Deep learning and machine learning based air pollution prediction model for smart environment design planning

Karthikeyan B.1, Mohanasundaram R.2, Suresh P.3 and Jagan Babu J.4
1Department of Information Technology, Panimalar Engineering College, Chennai, India
2School of Computer Science and Engineering, VIT, Vellore, India
3Department of Computer Science and Engineering, KPR Institute of Engineering and Technology, Coimbatore-641407, India
4Department of Electronics and Communication Engineering, R.M.D Engineering College, Kavaraipetaltai, Chennai, India
Received: 07/01/2023, Accepted: 31/01/2023, Available online: 04/02/2023
*to whom all correspondence should be addressed: e-mail: kartikeyanb953@gmail.com
https://doi.org/10.30955/gnj.004735

Graphical abstract

Abstract

For the past few decades, owing to human activities, urbanization, and industrialization, air pollution has become severe across several countries. Deep Learning (DL) and Machine Learning (ML) techniques had great contribution to the development of methods in various aspects of prediction, planning, and uncertainty analysis of smart cities and urban advancement in the current scenario. Many of the cities which are developed suffered from severe air quality (AQ) because of the rapid growth in industrialization and population. In this paper, we introduce a deep learning based air pollution prediction model for smart environment design planning (DLAPP-SEDP). The presented DLAPP-SEDP technique majorly intends to predict the level of air pollution in the smart environment. It follows a three-stage process namely data pre-processing, air pollution prediction, and hyperparameter tuning. At the initial stage, the presented DLAPP-SEDP technique performs various levels of data pre-processing such as missing value replacement, categorical value encoding, normalization, and feature selection. In the next stage, the DLAPP technique employs graph convolutional network (GCN) model. Finally, the DLAPP-SEDP technique utilizes atomic orbital search optimization (AOSO) algorithm for optimal hyperparameter tuning process, showing the novelty of the work. To demonstrate the enhanced predictive efficiency of the DLAPP-SEDP method, a wide-ranging experimental analysis can be carried out. The experimental values assured the enhancements of the DLAPP-SEDP method over other recent techniques.

Keywords: Air pollution monitoring, smart environment, sustainability, deep learning, parameter optimization

1. Introduction

Urban areas and people who were living in those areas are often affected by environmental factors. It imposes novel issues for urban planners, like enhancing the air quality (AQ) and minimizing sound levels for building a friendly and clean environment for the population of a city. Additionally for avoiding adverse effects on the enterprises and residents, like dense snowstorms or bullying, severe meteorological conditions in a city should be controlled appropriately (Iskandaryan et al., 2020). More stringent checks and tests, in the central areas of town car bans, Encouraging e-mobility, Greening of the city. Such things for enhancing AQ in cities and minimize the sound level. The vast and most unsatisfied capability of smart city technology for boosting the living standards (Ullo and Sinha, 2020). In the ecological sector-major variations have done and beyond the benefits of employment, safety, standard of housing, energy, interconnectivity, and wellness.

Air pollution contains an extensive range of impacts on humans which involve early death, breathing conditions, and pulmonary disease clinic (Masih, 2019). Ozone and nitrogen dioxide gas majorly affect individuals with conditions like liver cancer, asthma, and respiratory problems that make the disease severe. The Air Quality Index (AQI) can be utilized to measure AQ. These metrics
have a huge effect on the air pollution level and were computed by sensors in various SCs. The sensor systems connection in cities produces various data that are annotated in timely manner. Deep Learning (DL) forecasts large quantities of data (Sai et al., 2019). The important reason for success of DL was the enhanced chip processing capabilities, dramatically dropped cost of networking equipment, and current developments in data processing and artificial intelligence. As it is basically difficult for processing air pollution, its distribution pattern and temporal patterns were affected by several factors which include emissions of and accumulation of traffic flows, air pollutant, climatic conditions, and activities made by man, and many more (Ameer et al., 2019). The issue of utilizing conventional deep methods, mainly for providing quality representative of air pollution features, has enhanced.

The time series pollution data comprises long term dependence amongst all features. By the rapid advancement of machine learning (ML), artificial intelligence methods no longer stay as the existing methods (Al-Janabi et al., 2020). Numerous authors had carried out AQ modeling utilizing DL approaches and proved superior predictive methods compared to ML regarding temporal analysis of the air pollution data. DL methods displayed superior performance in medical image classification, sequential modeling, human detection, and other applications (Harishkumar et al., 2020). DL technique that is Gated Recurrent Unit (GRU), Recurrent Neural Network (RNN), LSTM, had served a significant part in predicting AQ. Some authors included a Convolutional Neural Networks (CNNs) layer with the shallow DL methods for capturing the spatial features in time-series dataset which is available that grants superior predictive performance through examining both the spatio-temporal features (Kabir et al., 2020).

Many prevailing predictive techniques will forecast air pollution levels for the following hours for a specific site. Forecasting air pollution levels for the entire research zone for a longer period could provide a benefit to receiving superior air pollution predictive outcomes (Castelli et al., 2020). Generally, air pollution forecasting performance for a longer time grants less accurateness compared to the shorter period. This may occur because the small number of samples will be performing long term AQ forecast (Janarthanan et al., 2021). Thus, it becomes necessary to advance air pollution predictive methods that could efficiently achieve air pollution forecasting for the whole research area at a more important time resolution.

In this paper, we introduce a deep learning based air pollution prediction model for smart environment design planning (DLAPP-SEDP). The presented DLAPP-SEDP technique performs various levels of data pre-processing such as missing value replacement, categorical value encoding, normalization, and feature selection. In the next stage, the DLAPP-SEDP technique employs graph convolutional network (GCN) model. Finally, the DLAPP-SEDP technique utilizes atomic orbital search optimization (AOSO) algorithm for optimal hyperparameter tuning process. To demonstrate the enhanced predictive efficiency of the DLAPP-SEDP algorithm, a wide-ranging experimental analysis is carried out.

The rest of the paper is organized as follows. Section 2 provides the related works and section 3 offers the proposed model. Then, section 4 gives the result analysis and section 5 concludes the paper.

2. Literature review

Kalajdjeski et al. (2020) present a new technique assessing 4 different structures for estimating the air pollution in those areas by using camera images. Such images were enhanced by meteorological data for boosting the classifier accuracy. The presented technique will exploit generative adversarial networks (GAN) integrated with data augmenting methods for mitigating the class imbalance issue. Du et al. (2019) introduces a new DL algorithm for AQ (mainly PM2.5) prediction that learns interdependence of multivariate AQ related time series data by hybrid DL structure and the spatial-temporal relation features. Owing to the dynamic and nonlinear features of multivariate AQ time sequence data, the base modules of this method add Bi-directional Long Short-term Memory networks (Bi-LSTM) and 1D-CNNs. Previously extracted the spatial correlation features local and trend features, and the later was to study spatial and temporal dependencies.

In (Rao et al., 2019), this work presents a DL technique for prediction and quantification of ambient AQ. RNN-related structure having special structured memory cells called LSTM can be presented for capturing the dependences in several pollutants and performing AQ forecasting. Ma et al. (2019) devises a DL-related technique like transferred bi-directional LSTM (TL-BLSTM) method for predicting AQ. The techniques use the bi-directional LSTM method for learning the longer period dependencies of PM2.5 and implement TL for transferring the knowledge learnt from smaller to larger temporal resolutions. Chang et al. (2020) designed an Aggregated LSTM method (ALSTM) related to the LSTM-DL approach. In this novel technique, the author integrates the stations for external pollution sources, local AQ monitoring stations, and the station in nearby industrial areas. For enhancing prediction accuracy, the author aggregates 3 LSTM methods into a prediction technique for initial forecasting related to exterior sources of pollution and information from neighboring industrial AQ stations.

Le et al. (2020) offer the use of Convolutional LSTM (ConvLSTM) method, the grouping of CNN and LSTM that automatically uses both the spatial-temporal features. Particularly, the author presents the conversion way of the air pollution data into series of images that uses ConvLSTM technique for interpolating and predicting AQ for the whole city in the due course. In (Heydari et al., 2022), a novel hybrid intellectual method related to multi-verse optimization algorithm (MVO) and LSTM was advanced for predicting and analyzing air pollution gained from Combined Cycle Power Plants. In the presented
method, LSTM method becomes forecaster engine for predicting the sum of produced SO2 and NO2 by the integrated Cycle Power Plant in which MVO technique can be employed for optimizing the LSTM variables for achieving a less prediction error.

3. The proposed model

In this paper, a novel DLAPP-SEDP algorithm was introduced to predict the level of air pollution in the smart environment. It follows a three-stage process namely data pre-processing, GCN based air pollution prediction, and AOSO based hyperparameter tuning. Figure 1 showcases the overall process of DLAPP-SEDP technique.

3.1 Data pre-processing

In the early phase, the proposed DLAPP-SEDP algorithm executes different levels of data pre-processing like normalization, missing value replacement, feature selection, and categorical value encoding (Abdellatif et al., 2021). The values that were missing are sorted by linear spline imputation. The SL(x) equation could adapt to local anomaly without affecting the interpolate values at other points. The equation of the spline linear interpolation function as follows:

\[
SL(x) = f(x) = \frac{x-x_0}{x_{i+1}-x_i} + f(x_i)
\]

(1)

whereas \(x\) denotes the independent variable, \(x_0, x_1, \ldots, x_n\) are well-known values of the spline and \(SL(x)\) represents the linear spline which will interpolate \(f\) at these points.

To enhance the estimation accuracy, the author makes a normalization of the values utilizing the Min-Max normalization. If many features are entering the network for training purposes, discovering the relation among the target output values and those features minimize the difficulty of training and enhances performance. The Pearson correlation becomes the well-known technique utilized for finding the relation between two variables.

3.2 Air pollution prediction using GCN model

For air pollution prediction process, the DLAPP-SEDP technique applied the GCN model. Based on the CNN, GCN refers to a multilayer neural network that functions straightway on graph and intends for extracting higher-level features via aggregating data from the neighborhood of graph node (Sofianos et al., 2021). In GCN, an undirected graph can be generally described by \(G=(V, E)\) with \(V\) and \(E\) representing the set of nodes and edges, correspondingly. The notation \(A\) signifies adjacency matrix of \(G\) that represents the presence of edges among all pairs of the node, and its \(i, j\)th component is evaluated by Eq. (2).

\[
A_{ij} = \begin{cases} 
    e^{-\gamma x_i x_j} & \text{if } x_i \in N(x_j) \text{ or } x_j \in N(x_i), \\
    0 & \text{otherwise} 
\end{cases}
\]

(2)

In Eq. (2), the variable \(\gamma\) is empirically fixed as 0.2 in the experiment, \(x_i\) and \(x_j\) characterize two graph nodes (viz., image region), and \(N(x_i)\) shows the set of neighbors of \(x_i\).

Firstly, to conduct node embedding’s for \(G\), spectral filtering on the graph can be determined that is formulated by the multiplication of signal \(x\) with filter \(g_\theta = \text{diag}(\theta)\) in the Fourier domain

\[
g_\theta \in x = U g_\theta U^T x,
\]

(3)

In Eq. (3), \(U\) indicates the matrix of eigenvector of normalized graph Laplacian \(L=I-D^{-\frac{1}{2}}AD^{-\frac{1}{2}}=U\Lambda U^T\), \(\Lambda\) means a diagonal matrix comprised of the eigenvalue of \(L\), \(\theta\) denotes degree matrix having diagonal component \(D_{ij} = \sum_i A_{ij}\), and \(I\) characterizes identity matrix with appropriate size. Next, \(g_\theta\) is understood as a function of eigenvalue of \(L\), viz., \(g_\theta(\Lambda)\). To decrease the computation cost of Eigen decomposition and estimated \(g_\theta(\Lambda)\) by means of truncated expansion interms of Chebyshev polynomial \(T_k(x)\) up to \(K\)th-order,

\[
g_\theta(\Lambda) = \sum_{k=0}^{K} \theta_k T_k(\Lambda),
\]

(4)

In Eq. (4), \(\theta\) indicates a vector of Chebyshev coefficients, and \(\bar{\Lambda} = \frac{2}{\lambda_{\max}} \Lambda - I\) with \(\lambda_{\max}\) being the maximum eigenvalues of \(L\). Based on Chebyshev polynomial is represented by \(T_k(x) = 2xT_{k-1}(x) - T_{k-2}(x)\), where \(T_0(x) = 1\) and \(T_1(x) = x\).

\[
g_\theta \in x = \sum_{k=0}^{K} \theta_k T_k(\bar{\Lambda}) x,
\]

(5)
In Eq. (5), \( \bar{L} = \frac{2}{\lambda_{\text{max}}} L - I \) show the scaled Laplacian matrix and simply verified based on the fact that \((LUU^T)^k = LU^k U^T \). It is noted that, \( K^m \) order polynomial w.r.t Laplacian (viz., \( K \)-localized). In another word, filtering relies merely on node viz., at \( k \) steps farther from the centralized node. In the CAD-GCN method, first-order neighborhood is taken into account, viz., \( K = 1 \), and therefore Eq. (5) becomes a linear function on graph Laplacian spectrum regarding \( L \). Figure 2 depicts the infrastructure of GCN technique.

![Figure 2. Framework of GCN](image)

Later, a neural network depends on GCN is made through stacking more than one convolution layer, where all the layers are followed by element-wise nonlinear function (viz., softplus ()). In that regard, different classes of convolution filter functions can be derived by stacking more than one layer of a similar configuration. By using linear formula, Kipf and Welling Reject? Reject estimated \( \lambda_{\text{max}} \approx 2 \), consider the network parameter is adapted to this change in scale at the training model as follows

\[
g_{\theta} \in x \approx \theta_0 x + \theta_1 (L - I)x = \theta_0 x - \theta_1 D^{-\frac{1}{2}} A D^{-\frac{1}{2}} x , \quad (6)
\]

In Eq. (6), \( \theta_0 \) and \( \theta_1 \) indicates two free parameters. Because decreasing the parameter count assists to prevent overfitting,

\[
g_{\theta} \in x \approx \theta_0 \left( I + D^{-\frac{1}{2}} A D^{-\frac{1}{2}} \right) x , \quad (7)
\]

By letting \( \theta = \theta_0 = -\theta_1 \). As \( I + D^{-\frac{1}{2}} A D^{-\frac{1}{2}} \) has the eigenvalue range within \([0,2]\), frequently employing these operators will lead to vanishing or exploding gradient and numerical instabilities. To resolve these shortage, the renormalization trick \( I + D^{-\frac{1}{2}} A D^{-\frac{1}{2}} \rightarrow D^{-\frac{1}{2}} A D^{-\frac{1}{2}} \) with \( \tilde{A} = A + I \) and \( \tilde{\theta}_0 = \sum \tilde{A}_0 \). Consequently, the convolutional function of GCN method is formulated by Eq. (8)

\[
H^{(i)} = \sigma \left( \tilde{A} H^{(i-1)} W^{(i)} \right) , \quad (8)
\]

Now \( H^{(i)} \) means the output of \( i \)-th layer, \( \sigma \) characterizes an activation function, namely softplus function applied in CAD-GCN, and \( W^{(i)} \) shows trainable weight matrixes included in the \( l \)-th layer.

3.3. AOSO based Hyperparameter Prediction

Finally, the DLAPP-SEDP technique makes use of AOSO algorithm for optimal hyperparameter tuning process. The AOS is a recently designed optimization algorithm that is stimulated from the laws of quantum technicians whereby the standard arrangement of electrons around the nucleus (Azizi et al., 2021). The AOS can be mathematically expressed in the following.

This study makes use of various solutions (X) as follows, and every solution (X) hold different decision parameters \( (x_{ij}) \).

\[
X = \begin{bmatrix}
x_1 & x_1' & \ldots & x_i & x_i' & \ldots & x_D & x_D' \\
x_2 & x_2' & \ldots & x_i & x_i' & \ldots & x_D & x_D' \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\
x_n & x_n' & \ldots & x_i & x_i' & \ldots & x_D & x_D' \\
\end{bmatrix}
\]

In Eq. (9), \( N \) signifies the used amount of solutions, and \( D \) specify the dimension length of the tested problem.

The initial solution is initialized arbitrarily in the following.

\[
x^{(i)}_j = x^{(i)}_{j,\text{min}} + \text{rand} \times \left( x^{(i)}_{j,\text{max}} - x^{(i)}_{j,\text{min}} \right) , \quad (10)
\]

In Eq. (10), \( x^{(i)}_j \) indicates the \( i \)-th location in the \( j \)-th solution, \( x^{(i)}_{j,\text{min}} \) and \( x^{(i)}_{j,\text{max}} \) specifies the lower and upper bounds of the \( i \)-th and \( j \)-th position.

A vector of energy value comprises the objective function of dissimilar solutions as given below.

\[
E = \begin{bmatrix}
E_1 \\
E_2 \\
\vdots \\
E_m \\
\end{bmatrix} \quad , \quad (11)
\]

In Eq. (11), \( E \) embodies a vector of objective value, and \( E_i \) denotes the energy level of solution \( i \)-th number.

The electron probability density chart describes solution position assessed by the Probability Density Function (PDF). Based on the particular definition of the individual by PDF, every imaginarily formulated layer comprises numerous solutions. With that regard, the mathematical models of the \( K \)-position and \( E_k \) of the individual used n imaginary course is shown in the following:
Now $X^i_k$ indicates the $i$-th number of solutions in the $k$-th IL number, and $n$ signifies the amount of the generated IL. $p$ show the solution number of $k$-th IL number. $E^i_k$ is the objective value of $i$-th number of solutions in the $k$-th IL number.

With that regard, the requisite energy and state are determined for the solution in every IL by examining each solution's average position and objective value and it is mathematically expressed in the following:

$$BS^k = \frac{\sum_{i=1}^{m} X^i_k}{p}$$

(14)

$$BE^k = \frac{\sum_{i=1}^{m} E^i_k}{p}$$

(15)

Here, $BS^k$ and $BE^k$ denotes the requisite state and energy of the layer number $k$, respectively. $X^i_k$ and $E^i_k$ stand for the position and fitness value of the solution number $i$ in $k$-th layer.

Based on the presented item, the required energy and state of atom are described by approximating the mean position and objective value of the solution used:

$$BS = \frac{\sum_{i=1}^{m} X^i}{m}$$

(16)

$$BE = \frac{\sum_{i=1}^{m} E^i}{m}$$

(17)

Let $BS$ and $BE$ be the requisite state and energy of the atom.

The energy level ($E^i_k$) of $TX^i_k$ in every IL is related to the requisite energy of layer ($BE^k$). Assume the energy ratio of existing solution in a specific layer is higher than the requisite energy (viz. $E^i_k \geq BE^k$) hence, the photon emission is assessed. In these rules, the individuals are handling to transfer a photon with a cost of energy assessed by $\gamma$ and $\beta$ to simultaneously provide the requisite location of the atom (BS) and the location of electron with the lowest energy ratio (LE) in the atom and it is shown below:

$$X^i_{k+1} = X^i_k + \frac{\alpha_i(\beta_i \times LE - \gamma_i \times BS)}{k}, \ k = 1, 2, \ldots, n, \ i = 1, 2, \ldots, p$$

(18)

In the above expression, $X^i_k$ and $X^i_{k+1}$ denotes the present and estimated values for $i$-th individuals at $k$-th layers. $\alpha_i$, $\beta_i$, and $\gamma_i$ denotes random vector. Supposing the energy ratio of solution in a specific layer is small when compared to the requisite energy ($E^i_k < BE^k$); then photon consumption is inspected and mathematically expressed in the following:

$$X^i_{k+1} = X^i_k + \frac{\alpha_i \times (\beta_i \times LE^k - \gamma_i \times BS^k)}{k}$$

(19)

While producing a random number ($\mathcal{O}$) for all the individuals and it is valued lesser than $PR$ (i.e., $\mathcal{O} < PR$), the photon number on the solution isn’t possible. As a result, the action of particles among different layers near the nucleus is assessed:

$$X^i_{k+1} = X^i_k + r_i$$

(20)

In Eq. (20), $r_i$ indicates a vector of arbitrary numbers.

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. In this section, the air pollution prediction outcomes of the DLAPP-SEDP model are examined in detail. Table 1 provides an overall prediction performance of the DLAPP-SEDP model under day 1 and day 7 with varying batches.

Figure 3 reports an average MAE inspection of the DLAPP-SEDP model under day 1 and day 7. The figure implied that the DLAPP-SEDP model has attained reduced values of MAE under all aspects. For instance, on batch 8, the DLAPP-SEDP model has attained MAE of 4.604 and 4.270 under days 1 and 7 respectively. Similarly, on batch 16, the DLAPP-SEDP approach has gained MAE of 3.492 and 4.260 under days 1 and 7 correspondingly. Also, on batch 24, the DLAPP-SEDP technique has gained MAE of 5.203 and 3.257 under days 1 and 7 correspondingly. Finally, on batch 32, the DLAPP-SEDP approach has reached MAE of 5.362 and 4.692 under days 1 and 7 correspondingly.
Likewise, on batch 24, the DLAPP and SEDP methodology has gained MSE of 103.565 and 98.970 under days 1 and 7 correspondingly. Also, on day 1 and day 7 of the SEDP approach with distinct measures and runs. The figure implied that the DLAPP approach with distinct runs 4.536 and 4.759 distinct runs.

Average MSE analysis of DLAPP and SEDP methodology has gained MSE of 103.565 and 98.970 under days 1 and 7 correspondingly.

At last, on batch 32, the DLAPP-SEDP method has gained MSE of 105.203 and 86.502 under days 1 and 7 correspondingly. At last, on batch 32, the DLAPP-SEDP method has gained MSE of 105.203 and 86.502 under days 1 and 7 correspondingly.

Figure 3. Average MAE analysis of DLAPP-SEDP approach with distinct runs.

Figure 4 reports an average MSE analysis of the DLAPP-SEDP method under day 1 and day 7. The figure implied that the DLAPP-SEDP approach has reached reduced values of MSE under all aspects. For example, on batch 8, the DLAPP-SEDP methodology has gained MSE of 103.565 and 86.218 under days 1 and 7 correspondingly. Also, on batch 16, the DLAPP-SEDP algorithm has achieved MSE of 97.289 and 92.555 under days 1 and 7 correspondingly. Likewise, on batch 24, the DLAPP-SEDP technique has achieved MAE of 98.773 and 92.448 under days 1 and 7 correspondingly. At last, on batch 32, the DLAPP-SEDP method has gained MSE of 105.203 and 86.502 under days 1 and 7 correspondingly.

Figure 5 reports an average RMSE inspection of the DLAPP-SEDP model under day 1 and day 7. The figure implied that the DLAPP-SEDP model has attained reduced values of RMSE under all aspects. For instance, on batch 8, the DLAPP-SEDP model has attained RMSE of 10.177 and 9.285 under days 1 and 7 respectively. Likewise, on batch...
16, the DLAPP-SED model has attained RMSE of 9.864 and 9.621 under days 1 and 7 correspondingly. Similarly, on batch 24, the DLAPP-SED approach has gained RMSE of 99.270 and 9.615 under days 1 and 7 correspondingly. At last, on batch 32, the DLAPP-SED algorithm has attained RMSE of 99.100 and 9.301 under days 1 and 7 correspondingly.

![Figure 5](image)

**Figure 5.** Average RMSE analysis of DLAPP-SED approach with distinct runs

An average R2 examination of the DLAPP-SED model under varying batches is given in Figure 6. The results inferred that the DLAPP-SED model has gained maximum prediction outcomes. For instance, on batch 8, the DLAPP-SED model depicted R2 of 98.740 and 99.400 under days 1 and 7 respectively. Besides, on batch 16, the DLAPP-SED algorithm has depicted R2 of 98.760 and 99.790 under days 1 and 7 correspondingly. Concurrently, on batch 24, the DLAPP-SED method has depicted R2 of 99.270 and 99.050 under days 1 and 7 correspondingly. Simultaneously, on batch 32, the DLAPP-SED approach has depicted R2 of 99.100 and 99.000 under days 1 and 7 correspondingly.

![Figure 6](image)

**Figure 6.** Average R2 analysis of DLAPP-SED approach with distinct runs

Table 2 offers a brief comparison predictive outcome of the DLAPP-SED model with existing models [18]. Figure 7 exhibits a comparative MAE and RMSE inspection of the DLAPP-SED model with recent models on day 1. The figure demonstrated that the DLAPP-SED model has shown enhanced performance with minimal MAE and RMSE values. With respect to MAE, the DLAPP-SED model has reached minimal MAE of 4.361 whereas the GRU, LSTM, Bi-LSTM, Bi-GRU, CNN, CNN-LSTM, and CNN-GRU models have resulted to maximum MAE of 9.841, 9.396, 9.216, 9.505, 9.622, 8.187, and 9.551 respectively. Likewise, With respect to RMSE, the DLAPP-SED approach has attained minimal RMSE of 9.565 whereas the GRU, LSTM, Bi-LSTM, Bi-GRU, CNN, CNN-LSTM, and CNN-GRU algorithms have resulted in maximum RMSE of 16.826, 16.314, 16.036, 16.453, 17.087, 15.385, and 17.344 correspondingly.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Day-1 MAE</th>
<th>Day-1 RMSE</th>
<th>Day-1 R2</th>
<th>Day-7 MAE</th>
<th>Day-7 RMSE</th>
<th>Day-7 R2</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSTM</td>
<td>9.396</td>
<td>16.314</td>
<td>98.000</td>
<td>11.728</td>
<td>19.198</td>
<td>97.100</td>
</tr>
<tr>
<td>Bi-LSTM</td>
<td>9.216</td>
<td>16.036</td>
<td>98.100</td>
<td>11.816</td>
<td>18.994</td>
<td>97.000</td>
</tr>
<tr>
<td>Bi-GRU</td>
<td>9.505</td>
<td>16.453</td>
<td>98.000</td>
<td>11.664</td>
<td>19.138</td>
<td>97.000</td>
</tr>
<tr>
<td>CNN</td>
<td>9.622</td>
<td>17.087</td>
<td>97.800</td>
<td>10.785</td>
<td>18.409</td>
<td>97.400</td>
</tr>
<tr>
<td>CNN-LSTM</td>
<td>8.187</td>
<td>15.385</td>
<td>98.300</td>
<td>9.327</td>
<td>17.062</td>
<td>97.800</td>
</tr>
<tr>
<td>CNN-GRU</td>
<td>9.551</td>
<td>17.344</td>
<td>97.800</td>
<td>9.743</td>
<td>18.251</td>
<td>97.100</td>
</tr>
</tbody>
</table>

A detailed R2 assessment of the DLAPP-SED model with other models is made in Figure 8. These results indicated the betterment of the DLAPP-SED model with higher R2 value of 99.148. At the same time, the other existing models such as GRU, LSTM, Bi-LSTM, Bi-GRU, CNN, CNN-LSTM, and CNN-GRU models have provided decreased R2 values of 97.800, 98.000, 98.100, 98.000, 97.800, 98.300, and 97.800 respectively.

Figure 9 displays a detailed MAE and RMSE analysis of the DLAPP-SED approach with recent algorithms on day 7. The figure established that the DLAPP-SED technique has shown enhanced performance with minimal MAE and RMSE values. With respect to MAE, the DLAPP-SED approach has gained minimal MAE of 4.153 whereas the GRU, LSTM, Bi-LSTM, Bi-GRU, CNN, CNN-LSTM, and CNN-GRU techniques have resulted in maximum MAE of 12.020, 11.728, 11.816, 11.664, 10.785, 9.327, and 9.743 correspondingly. Also, with respect to RMSE, the DLAPP-SED algorithm has attained minimal RMSE of 9.599.
whereas the GRU, LSTM, Bi-LSTM, Bi-GRU, CNN, CNN-LSTM, and CNN-GRU approaches have resulted in maximum RMSE of 19.751, 19.198, 18.994, 19.138, 18.409, 17.062, and 18.251 correspondingly.

Figure 7. MSE and RMSE analysis of DLAPP-SEDAP approach with existing algorithms under day 1

A comprehensive R2 assessment of the DLAPP-SEDAP approach with other models is made in Figure 10. These results indicated the betterment of the DLAPP-SEDAP method with higher R2 value of 99.048. In the meantime, the other existing methodologies such as GRU, LSTM, Bi-LSTM, Bi-GRU, CNN, CNN-LSTM, and CNN-GRU approaches have offered decreased R2 values of 96.600, 97.100, 97.000, 97.400, 97.800, and 97.100 correspondingly.

Table 3 and Figure 11 provides actual vs prediction outcomes of the DLAPP-SEDAP model under several time step. The experimental values indicated that the DLAPP-SEDAP model has effectually predicted the PM2.5 values. For instance, on 20 time step and actual value of 90.49, the DLAPP-SEDAP model has attained predicted value of 88.73.

Similarly, on 40-time step and actual value of 71.94, the DLAPP-SEDAP method has achieved predicted value of 64.94. In addition, on 60-time step and actual value of 14.08, the DLAPP-SEDAP approach has achieved predicted value of 6.89. Also, on 80-time step and actual value of 53.39, the DLAPP-SEDAP technique has gained predicted value of 57.65. At last, on 100-time step and actual value of 146.13, the DLAPP-SEDAP approach has achieved predicted value of 140.95. These values assured the enhanced air quality prediction performance of the DLAPP-SEDAP model.

Table 3. PM2.5 analysis of DLAPP-SEDAP model under several time steps

<table>
<thead>
<tr>
<th>Time Step</th>
<th>Actual</th>
<th>Predicted</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>5.17</td>
<td>14.03</td>
</tr>
<tr>
<td>20</td>
<td>90.49</td>
<td>88.73</td>
</tr>
<tr>
<td>40</td>
<td>71.94</td>
<td>64.94</td>
</tr>
<tr>
<td>60</td>
<td>14.08</td>
<td>6.89</td>
</tr>
<tr>
<td>80</td>
<td>53.39</td>
<td>57.65</td>
</tr>
<tr>
<td>100</td>
<td>146.13</td>
<td>140.95</td>
</tr>
<tr>
<td>120</td>
<td>2.95</td>
<td>2.54</td>
</tr>
<tr>
<td>140</td>
<td>83.07</td>
<td>79.46</td>
</tr>
<tr>
<td>160</td>
<td>20.75</td>
<td>25.32</td>
</tr>
<tr>
<td>180</td>
<td>27.43</td>
<td>32.61</td>
</tr>
<tr>
<td>200</td>
<td>91.97</td>
<td>94.74</td>
</tr>
<tr>
<td>220</td>
<td>66.75</td>
<td>61.25</td>
</tr>
<tr>
<td>240</td>
<td>47.46</td>
<td>45.32</td>
</tr>
<tr>
<td>250</td>
<td>38.56</td>
<td>47.33</td>
</tr>
</tbody>
</table>
5. Conclusion

In this paper, a novel DLAPP-SEDP technique has been introduced to predict the level of air pollution in the smart environment. It follows a three-stage process namely data pre-processing, air pollution prediction, and hyperparameter tuning. At the initial stage, the presented DLAPP-SEDP technique performs various levels of data pre-processing such as missing value replacement, categorical value encoding, normalization, and feature selection. In the next stage, the DLAPP-SEDP technique applied the GCN model. In the last stage, the DLAPP-SEDP technique makes use of AOSO algorithm for optimal hyperparameter tuning process. To demonstrate the enhanced predictive efficiency of the DLAPP-SEDP technique, a wide-ranging experimental analysis is carried out. The experimental values assured the enhancements of the DLAPP-SEDP technique over other recent approaches. In future, the proposed model can be extended to the IoT enabled air pollution monitoring system in real time.

References


