. The fundamental concept behind DTR is to find a solution to the problem that has been presented by employing a tree representation, in which the base component of the tree stands for the attribute. As depicted in figure 3 the central component symbolizes an inspection or evaluation of the property. The branch represents quality and the internal node displays the assessment outcome. The outcome node, which indicates the final judgment and symbolizes the classes, is sometimes called the leaf node in [20].

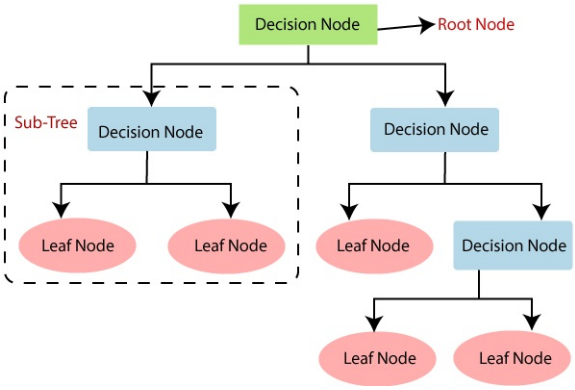

Fig. 3 Decision Tree Regressor.

3.3 Linear regression

One of the oldest and most popular forms of Machine Learning is linear regression. It's a method of predictive analysis that relies on statistics. The linear regression formula illustrates a straight-line relationship between an independent (X) and a dependent (Y) variable as shown in figure 4. Linear regression establishes a linear relationship between two variables to learn how one changes in response to a shift in the other [21].

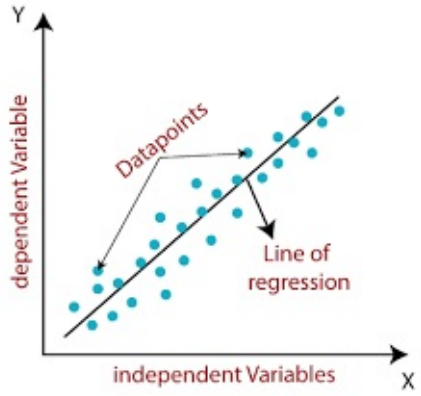

Fig. 4 Linear Regression.

Eighty percent of the data is used for training the algorithm and twenty percent is used for assessment. The training set is used to determine the optimal weights for the ANN model, the validation set is used to choose the model variant that offers the highest degree of generalization and the test set is used to assess the model's performance on unseen data.

4 Results and discussion

The performance of the proposed $PM_{2.5}$ prediction system is evaluated with different metrics in this section by running several experiments on the Python working platform.

4.1 Performance criteria

Coefficient of Determination (R2), Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) are some of the statistical measures used to assess the model's efficacy Chicco et al. [22]. The coefficient of determination, also known as R-squared (R2), is a statistical metric that measures the amount of variance in the dependent variable (the variable being predicted) in a regression model that can be explained by the independent variable(s) (the predictors). R-squared, in other words, reflects how well the regression line (or curve) matches the data. R-Squared values range from 0 to 1. Formula for determining the R2 is provided below.

$$R^2 = \left[\frac{1}{M\sigma_x\sigma_y}\Sigma_{j=1}^{M}[(Y_j - \bar{Y})(X_j - \bar{X})]\right], \tag{1}$$

where, M is the number of observations, σ_x is the standard deviation of the observation X, σ_Y is the standard deviation of Y, X_j is the observed values, $\bar{X}$ is the mean of the observed values, Y_j is the calculated values and $\bar{Y}$ is the mean of the calculated values. R^2 is an important metric in regression analysis since it helps determine the model's degree of goodness of fit. A higher R-squared value indicates a better model-data fit. Because R-squared provides a useful summary of how well the model fits the data, it is still commonly used in regression analysis. Despite a high R-squared, the model may be inaccurate or fail to demonstrate a causal relationship between the predictors and the dependent variable.

RMSE: The Root Mean Square Error (RMSE) is a typical method for calculating a model's error in predicting quantitative data. It computes the average difference between the anticipated and actual values of a model. It gives an estimate of how well the model predicts the target value (accuracy). A RMSE value of 0 implies that the model is perfectly fitted. The model and its predictions perform better when the RMSE is low.

MAE: The Mean Absolute Error (MAE) is a statistic used to assess a regression model's performance. It is defined as the average absolute difference between the model's predicted values and the data's true values. MAE just measures the magnitude of mistakes and is unconcerned about their direction. The lower the MAE, the more accurate the model.

Formulas for calculating RMSE and MAE are depicted below

$$RMSE = \sqrt{\frac{1}{n}\Sigma_{i=1}^{n}(X_i - \bar{Y})^2}, \tag{2}$$

and

$$MAE = \frac{1}{n}|X_i - Y_i|, \tag{3}$$

where X_i and Y_i are the ith observed and estimated values, respectively; $\bar{X}$ and $\bar{Y}$ are the average of X_i and Y_i and n is the total number of data.

4.2 Performance analysis

Using the aforementioned statistical measures, we compare the ANN model's success to that of LR and DTR. Input factors were the same for the LR and DTR as they were for the ANN model. The R^2 for the ANN model was 0.987 percent, which is relatively high and suggests that the model is quite accurate at making predictions. The R2 value for the linear regression model was 0.88, while that for the Decision Tree Regressor model was 0.89. The efficacy of ANNs for $PM_{2.5}$ prediction is shown in Table 4, which compares the performance of the LR, DTR and ANN models in terms of R2, RMSE and MAE. The ANN model clearly beat both models. Figure 5 is a scatter diagram depicting only how well the ANN model did at forecasting the $PM_{2.5}$ concentration, in which actual values are displayed along the X-axis and predicted values along the Y-axis. Three algorithms' abilities to predict $PM_{2.5}$ accurately are compared in Figure 6. The first two images display the observed and forecasted $PM_{2.5}$ concentrations for the training and testing sets used by the ANN model. Similarly, the following two images pertain to the DTR model, while the following two pertain to the LR model. In addition, the $PM_{2.5}$ predictions are in blue, while the observed numbers are shown in red. The best model from Table 4 is compared in Table 5 with other similar studies, such as Gradient Boosting Regressor [15].

Table 4 Comparison between LR, DTR and ANN.

Models	R^2	RMSE	MAE
Linear Regression (LR)	0.88	0.223	0.018
Decision Tree Regressor (DTR)	0.89	0.123	0.116
ANN	0.987	0.021	0.014

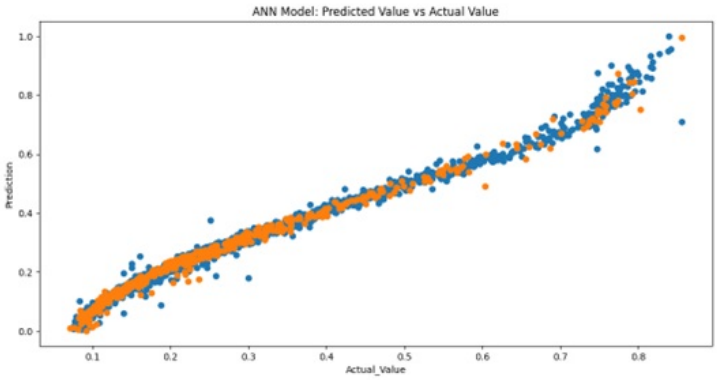

Fig. 5 $PM_{2.5}$ concentration prediction of ANN.

Table 5 Comparison results for the proposed model with the other models for forecasting of the $PM_{2.5}$.

Performance Metrics	Proposed Model	Harish et al. [16]
Linear Regression (LR)	0.88	0.223
Decision Tree Regressor (DTR)	0.89	0.123
ANN	0.987	0.021

sciendo

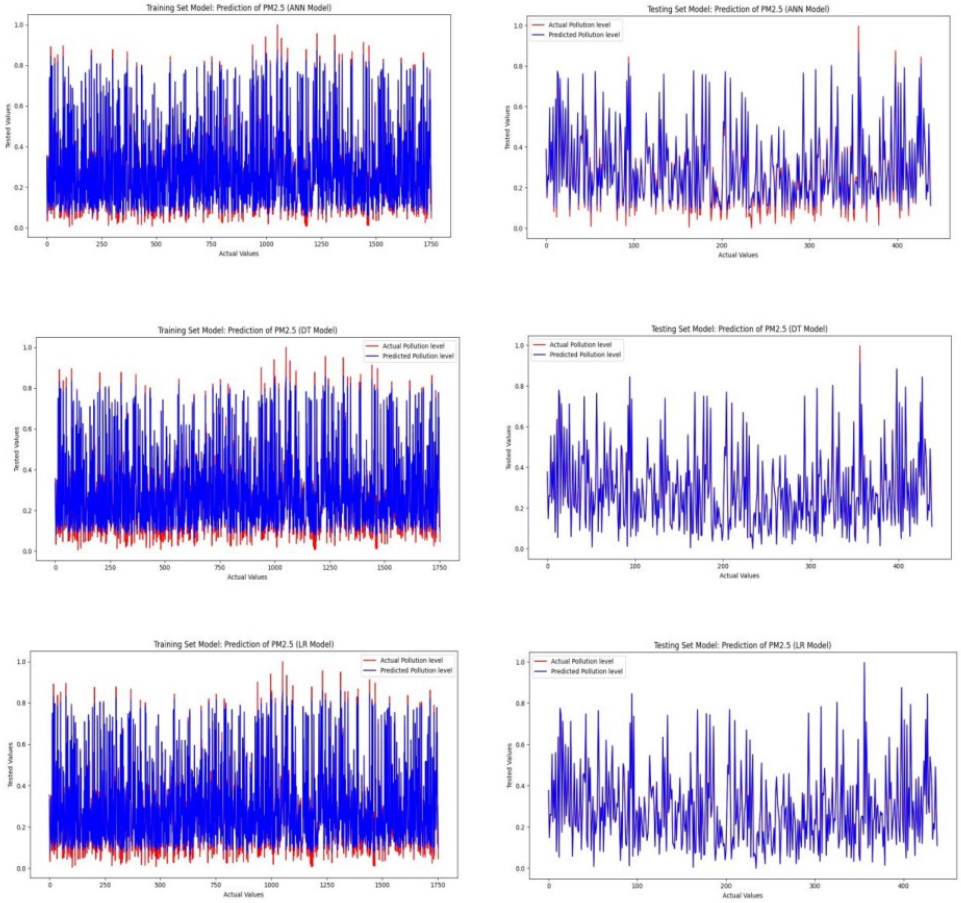

Fig. 6 $PM_{2.5}$ concentration prediction of ANN, DTR and LR model for training and testing set separately.

The studys ANN model utilises a wide range of input data, including meteorological and pollutant factors for forecasting $PM_{2.5}$ concentrations in a given area.The proposed model uses the standard performance metrics such as R-squared, RMSE and MAE which allows for a standardized evaluation of different models and facilitates comparison between them. This can help researchers to identify the best-performing models and to improve the state-of-the-art in the field. In addition, they provide interpretable results that help to understand the strengths and weaknesses of different models and guide us for the selection of the most appropriate model for a given problem and facilitate its deployment in practical applications. Further, the use of these performance metrics allows for a reproducible evaluation of the models, as the same metrics can be used to evaluate the models across different experiments or by different researchers. The results of the study reveal the ANN model's advantage over more conventional statistical models in terms of its capacity to forecast $PM_{2.5}$ concentrations with high precision and accuracy. The ANN model's high R2 value indicates that it has the potential to be a trustworthy tool for forecasting air quality and informing public health and environmental policy decisions. However, the limitation of this study is that it uses specific dataset for evaluation which may limit the generalization of the results to other datasets or real-world scenarios. Therefore, it is important to evaluate the models on multiple datasets and to consider their performance in practical applications.

5 Conclusion

Our study utilized ANN to develop a model for forecasting $PM_{2.5}$ concentration from meteorological data. The results indicate that the ANN model outperformed standard statistical models in accurately predicting $PM_{2.5}$ concentrations with an R2 value of 0.987, RMSE value of 0.021 and, MAE value of 0.014. These more precise predictions can potentially mitigate the adverse effects of air pollution on human health. ANNs could serve as a valuable tool for policy makers and stakeholders in making informed decisions related to air quality management. Further, the data provided here demonstrate the significance of this research and its usefulness as an early warning system for enhanced urban air quality management. We can acquire more insight into the intricate interaction between meteorological variables and $PM_{2.5}$ concentrations and create more potent air quality management plans by utilizing the ability of machine learning algorithms like ANNs.

6 Declarations

6.1 Conflict of interest:

The authors hereby declare that there is no conflict of interests regarding the publication of this paper.

6.2 Funding:

Not applicable.

6.3 Authors Contribution:

K.V.P.-Conceptualization, Investigation, Methodology, Modeling; H.V.-Supervision, Conceptualization, Methodology, Writing-original draft, review and editing; C.R.-Methodology, Review and editing; K.S.K.-Methodology, Writing-original draft, Formal analysis; R.S.-Methodology, Writing, Formal analysis.

6.4 Acknowledgement:

The authors sincerely appreciate the reviewers' insightful remarks, which resulted in a significant improvement in the manuscript.

6.5 Data availability statement:

All data that support the findings of this study are included within the article (and any supplementary files).

References

[1] Zhao P.S., Dong F., He D., Zhao X.J., Zhang X.L., Zhang W.Z., Yao Q., Liu H.Y., Characteristics of concentrations and chemical compositions for $PM_{2.5}$ in the region of Beijing, Tianjin, and Hebei, China, Atmospheric Chemistry and Physics, 13(9), 4631-4644, 2013.

[2] Xing Y.F., Xu Y.H., Shi M.H., Lian Y.X., The impact of $PM_{2.5}$ on the human respiratory system, Journal of thoracic disease, 8(1), E69-E74, 2016.

[3] Grantz D.A., Garner J.H.B., Johnson D.W., Ecological effects of particulate matter, Environment international, 29(2-3), 213-239, 2003.

[4] Pérez P., Trier A., Reyes J., Prediction of $PM_{2.5}$ concentrations several hours in advance using neural networks in Santiago, Chile, Atmospheric Environment, 34(8), 1189-1196, 2000.

[5] Ordieres J.B., Vergara E.P., Capuz R.S., Salazar R.E., Neural network prediction model for fine particulate matter ($PM_{2.5}$) on the USMexico border in El Paso (Texas) and Ciudad Juárez (Chihuahua), Environmental Modelling Software, 20(5), 547-559, 2005.

[6] Zhou Q., Jiang H., Wang J., Zhou J., A hybrid model for $PM_{2.5}$ forecasting based on ensemble empirical mode decomposition and a general regression neural network, Science of the Total Environment, 496, 264-274, 2014.

[7] Elbayoumi M., Ramli N.A., Yusof N.F.F., Development and comparison of regression models and feedforward backpropagation neural network models to predict seasonal indoor $PM_{2.5-10}$ and $PM_{2.5}$ concentrations in naturally ventilated schools, Atmospheric Pollution Research, 6(6), 1013-1023, 2015.
[8] Biancofiore F., Busilacchio M., Verdecchia M., Tomassetti B., Aruffo E., Bianco S., Tommaso D.S., Colangeli C., Rosatelli G., Carlo D.P., Recursive neural network model for analysis and forecast of PM_{10} and $PM_{2.5}$, Atmospheric Pollution Research, 8(4), 652-659, 2017.
[9] Chen B., Wang X., Yu L., Wang H., Li Y., Chen J., Chen B., Wang X., Yu L., Wang H., Li Y., Chen J., Zhu J., Nan H., Hou L., Prediction of $PM_{2.5}$ concentration in a agricultural park based on BP artificial neural network, Advance Journal of Food Science and Technology, 11(4), 274-280, 2016.
[10] Ong B.T., Sugiura K., Zettsu K., Dynamically pre-trained deep recurrent neural networks using environmental monitoring data for predicting $PM_{2.5}$, Neural Computing and Applications, 27, 1553-1566, 2016.
[11] Jiang P., Dong Q., Li P., A novel hybrid strategy for $PM_{2.5}$ concentration analysis and prediction, Journal of environmental management, 196, 443-457, 2017.
[12] Zhang G., Rui X., Fan Y., Critical review of methods to estimate $PM_{2.5}$ concentrations within specified research region, ISPRS International Journal of Geo-Information, 7(9), 368, 2018.
[13] Zhang Z., Wang J., Hart J.E., Laden F., Zhao C., Li T., Zheng P., Li D., Ye Z., Chen K., National scale spatiotemporal land-use regression model for $PM_{2.5}$, PM_{10} and NO_2 concentration in China, Atmospheric environment, 192, 48-54, 2018.
[14] Liu H., Duan Z., Chen C., A hybrid framework for forecasting $PM_{2.5}$ concentrations using multi-step deterministic and probabilistic strategy, Air Quality Atmosphere and Health, 12, 785-795, 2019.
[15] Doreswamy, Harishkumar K.S., Yogesh K.M., Gad I., Forecasting air pollution particulate matter ($PM_{2.5}$) using machine learning regression models, Procedia Computer Science, 171, 2057-2066, 2020.
[16] Kumar S., Mishra S., Singh S.K., A machine learning-based model to estimate $PM_{2.5}$ concentration levels in Delhi's atmosphere, Heliyon, e05618, 2020.
[17] Chang-Hoi H., Park I., Oh H.R., Gim H.J., Hur S.K., Kim J., Choi D.R., Development of a $PM_{2.5}$ prediction model using a recurrent neural network algorithm for the Seoul metropolitan area Republic of Korea, Atmospheric Environment, 245, 118021, 2021.
[18] Krogh A., What are artificial neural networks?, Nature biotechnology, 26, 195-197, 2008.
[19] Prasad K.V., Vaidya H., Swamy K.K., Renuka S., Pumpkin Seeds Classification: Artificial Neural Network and Machine Learning Methods, Journal of International Academy of Physical Sciences, 27(1), 22-23, 2023.
[20] Song Y.Y., Lu Y., Decision tree methods: applications for classification and prediction, Shanghai Arch Psychiatry, 27(2), 130, 2015.
[21] Maulud D., Abdulazeez A.M., A review on linear regression comprehensive in machine learning, Journal of Applied Science and Technology Trends, 1(4), 140-147, 2020.
[22] Chicco D., Warrens M.J., Jurman G., The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation, PeerJ Computer Science, 7, e623, 2021.