International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

Artificial Flora Optimization with Deep Learning Enabled Air Pollution Monitoring System in IoT Environment

Mr.K. Azhahudurai[1], Dr. V. Veeramanikandan[2]

Submitted: 04/11/2022 **Accepted:** 06/02/2023

Abstract: Air pollution monitoring is ever-growing, which provides increasing priority toward the consequence on human health. A higher level of pollution in air might have a lot of adverse effects on the health of fellow human beings. It increases the possibility of lung cancer, chronic disease, and cardiac disease. Meanwhile, pollution in the air impacts the health and leads to increased mortality and morbidity raised the toxicological analysis focuses on components influencing air pollution. Thus, there is a need for designing an automatic environment pollution monitoring system. With this background this research paper focuses on building an automated pollution monitoring system using Artificial Intelligence techniques. The aim is to construct a deep learning based prediction model that could forecast the level of air pollutants (PM2.5 concentration) with reference to the weather and climatic parameters. The deep learning model was trained using the dataset collected from the benchmarked and real time data. To overcome the problems and pitfalls in the earlier statistical and deep learning based approaches; this research employs Artificial Flora Optimization algorithm for optimization of the deep learning model's hyperparameters termed as AFODL. The experimental evaluation analyses the performance of the AFODL model under different experimental scenario. The simulation results reported the enhancements of the AFODL based air pollutant forecast model when compared to earlier approaches.

Keywords: *Air pollution; Deep learning; Artificial Flora Algorithm; Data acquisition; Prediction*

1. Introduction

Typically, the chemical substances that worsen quality of the air are termed as air pollutants. These substances can exist in the air in two main states: gaseous and solid (suspended in air), the latter of which is known as particular matter (PM). The health of all living things may be badly impacted, particularly by pollutants such dust, gases, smell, smoke, and fume that are not naturally present in the atmosphere. Particulate matter (PM) is an important air pollutants, despite the fact that there are many others that are detrimental to all ecosystems, including sulphur oxides (SOx), nitrogen oxides (NO), and carbon monoxide (CO).

Particulate matter (PM), which is categorised based on aerodynamic diameter, is one of the toxic pollutants that has an adverse effect on health. Because of the compounds present in PM's chemical makeup, the World Health Organization (WHO) classifies it as carcinogenic. In this study, deep learning techniques were used to forecast an approximate value of PM2.5 pollutants. Long-short term memory (LSTM) based Recurrent Neural Networks (RNN), and another variant of RNN Gated Recurrent Units (GRU) were tested as three deep learning techniques (GRU). Additionally, a CNN and recurrent model combination was evaluated.

In the field of artificial intelligence; deep learning is the newest area. Artificial neural networks, a subclass of machine learning, have been extensively employed to tackle challenging global issues. Unfortunately, because to issues with training big data sets and the loss of gradients, their prediction performance has not been very encouraging. A branch of machine learning called "deep learning" advances machine learning. In order to generate more accurate findings, deep learning may be used to solve issues by using more layers, larger volume of data samples, and simultaneous computation in all layers. Deep learning is more compatible for modelling and predicting the level of air pollutants because of all these advantageous characteristics. This work investigates how to combine RNN, GRU, and LSTM models to find the best method for predicting PM2.5 pollution.

This research aims at developing an efficient optimization approach using Artificial Flora Optimization (AFO) and use the same with Deep Learning model for prediction of

[1]*Research Scholar,Department of Computer and Information Science, Annamalai University, Annamalai Nagar , Email: jkmaz477@gmail.com*
[2]*Assistant Professor,Department of Computer Science,Thiru Kolanjiappar Government Arts College, Vridhachalam. , Email: klmvmani@gmail.com*

air pollutants. The presented prediction approach initially pre-processes the input data into useful format. In addition, the presented approach applies deep neural network model for predicting the value of PM2.5. The objective of using AFO algorithm is to optimize the hyperparameters of the DBN model in a way to enhance the accuracy of the model. The evaluation of the AFODL model is evaluated and the results are analysed under several scenario. The efficiency of the deep neural network optimized using the AFO algorithm is compared with other deep neural based predictive models including RNN, LSTM, GRU, and CNN-LSTM.

2. Literature Survey

Pollution is regarded as a serious concern in the present situation. We are moving towards smart city advancement and air pollution becomes a significant matter for human well beings and other living beings [1]. Some variations in the arrangement of air could be harmful to diverse life forms. Air pollution is the prevalence of few or more pollutants in the environment like toxin gases or smoking in excessive amounts which might be harmful to animals, human beings, and plants [2]. Air contaminants can be measured either in ppm or percentage form. The good air quality index (AQI) indicates the range starting from 0 to 50 denotes that air contamination does not occur. Fair AQI indicates the range starting with 51 upto 100 in that the sensitive individual must take into account decreasing long or heavy outdoor travelling. Moderately contaminated AQI indicates the range starting with 101 upto 200 and it is considered harmful for sensitive crews [3].

The forecasting of AQI for ordinary people seems to be more complex. Whenever an individual wants to move from one place to another place, they must go across numerous cities. The level of pollution in major cities is rising every day and whenever the individuals travel across non-polluted regions, it seems to be more complex for them to predict the hour quality index value for diverse regions [4]. Making them travel across the least polluted zone and IoT allowed environmental air pollution rerouting and monitoring system was suggested in this article with a machine learning (ML) method for viewing all the pollution levels. Current advancements in information technology allow the compilation of several monitoring applications into single complex system for the entire supply chain [5]. Generally, monitoring application implementations acts a vital role in improving production, forecasting diseases, reducing costs, and offering a prior warning system [6].

Recent technologies like Internet of Things (IoT)-related sensors are used and complied with monitoring structures. Researchers were carried out in the production industry and displayed prominent advantages from the utilization of IoT-related sensors for observing like quality prediction, working condition developments, fault diagnosis error, helping managers with better decision making and design prevention [7, 8]. Deep learning (DL) employs multi-layer substructure for the abstraction of inherent structures in layers. It derives the data from lower level to the higher level and could fix the characteristic outline in data. The temporal trends and spatial distribution of air quality processes are influenced by several elements like weather conditions, air pollution emissions, depositions, human activities, etc that making the process very difficult [9, 10]. This state leads to more trouble in using traditional trivial methods for acquiring a superior pattern. DL might result in good performances for air quality forecasting by deriving the air quality features without not being aware of past information.

In [11], an IoT based air pollution monitored and predicted methods are presented. This method is employed to monitor air pollutants of specific regions and to air quality analysis and predict the air quality. The presented method concentrates on the monitoring of air pollutants concentration with integration of IoT with ML technique shortly termed as RNN further definitely LSTM. In [12], the researchers presented a smart bin based on IoT, ML and DL techniques for handling the garbage disposal and for forecasting the pollutant existing from the neighboring smart-bin environments. The smart-bin is linked to IoT based server deployed inside a cloud environment that provides the computational support to predict the status and to forecast quality of the air dependent upon real-time data. Pushpam and Kavitha [13] presented the IoT based method in which sensors are employed for measuring carbon monoxide (CO), Particulate Matter (PM) level, and atmospheric situation as humidity and temperature. An important objective of this technique is for suggesting another route for users dependent upon pollution status and distance of all the routes that refers to a route free from pollution. The browser based software established has navigation map API (Google) whereas the status of the pollution and alternate travel path are recommended. Based on the collected temporal data, the forecast method is complete to PM with NN, MLP, and SVM regression (SVMR) learning techniques.

Asha et al. [14] presented the ETAPM-AIT algorithm containing a fixed IoT based sensor array for sensing 8 pollutants. To classify pollutants in the air and define quality of the air, Artificial Algae Algorithm (AAA) based Elman Neural Network (ENN) techniques are utilized as classification that forecasts the quality of the air with reference to the future time steps. The authors in [15] revealed the progress of a new minimum-cost sensor node which employs cost-effectual electrochemical sensor for measuring CO and NO_2 focuses and an infrared sensor for measuring PM level. The node is controlled by both solar-recharge battery and mains supply. It can be accomplished of long distance, minimal transmission on public or private long distant wide area network (LoRaWAN) IoT networks and short distant superior data transmission rate on Wi-Fi.

2.1 Deep Learning based PM2.5 prediction

In order to forecast PM2.5, Xiang Li et al. [20] employed LSTM. They also compared the prediction effects of LSTM and other regression techniques. The training variables for this prediction included climatic variables as additional data, such as temperature, humidity, and wind speed, in addition to chemical factors as models. However, the prediction model used in this study, MAPE, provides a more accurate prediction performance with an evaluation result value of 11. As a result, in our study, we carried out modelling, assessed whether or not to include climatic parameters, and adjusted the subsequent model design in accordance with the experimental results.

Two neural network building models were utilised by Yi Ting Tsai et al. [21] to forecast PM2.5. The Taiwan Environmental Protection Agency offers two neural networks: RNN and LSTM. From 2012 to 2016, the historical data were used. The forecasted time was four hours away. RMSE is used in the evaluation model technique. The experimental findings in this research demonstrate that the LSTM outperforms the RNN in terms of RMSE performance. The accuracy of different regions' predictions is likewise extremely high. In order to compare multiple neural network models in their tests, including CNN, RNN, LSTM, and GRU, they analysed to identify which type of neural architecture was best for this particular problem.

Using a fresh deep learning hybrid model, Jun Ma et al. This model is a combination of the BLSTM and the IDW. The temporal sequence of data characterizing air pollution can be successfully captured using BLSTM. Using the IDW connected, it can be interpolated with a spatial distribution with reference to the temporal similarity of pollutants in the air. The MAPE evaluation value is 9, and the experimental results of the paper show that the analysis values offered by their suggested architecture perform remarkably well. Therefore, in our investigation, PM2.5 diffusion estimates were performed using the IDW technique. To understand the stature of the air quality in the area, they approximated the regions lacking air quality observation stations based on their features.

In South Korea, Kim et al. [23] applied LSTM-based daily PM10 and PM2.5 predictions. They contrasted the simulations of a chemistry-transport model (CTM) with observations of ground-level PM. According to the findings, the index of agreements (IOA) raised from 0.3 to 0.7 and from 0.6 to 0.7 based on the LSTM model for 3-D CTM simulations. Three deep neural architectures; LSTM, RNN, and GRU, were used by Ayturan et al. [24] to construct a PM2.5 short term forecast in Turkey. For 1, 2, and 3 h forecasts, the combined RNN and GRU model produced the efficient results, with R^2 values of 0.8, 0.72, and 0.6.

The 24-hour prediction of PM2.5 employing attention based neural network generated by combining CNN and LSTM was proposed by Li et al. [25]. Support Vector Regression, Random Forest Regression, Multi-Layer Perceptron, Simple RNN, LSTM, CNN-LSTM, and AC LSTM were the seven models they compared. According to the results, the AC-LSTM model performed better in the prediction tasks at multiple scale. In comparison to the other models, the RMSE value for the 13–24 h measurement was the lowest at 14.83[31].

The prediction of PM10 and PM2.5 utilising composite approaches was established by Wu et al. [26]. The spectral data from MODIS and using the dense dark vegetation method was employed. The planetary boundary layer height (PBLH) and relative humidity calculations were used to optimise the aerosol optical depth (AOD) (RH). The PM2.5 and PM10 connection with the highest value was chosen. Based on the mean, max, and min accuracy as well as the stableness of PM10/PM2.5 forecast, the LSTM model demonstrated outstanding performance. Yang et al hybrid's model of CNN combined with LSTM and CNN combined with GRU was presented in order to forecast PM10 and PM2.5 in South Korea, for a period of 07 days [27]. Individual and combination models were contrasted. Their research indicates that combination models perform better than individual models [32].

Xayasoux et al. [28] examined the RMSE model outputs while developing LSTM and auto-encoder models for predicting PM concentrations [33][35][37]. They used the algorithms to analyse hourly quality of the air from 25 sites in South Korea, between 01.01.2015 and 31.12.2018. PM concentration was forecast using the LSTM model with an ideal learning rate of 0.01 with 100 iterations, and min-batch size of 32 for the ensuing 10 days. A batch size of 64 produced the best results for the DAE model. The LSTM units based neural architecture performed significantly higher than the other proposed architectures in forecasting the concentrations of the PM25 [34][36].Further the meta-heuristic algorithms were used in combination with machine learning and deep learning approaches for building efficient forecast models [29-32].

3. The Proposed Model

In this research, a new AFODL model has been developed for the prediction of PM2.5 concentration by the detection of air pollutants. The presented AFODL-APM model initially uses IoT devices for data collection and pre-processes the input data into useful format. In addition, the presented AFODL-APM model applies DBN model for predicting air quality. Followed by, the AFO algorithm is employed to fine tune the hyperparameters of the network to enhance the predictive performance.

3.1. Dataset

The presented AFODL model was trained on pre-processed benchmarked dataset. The data samples considered were multivariate and time series in nature. In total 43824 samples were considered for training and testing the model wherein 80% of the data was used for training and 20% for evaluating the performance of the proposed approach. The dataset consists of the PM2.5 concentration, temperature, pressure, wind direction, wind speed, snow duration and rain duration. In the experiments the PM2.5 concentration is considered as the dependent variables and other seven attributes were considered as the independent variables.

3.2. DBN based Predictive Model

The proposed AFODL model applies DBN model for predicting air quality. For probability generation module, DBN is stacked generally by restricted Boltzmann machine (RBM). The bottommost layer is primarily utilized for receiving the input dataset and converting that input dataset into the hidden neuron via RBM, especially, the input of larger layer RBM derives from the output of bottom layer RBM [16]. The RBM is a specific context that depends on the generation module that might offer a learning mechanism for arbitrarily distributing dataset. In general, every RBM comprises hidden and visual layers, also there is a bi-direction connectivity weight amongst hidden and visual units. However, there exist no connections among the visual and the hidden units. Assume the condition of visual layer v, hidden layer h, weight matrixes w, visual and hidden unit thresholds b, the function of the joint state of every visual and hidden unit (v, h) are demonstrated in the following.

$$\mathrm{v} = \{\mathrm{v}_1, \mathrm{v}_2, \mathrm{v}_3, \dots, \mathrm{v}_\mathrm{I}\} \in \{0,1\} \tag{1}$$

$$\mathrm{h} = \{\mathrm{h}_1, \mathrm{h}_2, \mathrm{h}_3, \dots, \mathrm{h}_\mathrm{I}\} \in \{0,1\} \tag{2}$$

$$\mathrm{E}(\mathrm{v}, \mathrm{h}) = -\sum_{\mathrm{i}=1}^{\mathrm{I}} \mathrm{a}_\mathrm{i}\, \mathrm{v}_\mathrm{i} - \sum_{\mathrm{j}=1}^{\mathrm{J}} \mathrm{b}_\mathrm{i}\, \mathrm{h}_\mathrm{i} - \sum_{\mathrm{j}=1}^{\mathrm{J}} \sum_{\mathrm{i}=1}^{\mathrm{I}} \mathrm{w}_{\mathrm{ji}}\, \mathrm{v}_\mathrm{i} \mathrm{h}_\mathrm{j} \tag{3}$$

In Eq. (3): I indicate the amount of visual layers. i signifies the i-th numbers in the visual layer. J denotes the amount of hidden units. j characterizes the j-th amount in the visual layer. Consequently, the joint likelihood distributed among the visual and hidden layers is attained.

$$\mathrm{p}(\mathrm{v}, \mathrm{h}) = \frac{\mathrm{e}^{-\mathrm{E}(\mathrm{v},\mathrm{h})}}{\mathrm{Z}} \tag{4}$$

$$\mathrm{Z} = \sum_{\mathrm{v}} \sum_{\mathrm{h}} \mathrm{e}^{-\mathrm{E}(\mathrm{v},\mathrm{h})} \tag{5}$$

In Eq. (5): Z indicates the standardized constant of simulating physical scheme that is attained by including the energy values among each visible and hidden layer. Likewise, assume that haphazard input hidden layers h, the likelihood of the visual layers v assume the hidden layers h is accomplished by.

$$\mathrm{p}(\mathrm{h}/\mathrm{v}) = \prod_{\mathrm{j}} \mathrm{p}\,(\mathrm{h}_\mathrm{j} = 1/\mathrm{v}) \tag{6}$$

$$\mathrm{p}(\mathrm{v}/\mathrm{h}) = \prod_{\mathrm{i}} \mathrm{p}\,(\mathrm{v}_\mathrm{i} = 1/\mathrm{h}) \tag{7}$$

But the RBM units are in a dual state, the activation possibility is achieved on the basis of determining the sigmoid function sig(x).

$$\mathrm{sig}(\mathrm{x}) = \frac{1}{1 + \mathrm{e}^{-\mathrm{x}}} \tag{8}$$

$$\mathrm{p}\left(\mathrm{h}_\mathrm{j} = \frac{1}{\mathrm{v}}\right) = \mathrm{sig}\left(\mathrm{b}_\mathrm{i} + \sum_{\mathrm{i}}^{\prime} \mathrm{v}_\mathrm{i}\, \mathrm{w}_{\mathrm{ji}}\right) \tag{9}$$

$$\mathrm{p}\left(v_i = \frac{1}{h}\right) = sig\left(a_i + \sum_{j}^{J} h_j\, w_{ji}\right) \tag{10}$$

From the abovementioned equations, the RBM has eventually accomplished the determination of feature extraction. The trained DBN is largely comprised of: supervised finetuning and unsupervised layer-wise pertaining. Fig. 1 depicts the architecture of DBN technique.

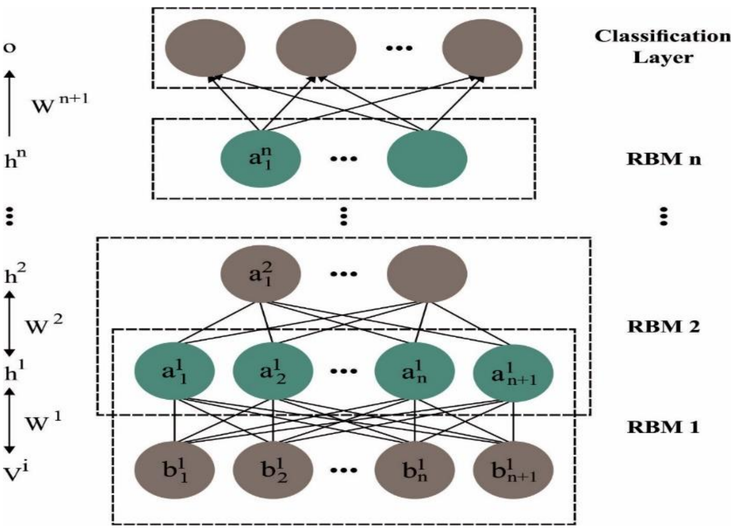

Fig. 1. Architecture of DBN

3 Hyperparameter Optimization

In this work, the AFO algorithm is employed to fine tune the hyperparameters related to the DBN model to enhance the predictive performance. The AFO algorithm follows the process of identifying an optimum survival position and is exploited for discovering the optimal solution set of challenges. The original plant, Offspring plant, plant place, and propagation distance are the four basic units in AFO approach [17]. Evolution, spreading, and select actions are 3 most important behavioural patterns.

Each plant location represents to solution, and fitness of position is exploited for symbolizing quality of the solution. At first, this technique arbitrarily generates the actual plants. Finally, roulette is exploited for determining survival seed. Survival seed turns out to be a novel plant. Recurrent iteration until the termination conditions are satisfied. This technique comprises external documents to save the finest solution. Fig. 2 demonstrates the flowchart of AFO algorithm.

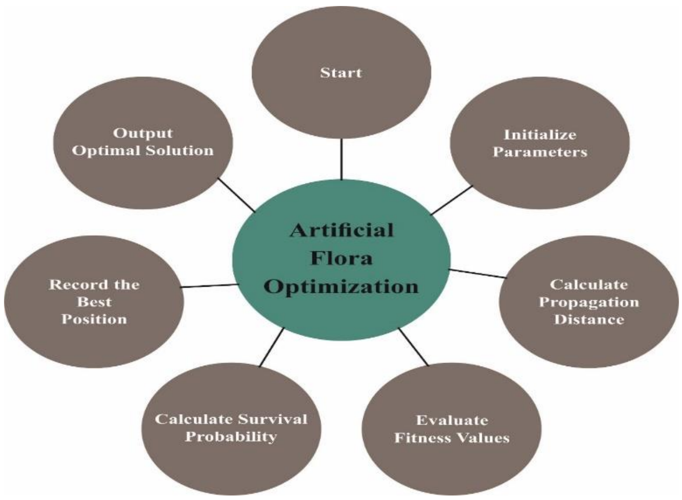

Fig. 2. Flowchart of AFO technique

In this study, all the decision parameters of testing function encompass upper bound $\vec{X}^{max} = [X_1^{max}, X_2^{max}, \dots, X_D^{max}]^T$ & lower bound $\vec{X}^{min} = [X^{min}, X^{min}, \dots, X_D^{min}]^T$. Primarily, this technique produces N unique plant according to the lower & upper limits of the decision variable. This technique makes use of i rows and j columns matrixes P_{ij} for demonstrating the position of novel plant, while $i = 1,2, \dots, D$ symbolizes the dimension, $j = 1,2, \dots, N$ specifies the quantity of new plants:

$$P_{ij} = rand(0,1) \cdot (X_i^{max} - X_i^{mm}) + X_i^{min} \quad (11)$$

In Eq. (11), $rand(0, 1)$ designates the uniformly distributed number within $zero$ and one. The novel plant spread its offspring with a certain scope using radius i.e., propagative distance, original propagative distance imitates the grandparent and parent plants:

$$d_j = d_{lj} \cdot rand(0,1) \cdot c_1 + d_{2j} \cdot rand(0,1) \cdot c_2 \quad (12)$$

In Eq. (12), $c_1 \& c_2$ represents learning coefficient, d_{lj} and d_{2j} correspondingly signifies propagation distance of grandparent and parent plants, rand $(O, 1)$ designates uniformly distributed number within *zero* and *one* (Wu, 2019). The parent propagation distance becomes an original grandparent propagative distance:

$$d'_{1j} = d_{2j} \quad (13)$$

The representative AF optimization technique exploits SD amongst the location of offspring and original plants as parent propagative distance:

$$d'_{2j} = \sqrt{\sum_{i=1}^{N} (P_{ij} - P'_{ij})^2 / N} \quad (14)$$

In order to preserve the information of the finest solution completed previously, AFO optimization method exploits plant from the external document. The parent propagative distance is the variations amongst the location of plant in the offspring plants $P'_{i_d}$and external document $P^*_{i_d}$:

$$d'_{2j} = P^*_{ij} - P'_{ij} \quad (15)$$

This method generates offspring plants depending on novel propagative and plant distance:

$$P'_{i,j \cdot b} = G_{i,j \cdot b} + P_{i,j} \quad (16)$$

In Eq. (16), $b = 1,2, \ldots, B, B$ represents amount of offspring plants that unique plant might propagate, $P'_{i,j \cdot b}$ characterizes position of offspring plants, $P_{i,j}$ designates position of novel plant, $G_{i,j \cdot b}$ symbolizes arbitrary quantity utilizing Gaussian distribution with variance j and mean 0. Producing new plants depends on there being no offspring plants continues. In typical AF methodology, the survival possibility determines the survived offspring plants. It can be determined in the following equation:

$$p = \left| \sqrt{F(P'_{i,j \cdot b}) / F_{max}} \right| \cdot Q_x^{(j \cdot b - 1)}. \quad (17)$$

In Eq. (17), Q_X is probability selection ranges from zero&one. F_{max} indicates fitness of offspring plants with maximal fitness. $F(P'_{i,j \cdot b})$ signifies fitness of $(j \cdot b)$th solution. The assessment equation of fitness is the function of objective problem. In this presented AFO approach, the Pareto dominance relationship has been exploited. The survival possibility can be defined as follows:

$$p = 0.9 \cdot \frac{domi(j \cdot b)}{B} + 0.1 \quad (18)$$

In Eq. (18), domi $(j \cdot b)$ represents amount of solution that subjected to solution $(j \cdot b)$. B designates amount of offspring plants that one unique plant might propagate. Here,Thestudy presented an AFO mechanism for selecting DBN hyperparameter appropriately by diminishing the mean square error (MSE). In the following, the mathematical expression of The MSEis given.

$$MSE = \frac{1}{T} \sum_{j=1}^{L} \sum_{i=1}^{M} \left(y_j^i - d_j^i\right)^2, \quad (19)$$

From the above equation, M and L describe the output value of layers and the data, correspondingly, and y_j^i and d_j^i represents the accomplished and the appropriate magnitude for j-th units in the output layer of the network in t time.

4. Results And Discussion

For the purpose of predicting the concentration of air pollutants, this experiment use the regression prediction method. As a result, the model performance was assessed using MAE, MSE, and RMSE. After comparing the results of the trials, it was shown that MAE has a better effect than the other two. In order to illustrate our prediction findings and train the model, MAE loss function was used. The data samples were standardized using z-score method before using them for training the model.

$$z = \frac{x - \mu}{\sigma}$$

where μ – mean, and σ – standard deviation.

The simulation analysis of the AFODL-APM model is tested using a dataset comprising 420768 samples with 18 features from the UCI repository [18]. The dataset holds concentration of air pollutants and air quality at 12 sites.

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |f_i - y_i|$$

$$RMSE = \sqrt{MSE}$$

Before building a neural network, it becomes essential to determine the relationship among the different elements and the PM2.5 concentration. This will guarantee that the network employs the appropriate input diagnostic features for forecast. There are numerous observable factors that affect PM2.5, but not all of them are useful for making predictions, and the model will be burdened by the ineffective factors. Therefore, it is crucial to determine the correlation coefficient between each element and the target characteristic and to infer from this value the indirect relationship between PM2.5 concentration and the attributes of interest.

The similarity between two time series vectors represented as $X = (x_1, x_2, \dots \dots, x_n)$ and $Y = (y_1, y_2, \dots \dots, y_n)$ respectively can be expressed mathematically as;

$$r = \frac{n \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{n \sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2} \sqrt{n \sum_{i=1}^{n} y_i^2 - (\sum_{i=1}^{n} y_i)^2}}$$

Table. Pearson Correlation co-efficient between features and PM2.5 concentration.

	PM2.5	Dew Point	Temperature	Pressure	Wind Direction	Wind Speed	Rain
PM2.5	1.00	0.178	-0.088	-0.059	0.189	-0.238	-0.049
Dew Point	0.178	1.00	0.816	-0.738	0.227	-0.298	0.128
Temperature	-0.088	0.816	1.00	-0.786	0.177	-0.148	0.049
Pressure	-0.059	-0.738	-0.786	1.00	-0.158	0.179	-0.068
Wind Direction	0.189	0.227	0.177	-0.158	1.00	-0.19	-0.048
Wind Speed	-0.238	-0.298	-0.148	0.179	-0.19	1.00	-0.008
Rain	-0.049	0.128	0.049	-0.068	-0.048	-0.008	1.00

Table 1 presents a comparison of the AFODL model with other DL models on 1st day and 7th day [19]. Fig. 3 illustrates the comparative RMSE investigation of the AFODL-APM model with other models on 1st day and 7th day with distinct BSs. The figure represented that the AFODL-APM model has accomplished improved outcomes with lower values of RMSE. For instance, on 1st day and BS of 24, the AFODL-APM model has offered reduced RMSE of 14.610 whereas the GRU, LSTM, BiLSTM, CNN-LSTM, and CNN-GRU models have obtained increased RMSE of 16.671, 16.249, 15.320, 16.603, and 17.654 respectively. At the same time, on 7st day and BS of 24, the AFODL-APM approach has obtainabledecreased RMSE of 16.194 whereas the GRU, LSTM, BiLSTM, CNN-LSTM, and CNN-GRU algorithms have reached maximal RMSE of 21.316, 19.205, 20.090, 16.824, and 18.494 correspondingly.

The R^2 score ranges from 0% to 100%. Although not the same, it is closely connected to the MSE. It is "the fraction of the variance in the dependent variable that is predicted from the independent variable(s)," according to one definition. "Total variance explained by model divided by total variance" is another definition. If it is 100%, the 02 variables are densely connected, so that there is absolute no variation. A lower value shall represent a less similarity, which would represent that the regression is invalid.

Table 1 Comparative analysis of AFODL approach with various measures and Batch size

Methods	**MAE**		**RMSE**		**R^2**	
	1st Day	**7th Day**	**1st Day**	**7th Day**	**1st Day**	**7th Day**
Batch Size = 24						
GRU	10.180	12.558	16.671	21.316	0.984	0.956
LSTM	8.982	11.726	16.249	19.205	0.980	0.971
Bi-LSTM	9.337	12.014	15.320	20.090	0.986	0.972
CNN-LSTM	9.148	9.553	16.603	16.824	0.985	0.980
CNN-GRU	9.612	9.592	17.654	18.494	0.977	0.963
AFODL-APM	8.382	8.883	14.610	16.194	0.994	0.989
Batch Size = 32						

GRU	9.811	10.955	16.429	18.232	0.977	0.973
LSTM	9.663	11.373	16.337	18.904	0.988	0.970
Bi-LSTM	8.738	11.223	15.627	18.308	0.985	0.982
CNN-LSTM	6.672	8.804	13.041	16.345	0.989	0.984
CNN-GRU	9.489	9.609	17.390	18.720	0.981	0.967
AFODL-APM	5.912	8.014	12.291	15.615	0.997	0.994
Batch Size = 64						
GRU	10.191	11.998	16.860	20.007	0.985	0.963
LSTM	9.512	11.611	16.374	18.945	0.986	0.975
Bi-LSTM	9.181	12.405	16.730	19.153	0.988	0.976
CNN-LSTM	8.049	9.615	15.437	18.493	0.988	0.982
CNN-GRU	9.583	9.873	17.166	18.156	0.987	0.985
AFODL-APM	7.259	9.105	14.707	17.426	0.998	0.991
Batch Size = 128						
GRU	9.502	11.739	16.624	19.160	0.987	0.978
LSTM	9.575	11.953	16.767	19.087	0.978	0.973
Bi-LSTM	9.228	11.353	16.646	17.853	0.978	0.969
CNN-LSTM	8.740	8.867	16.467	16.844	0.986	0.983
CNN-GRU	9.659	10.219	17.085	17.675	0.980	0.975
AFODL-APM	8.240	8.217	15.677	16.214	0.996	0.989

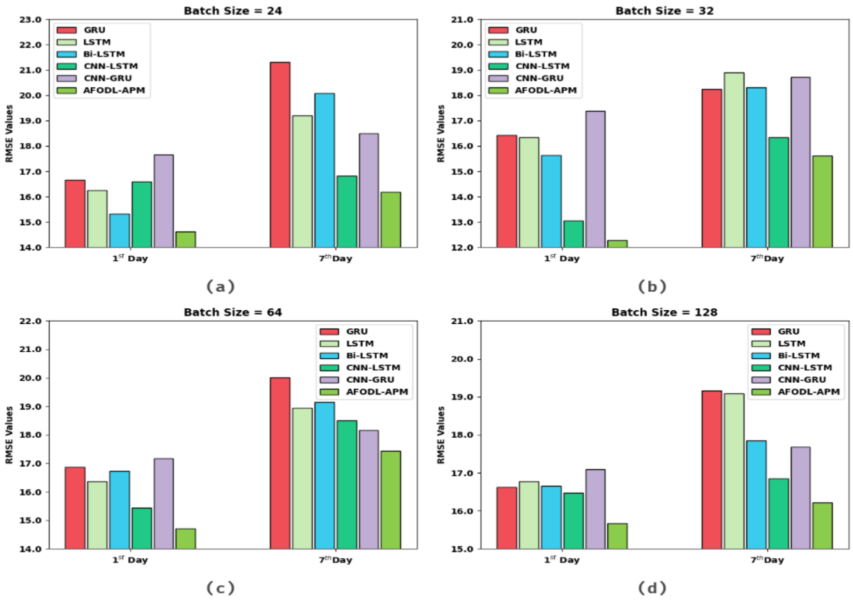

Fig. 3. RMSE analysis of AFODL-APM technique on 1st and 7th day (a) BS=24, (b) BS= 32, (c) BS=64, and (d) BS=128

Along with that, on 1st day and BS of 32, the AFODL-APM model has offered decreased RMSE of 12.291 whereas the GRU, LSTM, BiLSTM, CNN-LSTM, and CNN-GRU systems have obtained increased RMSE of 16.429, 16.337, 15.627, 13.041, and 17.390 correspondingly. Meanwhile, on 7st day and BS of 32, the AFODL-APM system has obtainable lower RMSE of 15.615 whereas the GRU, LSTM, BiLSTM, CNN-LSTM, and CNN-GRU methodologies have obtained increased RMSE of 18.232, 18.904, 18.308, 16.345, and 18.720 respectively. Eventually, on 1st day and BS of 64, the AFODL-APM approach has accessible reduced RMSE of 14.707 whereas the GRU, LSTM, BiLSTM, CNN-LSTM, and CNN-GRU techniques have obtained maximal RMSE of 16.860, 16.374, 16.730, 15.437, and 17.166 correspondingly. Along with that, on 1st day and BS of 128, the AFODL-APM model has offered reduced RMSE of 15.677 whereas the GRU, LSTM, BiLSTM, CNN-LSTM, and CNN-GRU models have obtained increased RMSE of 16.624, 16.767, 16.646, 16.467, and 17.675 respectively. At last, on 7st day and BS of 128, the AFODL-APM method has presented lesser RMSE of 16.214 whereas the GRU, LSTM, BiLSTM, CNN-LSTM, and CNN-GRU algorithms have obtained enhanced RMSE of 19.160, 19.087, 17.853, 16.844, and 17.675 correspondingly.

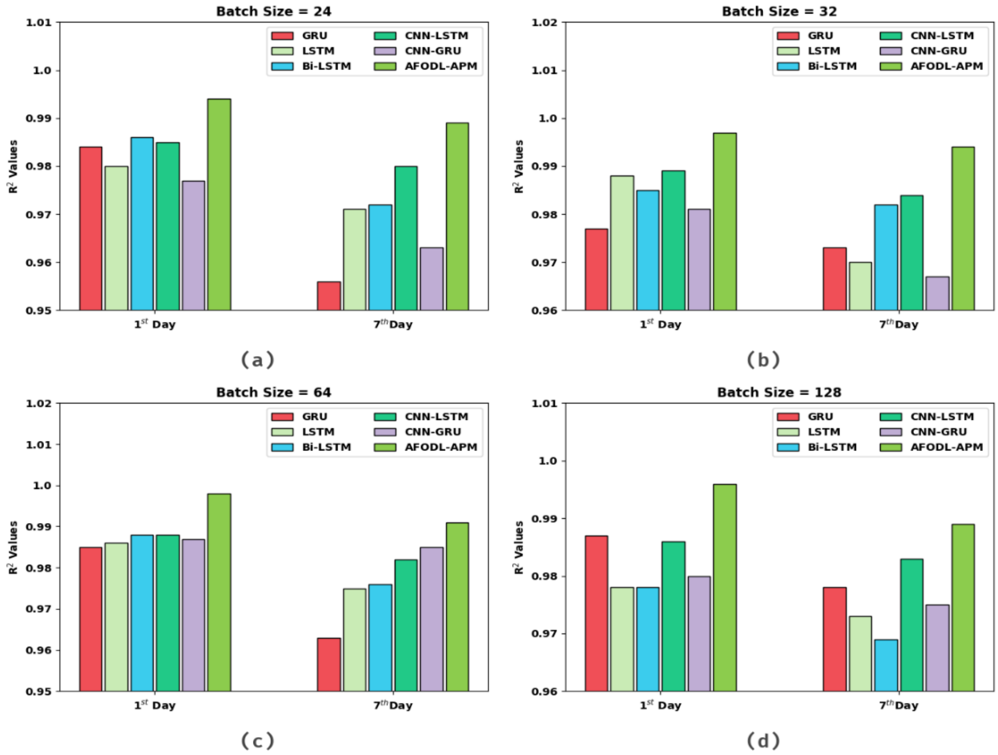

Fig. 4. R^2 values analysis of AFODL-APM technique on 1st and 7th day (a) BS=24, (b) BS= 32, (c) BS=64, and (d) BS=128

Next, Fig. 4 showcases the R^2 examination of the AFODL-APM model with other existing approaches. For sample, on 1st day and BS of 24, the figure pointed out that the AFODL-APM model has gained maximum R^2 value of 0.994 whereas the GRU, LSTM, BiLSTM, CNN-LSTM, and CNN-GRU models have resulted to minimal R^2 values of 0.984, 0.980, 0.986, 0.985, and 0.977 respectively. In addition, on 7st day and BS of 24, the figure mentioned that the AFODL-APM model has gained higherR^2 value of 0.989 whereas the GRU, LSTM, BiLSTM, CNN-LSTM, and CNN-GRU models have resulted to lower R^2 values of 0.956, 0.971, 0.972, 0.980, and 0.963 correspondingly. Also, on 1st day and BS of 32, the figure indicated that the AFODL-APM algorithm has gained enhancedR^2 value of 0.997 whereas the GRU, LSTM, BiLSTM, CNN-LSTM, and CNN-GRU approaches have resulted indecreased R^2 values of 0.977, 0.988, 0.985, 0.989, and 0.981 correspondingly. Concurrently, on 7st day and BS of 32, the figure exposed that the AFODL-APM model has gained maximum R^2 value of 0.994 whereas the GRU, LSTM, BiLSTM, CNN-LSTM, and CNN-GRU models have resulted to lower R^2 values of 0.973, 0.970, 0.982, 0.984, and 0.967 respectively. In the same way, on 1st day and BS of 64, the figure revealed that the AFODL-APM model has gained improvedR^2 value of 0.998 whereas the GRU, LSTM, BiLSTM, CNN-LSTM, and CNN-GRU models have resulted to lesser R^2 values of 0.985, 0.986, 0.988, 0.988, and 0.987 correspondingly. Followed by, on 1st day and BS of 128, the figure noted that the AFODL-APM approach has reached superiorR^2 value of 0.996 whereas the GRU, LSTM, BiLSTM, CNN-LSTM, and CNN-GRU methodologies have resulted to reduced R^2 values of 0.987, 0.978, 0.986, 0.980, and 0.996 correspondingly.

Table 2 and Fig. 5 offers actual vs predicted PM2.5 values of the AFODL-APM model under distinct time step. The experimental results portrayed that the AFODL-APM model has shown closer predicted outcomes over actual ones. For instance, with time step of 10hrs and actual PM2.5 value of 103.52, the AFODL-APM model has predicted value of 103.94. Then, with time step of 40hrs and actual PM2.5 value of 162.08, the AFODL-APM approach has predicted value of 158.88. Followed by, with time step of 80hrs and actual PM2.5 value of 33.55, the AFODL-APM methodology has predicted value of 37.07.

Table 2 PM2.5analysis of AFODL-APM technique under distinct time steps

PM2.5		
Time Step (hrs)	**Actual**	**Predicted**
0	5.41	8.29
10	103.52	103.94
20	102.76	106.85
30	162.08	160.23
40	162.08	158.88
50	5.41	6.43
60	51.04	51.70
70	84.50	87.66
80	33.55	37.07
90	60.93	58.29
100	108.84	104.04
110	152.19	152.61
120	111.88	115.29
130	13.77	16.55
140	8.45	4.91
150	39.63	44.13
160	64.73	63.75
170	69.29	70.25
180	20.62	16.80
190	13.77	18.61
200	6.93	4.22
210	9.21	12.11
220	19.86	23.74
230	25.94	21.35
240	55.60	53.21

Next, with time step of 100hrs and actual PM2.5 value of 108.84, the AFODL-APM model has predicted value of 104.04. Afterward, with time step of 150hrs and actual PM2.5 value of 39.63, the AFODL-APM system has predicted value of 44.13. In line with, with time step of 200hrs and actual PM2.5 value of 6.93, the AFODL-APM algorithm has predicted value of 4.22. At last, with time step of 240hrs and actual PM2.5 value of 55.60, the AFODL-APM model has predicted value of 53.21.

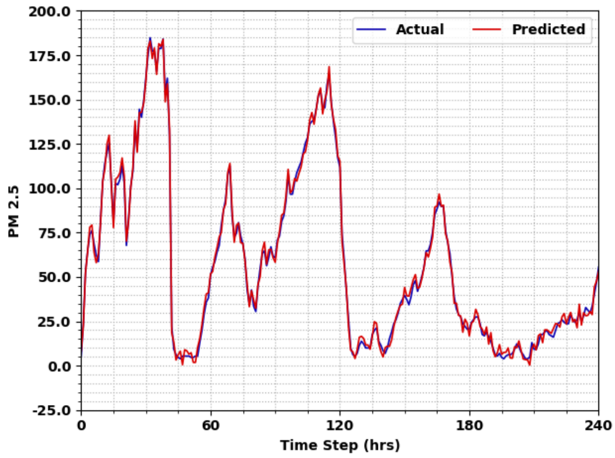

Fig. 5.PM2.5analysis of AFODL-APM technique under distinct time steps

Based on the results from the experiments and analysis, it can be inferred that the proposed neural network has the ability of predicting AQI effectively over the other existing AQI prediction models.

Table 3. Model Performance analysis with different feature set

Features	MAE	RMSE	R^2
All Eight Features	8.217	16.214	0.979
05 Features excluding rain	8.104	15.989	0.982
04 Features excluding rain, and pressure	8.087	15.786	0.988
03 Feature excluding temperature, rain, and pressure	7.956	15.156	0.992

The correlation coefficient between each attribute and PM2.5 concentration was determined for the Beijing PM2.5 dataset. Table 1 demonstrates a positive link between dew point, and wind direction, and a negative correlation between temperature, air pressure, wind speed, and rainfall and PM2.5 concentration. All of the meteorological variables are discovered to have modest correlations with one another, proving that there is no information duplication among them and allowing them to be used directly as input for the prediction model.

5. Conclusion

In this study, a novel AFODL model has been developed for the prediction of value of air pollutant by using hyperparameter optimized deep learning model. The presented AFODL model was trained on pre-processed input data. The presented AFODL model utilized DBN for performing the predictive analytics. The AFO algorithm is employed for optimizing the hyperparameters value of the DBN model which shall help the model to enhance the predictive performance. The experimental evaluation of the AFODL based network is evaluated and the results are inspected under different scenario. The simulation results reported the enhancements of the AFODL model over existing methods. Thus, the proposed AFODL model can be utilized for efficient air pollutant prediction in real time. In future, feature selection process can be combined with the prediction network to enhance the overall efficiency of the AFODL model.

References

[1] Shetty, C., Sowmya, B.J., Seema, S. and Srinivasa, K.G., 2020. Air pollution control model using machine learning and IoT techniques. In Advances in Computers (Vol. 117, No. 1, pp. 187-218). Elsevier.

[2] Srivastava, C., Singh, S. and Singh, A.P., 2018, September. Estimation of air pollution in Delhi using machine learning techniques. In 2018 International Conference on Computing, Power and Communication Technologies (GUCON) (pp. 304-309). IEEE.

[3] Desai, N.S. and Alex, J.S.R., 2017, March. IoT based air pollution monitoring and predictor system on Beagle bone black. In 2017 International Conference on Nextgen Electronic Technologies: Silicon to Software (ICNETS2) (pp. 367-370). IEEE.

[4] Dhanalakshmi, M., 2021. A Survey paper on Vehicles Emitting Air Quality and Prevention of Air Pollution

by using IoT Along with Machine Learning Approaches. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12(11), pp.5950-5962.

[5] Montanaro, T., Sergi, I., Basile, M., Mainetti, L. and Patrono, L., 2022. An IoT-Aware Solution to Support Governments in Air Pollution Monitoring Based on the Combination of Real-Time Data and Citizen Feedback. Sensors, 22(3), p.1000.

[6] Jo, B. and Khan, R.M.A., 2018. An internet of things system for underground mine air quality pollutant prediction based on azure machine learning. Sensors, 18(4), p.930.

[7] Gunasekar, S., 2021. Air pollution forecasting in urban environment using IoT and machine learning scheme. SPAST Abstracts, 1(01).

[8] Kök, I., Şimşek, M.U. and Özdemir, S., 2017, December. A deep learning model for air quality prediction in smart cities. In 2017 IEEE International Conference on Big Data (Big Data) (pp. 1983-1990). IEEE.

[9] Duangsuwan, S., Takarn, A. and Jamjareegulgarn, P., 2018, September. A development on air pollution detection sensors based on NB-IoT network for smart cities. In 2018 18th International Symposium on Communications and Information Technologies (ISCIT) (pp. 313-317). IEEE.

[10] Mandava, T., Chen, S., Isafiade, O. and Bagula, A., 2018, May. An IoT Middleware for Air Pollution Monitoring in Smart Cities: A Situation Recognition Model. In Proceedings of the IST Africa 2018 Conference, Gabarone, Botswana (pp. 9-11).

[11] Ayele, T.W. and Mehta, R., 2018, April. Air pollution monitoring and prediction using IoT. In 2018 second international conference on inventive communication and computational technologies (ICICCT) (pp. 1741-1745). IEEE

[12] Hussain, A., Draz, U., Ali, T., Tariq, S., Irfan, M., Glowacz, A., Antonino Daviu, J.A., Yasin, S. and Rahman, S., 2020. Waste management and prediction of air pollutants using IoT and machine learning approach. Energies, 13(15), p.3930

[13] Pushpam, V.E. and Kavitha, N.S., 2019, January. Iot enabled machine learning for vehicular air pollution monitoring. In 2019 International Conference on Computer Communication and Informatics (ICCCI) (pp. 1-7). IEEE

[14] Asha, P., Natrayan, L., Geetha, B.T., Beulah, J.R., Sumathy, R., Varalakshmi, G. and Neelakandan, S., 2022. IoT enabled environmental toxicology for air pollution monitoring using AI techniques. Environmental research, 205, p.112574

[15] S. Ali, T. Glass, B. Parr, J. Potgieter and F. Alam, "Low Cost Sensor With IoT LoRaWAN Connectivity and Machine Learning-Based Calibration for Air Pollution Monitoring," in IEEE Transactions on Instrumentation and Measurement, vol. 70, pp. 1-11, 2021, Art no. 5500511, doi: 10.1109/TIM.2020.3034109.

[16] Keyvanrad, M.A. and Homayounpour, M.M., 2014. A brief survey on deep belief networks and introducing a new object oriented toolbox (DeeBNet). arXiv preprint arXiv:1408.3264.Kumar, A.R., Vinoth, R. and Kasinathan, P., 2021.

[17] Opposition Based Artificial Flora Algorithm for Load Balancing in LTE Network. Wireless Personal Communications, 118(1), pp.141-159.Zhang S, Guo B, Dong A, He J, Xu Z, Chen SX. Cautionary tales on air-quality improvement in Beijing. The Royal Society Publishing; 2017. https://archive.ics.uci.edu/ml/datasets/Beijing+Multi-Site+Air-Quality+Data.Bekkar

[18] A., Hssina, B., Douzi, S. and Douzi, K., 2021. Air-pollution prediction in smart city, deep learning approach. Journal of big Data, 8(1), pp.1-21.Li, X.; Peng, L.; Yao, X.; Cui, S.; Hu, Y.; You, C.; Chi, T. Long short-term memory neural network for air pollutant concentration predictions: Method development and evaluation. Environ. Pollut. 2017, 231, 997–1004.

[19] Tsai, Y.-T.; Zeng, Y.-R.; Chang, Y.-S. Air Pollution Forecasting Using RNN with LSTM. In Proceedings of the 2018 IEEE 16th Intl Conf on Dependable, Autonomic and Secure Computing, 16th Intl Conf on Pervasive Intelligence and Computing, 4th Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress(DASC/PiCom/DataCom/CyberSciTech), Athens, Greece, 12–15 August 2018; pp. 1074–1079.

[20] Ma, J.; Ding, Y.; J, L.G.V.; Lin, C.; Wan, Z. Spatiotemporal Prediction of PM2.5 Concentrations at Different Time Granularities Using IDW-BLSTM. IEEE Access 2019, 7, 107897–107907.

[21] Burada, S., Swamy, B. E., & Kumar, M. S. (2022). Computer-Aided Diagnosis Mechanism for Melanoma Skin Cancer Detection Using Radial Basis Function Network. In Proceedings of the International Conference on Cognitive and Intelligent Computing (pp. 619-628). Springer, Singapore.

[22] Prasad, T. G., Turukmane, A. V., Kumar, M. S., Madhavi, N. B., Sushama, C., & Neelima, P. (2022). CNN BASED PATHWAY CONTROL TO PREVENT COVID SPREAD USING FACE MASK AND BODY TEMPERATURE DETECTION. Journal of Pharmaceutical Negative Results, 1374-1381.

[23] M. S Kumar. "Forecasting of Wind Power using LSTM Recurrent Neural Network." Journal of Green Engineering 10 (2020).

[24] Kumar, M. S.& Chevdumoi, R. M. (2022). PREDICTION OF SOIL PH FROM REMOTE SENSING DATA USING GRADIENT BOOSTED REGRESSION ANALYSIS. Journal of Pharmaceutical Negative Results, 29-36.

[25] Kumar, M. Sunil, et al. "Deep Convolution Neural Network Based solution for Detecting Plant Diseases." Journal of Pharmaceutical Negative Results (2022): 464-471.

[26] M. Sunil Kumar, Rizwan Patan and Mohamed Yasin Noor Mohamed. "Segmentation of nuclei in histopathology images using fully convolutional deep neural architecture." In 2020 International Conference on computing and information technology (ICCIT-1441), pp. 1-7. IEEE, 2020.

[27] Sreedhar, B., BE, M. S., & Kumar, M. S. (2020, October). A comparative study of melanoma skin cancer detection in traditional and current image processing techniques. In 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC) (pp. 654-658). IEEE.

[28] Kim, H.S.; Park, I.; Song, C.H.; Lee, K.; Yun, J.W.; Kim, H.K.; Jeon, M.; Lee, J.; Han, K.M. Development of a daily PM10 and PM2.5 prediction system using a deep long short-term memory neural network model. Atmos. Chem. Phys 2019, 19, 12935–12951.

[29] Ayturan, Y.A.; Ayturan, Z.C.; Altun, H.O.; Kongoli, C.; Tuncez, F.D.; Dursun, S.; Ozturk, A. Short-term prediction of PM2.5 pollution with deep learning methods. Glob. Nest J. 2020, 22, 126–131.

[30] Li, S.; Xie, G.; Ren, J.; Guo, L.; Yang, Y.; Xu, X. Urban PM2.5 Concentration Prediction via Attention-Based CNN–LSTM. Appl. Sci. 2020, 10, 1953.

[31] Wu, X.; Wang, Y.; He, S.; Wu, Z. PM2.5/PM10 ratio prediction based on a long short-term memory neural network in Wuhan, China. Geosci. Model Dev. 2020, 13, 1499–1511.

[32] Yang, G.; Lee, H.; Lee, G. A Hybrid Deep Learning Model to Forecast Particulate Matter Concentration Levels in Seoul, South Korea. Atmosphere 2020, 11, 348.

[33] Xayasouk, T.; Lee, H.; Lee, G. Air Pollution Prediction Using Long Short-Term Memory (LSTM) and Deep Autoencoder (DAE) Models. Sustainability 2020, 12, 2570.

[34] Veeramanikandan, V., and M. Jeyakarthic. "Forecasting of Commodity Future Index using a Hybrid Regression Model based on Support Vector Machine and Grey Wolf Optimization Algorithm." International Journal of Innovative Technology and Exploring Engineering (IJITEE) 10.10 (2019): 2278-3075.

[35] Veeramanikandan, V., and M. Jeyakarthic. "A Futuristic Framework for Financial Credit Score Prediction System using PSO based Feature Selection with Random Tree Data Classification Model." 2019 International Conference on Smart Systems and Inventive Technology (ICSSIT). IEEE, 2019.

[36] Veeramanikandan, Varadharajan, and Mohan Jeyakarthic. "Parameter-Tuned Deep Learning Model for Credit Risk Assessment and Scoring Applications." Recent Advances in Computer Science and Communications (Formerly: Recent Patents on Computer Science) 14.9 (2021): 2958-2968.

[37] Veeramanikandan, V., and M. Jeyakarthic. "An ensemble model of outlier detection with random tree data classification for financial credit scoring prediction system." International Journal of Recent Technology and Engineering (IJRTE) 8.3 (2019): 2277-3878.