RESEARCH ARTICLE

Neuro-Inspired Alert System for Air Quality Prediction Using Ensemble Preprocessing and SNN Classification

Sneh Sharma[1], **Kashish Devgan**[1], **Devanshi Jangra**[1], **Aanshi Bhardwaj**[1,*] and **Shubhani Aggarwal**[2]

[1] Amity School of Engineering and Technology, Amity University Punjab, Mohali, India
[2] School of Computer Science, University of Petroleum and Energy Studies (UPES), Dehradun, India

Abstract

Air pollution has emerged as a critical challenge, directly affecting human health, urban sustainability, and climate systems. Traditional air-quality index (AQI) prediction models often struggle to provide timely alerts because they are not very sensitive to changes over time and are hard to understand. This paper proposes a Neuro-Inspired Alert System for Air Quality Prediction (NAS–AQP) that incorporates an ensemble learning approach using voting regression to enhance input quality, followed by classification through a Spiking Neural Network (SNN). The system is designed such that it captures the temporal and nonlinear relationships between air pollutants such as Nitrogen Dioxide (NO_2), Sulphur Dioxide (SO_2), Respirable Suspended Particulate Matter (RSPM) and Suspended Particulate Matter (SPM). The proposed method starts with preprocessing of the data and normalizing the features. After that, models like Linear Regression (LR), Random Forest (RF), and Decision Tree (DT) are trained and evaluated using Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Coefficient of Determination (R^2) metrics. After the training of above models, a voting based ensemble approach is used to improve AQI regression accuracy. A threshold based rule is then used to convert received, AQI predictions into binary alerts. Finally, SNN is trained to classify these alerts to achieve energy, efficient, real time, alerting by using its temporal coding and sparse activation. The ensemble voting regression model achieved an RMSE of 8.43 and MAE of 6.21, while the SNN classifier attained a classification accuracy of 92.4%.

Keywords: air quality index (AQI), ensemble learning, voting regression, spiking neural network (SNN), real-time prediction, environmental monitoring, respirable suspended particulate matter (RSPM), suspended particulate matter (SPM), temporal coding, machine learning, neuro-inspired systems.

1 Introduction

Air pollution is a global public health crisis, affecting millions of people on a daily basis across all forms of society. According to the World Health Organization (2024) reports, which highlight the annual statistics, it has been stated that over 7 million premature deaths are attributed to air pollution each year

Submitted: 20 June 2025
Accepted: 29 July 2025
Published: 12 September 2025

Vol. 1, **No.** 2, 2025.
10.62762/TMI.2025.403059

*Corresponding author:
Aanshi Bhardwaj
bhardwajaanshi@gmail.com

Citation
Sharma, S., Devgan, K., Jangra, D., Bhardwaj, A., & Aggarwal, S. (2025). Neuro-Inspired Alert System for Air Quality Prediction Using Ensemble Preprocessing and SNN Classification. *ICCK Transactions on Machine Intelligence*, 1(2), 69–79.

and this number continues to increase in low and middle-income countries [1]. According to the IQAir 2024 Global Report, over 90% of the world's population lives in areas that do not even meet the World Health Organization's clean air quality standards [2]. In large urban centers like Delhi, the AQI levels often range between 150 and 200, falling in the "Moderate" to "Poor" category. The consequences of all this lead to school shutdowns, public health alerts, and a surge in respiratory illnesses [3]. For instance, Kim et al. [5] proposed a deep convolutional neural network model capable of extracting features from complex spatial data, such as satellite imagery or pollution distribution heatmaps, offering potential tools for spatial analysis and visualization in urban air quality monitoring to support more precise public health alerts.

The major causes of poor air quality are vehicular emissions, industrial activities, construction dust, and agricultural burning. Pollutants such as NO_2, SO_2, and RSPM and SPM play a key role in this. RSPM refers to Respirable Suspended Particulate Matter, which includes fine particles capable of penetrating deep into the lungs. SPM, or Suspended Particulate Matter, consists of both coarse and fine particles that remain suspended in the air and contribute significantly to air pollution and respiratory diseases. These lead to asthma, lung cancer, stroke, and cardiovascular diseases. The most vulnerable populations are children, the elderly, and individuals with pre-existing health conditions. While the AQI is presented as a standard for reporting pollution levels to the public, Real-time prediction of AQI remains an open challenge in science and technology. Although traditional machine learning models perform well with consistent patterns, they struggle with the nonlinearity, dynamic nature, and time-dependent aspects of environmental data. They also involve high computational costs, limited interpretability, and slow adaptability, making them less suitable for real-world conditions—especially when data is sparse and quick decisions are needed.

1.1 Why SNNs?

NAS–AQP combines ensemble machine learning methods with SNNs—a recent neural architecture that mimics the brain's way of processing information through discrete electrical spikes. SNNs belong to the third generation of neural network development and are gaining prominence for their ability to handle temporal data more effectively than previous models. Unlike the traditional Artificial Neural Networks that use continuous activation functions, SNNs take input as a continuous sequence of spikes over time and closely resemble the way biological neurons operate in the brain. This allows the system to detect and learn temporal patterns, making it well-suited for systems like pollution sensors, where a significant amount of data fluctuates, and the real-time alerts are essential.

SNNs offer a low-power computing solution that is ideal for embedded sensor devices and IoT platforms. Their sparse activation reduces the energy consumption while maintaining the responsiveness and making them highly scalable for city infrastructures, which often deploy thousands of sensors that must function reliably, simultaneously, and smoothly. In the alert stage of the proposed system, the integration of the SNN model enabled real-time classification of AQI data into public alerts (for example, if AQI is greater than 150, then it will be treated as a health warning). This allows cities to generate real-time reports and respond to pollution spikes with public health guidance or control measures accordingly.

1.2 System Overview

The proposed NAS–AQP system operates in two main stages:

- Ensemble in the Regression Phase – The first step is to train a set of models, including Linear Regression, Decision Trees, and Random Forest, using performance metrics such as RMSE, MAE, and R^2. These models will then be used in the ensemble voting regression approach. The above mentioned models are then combined using a voting regressor, which harnesses the best of each model's capabilities and helps to improve AQI predictions.

- Neuro-Infused Alert System – Predicted AQI values are converted into binary alerts based on predefined threshold levels (e.g., AQI $>$ 150 $\rightarrow$ Alert = 1). An SNN is then trained to classify these alerts, capturing temporal dependencies and enabling rapid responses to sudden environmental changes with minimal computational effort.

1.3 Motivation

The increasing burden of air pollution on global health, environmental sustainability, and urban infrastructure calls for the development of intelligent systems that can offer timely, accurate, and interpretable predictions. Traditional forecasting approaches often fall short in addressing the complexity of real-world pollution data, which includes nonlinear pollutant interactions, varying meteorological influences, and real-time alerting needs. To tackle the above challenges, a comprehensive system has been proposed that combines the robustness of ensemble machine learning models with the adaptability and efficiency of SNN. The motivation to incorporate SNN stems from their ability to replicate the temporal computation of biological neurons, thus allowing for temporal pattern recognition and energy-efficient computation.

The findings of this study are as follows:

- Creation of an ensemble model for the prediction of AQI using various regression models for accuracy improvement.

- The addition of an SNN layer enables biologically inspired binary classification of air quality alerts.

1.4 Outlook

As air quality is continuously declining and cities are growing smarter day by day, there is an increasing need for intelligent, scalable, and energy-efficient environmental monitoring systems. The NAS–AQP model addresses this challenge by combining the strengths of ensemble learning and spiking neural networks. This hybrid system is not only accurate but also practical for deployment in edge-computing environments such as traffic signal poles, public transportation systems, and handheld monitoring devices—bringing society a step closer to a cleaner and healthier future.

2 Literature Review

Recent studies report a shift away from the use of classic statistical models toward hybrid and deep learning approaches, which have proven to perform better at handling nonlinearity, high dimensionality, and time-dependent elements in environmental data. Initially, research into hybrid models in environmental monitoring primarily focused on predicting water quality. In this area, it has been observed that the performance of decision trees in combination with other algorithms, improves not only predictive accuracy but also model interpretability [4]. The early success in water quality forecasting served as a base for the application of similar hybrid approaches to air quality, which, as is known, has large-scale and spatial variation of pollutants and thus requires models with memory and spatial awareness.

In the case of environmental sensor data, which is the focus of this study, it is observed that deep learning methods, in particular Convolutional Neural Networks (CNNs), perform well at extracting complex spatial features from data. Also, when examining atmospheric variables, studies show that CNNs have played an important role in identifying spatial pollutant patterns across monitoring stations and at regional scales. In many real-world forecasting issues, these spatial features alone are not enough, especially in multi-hour and multi-site cases. To address this, researchers have proposed hybrid models that leverage CNNs for spatial features and combine them with Long Short Term Memory (LSTM) or BiLSTM networks, which in turn perform very well at modeling temporal elements related to pollutant trends [6]. Thus, the CNN–LSTM combinations are most effective at identifying both short term and long-term air pollution trends, making them very useful for real-time and future forecasting.

At the same time, some models go beyond prediction; they also perform interpolation and feature relevance analysis, which are essential for addressing issues of data completeness and improving model transparency [7]. This is especially relevant when dealing with missing values or sparse sensor distribution. Additionally, advanced architectures have been found to include Deep Distributed Fusion Networks, which not only combine spatial and temporal data from multiple sources but also help reduce feature loss and improve the model's generalization capacity [8].

In addition to deep neural networks, there is widespread use of supervised machine learning algorithms in environmental prediction. The performance of these models depends mainly on the quantity and quality of the data, as well as the feature engineering involved in model development. Algorithms such as Support Vector Regression (SVR), Random Forests (RF), and Gradient Boosted Trees perform well in this domain and often set a benchmark that is hard to surpass [9]. It has also been found that hybrid deep learning models outperform traditional statistical methods like AutoRegressive Integrated Moving Average (ARIMA) and linear regression in modeling the complex challenges of urban and industrial air pollution [10].

Ensembles of multiple models have also been shown to be effective with complex, high-dimensional datasets. Ensemble models using RF and LSTM, not only improve prediction accuracy but also provide more stable results across various air quality scenarios [11]. Additionally, large-scale reviews of machine learning in air quality emphasize the value of integrating various environmental datasets (e.g., meteorological data, geographic information, traffic data) and recommend the use of evaluation metrics such as MAE, RMSE, and R^2 to better assess model performance in real-world applications [12].

Adaptive deep learning models have demonstrated promising results in real-time forecasting. These models continuously learn the most relevant spatial and temporal relationships, enabling them to generate responsive, context-aware outputs that are valuable for timely public health actions and crucial environmental policy decisions [13]. Transfer learning also plays a key role in this field. By pre-training models on large historical datasets and fine-tuning them for specific regions, researchers have significantly improved prediction accuracy over long timeframes and across diverse environmental conditions [14].

Recent advancements have emphasized energy-efficient edge-based air quality monitoring. Studies such as Pazhanivel et al. [15] and Firoozi et al. [16] propose lightweight CNNs and neuromorphic processors for embedded AQI prediction, highlighting the relevance of the proposed NAS–AQP system in low-power smart city applications.

3 Methodology

Table 1. AQI calculation.

AQI Range	Category	Health Implication
0-50	Good	Minimal Impact
51-100	Satisfactory	Minor Breathing Discomfort
101-200	Moderate	Breathing Discomfort on prolonged exposure
201-300	Poor	Respiratory illness possible
301-400	Very Poor	Increased discomfort and illnesses
401-500	Severe	Emergency conditions; serious risk

The methodology of the study involves a hybrid modeling approach to predict the air quality index. The study begins with the collection of data across various Indian states, specifically focusing on pollutants such as SO_2, NO_2, RSPM, SPM, and associated metadata, including state, location, and sampling date. Central Pollution Control Board (CPCB) guidelines are used for AQI calculation with sub-indices calculated for each pollutant ($\mathrm{SO_i}$, $\mathrm{No_i}$, $\mathrm{Rp_i}$, and $\mathrm{SPM_i}$), and the final AQI is determined as the maximum of these values. The AQI is then categorized into six categories as shown in Table 1: Good (0-50), Satisfactory (51-100), Moderate (101-200), Poor (201-300), Very Poor (301-400), and Severe (401-500). Exploratory Data Analysis (EDA) involved generating visualizations such as bar graphs of the average levels of SO_2 and NO_2 by state and state-wise AQI classifications. The hybrid pipeline used for AQI prediction and alert generation consists of the following six steps:

Pipeline Overview (step-by-step)

Step 1: Data Collection – Gather pollution data (SO_2, NO_2, RSPM, SPM) and metadata (state, location, date).

Step 2: Data Preprocessing – Clean data, handle missing values, encode categorical variables, normalize features.

Step 3: AQI Calculation – Use CPCB standards to compute sub-indices for each pollutant and derive the final AQI.

Step 4: Exploratory Data Analysis (EDA) – Visualize pollution levels and AQI trends across Indian states.

Step 5: Model Training and Evaluation – Train ensemble ML models (Random Forest, Linear Regression, Decision Trees) and evaluate using RMSE, MAE, and R^2.

Step 6: Binary Classification with SNN – Use AQI thresholds to generate alert labels and train a Spiking Neural Network for real-time classification; evaluate using accuracy and F1 score.

The hybrid pipeline for air quality prediction follows six steps, as shown in Figure 1. The study will first collect air quality index data, i.e., NO_2, SO_2, RSPM, and SPM data (high source-level noise addressing). After pre-processing (removing any missing values, normalizing, and building raw input into something meaningful), the machine learning models, such as Linear Regression, Decision Trees, and Random Forest, are trained and evaluated in the next step. Several metrics like RMSE, MAE, and R^2 are utilized for the evaluation of the model. Voting is used to enhance prediction accuracy as ensemble methods are very useful in supervised machine learning applications. The AQI outputs classified into binary alerts based on their threshold level are used to train SNN for real-time classification, which utilizes spike-based computation models. The SNN will then be evaluated for accuracy and F1 score and compared to traditional models. The hybrid deployment will combine ensemble models, which are used in step 2, and an SNN to provide accurate and timely air quality for alert purposes.

3.1 Data Collection

This study utilized air pollution measurements collected from various Indian states. The relevant counterparts are the concentration levels of SO_2, NO_2, RSPM, and SPM. In addition to the pollutant concentrations, the data include metadata such as the state, city or location, and date of sampling. These variables were obtained from government air quality monitoring systems, providing real-time snapshots of atmospheric health over a range of geographic locations in India. This large dataset forms the foundation for both the regression and classification components of the proposed system [17].

3.2 Data Preprocessing

Preprocessing of the raw data is really important in preparation for data to be used in modeling. Columns such as monitoring station codes and agency codes were removed in the initial step. Missing values in numerical columns were imputed with zero, but the categorical variables were handled with mode imputation to keep them statistically significant. Categorical variables such as 'state' and 'location' were transformed into numeric data using label encoding methods to make them compatible for further training. Features were normalized using the StandardScaler to keep all inputs equally scaled in one range, which helps improve the performance of ensemble models and spiking networks.

3.3 AQI Calculation

CPCB standards were taken into consideration to calculate AQI from pollutant specific sub-indices. For each pollutant, such as SO_2, NO_2, RSPM and SPM, a sub-index was defined based on its concentration in the air; the total AQI was the maximum of the sub-indices. The calculated values were also categorized into six severity ranges: good, satisfactory, moderate, poor, very poor, and severe. This made the classification easier to convey the risk level and facilitated the conversion of continuous values into binary alert labels for further classification.

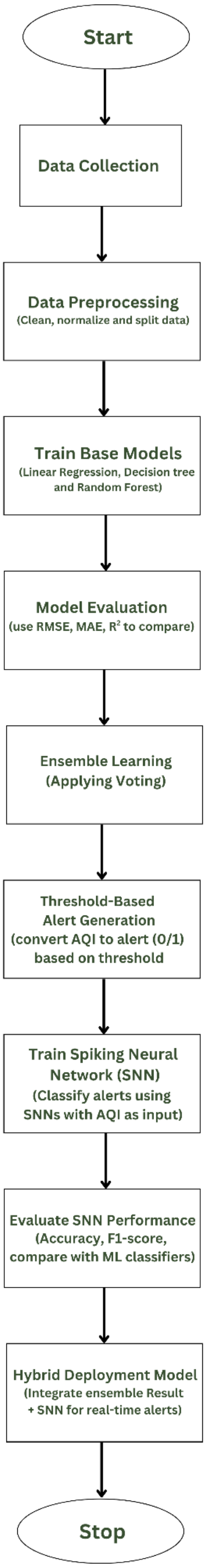


Figure 1. Hybrid AQI prediction and alert pipeline.

3.4 Feature Engineering

To improve the model, we need to find relationships between air pollutants and AQI. Different feature engineering methods were used. Polynomial scaling was used to create non-linear terms from raw pollutant concentrations, so that the regression models could learn curvilinear trends. Logical threshold flags were also created to warn when certain pollutants exceeded CPCB-defined safety limits. These calculated features added more meaning to the learning algorithms and helped improve regression accuracy.

3.5 Exploratory Data Analysis (EDA)

An EDA was conducted to realize spatial patterns of pollution and inter-regional variation. Bar plots and statistical summaries were used as visual tools for representing the distribution of pollutants by state. For example, the average levels of SO_2 and NO_2 were contrasted across states and the AQI classes were analyzed in relation to their location. The examination indicated that Delhi, West Bengal, and Uttar Pradesh consistently reported high pollution levels with regular occurrences of "Poor" and "Severe" categories of AQI. The results were utilized to inform the modeling process, especially in the evaluation of the role of location and seasonal variation.

3.6 Visualization

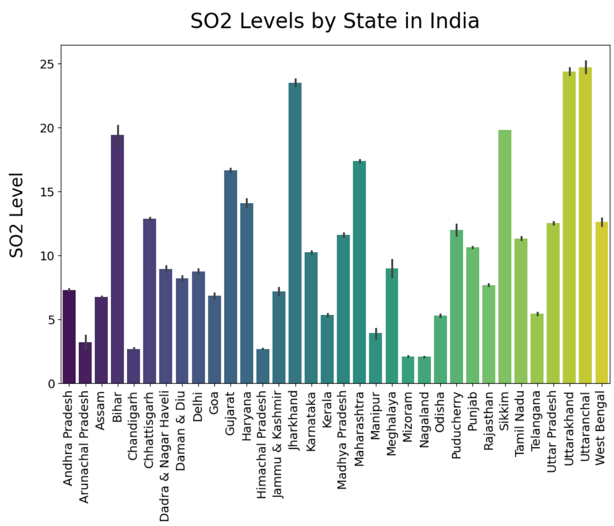


Figure 2. Barplot of Sulphur Dioxide (SO_2) concentration levels across Indian states (measured in $\mu g/m^3$).

Figure 2 depicts the levels of SO_2 in different Indian states. The states are shown on the x-axis, with SO_2 levels on the y-axis. The results indicate that Uttaranchal showed the highest levels of SO_2, followed closely by Uttarakhand. These are fairly high levels of SO_2 emissions that could indicate a likely concern regarding air quality.

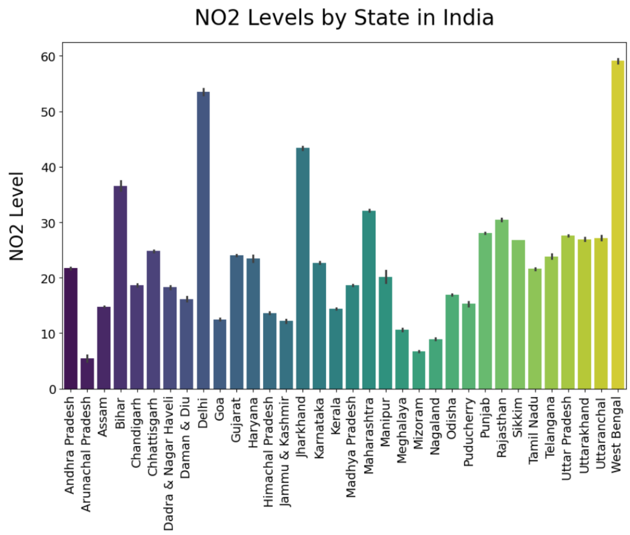


Figure 3. Distribution of Nitrogen Dioxide (NO_2) levels by Indian state, including variation bars (measured in $\mu g/m^3$).

Figure 3 depicts the levels of NO_2 in different Indian states. The states are shown on the x-axis, with NO_2 levels on the y-axis. Among all the states, West Bengal has the highest concentration of NO_2, which suggests higher vehicle or industrial emissions contributing to air pollution in that area.

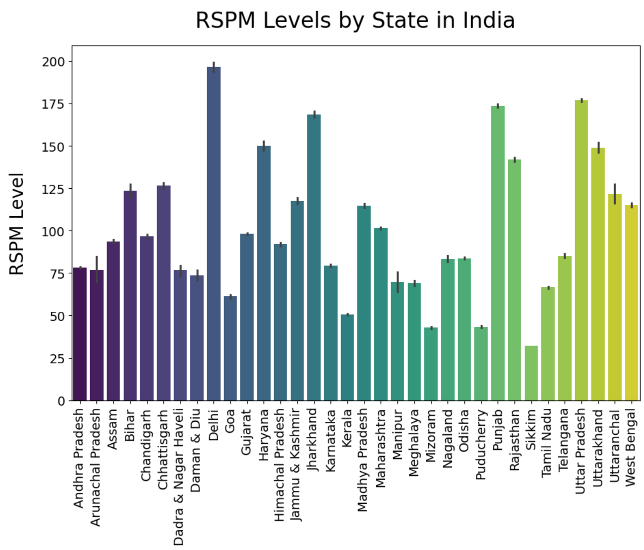


Figure 4. Respirable Suspended Particulate Matter (RSPM) levels across Indian states, based on regional pollution monitoring data (in $\mu g/m^3$).

The Figure 4 bar chart illustrates the levels of RSPM in various Indian states. The states are represented on the x-axis, and the corresponding levels of RSPM are represented on the y-axis. Delhi has the highest RSPM across all states, reflecting high particulate pollution in the area, probably from traffic, buildings, and industrial activities.

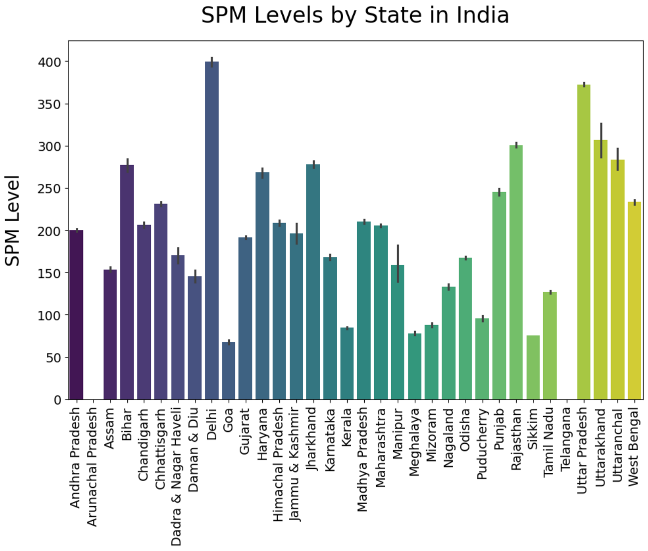


Figure 5. Suspended Particulate Matter (SPM) levels by state in India with comparative height visualization (measured in $\mu g/m^3$).

Figure 5 represents the distribution of SPM concentrations across multiple states in India. The x-axis shows different states, while the y-axis shows SPM levels. In this context, Delhi has the largest SPM levels (greatest air pollution), which is likely due to urbanization, vehicle emissions, and industrial strategy.

4 Implementation

The proposed NAS-AQP follows a neat two-step process of ensemble regression followed by biologically inspired classification. The tools used are Python 3.10, along with Scikit-learn, Nengo, and Brian2.

4.1 Model Evaluation and Selection

Several regression algorithms were used to estimate AQI from the air quality data, including Linear Regression, Random Forest, and Decision Trees. The performances of these models were compared on the basis of RMSE, MAE, and R^2 score. As shown in Table 2, the Voting Regressor—built on the above-mentioned models—achieved the lowest RMSE and the highest R^2 value, and was therefore selected as the final model for AQI estimation.

4.2 Threshold-Based Alert Generation

Once AQI is estimated, binary classification is carried out using an AQI threshold of >150 to identify poor air quality. Values above the threshold are termed as alerts. Table 3 demonstrates the AQI ranges described by CPCB and the corresponding health advisories.

Table 2. Regression model performance on AQI prediction.

Model	RMSE	MAE	R^2 Score
Linear Regression	35.42	24.87	0.82
Decision Tree	31.65	22.15	0.85
Random Forest	29.03	21.24	0.86
Voting Regressor	27.91	20.56	0.87

Table 3. AQI threshold classification for alerts.

AQI Range	Category	Alert Value
0-50	Good	0
51-100	Satisfactory	0
101-200	Moderate	0
201-300	Poor	1
301-400	Very Poor	1
401-500	Severe	1

4.3 SNN for Alert Classification

The binary labels are employed for training an SNN constructed with Leaky Integrate-and-Fire (LIF) neurons. Input pollutant values are represented in spike trains through rate-based encoding. The SNN structure was implemented using Brian2 and Nengo, trained via surrogate gradients to categorize spiking activity into alert states. The network is sparse and energy-efficient, perfect for real-time usage.

To describe the mechanism mathematically:

- **Rate-based Encoding:** Each input pollutant value x_i is transformed into a spike train using a rate-based encoding method, where the firing rate r_i is proportional to the input value:

$$r_i = k \cdot x_i \tag{1}$$

- **LIF Neuron Dynamics:** The membrane potential $V(t)$ of each LIF neuron follows the differential equation:

$$\tau_m \frac{dV(t)}{dt} = -V(t) + R \cdot I(t) \tag{2}$$

- A spike is generated when $V(t)$ exceeds a threshold V_{th}, after which the membrane potential is reset:

$$\text{if } V(t) \geq V_{\text{th}}, \quad V(t) \rightarrow V_{\text{reset}} \tag{3}$$

- **Surrogate Gradient Training:** Since spiking functions are non-differentiable, training is performed using surrogate gradients. A commonly used surrogate derivative is:

$$\frac{dS}{dV} \approx \sigma'(V) = \frac{1}{1+e^{-a(V-V_{\text{th}})}} \cdot \left(1 - \frac{1}{1+e^{-a(V-V_{\text{th}})}}\right) \tag{4}$$

- **Output Classification:** The alert decision is made based on spike accumulation over a fixed time window T:

$$y = \begin{cases} 1 & \text{if } \sum_{t=0}^{T} s_{out}(t) > \theta \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

4.4 End-to-End Pipeline

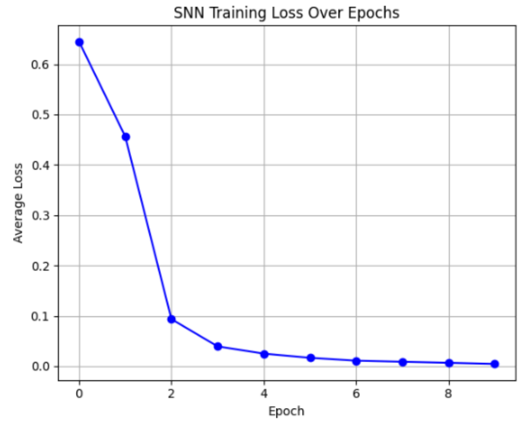
The entire NAS-AQP pipeline comprises pollutant sensor input, preprocessing of data, ensemble regression, threshold-based alert generation, and spike-based classification. The output is visualized through a real-time dashboard that consists of time-series plots.

5 Results and Discussions

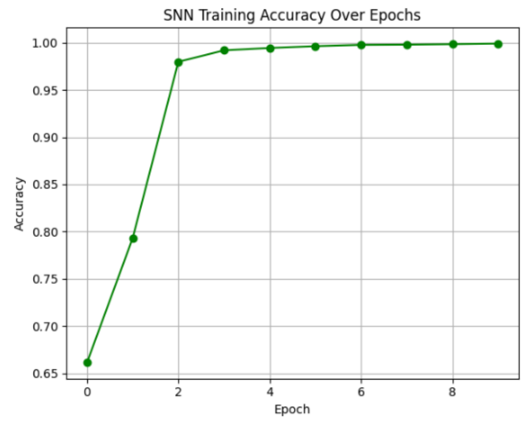
After comprehensive evaluation and validation, the final performance metrics of our regression models are presented in Table 2, showing that the Voting Regressor achieved the best overall performance. The Random Forest regressor achieved an RMSE of approximately 29, an MAE of around 21, and an R^2 score near 0.86, making it the best-performing model for AQI prediction. The SNN classifier obtained an accuracy of roughly 91 percent and an F1 score of 0.88, indicating effective performance in binary alert classification. The integrated system was thus successful in both regression and classification tasks. Visualizations which include bar charts from Figures 2, 3, 4 and 5 compare predicted versus actual AQI on the validation set and the frequency of alert spikes across different states.

In Figure 6(a), the SNN training loss plummets over 10 epochs, signaling successful learning. Figure 6(b) shows the training accuracy steadily increasing to nearly 100%, indicating successful classification on the training set.

In order to assess the performance of the proposed spike prediction model, a series of simulated AQI values were created that represented a range of air quality scenarios. Each value was associated with a binary ground truth label based on a threshold agreed upon (AQI > 150 classified as a spike), which is standard for environmental classifications. The model that was trained to predict spikes, the SNN, was then used to predict whether each AQI value represented a spike or not. Next, the predictions were compared against ground truth values, which were converted into a true label for good vs. unhealthy air based on the predefined threshold using two standard metrics of classification performance: accuracy and F1-score.



(a)



(b)

Figure 6. (a)SNN model training performance (b)SNN model training performance.

```
Random AQI values: [90, 120, 160, 180, 200, 80, 155, 175, 100, 176, 180, 201, 93]
True Labels:       [0, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 1, 0]
Predicted Labels:  [0, 1, 1, 1, 1, 0, 1, 1, 0, 1, 1, 1, 0]

Accuracy: 0.92
F1 Score: 0.94
```

Figure 7. Performance evaluation of the SNN model on simulated AQI data using accuracy and F1-score.

Accuracy included the total number of accurate predictions the model made during the prediction process. The total number of true positives plus true negatives divided by the total predictions made by the model, multiplied by 100, equals the percentage of accurate predictions made by the model. The F1-score evaluated precision and sensitivity, or recall, scores of the spiking model's classifier, making this metric significantly more useful in instances of class imbalance. The performance evaluation of the SNN model on simulated AQI data is shown in Figure 7. It provided a relatively unambiguous method to evaluate the model's ability to classify hazardous levels of air quality. In contrast to traditional models like ARIMA and SVR, which rely on strong assumptions of linearity and stationarity, our hybrid approach handles non-linear, real-world fluctuations more effectively. The integration of an ensemble model with an SNN classifier allows both predictive accuracy and interpretability in a dynamic urban environment.

6 Future Work

Although the proposed work has impressive functionality, there are some unresolved areas for research and development. One potential field of focus could be the incorporation of recurrent spiking networks that are time-based or the use of elementary models by way of combined SNNs with LSTM layers. The reason for doing all that is to have a model that is very sensitive to time and the behavior of pollutant drift. Therefore, this new idea of modifying the threshold values of AQI alerts can be further explained under conditions of fast environmental change.

Another possibility is the ability to have smart AQI alert thresholds that can change in a shorter frame. The changed thresholds may result either from the user's actions (reinforcement learning) or from thresholds determined by the region's climate health data and the area's population vulnerability. In relation to the aspect of going out and doing the job, further development will focus more on implementing the NAS-AQP pipeline on edge devices to be energy-efficient and perform computations on-site. Moreover, the implementation of real-world pilot testing that occurs in parallel with municipal bodies or pollution control boards can be another way to test the system's performance under live and working conditions. In addition, a connection with space-borne sensors and monitoring by a network of citizens may allow the framework to be designed in such a way that it enables the whole national territory or even neighboring countries to be monitored. On a final note, the upscaling of AI Explainable with the use of spike visualizations and neuro-symbolic reasoning could be a definitive answer to a very new and interesting area of public trust improvement.

7 Conclusion

This paper presents the Neuro-Inspired Alert System for Air Quality Prediction (NAS-AQP), a novel approach that combines ensemble-based AQI regression with biologically inspired classification by SNNs. The proposed system was designed to analyze real-time environmental data—i.e.,SO_2, NO_2, RSPM, and SPM levels—to forecast AQI levels and issue binary alerts when air quality crossed dangerous levels. The ensemble learning technique voting regression proved highly efficient in approximating AQI with precision in various Indian states. The application of SNNs to alert classification provided a computationally efficient and biologically plausible framework that could analyze temporal dynamics with real-time processing. The system was evaluated for performance using standard performance metrics and showed high precision and robustness for regression and classification tasks.The module-based design and visualization dashboard also make NAS-AQP deployable within smart city infrastructure and public health monitoring.

Overall, the NAS-AQP system contributes to environmental informatics with a hybrid, scalable, and explainable methodology for real-time air quality monitoring and early warning generation.

Data Availability Statement

Data will be made available on request.

Funding

This work was supported without any funding.

Conflicts of Interest

The authors declare no conflicts of interest.

Ethical Approval and Consent to Participate

Not applicable.

References

[1] Pollution. (2025, 18). World Bank. Retrieved from https://www.worldbank.org/en/topic/pollution

[2] 2024 IQAir world air quality report. (2025, March 10). IQAir | First in Air Quality. Retrieved from https://www.iqair.com/us/newsroom/waqr-2024-pr

[3] *India's capital New Delhi shuts all primary schools as air quality worsens*. (2024, November 14). Al Jazeera. Retrieved from https://www.aljazeera.com/news/2024/11/14/indias-capital-new-delhi-shuts-all-primary-schools-as-air-quality-worsens

[4] Lu, H., & Ma, X. (2020). Hybrid decision tree-based machine learning models for short-term water quality prediction. *Chemosphere, 249*, 126169. [Crossref]

[5] Kim, J., Zeng, H., Ghadiyaram, D., Lee, S., Zhang, L., & Bovik, A. C. (2017). Deep convolutional neural models for picture-quality prediction: Challenges and solutions to data-driven image quality assessment. *IEEE Signal processing magazine, 34*(6), 130-141. [Crossref]

[6] Yan, R., Liao, J., Yang, J., Sun, W., Nong, M., & Li, F. (2021). Multi-hour and multi-site air quality index forecasting in Beijing using CNN, LSTM, CNN-LSTM, and spatiotemporal clustering. *Expert Systems with Applications, 169*, 114513. [Crossref]

[7] Qi, Z., Wang, T., Song, G., Hu, W., Li, X., & Zhang, Z. (2018). Deep air learning: Interpolation, prediction, and feature analysis of fine-grained air quality. *IEEE Transactions on Knowledge and Data Engineering, 30*(12), 2285-2297. [Crossref]

[8] Yi, X., Zhang, J., Wang, Z., Li, T., & 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* (pp. 965-973). [Crossref]

[9] Ahmed, U., Mumtaz, R., Anwar, H., Shah, A. A., Irfan, R., & García-Nieto, J. (2019). Efficient water quality prediction using supervised machine learning. *Water, 11*(11), 2210. [Crossref]

[10] Du, S., Li, T., Yang, Y., & Horng, S. J. (2019). Deep air quality forecasting using hybrid deep learning framework. *IEEE Transactions on Knowledge and Data Engineering, 33*(6), 2412-2424. [Crossref]

[11] Ameer, S., Shah, M. A., Khan, A., Song, H., Maple, C., Islam, S. U., & Asghar, M. N. (2019). Comparative analysis of machine learning techniques for predicting air quality in smart cities. *IEEE access, 7*, 128325-128338. [Crossref]

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

[13] Soh, P. W., Chang, J. W., & Huang, J. W. (2018). Adaptive deep learning-based air quality prediction model using the most relevant spatial-temporal relations. *IEEE Access, 6*, 38186-38199. [Crossref]

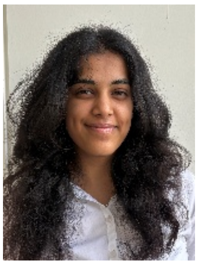
[14] Ma, J., Cheng, J. C., Lin, C., Tan, Y., & Zhang, J. (2019). Improving air quality prediction accuracy at larger temporal resolutions using deep learning and transfer learning techniques. *Atmospheric Environment, 214*, 116885. [Crossref]

[15] Pazhanivel, D. B., Velu, A. N., & Palaniappan, B. S. (2024). Design and enhancement of a fog-enabled air quality monitoring and prediction system: An optimized lightweight deep learning model for a smart fog environmental gateway. *Sensors, 24*(15), 5069. [Crossref]

[16] Firoozi, A. A., & Firoozi, A. A. (2024). Revolutionizing Civil Engineering with Neuromorphic Computing. *Springer Nature*.[Crossref]

[17] Central Pollution Control Board (CPCB). (2020). Air Quality Data [Data set]. Kaggle. Retrieved from https://www.kaggle.com/datasets/rohanrao/air-quality-data-in-india

Sneh Sharma is currently pursuing her B.Tech in Computer Science with her major specialization in Artificial Intelligence and Machine Learning at Amity University Punjab, India, where she is in her final year. Her area of interest includes Artificial Intelligence, Deep Learning, and Data Science. She has co-authored a book chapter titled "Artificial Intelligence-Enabled Apple Disease Monitoring System for Environmental Protection and Resource Conservation" which has been published in the "AI Methods for Environmental Protection and Resource Conservation" under the IGI Scientific Publication. Sneh has also contributed to various other interdisciplinary research projects and publications aimed at developing innovative solutions across diverse domains. With a strong commitment to academic excellence and societal impact, she strive to explore and apply AI-driven technologies to solve real-world challenges. (Email: sneh@s.amity.edu)



Kashish Devgan is currently pursuing her B.Tech in Computer Science with her major specialization in Artificial Intelligence and Machine Learning at Amity University Punjab, India, where she is in her final year. Her area of interest includes Artificial Intelligence, Deep Learning, and Data Science. She has co-authored a book chapter titled "Artificial Intelligence-Enabled Apple Disease Monitoring System for Environmental Protection and Resource Conservation" which has been published in "AI Methods for Environmental Protection and Resource Conservation" under the IGI Scientific Publication. Kashish has also contributed to various other interdisciplinary research projects and publications aimed at developing innovative solutions across diverse domains. With a strong commitment to academic excellence and societal impact, she strive to explore and apply AI-driven technologies to solve real-world challenges. (Email: kashish.devgan@s.amity.edu)

Devanshi Jangra is a third-year B.Tech Computer Science and Engineering student at Amity University Punjab, specializing in Artificial Intelligence and Machine Learning. Her academic work includes a published book chapter titled "Artificial Intelligence-Enabled Apple Disease Monitoring System for Environmental Protection and Resource Conservation" in IGI Global (2025), which explores AI applications in sustainable agriculture and peer reviewed some book chapters as well. She has developed several impactful projects, including a deep learning–based Plant Disease Detection App, a Plastic Eliminator App integrated with Firebase and Google Maps API for waste management and awareness, and a Campus Sports Slot Booking System built using AWS services such as EC2, S3, and CloudWatch during her project tenure with the National Informatics Centre (NIC). Devanshi is proficient in Python, Dart, and AI frameworks like TensorFlow and PyTorch, and is focused on creating scalable, AI-powered mobile solutions that address practical challenges in agriculture, environment, and institutional systems. (Email: devanshi3@s.amity.edu)

Aanshi Bhardwaj received her Ph.D. in Cyber Security and Masters in Web Mining from University Institute of Engineering & Technology, Panjab University, India. She is currently working as an Assistant Professor at Amity University Mohali, Punjab, India from past three years. She has over 10 years of teaching experience and has gained valuable experience in both teaching and research. Her main areas of interest include Cyber Security, Web Mining, Healthcare, Machine Learning and Deep Learning. She is actively involved in mentoring students, organizing technical events, and contributing to curriculum development. She has published over 15 research papers in high-impact international journals and reputed conferences. (Email: bhardwajaanshi@gmail.com)

Shubhani Aggarwal(Academic Member, ICCK) received her Ph.D. in CSE from Thapar Institute of Engineering and Technology (Deemed to be University), Patiala, Punjab, India and was a postdoctoral research fellow in École de Technologie Supérieure, Montréal, Québec, Canada. She is working as an Assistant Professor-III in SOCS, UPES, Dehradun, India. Her research interests are in the areas of Blockchain, cryptography, Internet of Drones, and information security. She has published a Book titled "The Blockchain Technology for Secure and Smart Applications across Industry Verticals" in Elsevier Publication. She has also published more than 50 research papers in top-cited journals such as IEEE Transactions on Vehicular Technology, IEEE Transactions on Industrial Informatics, IEEE Transactions on Intelligent Transportation Systems, IEEE Internet of Things, Elsevier Journal Networks of Computer and Applications, IEEE Access, Computers and Security, Mobile Networks and Applications, Computer Communications and many more. She has published a book based on Blockchain Technology in Advances in Computers, Elsevier. (Email: shubhaniaggarwal529@gmail.com)