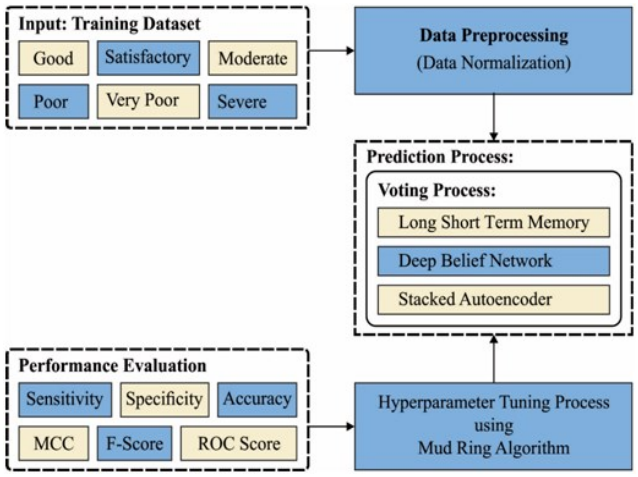
Global NEST Journal, Vol 25, No 6, pp 100-108

Air quality prediction using ensemble voting based deep learning with mud ring algorithm for intelligent transportation systems

Sivanesh S.[1,*], Mani G.[2], Venkatraman S.[3] and Nandhini R.[4]
[1]Department of computer science engineering, BIT Campus, Anna University, Tiruchirappalli,
[2]Department of Computer Science and Engineering, University College of Engineering, Kanchipuram-631 552
[3]School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, India
[4]Department of Computer Science and Engineering, K.Ramakrishnan College of Engineering, Trichy
Received: 15/02/2023, Accepted: 02/04/2023, Available online: 06/04/2023
*to whom all correspondence should be addressed: e-mail: sivaneshs1989@gmail.com
https://doi.org/10.30955/gnj.004810

Graphical abstract



Abstract

In recent times, advanced technologies in transportation are developing like connected and automated vehicles and shared mobility services. A rapidly increasing number of vehicles in intelligent transportation system (ITS) and smart cities causes pollution and degrade the quality of air. Owing to the incredible effect of air quality on individual lives, it is indispensable to design a system by which air pollutants (PM2.5, NOx, COx, SOx) are predicted. But predicting air quality and its pollutants was complex since air quality relies on various elements like power plants, weather, and vehicular emissions. Deep learning (DL) and Machine learning (ML) techniques are leveraged for developing an air quality predictive method. This study develops an Air Quality Prediction utilizing Ensemble Voting based Deep Learning with Mud Ring Algorithm (AQP-EDLMRA) technique. The presented AQP-EDLMRA technique follows the ensemble voting model, which exploits three DL classification methods like long short-term memory (LSTM), deep belief network (DBN), and stacked autoencoder (SAE). Then, the new data can be classified by the weighted vote of their prediction outcomes. To adjust the hyperparameter values of the DL methods, the MRA was exploited, showing the novelty of the work. The simulation values of the AQP-EDLMRA approach are tested using a series of air quality data and the comprehensive comparative results demonstrated that the AQP-EDLMRA technique has reached improved forecasting performance.

Keywords: Intelligent transportation system, air quality, deep learning, ensemble models, mud ring algorithm, normalization, pollution monitoring

1. Introduction

Intelligent Transportation Systems (ITS) have emerged as a vital component to improve human life and the modern economy, to optimize road traffic by handling the road capacity, enhancing driver safety, decreasing energy consumption and improving the quality of the environment, among many other things. Among several natural resources, Air is one significant natural resource for the survival of the entire life on this planet (Zhu *et al.*, 2018). The basic survival of all organisms like animals and plants relies upon the air. Hence, all living creatures require good quality air free from detrimental gases for a healthy life. As per the report of the Blacksmith Institute in 2008, the two worst pollution issues in the world were indoor air pollution and urban air quality (Ke *et al.*, 2022). The rising population, industries, and vehicles were polluting the air at an alarming rate. Air pollution may cause short or long-term health disorders. It is found that young children and the elderly were most affected by air pollution (Wu and Lin, 2019). Air quality assessment will be considered a significant means for controlling and monitoring air pollution. Some air pollutants known as criteria air pollutants are commonly seen in the US. Such pollutants can be harmful and affect the health and environment. In addition to monitoring, precise air quality prediction is highly significant, which will be advantageous to governments' pollution controls and individual activity arrangements (Zeng *et al.*, 2022). If air pollution is informed to be severe in the future, an individual can postpone meetings and cancel outdoor activities to ignore exposure

to detrimental air pollution. The ruling classes make initiatives to thwart the quality of air from worsening by constraining factories of high pollution and executing traffic control (Rybarczyk and Zalakeviciute, 2018). Furthermore, it is practical to make citywide air quality forecasting, which allows flexible options like reorganizing outdoor actions to somewhere free from air pollution instead of postponing or cancelling. The main aim of this work is to forecast future air quality for city air quality monitoring stations (Cabaneros *et al.*, 2019).

The statistical prediction methods depend on statistics and leverage historic time sequence data for forecasting upcoming air quality (Xayasouk *et al.*, 2020). Through the comparison made with physical predictive techniques, statistical predictive approaches not only avoid complicated systems thereby the cost of each estimation will be reduced and the calculation is faster, but even attain an equivalent level of PM2.5 concentration estimation precision of physical predictive methods. Owing to such benefits, the implementation of statistical predictive techniques was more extensive (Tao *et al.*, 2019). Typically employed statistical predictive techniques involved artificial neural network (ANN) and machine learning (ML) methods. ML techniques have clear mathematical logic, where the relation between output and input was relatively definite, and structures were simple. But such methods avoid the influences of spatial factors and external variables and are vulnerable to overfitting because of improper index selection (Wang *et al.*, 2022). With the advances in artificial intelligence (AI), ANN techniques, particularly DL techniques, which consider the non-linear relationships among the external variables and prediction targets, have outstanding performance in predictive tasks due to their robustness and adaptiveness.

With the rapid growth of DL and AI techniques, the method performance of conventional ML shallow neural networks no longer exists. Different types of DL techniques were devised for enhancing the predictive performance of air quality (Mao *et al.*, 2021). In addition, the existing models does not focus on the hyperparameter selection process which mainly influence the performance of the classification model. Since the trial and error method for hyperparameter tuning is a tedious and erroneous process, metaheuristic algorithms can be applied. Therefore, in this work, we employ metaheuristic algorithm for the parameter selection of the DL model.

This study develops an Air Quality Prediction utilizing Ensemble Voting based Deep Learning with Mud Ring Algorithm (AQP-EDLMRA) technique. The presented AQP-EDLMRA technique follows the ensemble voting model, which exploits three DL classification methods like long short-term memory (LSTM), deep belief networks (DBNs), and stacked autoencoder (SAE). Then, the new data can be classified by the weighted vote of their prediction outcomes. To adjust the hyperparameter values of the DL methods, the MRA was used in this article. The simulation values of the AQP-EDLMRA approach are tested utilizing a sequence of air quality data.

2. Related works

In (Heydari *et al.*, 2022), a novel hybrid intelligent system dependent upon LSTM and multi-verse optimization (MVO) system was established for analyzing and predicting air pollution achieved in Combined Cycle Power Plants. During the presented method, the LSTM technique has an analyst engine for predicting the count of produced NO_2 and SO_2 by the Combined Cycle Power Plant, whereas the MVO system was utilized for optimizing the LSTM parameter for achieving a lesser predictive error. Gu *et al*. (2022) examine a novel Hybrid Interpretable Predictive ML approach for Particulate Matter 2.5 predictive that takes 2 novelties. Primarily, a hybrid method infrastructure was generated with DNN and Non-linear Auto Regressive Moving averages with an Exogenous Input method. Second, the automatic feature generation and FS methods were combined as in this hybrid system. In [13], a new DL infrastructure combining multiple nested LSTM networks (MTMC-NLSTM) was presented to correct AQI predicting learned with federated learning. The presented MTMC-NLSTM systems performance was related to typical ML, DL, and hybrid DL approaches.

Chang *et al*. (2020) examine an Aggregated LSTM method (ALSTM) dependent upon the LSTM using the DL approach. During this novel approach, the authors integrate local air quality monitor places, stations for external pollution causes, and stations in neighboring industrial regions. For improving predicting accuracy, the authors combined 3 LSTM approaches to predicting systems to primary predictive dependent upon external sources of pollution and data in neighboring industrial air quality stations. Gilik *et al*. [15] establish a supervised method for forecasting air pollution by employing real sensor information and for transferring the method betwixt cities. An integration of CNN and LSTM-DNN approach has been presented for predicting the focus on air pollutants in several places of cities by utilizing spatial-temporal connections.

A novel hybrid system utilizing outlier recognition and correction techniques and a heuristic intelligent optimized system was presented in (Wang *et al.*, 2020). Primary, data pre-processed systems were conducted for detecting and correcting outliers, and excavating the essential features of a novel time series; secondary, an extremely utilized heuristic intelligent optimized system was implemented for optimizing the parameter of ELM for obtaining the predicting outcomes of all the subseries with enhancement in accuracy. Yi *et al*. (2018) introduce a DNN depending on the method (allowed DeepAir) that contains spatial transformation element and deep distributed fusion network. Assuming air pollutant spatial correlations, the former element adapts the spatial sparse air quality information as to reliable input for simulating the pollutant source. The latter network implements a neural distributed infrastructure for fuse heterogeneous city information to concurrently capture the factors that affect air quality.

3. The proposed model

Automated air quality forecasting using the AQP-EDLMRA technique has been introduced in this study. The presented AQP-EDLMRA technique employed the ensemble voting

model by the use of three DL methods namely LSTM, DBN, and SAE. In the ensemble voting model, the new data can be classified by the weighted vote of their prediction outcomes. Figure 1 illustrates the workflow of the AQP-EDLMRA system.

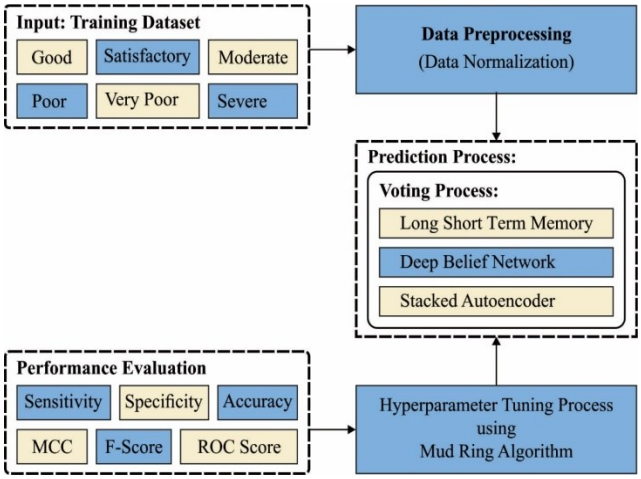


Figure 1. Workflow of AQP-EDLMRA system

3.1. Data normalization

In this work, the input data is initially pre-processed by the use of a min-max normalization process. The min-max technique is renowned as and simple normalization technique in medical imaging. Depending on this method, above and over the unifying data scales, the data changing edges are dispersed in the range between 1 and 0. By considering attribute X, it has a mapping from the dataset amongX_{max} and X_{min}, the min-max normalization (X_{norm}) is achieved by the following:

$$X_{norm} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (1)$$

3.2. Ensemble voting based prediction model

For the air quality prediction process, the ensemble voting process is performed in this study. The results of separate classification techniques were integrated through a majority vote utilizing a voting method. Class labels were allocated to test data by every technique separately after the outcomes were collected through voting, and the final class prediction was made by considering the class with a maximum number of votes. The following equation was employed to execute majority voting on a dataset:

$$\sum_{c=1}^{c} D_{c,i}(x) = \max_{i=1,2,3,\ldots n} \sum_{c=1}^{C} (D_{c,i}) \quad (2)$$

where C denotes the total number of classification models, and(c,i) indicates the classifier decision and class labels.

3.2.1. LSTM model

LSTM is a development over the RNN that was formerly applied in the analyses of EEG. In comparison with traditional RNN, the innovation of the LSTM network includes: (1) an output gate, (2) a forget gate, and (3) an input gate (Sun *et al.*, 2021).

Forget gate: It will decide what preceding data need to be forgotten. The existing input x_t and hidden state h_{t-1} from previous units were concatenated into novel vectors. Multiplier by the W_f weight parameter of the gate, all element values of the output vectors f_t were scaled from 0-1 through component-wise sigmoid function σ A $0'$ component allows the respective data in C_{t-1} to be disregarded, whereas a $1'$ shows that the respective data is allowable to be passed through.

$$f_t = \sigma\left(W_f \cdot h_{t-1}, x_t + b_f\right) \quad (3)$$

Input gate: It decides how much of input x_t is saved to unit state C_t. The fulfilment of this gate needs co-operation amongst 2 parallel layers. The tangent layer output candidate's data C_t for selection, whereas the sigmoidal layer, performs as the forget gate and determines what candidate data would be chosen by outputting the decision vector . Afterwards, the component-wise multiplication of the candidate data through decision vector $C_t \times i_t$ is implemented, and final update data which must be included in the cell state can be defined.

$$i_t = \sigma\left(W_i \cdot h_{t-1}, x_t + b_i\right) \quad (4)$$

$$\bar{C}_t = \tan\left(W_c \cdot h_{t-1}, x_t + b_c\right) \quad (5)$$

Thus, the C_t cell state of the existing chain was an integration of the reserved past data of C_{t-1}, and the updated data were chosen from C_t (Eq. 6).

$$C_t = C_{t-1} \times f_t + \bar{C}_t \times i_t \quad (6)$$

Output gate: It determined which hidden state h_t in the existing chain to output through the multiplication of decision vector)0_t by the candidate data chosen from C_t, as follows.

$$0_t = \sigma\left(W_0 \cdot h_{t-1}, x_t + b_0\right) \quad (7)$$

$$h_t = \tan\left(C_t\right) \times 0_t \quad (8)$$

Depending on the LSTM structure, context time sequence data can be learned from the feature sequence derived from the CNN and later defines the total classification.

3.2.2. DBN Model

The DBN of NN is regarded as a generative mechanism which applies a set of Boltzmann machines as fundamental components (Ali *et al.*, 2022). All layers of the DBN contain an RBM. DBN will extract the H.L. feature from the dataset slated for training to increase between-classes separation power. The training can be implemented on each layer in a supervised manner, and the backpropagation method alters the weight in the network for reducing overfitting. This study presents a DBN method well-trained through greedy layer-wise learning by stacking up RBM. The RBM focuses on a specific layer during its learning process and neglects others. In this work, a set of hidden layers as

$h=\{hl,h2,h3\}$ and the set of weight matrices is allocated as $W=\{W1,W2,W3\}$. The weight matrices among i - th and $(i+1)$ layers, can be represented as Wi, whereas the j-th hidden layers are represented as h_j:

$$p(v,h)=\frac{e^{-E(v,h)}}{\sum_{v,h}e^{-E(v,h)}} \tag{9}$$

Whereas $E(v,h)$ signifies energy function of RBM,

$$E(v,h)=-\sum_{i=1}a_i v_i-\sum_{j=1}b_j h_j-\sum_{i,j}v_i h_j W_{ij} \tag{10}$$

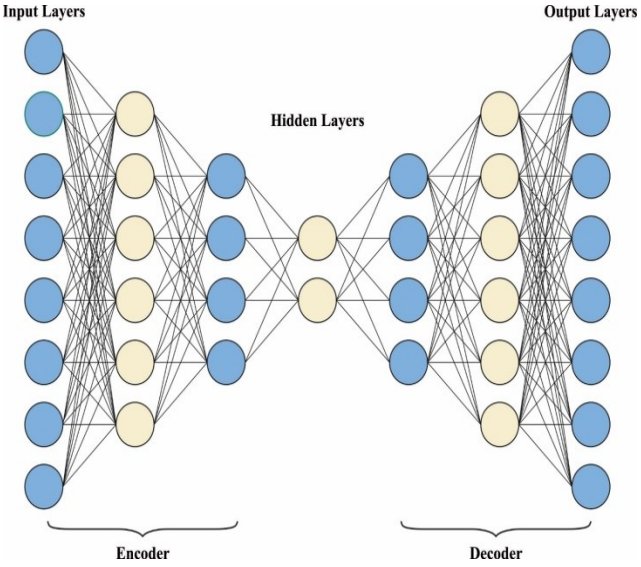


Figure 2. Architecture of DBN

W_{ij} signifies the weight among the visible layer (VL) and hidden layer (HL), a_i and b_j define the coefficients of VB and HL. This work applies the SGD technique and log-likelihood (L.L.) to achieve optimum training. This can be attained by enhancing RBM parameters, b, and wij. The derivative of log $p(v,h)$ concerning W_{ij} a_i and b_j should be calculated for upgrading the biases and weights.

$$W^{t+1}=W^t+\eta\left(p(h|v)v^T-p(h|v)v^T\right)-\lambda W^t +\alpha\Delta W^{t-1} \tag{11}$$

$$a^{t+1}=a^t+\eta(v-\tilde{v})+\alpha\Delta a^{t-1} \tag{12}$$

$$b^{t+1}=b^t+\eta\left(p(h|v)-p(\tilde{h}|\tilde{v})\right)+\alpha\Delta b^{t-1} \tag{13}$$

$p(h_j=1|v)=\sigma(\Sigma_{i=1}^{m}w_{ij}v_i+b_j)$,

$p(v_i=1|h)=\sigma(\Sigma_{j=1}^{n}w_{ij}h_j+a_i)$, and $\sigma(\square)$

Where signifies logistic sigmoid function. $\tilde{v}$) and $\tilde{h}$ represent reconstructed v and h, correspondingly. Figure 2 exhibits the framework of DBN. N denotes the hidden nodes count, η, the learning ratio, α is momentum weight, and λ signifies weight decay. The weighted matrices and accompanying bias vectors of hidden and visible nodes were studied through persistent contrastive divergence (PCD) and contrastive divergence (CD). This optimization technique employs BP with SGD for tuning the weight matrix to an optimum value. The optimization technique considers the outcomes of the additional layer constructed on DBN afterwards and it's preceding greedy training to minimalize error metrics. Logistic units, or Softmax, are widely applied in this layer.

3.2.3. SAE model

An Autoencoder (AE) comprises interlayer connection, input layer, intermediate layer, and reconstruction layer neurons (Sheng *et al.*, 2022). The dataset $(x_1,x_2,\ldots,x_n)$ in the input layer was encoded to attain $(y_1,y_2,\ldots,y_m)$ in the intermediate layer. Next, the dataset $(y_1,y_2,\ldots,y_m)$ is decoded to attain $(\tilde{x}_1,\tilde{x}_2,\ldots,\tilde{x}_n)$ from the intermediate layer to the reconstructed layer. In the supervised learning technique, every sample is encompassed by an input object and expected target values (named supervised signal). In contrast, when the input objects are included without target values are named unsupervised learning and it is used to train the AEs and supervised learning can be used to fine-tune SAE. Beforehand establishing the SAE, various AEs must be trained based on the abovementioned technique. In the data-training method, the output values produced by the intermediate layer of the initial AE can be applied as the input values to train the next AEs, and the succeeding AEs are sequentially trained based on the abovementioned rules. At last, the SAE was fine-tuned through supervised learning:

(1) Train initial AEs with the trained set as input, and stop training while error becomes minimal.

(2) Train second AEs with the encoding outcome of the initial AE as input. Stop training while the error was minimalized. Likewise, the remaining AEs were trained.

(3) Train the softmax classification layer with the encoding outcome of the final AE as input. Halt training while the error becomes minimal um

(4) Stack the output layer intermediate layer of each AE input layer, and the softmax classification layer to attain SAE. Fine-tuning the parameter of each layer of SAE must be implemented using the backpropagation in a supervised manner.

3.3. Hyperparameter tuning

To adjust the hyperparameter values of the DL models, the MRA is exploited in this work. The mathematical model of the search for prey and mud ring feeding was projected (Desuky *et al.*, 2022). According to the subsequent, the basic steps of the MRA are represented

3.3.1. Exploitation phase

When it can be idealized one of the echolocation features of dolphins, it can be utilizing the subsequent rules: Each dolphin employs echolocation for measuring distance, in seeking prey; dolphins swim arbitrarily but utilise velocity $\vec{V}$ at position $\vec{D}$ with sound loudness $\vec{K}$ to search for prey. Every dolphin is mechanically changing its created sounds loudness dependent upon the closeness of its prey;

although loudness can alteration in several approaches, it is assumed that loudness variations depend on the time step and rate of pulse 'r' that differs betwixt Zero and one, whereas Zero defines the no emission pulses, and one signifies the highest rate of pulse emission. The computations of vectors $\vec{K}$ were:

$$\vec{K} = 2\vec{a} \cdot \vec{r} - \vec{a} \tag{14}$$

whereas $\vec{r}$ implies the arbitrary vector betwixt Zero and one, and

$$\vec{a} = 2\left(1 - \frac{t}{T\_\max}\right) \tag{15}$$

For determining prey (exploration), it can be utilized virtual dolphin (searching agent) naturally. In a d -dimension parameter space, the dolphin searches in an arbitrary position, defined as its comparative places to one another. Thus, it can be employed $\vec{K}$ which differs arbitrarily with values superior to 1 or lesser than -1 for driving the dolphin for diverging in everyone and attempting to determine the fittest prey. Therefore, an arbitrarily selective dolphin was selected rather than an optimum dolphin. This selective process and $|\vec{K}| \geq 1$ encourage exploration and allow the MRA technique for undertaking global searching. Next mathematical process of the MRA technique. The workability $\vec{D}$ dependent upon velocity $\vec{V}$ at time step t is offered as:

$$\vec{D}^{\mathrm{t}} = \vec{D}^{\mathrm{t}} + \vec{V}^{t}, \tag{16}$$

whereas V has been established as an arbitrary vector. Primarily, all the dolphins are assigned an arbitrary velocity in [Vmin, Vmax] which was chosen depending on the size of the interest issue.

3.3.2. Exploitation phase

Afterwards identifying the prey, the dolphin is placed and surrounds it. The MRA approach assumes the targeted prey (optimal or adjacent to it) as present optimum solutions as a place of an optimum plan in searching space was not recognized a priori. The other dolphins are thus effort for updating their places based on an optimum dolphin position when the optimum searching agents were defined. The subsequent formulas define this performance:

$$\vec{A} = \left|\vec{C}\vec{D}^{*t} - \vec{D}^{t-1}\right| \tag{17}$$

$$\vec{D}^{\mathrm{t}} = \vec{D}^{*t} \cdot \sin(2\pi l) - \vec{K} \cdot \vec{A} \tag{18}$$

whereas t marks the present time step, 1 defines the arbitrary number, $\vec{C}$ and $\vec{K}$ were co-efficient vectors, $\vec{D}$ implies the dolphin position vector, and $\vec{D}$ represents position vectors of the optimum dolphin position attained. Monitored that $\vec{D}^{*}$ can be altered in all the time steps when there is an optimum position. Noticeably the optimum dolphin moves in a circle but moves its tail swiftly in the sand generating the shape of a sine wave for producing a plume but other dolphins enclose prey.

The computation of vector, $\vec{C}$ is given below:

$$\vec{C} = 2 \cdot \vec{r} \tag{19}$$

By defining the arbitrary vector r , some positions are attained in the searching region. Therefore, Eq. (18) put on the prey surrounding and supports some dolphins for justifying their place nearby the current optimum positions. The search procedure of MRA begins with a population of arbitrary solutions (dolphin position). At all times steps, dolphins validate their position concerning both the optimum position placed so far and arbitrarily selective dolphins. So, the parameter is dependent upon the time step for transferring betwixt the exploration as well as exploitation. If, $|\vec{K}| < 1$ the optimum dolphin position was chosen, but if, $|\vec{K}| \geq 1$, arbitrary dolphins are chosen to justify the dolphin positions.

Algorithm 1: Pseudocode of MRA

Arbitrary Dolphin population initiation, D_i, $i \in [1, 2, \ldots, n]$
and Velocity vi
Determine the fitness value of all Dolphins
D^* = optimal Dolphin location
While $(t < T\_\max)$
for i=1 to n
Change K, C a, and l
if $|K| >= 1$ Then
Produce New Solution by Changing Velocity vi
Else
/* Form Mud ring */
Upgrade the present place of dolphins
end If
end for
Upgrade the limits of the Dolphin exterior to the search area
Obtain the fitness function of a dolphin
Upgrade D^* for improved location
Set $t \to t+1$
end While
Display D^* (Optimal location position)

4. Experimental validation

The proposed model is simulated using Python 3.6.5 tool on PC i5-8600k, GeForce 1050Ti 4GB, 16GB RAM, 250GB SSD, and 1TB HDD. The parameter settings are given as follows: learning rate: 0.01, dropout: 0.5, batch size: 5, epoch count: 50, and activation: ReLU.

The air quality prediction results of the AQP-EDLMRA method are tested using the air quality dataset. The dataset comprises 3000 samples with six classes as shown in Table 1. Each class label holds a set of 500 samples.

Table 1. Details of the dataset

Class	No. of Instances
Good	500
Satisfactory	500
Moderate	500

Poor	500
Very Poor	500
Severe	500
Total Number of Instances	3000

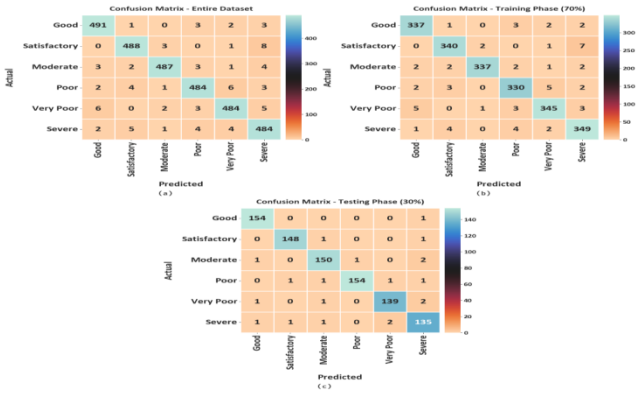


Figure 3. Confusion matrices of AQP-EDLMRA system (a) Entire database, (b) 70% of TR database, and (c) 30% of TS database

The confusion matrices of the AQP-EDLMRA method obtained under the applied dataset are given in Figure 3. The outcomes showcased that the AQP-EDLMRA method has effectually forecast different classes of air quality under all cases.

Table 2 and Figure 4 report a detailed results analysis of the AQP-EDLMRA method on the entire database. The AQP-EDLMRA method has classified samples under the 'good' class with an $accu_y$ of 99.27%, $spec_y$ of 98.20%, $spec_y$ of 99.48%, F_{score} of 97.81%, ROC_{score} of 98.84%, and MCC of 97.37%. Also, the AQP-EDLMRA method has classified samples under the 'poor' class with an $accu_y$ of 99.03%, $sens_y$ of 96.80%, $spec_y$ of 99.48%, F_{score} of 97.09%, ROC_{score} of 98.14%, and MCC of 96.51%. Moreover, the AQP-EDLMRA technique has classified samples under the 'severe' class with *an* $accu_y$ of 98.70%, $sens_y$ of 96.80%, $spec_y$ of 99.08%, F_{score} of 96.13%, ROC_{score} of 97.94%, and MCC of 95.35%. %.

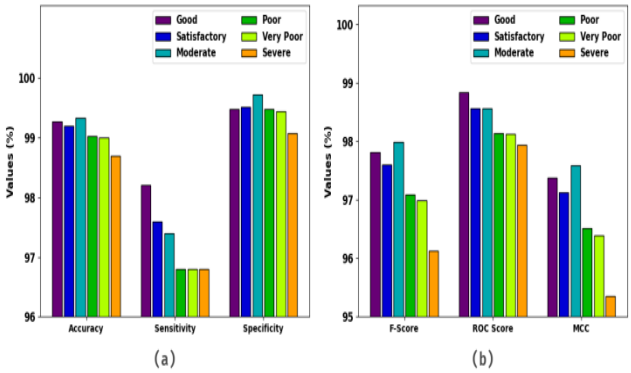


Figure 4. Result in the analysis of the AQP-EDLMRA method under the Entire Database

Table 2. Result analysis of the AQP-EDLMRA scheme with various classes under the Entire dataset

Entire dataset						
Class	$Accu_y$	$Sens_y$	$Spec_y$	F_{score}	ROC_{score}	**MCC**
Good	99.27	98.20	99.48	97.81	98.84	97.37
Satisfactory	99.20	97.60	99.52	97.60	98.56	97.12
Moderate	99.33	97.40	99.72	97.99	98.56	97.59
Poor	99.03	96.80	99.48	97.09	98.14	96.51
Very Poor	99.00	96.80	99.44	96.99	98.12	96.39
Severe	98.70	96.80	99.08	96.13	97.94	95.35
Average	99.09	97.27	99.45	97.27	98.36	96.72

Table 3. Result analysis of the AQP-EDLMRA approach with various classes under 70% of the TR dataset

Training Phase (70%)						
Class	$Accu_y$	$Sens_y$	$Spec_y$	F_{score}	ROC_{score}	**MCC**
Good	99.14	97.68	99.43	97.40	98.56	96.89
Satisfactory	99.05	97.14	99.43	97.14	98.29	96.57
Moderate	99.43	97.40	99.83	98.25	98.61	97.91
Poor	98.86	96.49	99.32	96.49	97.90	95.81
Very Poor	98.90	96.64	99.37	96.77	98.00	96.11
Severe	98.71	96.94	99.08	96.28	98.01	95.50
Average	99.02	97.05	99.41	97.06	98.23	96.47

Table 3 and Figure 5 illustrate a brief results analysis of the AQP-EDLMRA method on 70% of the TR database. The AQP-EDLMRA method has classified samples under the 'good' class with an $accu_y$ of 99.14%, $sens_y$ of 97.68%, $spec_y$ of 99.43%, F_{score} of 97.40%, ROC_{score} of 98.56%, and MCC of 96.89%. Similarly, the AQP-EDLMRA technique has classified samples under the 'poor' class with *an* $accu_y$ of 98.86%, $sens_y$ of 96.49%, $spec_y$ of 99.32%, F_{score} of 96.49%, ROC_{score} of 97.90%, and MCC of 95.81%. Furthermore, the AQP-EDLMRA algorithm has classified samples under the 'severe' class with an $accu_y$ of 98.71%, $sens_y$ of 96.94%, $spec_y$ of 99.08%, F_{score} of 96.28%, ROC_{score} of 98.01%, and MCC of 95.50%.

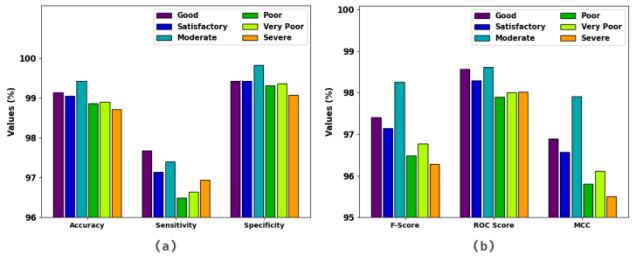


Figure 5. Result in the analysis of the AQP-EDLMRA system under 70% of the TR database

Table 4 and Figure 6 exhibits a detailed results analysis of the AQP-EDLMRA method on 30% of the TS database. The AQP-EDLMRA algorithm has classified samples under the 'good' class with an $accu_y$ of 99.56%, $sens_y$ of 99.35%, $spec_y$ of 99.60%, F_{score} of 98.72%, ROC_{score} of 99.48%, and MCC of 98.45%. Similarly, the AQP-EDLMRA approach has classified samples under the 'poor' class with $anaccu_y$ of 99.44%, $sens_y$ of 97.47%, $spec_y$ of 99.87%, F_{score} of 98.40%, ROC_{score} of 98.67%, and MCC of 98.07%. Furthermore, the AQP-EDLMRA method has classified samples under the 'severe' class with an $accu_y$ of 98.67%, $sens_y$ of 96.43%, $spec_y$ of 99.08%, F_{score} of 95.74%, ROC_{score} of 97.75%, and MCC of 94.96%.

Table 4. Result analysis of AQP-EDLMRA approach with various classes under 30% of the TS dataset

Testing Phase (30%)						
Class	$Accu_y$	$Sens_y$	$Spec_y$	F_{score}	ROC_{score}	**MCC**
Good	99.56	99.35	99.60	98.72	99.48	98.45
Satisfactory	99.56	98.67	99.73	98.67	99.20	98.40
Moderate	99.11	97.40	99.46	97.40	98.43	96.87
Poor	99.44	97.47	99.87	98.40	98.67	98.07
Very Poor	99.22	97.20	99.60	97.54	98.40	97.08
Severe	98.67	96.43	99.08	95.74	97.75	94.96
Average	99.26	97.75	99.56	97.75	98.66	97.31

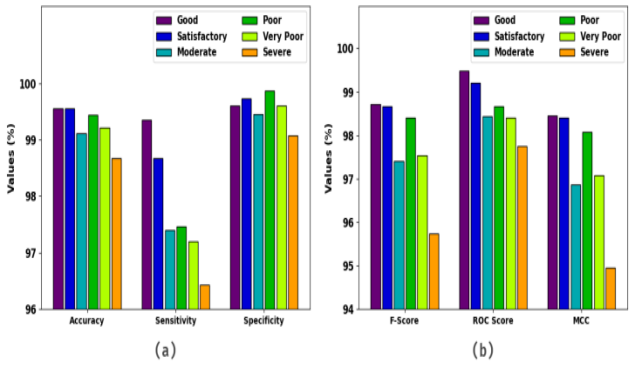


Figure 6. Result in the analysis of the AQP-EDLMRA system under 30% of the TS database

Figure 7 reveals an average result of the AQP-EDLMRA method on the entire dataset. On the entire dataset, the AQP-EDLMRA method has attained an average $accu_y$ of 99.09%, $sens_y$ of 97.27%, $spec_y$ of 99.45%, F_{score} of 97.27%, ROC_{score} of 98.36%, and MCC of 96.72%. Additionally, on 70% of the TR database, the AQP-EDLMRA approach has achieved an average $accu_y$ of 99.02%, $sens_y$ of 97.05%, $a spec_y$ of 99.41%, $a F_{score}$ of 97.06%, $_{a}score$ of 98.23%, and MCC of 96.47%. Also, on 30% of the TS database, the AQP-EDLMRA approach has gained an average $accu_y$ of 99.26%, $sens_y$ of 97.75%, $spec_y$ of 99.56%, F_{score} of 97.75%, ROC_{score} of 98.66%, and MCC of 97.31%.

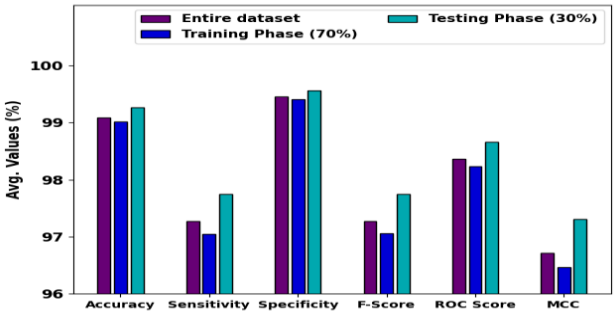


Figure 7. Average analysis of the AQP-EDLMRA system with various measures

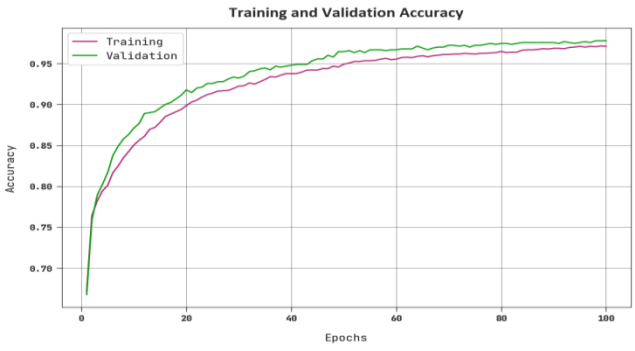


Figure 8. TACY and VACY analysis of AQP-EDLMRA approach

The TACY and VACY of the AQP-EDLMRA system are inspected on air quality prediction outcome in Figure 8. The figure exhibited that the AQP-EDLMRA approach has displayed improvized performance with increased values of TACY and VACY. Visibly, the AQP-EDLMRA technique has reached maximum TACY values.

The TLOS and VLOS of the AQP-EDLMRA approach are tested on air quality prediction performance in Figure 9.

The figure shows the AQP-EDLMRA method has exposed superior outcomes with the least values of TLOS and VLOS. Especially, the AQP-EDLMRA technique has lesser VLOS values.

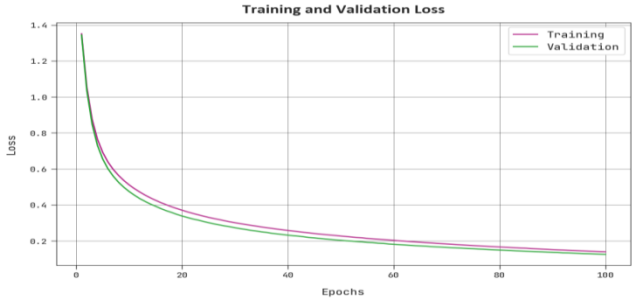


Figure 9. TLOS and VLOS analysis of AQP-EDLMRA scheme

A comprehensive precision-recall investigation of the AQP-EDLMRA methodology under the test database is exhibited in Figure 10. The figure highlighted that the AQP-EDLMRA approach has enhanced values of precision-recall values in every class label.

The brief ROC study of the AQP-EDLMRA methodology under the test database is shown in Figure 11. The outcomes signified the AQP-EDLMRA technique has emphasized its ability in classifying distinct classes under the test database.

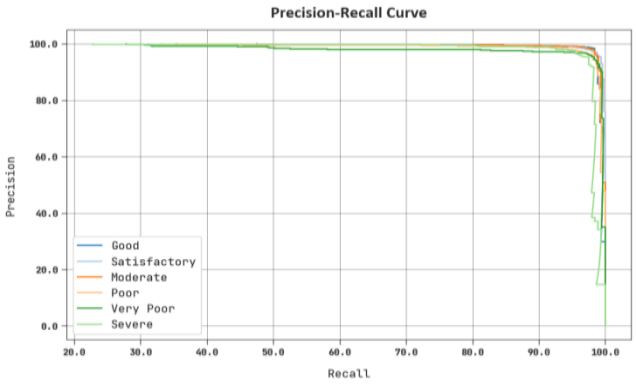


Figure 10. Precision-recall analysis of AQP-EDLMRA approach

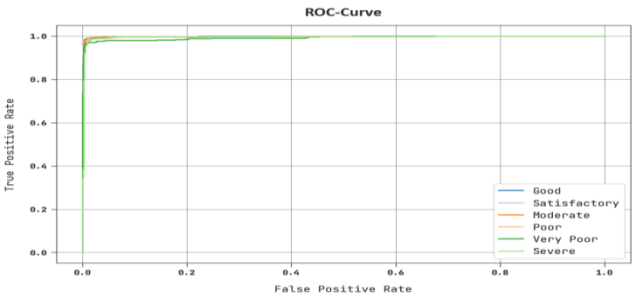


Figure 11. ROC curve analysis of AQP-EDLMRA approach

The air quality forecasting outcomes of the AQP-EDLMRA method were compared with recent ML methods in Table 5.

A comparative $Accu_y$ and F_{score} investigation of the AQP-EDLMRA model with recent methods is offered in Figure 12. The outcomes indicated that the LOR, SGD, and MLP methods have obtained poor performance with minimal values of $Accu_y$ and F_{score} Next to that, the GBR model has reached slightly improvised $Accu_y$ and F_{score} of 80.67% and 81.22% respectively. Although the XGBoost model has gained reasonable $Accu_y$ and F_{score} of 98.34% and 97.14%, the AQP-EDLMRA model has shown maximum $Accu_y$ and F_{score} of 99.26% and 97.75% respectively.

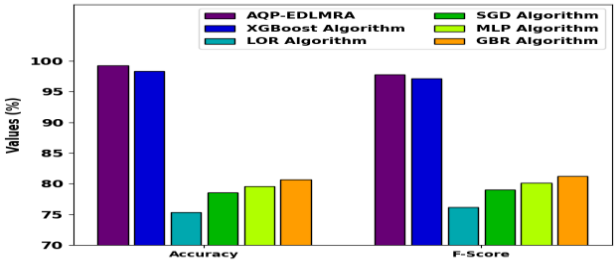


Figure 12. $Accu_y$ and F_{score} analysis of AQP-EDLMRA system with existing approaches

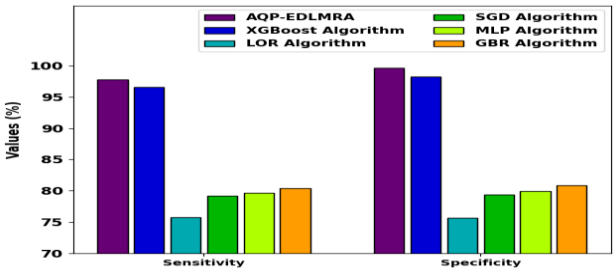


Figure 13. $sens_y$ and $spec_y$ analysis of AQP-EDLMRA system with existing approaches

Table 5. Comparative analysis of AQP-EDLMRA approach with existing methods

Methods	$Accu_y$	$Sens_y$	$Spec_y$	F_{score}
AQP-EDLMRA	99.26	97.75	99.56	97.75
XGBoost Algorithm	98.34	96.55	98.23	97.14
LOR Algorithm	75.32	75.80	75.70	76.15
SGD Algorithm	78.61	79.17	79.42	79.04
MLP Algorithm	79.54	79.68	79.92	80.15
GBR Algorithm	80.67	80.42	80.89	81.22

A comparative $sens_y$ and $spec_y$ study of the AQP-EDLMRA technique with current methods is given in Figure 13. The outcomes exhibited that the LOR, SGD, and MLP techniques have gained poor performance with minimal values of $sens_y$ and $spec_y$. Then, the GBR technique has reached slightly improvised $sens_y$ and $spec_y$ of 80.42% and 80.89% correspondingly. Although the XGBoost approach has obtained reasonable $sens_y$ and $spec_y$ of 96.55% and 98.23%, the AQP-EDLMRA methodology has shown maximum $sens_y$ and $spec_y$ of 97.75% and 99.56% correspondingly. These outcomes ensured the enhanced predictive outcomes of the AQP-EDLMRA method.

5. Conclusion

Automated air quality forecasting using the AQP-EDLMRA technique has been introduced in this study. The presented AQP-EDLMRA technique employed the ensemble voting model by the use of three DL methods namely LSTM, DBN, and SAE. In the ensemble voting model, the new data can be classified by the weighted vote of their prediction outcomes. To adjust the hyperparameter values of the DL methods, the MRA was exploited in this work. The experimental evaluation of the AQP-EDLMRA approach is tested using a series of air quality data and the comprehensive comparative results demonstrated that the AQP-EDLMRA technique has reached improved forecasting performance. In future, the performance of the proposed model can be boosted by the design of feature selection approaches.

References

Ali M.A., Balasubramanian K., Krishnamoorthy G.D., Muthusamy S., Pandiyan S., Panchal H., Mann S., Thangaraj K., El-Attar N.E., Abualigah L. and Abd Elminaam D.S. (2022). Classification of Glaucoma Based on Elephant-Herding Optimization Algorithm and Deep Belief Network. *Electronics*, ***11***(11), 1763.

Cabaneros S.M., Calautit J.K. and Hughes B.R. (2019). A review of artificial neural network models for ambient air pollution prediction. *Environmental Modelling & Software*, ***119***, 285–304.

Chang Y.S., Chiao H.T., Abimannan S., Huang Y.P., Tsai Y.T. and Lin K.M. (2020). An LSTM-based aggregated model for air pollution forecasting. *Atmospheric Pollution Research*, ***11***(8), 1451–1463.

Desuky A.S., Cifci M.A., Kausar S., Hussain S. and El Bakrawy L.M. (2022). Mud Ring Algorithm: A new meta-heuristic optimization algorithm for solving mathematical and engineering challenges. *IEEE Access*, ***10***, 50448–50466.

Gilik A., Ogrenci A.S. and Ozmen A. (2022). Air quality prediction using CNN+ LSTM-based hybrid deep learning architecture. *Environmental science and pollution research*, ***29***(8), 11920–11938.

Gu Y., Li B. and Meng Q. (2022). Hybrid interpretable predictive machine learning model for air pollution prediction. *Neurocomputing*, ***468***, 123–136.

Heydari A., Majidi Nezhad M., Astiaso Garcia D., Keynia F. and De Santoli L. (2022). Air pollution forecasting application based on deep learning model and optimization algorithm. *Clean Technologies and Environmental Policy*, ***24***(2), 607–621.

Jin N., Zeng Y., Yan K. and Ji Z. (2021). Multivariate air quality forecasting with nested long short term memory neural network. *IEEE Transactions on Industrial Informatics*, ***17***(12), 8514–8522.

Ke H., Gong S., He J., Zhang L., Cui B., Wang Y., Mo J., Zhou Y. and Zhang H. (2022). Development and application of an automated air quality forecasting system based on machine learning. *Science of The Total Environment*, ***806***, 151204.

Mao W., Wang W., Jiao L., Zhao S. and Liu A. (2021). Modeling air quality prediction using a deep learning approach: Method optimization and evaluation. *Sustainable Cities and Society*, ***65***, 102567.

Rybarczyk Y. and Zalakeviciute R. (2018). Machine learning approaches for outdoor air quality modelling: A systematic review. *Applied Sciences*, ***8***(12), 2570.

Sheng D., Yu J., Tan F., Tong D., Yan T. and Lv J. (2022). Rock mass quality classification based on deep learning: A feasibility study for stacked autoencoders. *Journal of Rock Mechanics and Geotechnical Engineering*.

Sun J., Cao R., Zhou M., Hussain W., Wang B., Xue J. and Xiang J. (2021). A hybrid deep neural network for classification of schizophrenia using EEG Data. *Scientific Reports*, ***11***(1), 1–16.

Tao Q., Liu F., Li Y. and Sidorov D. (2019). Air pollution forecasting using a deep learning model based on 1D convnets and bidirectional GRU. *IEEE access*, ***7***, 76690–76698.

Wang J., Du P., Hao Y., Ma X., Niu T. and Yang W. (2020). An innovative hybrid model based on outlier detection and correction algorithm and heuristic intelligent optimization algorithm for daily air quality index forecasting. *Journal of environmental management*, ***255***, 109855.

Wang W., An X., Li Q., Geng Y.A., Yu H. and Zhou X. (2022). Optimization research on air quality numerical model forecasting effects based on deep learning methods . *Atmospheric Research*, ***271***, 106082.

Wu Q. and Lin H. (2019). A novel optimal-hybrid model for daily air quality index prediction considering air pollutant factors. *Science of the Total Environment*, ***683***, 808–821.

Xayasouk T., Lee H. and Lee G. (2020). Air pollution prediction using long short-term memory (LSTM) and deep autoencoder (DAE) models. *Sustainability*, ***12***(6), 2570.

Yi X., Zhang J., Wang Z., Li T. and Zheng Y. (2018), July. Deep distributed fusion network for air quality prediction. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining* (965–973).

Zeng Y., Chen J., Jin N., Jin X. and Du Y. (2022). Air quality forecasting with hybrid LSTM and extended stationary wavelet transform. *Building and Environment*, *213*, 108822.

Zhu D., Cai C., Yang T. and Zhou X. (2018). A machine learning approach for air quality prediction: Model regularization and optimization. *Big data and cognitive computing*, ***2***(1), 5.