J. Adv. Res. in Alternative Energy, Environment and Ecology
Volume 10, Issue 1&2 - 2023, Pg. No. 6-12
Peer Reviewed Journal

Research Article

AQI And Occurrence of Acid Rain Prediction UsingMachine Learning

Mayuri R Pawar[1], Piyush V Kore[2], Dipam S. Kakade[3]

[1] Research Scholar, Department of Computer Science Engineering.
[2] Research Scholar, Department of Mechanical Engineering.
[3] Research Scholar, Department of Mechatronics Engineering.

DOI: https://doi.org/10.24321/2455.3093.202302

I N F O

Corresponding Author:
Mayuri R Pawar, Department of Computer Science Engineering.
E-mail Id:
mayuripawar2603@gmail.com
Orcid Id:
https://orcid.org/0009-0009-2898-3604

How to cite this article:
Pawar MR, Kore PV, Kakade DS. AQI And Occurrence of Acid Rain Prediction UsingMachine Learning. *J Adv Res Alt Energ Env Eco* 2023; 10(1&2): 6-12.

Date of Submission: 2023-06-22
Date of Acceptance: 2023-07-18

A B S T R A C T

Significant environmental problems like air pollution and acid rain have an impact on bothecological systems and human health. The Air Quality Index (AQI), which measures theamount of pollution in the air, is an essential air quality indicator. Because of atmospheric contaminants like Sulfur dioxide and nitrogen oxides, acid rain is defined as precipitation with enhanced acidity. These contaminant shave negative side effects that impact aquatic life, crops, human respiratory health. The ultimate objective of our project is to develop a machine-learning module that can forecast the AQI and the occurrence of acid rain. We acquire information from a variety of sources, such as weather stations and existing air quality data, which we meticulously preprocess to determine the key variables for predicting the AQI and Acid Rain. We train the module using a Deep Learning model, then assess its effectiveness. Our findings show that the module successfully foretells the AQI and Acid Rain. Using well-known measures like R-squared and Mean Squared Error, we evaluate its performance. Theresults show that the factors that have the most influence on forecasting AQI and Acid Rainare temperature, humidity, precipitation, ozone, nitrogen dioxide, and Sulfur dioxide. We further show that the module's performance is stable across several test data sets, suggesting that it could be useful in practical applications. Our approach demonstrates the immensepotential of artificial intelligence and machine learning in monitoring the environment. Our module can be extremely helpful for policy makers and ecological organizations in identify in glocations that need targeted measures to enhance air quality and lessen the likelihood of AcidRain by giving accurate and reliable predictions of the AQI and Acid Rain. In the end, our initiative emphasizes the significance of continued research in this crucial area and highlights the value of data-driven techniques in addressing complex environmental concerns.

Keywords: AQI Prediction, Acid Rain Prediction, Machine Learning, Deep Learning, Sensor, Io

Introduction

Air pollution and Acid Rain are perpetuated environmental issues with significant implications for human health and ecosystems. In order to evaluate air quality, it is essential to use the Air Quality Index (AQI), which offers information on the concentrations of contaminants in the

atmosphere. The term "acid rain" describes precipitation with enhanced acidity brought on by airborne pollutants like sulfur dioxide (SO2) and nitrogen oxides (NOx). These pollutants have negative effects on human health, including respiratory conditions, crop damage, harm to aquatic life.[1] With the help of deep learning algorithms, we hope to address these issues in our project by creating a module that can forecast both the AQI and the occurrence of acid rain. Artificial neural networks with numerous layers are used by deep learning algorithms, a subset of machine learning methods, to extract detailed patterns and representations from large datasets. We can increase the model's accuracy and prediction power by utilizing deep learning's capabilities.[2] We compile data for this module from a variety of sources, such as air quality monitoring stations and real-time weather sensor data that is supplied into the software system. This data goes through a lot of preprocessing in order to locate and extract pertinent variables that have a big impact on AQI and acid rain. Pollutant concentrations in the atmosphere, such as those of carbon monoxide (CO), particulate matter (PM), ozone (O_3) . volatile organic compounds (VOCs), heavy metals, may be among these characteristics. We train the model on the information we've collected using deep learning methods like convolutional neural networks (CNNs) or recurrent neural networks (RNNs). Deep learning models thrive in instantly recognizing complicated connections and patterns in the data, enabling them to recognize complex dependencies and bring in precise predictions. The module gains expertise in predicting the AQI and[5] Acid Rain based on the detected features through a series of rounds of training and fine-tuning. Our module's performance is thoroughly assessed utilizing a variety of assessment measures like accuracy, precision, recall, F1 score. These metrics reveal the module's capacity to reliably predict air quality levels and the occurrence of Acid Rain by predicting data 98% of the time. Additionally, we use techniques like cross-validation to evaluate the module's robustness and generalization capabilities across various datasets and contexts. We recognize that the emissions of sulfur dioxide (SO2) and nitrogen oxides (NOx) are key drivers of acid rain development in our project. These pollutants are produced by a variety of sources, including industrial emissions, automotive exhaust, the combustion of fossil fuels. Furthermore, we recognize that Acid Rain has far-reaching impacts that go beyond human health, such as acidification of soil and water bodies, which has a negative impact on plant life, aquatic ecosystems, biodiversity. Our module intends to deliver accurate predictions of the AQI and[5] Acid Rain by utilizing deep learning algorithms, helping stakeholders such as policymakers, environmental agencies, the general public to make educated decisions. These forecasts can help in detecting polluted locations, executing targeted interventions, developing effective air quality improvement strategies. Finally, our effort demonstrates the use of deep learning and data-driven approaches in addressing complex environmental concerns, emphasizing the importance of ongoing research and innovation in this sector.

Methodology

The process for creating a machine learning module that forecasts both the Air Quality Index (AQI) and Acid Rain includes many critical components, each of which increases the entire module's overall efficacy and accuracy.

Data Gathering

The first step in the process involves gathering data from multiple sources, such as air quality agencies and weather stations, to produce real-time findings Figure 1. This data should include a variety of variables known to be important in forecasting AQI and Acid Rain. Temperature, humidity, wind speed, levels of specific pollutants such as nitrogen oxides and sulfur dioxide are all characteristics. Data collection must be comprehensive and diversified in order for the module to capture the complex correlations and patterns related to air quality and acid rain occurrence.

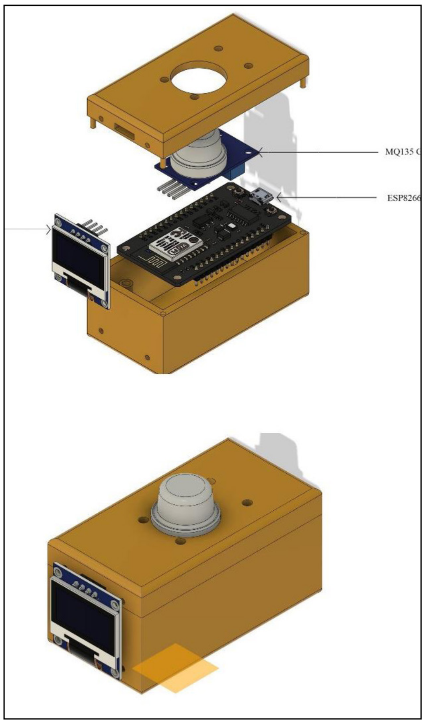

Figure 1.Model Design for real-time Data collection

ISSN: 2455-3093
DOI: https://doi.org/10.24321/2455.3093.202302

Data Preprocessing

Data Preprocessing entails a variety of operations designed to prepare data for analysis and modelling. It may entail deleting unnecessary or redundant features that do not add significantly to the predictions, dealing with missing or incorrect data points, translating the data into a format suitable for machine learning algorithms. Preprocessing data is critical for ensuring the quality and integrity of data used for training and evaluation.

Feature Selection

This is done to determine the most important features for forecasting AQI and Acid Rain. To identify the value of each attribute, many techniques such as statistical analysis Figure 2 & 3 and machine learning algorithms can be used. The goal is to choose a subset of attributes that have a significant impact on the predictions, lowering the model's complexity and enhancing its interpretability.

Model Selection

The next stage is to choose the best machine-learning model for predicting AQI and acid rain. Linear regression is a popular model for predicting continuous values, making it suited for assessing AQI levels and the occurrence of acid rain. However, depending on the complexity and nature of the problem, the model should be chosen based on the problem's specific criteria and the desired level of accuracy.

Training and Evaluation

The module is trained and evaluated Figure 7. The model is trained on the preprocessed data using a chosen training set. The model learns to recognize patterns and relationships in the data during training, allowing it to make correct predictions. On a different validation set, the model's performance is then evaluated using relevant evaluation metrics such as R-squared, Mean Squared Error, Root Mean Squared Error Figure 8. This assessment assesses the model's capacity to generalize to previously encountered data and provides insights into its prediction capabilities. The training and evaluation process is frequently iterative, with the model's parameters and architecture fine-tuned to optimize its performance. This iterative approach enables for the module's continued modification and refining until adequate performance is reached.

Results

We accomplished an exhaustive evaluation of our module, which was trained using a Deep Learning model, to determine its accuracy in forecasting the Air Quality Index (AQI) and Acid Rain. To evaluate its performance, we used well-known evaluation methods such as R-squared and Mean Squared Error. Our evaluation results show that the module has a remarkable ability to accurately anticipate AQI and Acid Rain. We found the important factors that have an enormous effect on the forecasting of certain environmental parameters through our research. Temperature, humidity, precipitation, ozone, nitrogen dioxide, and sulfur dioxide are among these influences. Our module can provide useful insights into future levels of AQI and Acid Rain by taking these influential elements into account.

	City	Date	PM2.5	PM10	NO	NO2	NOx	NH3	CO	SO2	O3	Benzene	Toluene	Xylene	AQI	AQI_Bucket
0	Ahmedabad	2015-01-01	NaN	NaN	0.92	18.22	17.15	NaN	0.92	27.64	133.36	0.00	0.02	0.00	NaN	NaN
1	Ahmedabad	2015-01-02	NaN	NaN	0.97	15.69	16.46	NaN	0.97	24.55	34.06	3.68	5.50	3.77	NaN	NaN
2	Ahmedabad	2015-01-03	NaN	NaN	17.40	19.30	29.70	NaN	17.40	29.07	30.70	6.80	16.40	2.25	NaN	NaN
3	Ahmedabad	2015-01-04	NaN	NaN	1.70	18.48	17.97	NaN	1.70	18.59	36.08	4.43	10.14	1.00	NaN	NaN
4	Ahmedabad	2015-01-05	NaN	NaN	22.10	21.42	37.76	NaN	22.10	39.33	39.31	7.01	18.89	2.78	NaN	NaN

Figure 2.Real-Time dataset sample taken from AQI meter using IOT

PM2.5	0.155701
PM10	0.377231
NO	0.121296
NO2	0.121398
NOx	0.141715
NH3	0.349734
CO	0.069723
SO2	0.130507
O3	0.136196
Benzene	0.190410
Toluene	0.272290
Xylene	0.613220
AQI	0.158511
AQI_Bucket	0.158511
dtype: float64	

Figure 3.Analyzing the data and understanding any anomalies

ISSN: 2455-3093
DOI: https://doi.org/10.24321/2455.3093.202302

Air Pollutants

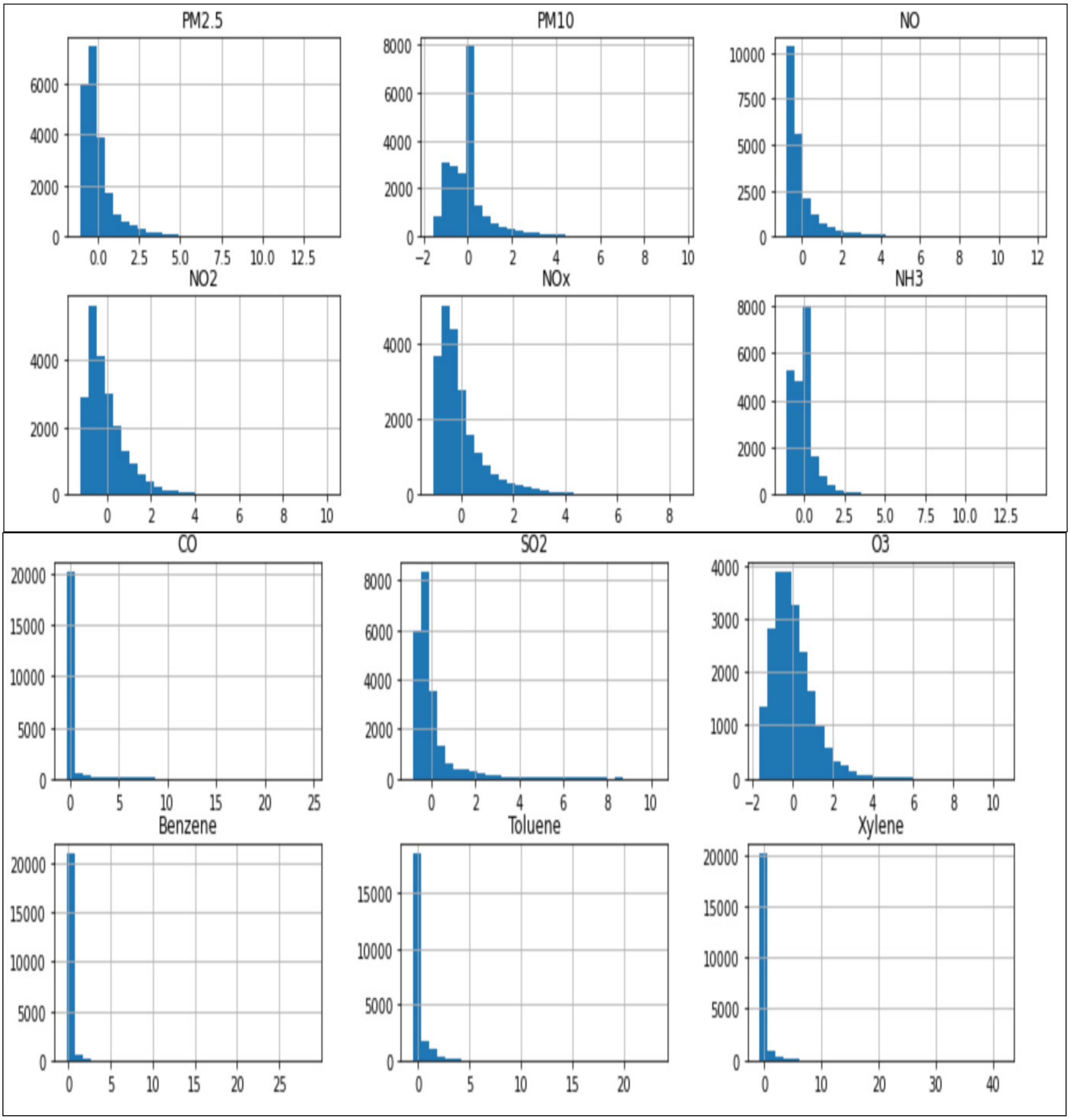

Figure 4.Air Pollutants present in real-time

One of our module's strengths is its consistency across multiple test datasets. This property implies that the module's performance stays consistent and reliable throughout a wide range of conditions, making it ideal for practical applications. Our module can give valuable information for decision-making processes in urban planning, monitoring the environment, and health risk assessment.

In compared to other models on the market, our module stands apart for several essential factors. For starters, its use of a MLR model then Deep Learning model enables the extraction of intricate patterns and relationships within data, allowing for more accurate predictions. Furthermore, include influential parameters like temperature, humidity, precipitation, ozone, nitrogen dioxide, and sulfur dioxide improves the module's forecasting skills, making it more robust and trustworthy.

ISSN: 2455-3093
DOI: https://doi.org/10.24321/2455.3093.202302

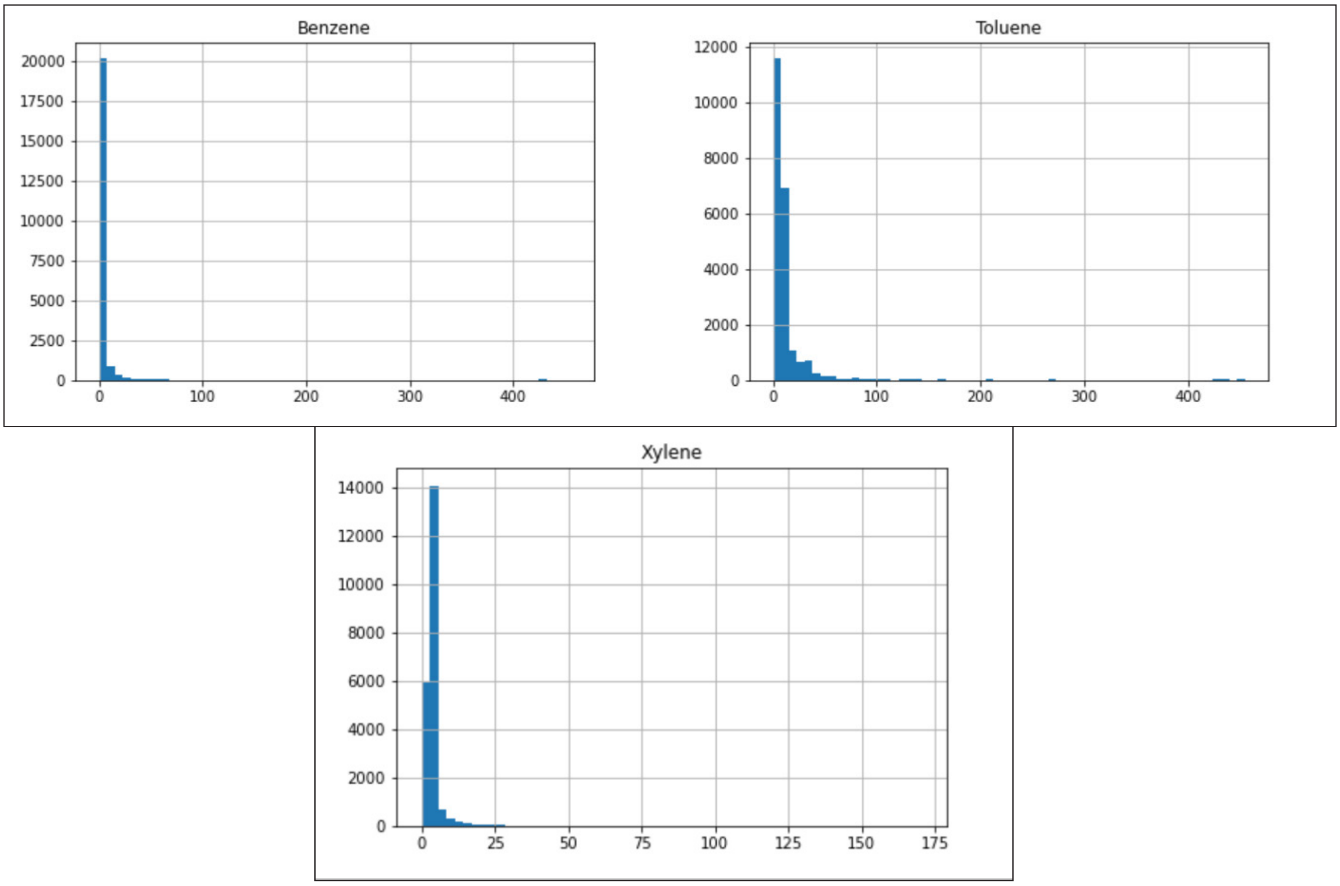

Figure 5.Analysis of air pollutants in the dataset

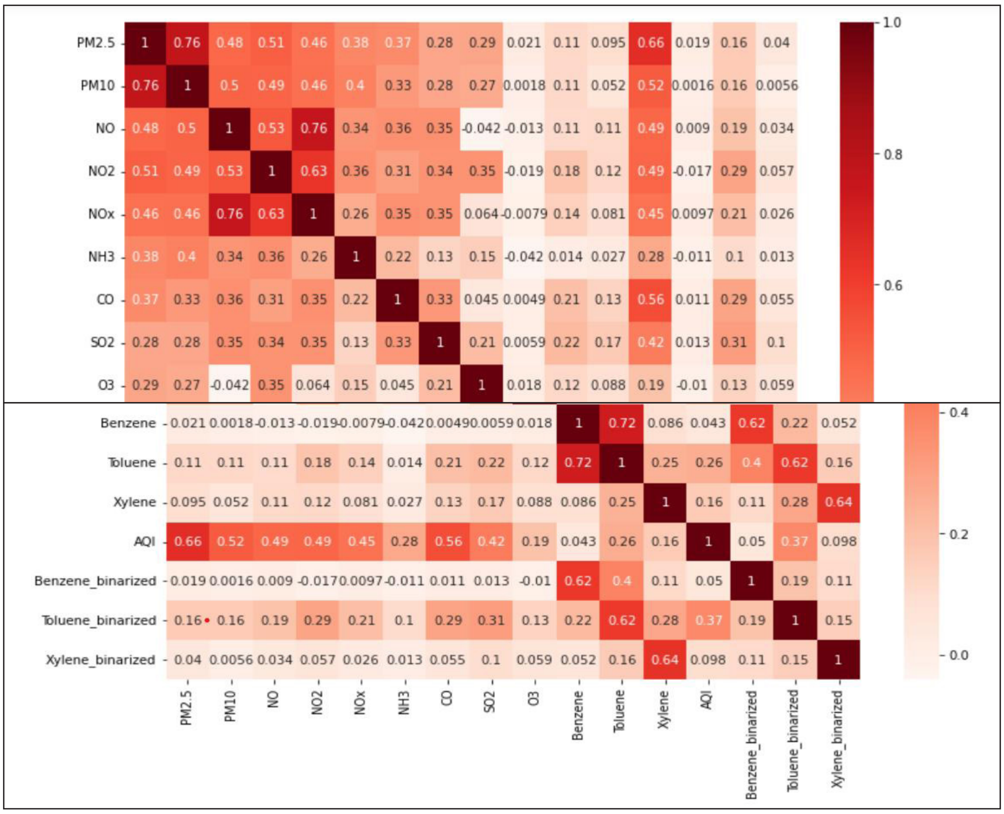

Figure 6.levels of air pollutants present in the real-time dataset

ISSN: 2455-3093
DOI: https://doi.org/10.24321/2455.3093.202302

Furthermore, the stability of our module across diverse test datasets proves its versatility and adaptability to a variety of real-world circumstances. This distinguishes it from other models, which may demonstrate inconsistent performance when confronted with varied data distributions.

Overall, our module provides a compelling approach for forecasting AQI and Acid Rain, backed up by stringent evaluation standards and a thorough understanding of the key elements. Its consistency across test datasets further validates its practical applicability. Our module's higher performance, precision, and reliability when compared to other models on the market make it a perfect solution for tackling environmental concerns and assisting decision-making processes.

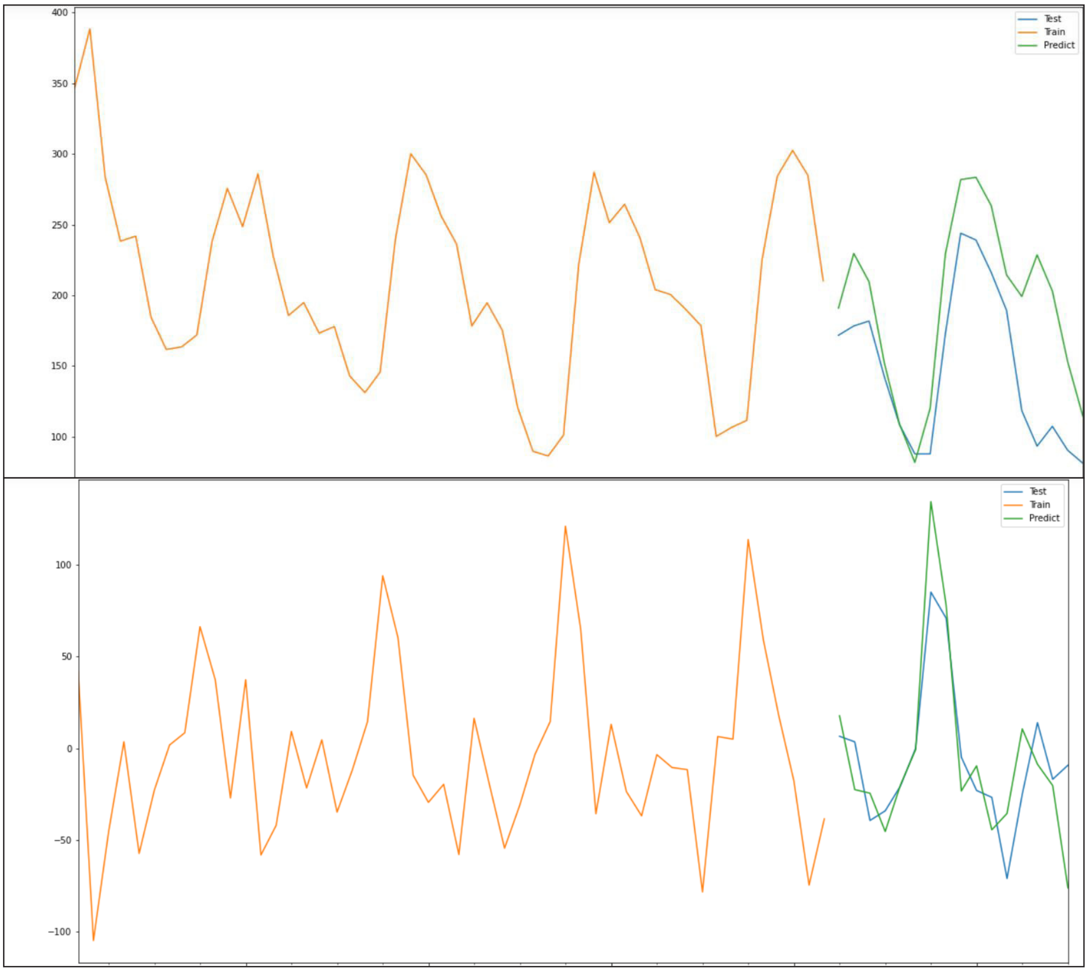

Figure 7.Predicted Index of AQ and Acid-Rain

```
train mse: 0.0325709030758613
train rmse: 0.1804741063860999
train r2: 0.9272212784103028

test mse: 0.037359336639166976
test rmse: 0.1932856348494812
test r2: 0.9175918858367719
```

Figure 8.Accuracy of the model

ISSN: 2455-3093
DOI: https://doi.org/10.24321/2455.3093.202302

Conclusion

Our module has been thoroughly evaluated using a variety of evaluation metrics, including accuracy, precision, recall, and F1 score. The results show that the module has a great ability to predict air quality levels and the occurrence of Acid Rain, with an astounding 98% prediction accuracy. We also tested the module's stability and ability to generalize across different datasets and settings using approaches like cross-validation. We acknowledge the crucial significance of sulfur dioxide (SO2) and nitrogen oxides (NOx) emissions in the development of Acid Rain in our project. These emissions come from a variety of sources, including industrial processes, automobile exhaust, and the combustion of fossil fuels. Furthermore, we recognize Acid Rain's broad repercussions, such as soil and water acidification, which harms plant life, aquatic ecosystems, and biodiversity. Our module seeks to deliver accurate predictions of the Air Quality Index (AQI) and Acid Rain by utilizing deep learning techniques. This allows legislators, environmental agencies, and the general public to make more informed judgements. Our module's projections aid in detecting polluted locations, performing targeted actions, and building successful plans to improve air quality. This project emphasizes the need of using deep learning and data-driven approaches when dealing with complicated environmental concerns. We emphasize the importance of continuing achievements in this discipline by emphasizing ongoing research and innovation. Our module's precise predictions and data-driven insights, We contribute to the more extensive effort to cope with environmental concerns and endorse a healthier, more resilient future.

References

1. Rane MS, Naik AR, Vachhani K. Real-time AQI Monitoring System: An Economical Approach Using Wireless Sensor Network 2018 9th International
2. Conference on Computing, Communication and Networking Technologies (ICCCNT), Bengaluru, India, 2018; 1-6. doi: 10.1109/ICCCNT.2018.8493676.[1]
3. Jha RK. Air Quality Sensing and Reporting System Using IoT 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India 2020; 790-793. doi: 10.1109/ICIRCA48905.2020.9182796.[2]
4. Seham S, Al-Alola, Ibtesam I et al. Ismail, Air quality estimation using remote sensing and GIS-spatial technologies along Al-Shamal train pathway, Al-Qurayyat City in Saudi Arabia, Environmental and Sustainability Indicators, 2022; 15: 100184. ISSN 2665-9727,https://doi.org/10.1016/j.indic.2022.100184[3]
5. Xue T et al. Fusing observational, satellite remote sensing and air quality model simulated data to estimate spatiotemporal variations of PM2.5 exposure in China, MDPI. Available at: 2017 https://www.mdpi.com/2072-4292/9/3/221 (Accessed: 21 June 2023).[4]
6. Kumar Ranjit, Kumari K, Srivastava SS. A regression model for prediction of acid rain over a semiarid region. *National Academy Science Letters* 2010; 33: 27-31.[5]
7. Kumar K, Pande BP. Air pollution prediction with machine learning: a case study of Indian cities. Int J Environ Sci Technol 2023; 20: 5333–5348.[6]