IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.53

Volume 4, Issue 1, September 2024

Air Quality Monitoring Using Machine Learning

Amisha Salian[1], Chethan K R[2], Kuladeep[3], Shreya Shetty[4], Sri Krishna Shastri C[5]

UG Students, Department of Electronics and Communication Engineering[1-4]

Associate Professor, Department of Electronics and Communication Engineering[5]

Mangalore Institute of Technology & Engineering, Mangalore, India

Abstract*: The combination of Internet of Things (IoT) sensors and machine learning algorithms has revolutionized the monitoring of air quality, offering more precise and up-to-date data for better environmental management. This study examines current advancements in this field, focusing especially on the development and implementation of Internet of Things (IoT)-based air quality monitoring systems that employ a range of sensors to detect pollutants such as particulate matter, CO, NO2, and CO2. These systems forecast pollution levels and look at trends in air quality using machine learning models like Random Forest and Linear Regression. They also collect data in real time via sensor networks. Notable advancements include low-cost sensors and wireless networks for industrial air quality prediction, Internet of Things platforms for continuous environmental monitoring, and real-time in-car monitoring systems to ensure driver safety. The examined studies highlight some problems with power consumption, data latency, and sensor accuracy but also demonstrate how IoT and machine learning may be utilized to increase the precision and effectiveness of air quality monitoring. Future directions point towards the integration of artificial intelligence, improved sensor technologies, and renewable energy solutions to further improve air quality management and forecasting capacities. This in-depth analysis provides recommendations for developing more dependable and scalable air quality monitoring systems, as well as highlighting the potential of IoT and machine learning technologies in addressing the issues related to global air pollution.*

Keywords: Air quality, Random Forest, Real-time analysis, Machine Learning

I. INTRODUCTION

In the area of designing electronic circuits is a need to test their power supplies. Particulate matter, hazardous substances, and biological molecules all contribute to air pollution in earth's atmospheres; however, despite the rapid development and widespread application of machine learning algorithms in a wide range of fields, there remains much to be discovered in the realm of airpollution prediction. Particulate matter is one crucial measure that is used as a key indicator of the levels of pollution in various regions at particular times. The importance of air quality monitoring has grown along with concerns about pollution and its effects on human health and the environment. In this project, integrated hardware and software components to develop an accurate air quality index measurement and forecast system. The hardware configuration is constructed using an Arduino microcontroller interfaced with multiple sensors (DHT11 for temperature and humidity, MQ2 for smoke, MQ135 for hazardous chemicals, and MQ7 for carbon monoxide). The real-time data collected by these sensors is displayed on a 16x2 LCD panel, providing users with immediate feedback regarding the air quality. By applying machine learning approaches, were able to enhance the software component's prediction capabilities. In order to forecast the air quality based on the sensor inputs, trained a model using a huge dataset using the Random Forest Regression technique. The dataset was handled and analyzed using Jupyter Notebook, ensuring a dependable and reproducible analytical pipeline. Used the same dataset for both training and testing in an attempt to assess our model's performance and increase its accuracy.

II. LITERATURE SURVEY

Design A systematic approach to locating, assessing, and understanding the work done by academics, practitioners, and researchers is provided by this assessment of the literature. In order to avoid duplication and to give credit to other researchers, the goal of a literature review is to obtain an understanding of the current research and debates pertinent to a specific topic or area of previous scholarship.

Rajasekar et al., [1]: Proposes a method to foresee potential health risks or diseases associated with exposure to air pollution by combining machine learning algorithms with Internet of Things (IoT) sensors for air quality monitoring. Air quality data may be more precisely and individually tailored by including more sophisticated sensors, such as wearable and multi-pollutant ones. The system will respond more quickly when advanced predictive analytics and real-time data processing are used since they produce forecasts and feedback promptly. The accuracy of illness prediction can be increased by combining machine learning models with deep learning techniques and hybrid models.

Goh et al., [2]: involves integrating many components to enable real-time air quality monitoring in cars as well as the ability to predict changes in air quality. concentrating on making these sensors more sensitive, precise, and small so that they can be more effectively integrated with other automobile parts. The creation of intricate machine learning algorithms is another crucial field. These include online learning algorithms that adjust in real-time when new data becomes available, as well as deep learning models like convolutional neural networks (CNNs) and recurrent neural networks (RNNs) for more accurate predictions. the combination of these systems with smart car autonomous driving technology and Internet of Things networks.

Kang et al., [3]: It looks at published research that use artificial intelligence, decision trees, deep learning, and other cutting edge methods to evaluate the quality of the air. The research focuses on the use of big data analytics, enabled by advancements in environmental sensing networks and the availability of sensor data, to collect information on air pollutant concentrations and provide assessments of air pollution across various areas. Utilizing health data to forecast the possible consequences of poor air quality on specific demographic groups can lead to targeted public health initiatives. increasing the range of high-resolution data sources used, including social media feeds, mobile sensors, and satellite photos, to increase forecast accuracy.

Bhalgat et al., [4]: This article proposes a method for predicting the amount of sulfur dioxide (SO_2) in the environment by using machine learning techniques. In particular, time series models like the autoregressive (AR) model are used. A more thorough evaluation of the quality of the air will be possible with the addition of more sensors and monitoring for pollutants other than particulate matter (PM2.5 and PM10), carbon monoxide (CO), and nitrogen dioxide (NO2). To give consumers a more comprehensive picture of the environmental conditions, this may need adding sensors for pollutants like sulfur dioxide (SO2), ozone (O3), volatile organic compounds (VOCs), and other harmful gases.

Mishra et al., [5]: An Internet of Things air quality monitoring system was created using a MQ Series sensor connected to a NodeMCU via an ESP8266 WLAN adaptor, which transmitted data to the ThingSpeak cloud. The air quality was then predicted using time-series forecasting and predictive modeling using three machine learning algorithms (Linear Regression, Random Forest, XGBoost), as well as the ARIMA model. Random Forest outperformed other models in terms of predicting CO, NO2, and SO2 air quality index (AQI) levels. It aims to enhance the models' anticipated accuracy and resilience, expand the system's deployment to more high-pollution areas, and incorporate additional environmental characteristics for a more comprehensive pollution analysis.

Ameer et al., [6]: The researchers used a hybrid computational intelligence system to incorporate machine learning and find correlations between weather patterns and air pollution levels in an attempt to identify the underlying cause of pollution. Its objective is to investigate how well these techniques perform in the multi-core context of Spark. The paper also intend to examine the different factors that contribute to air pollution.

Vito et al., [7]: The goal of the widespread deployment was to provide high resolution AQ mapping permanently in urban settings and to use AQMS as a backup system for AQRMS, providing data in the event that AQRMS data is unavailable due to scheduled maintenance or faults. It places special emphasis on network calibration and the use of online machine learning tools. To address this non-stationary framework, we employ adaptive learning strategies and extend the validity of multi-sensor calibration models that facilitate continuous learning.

III. DESIGN AND IMPLEMENTATION

The The system's software and hardware components work together to forecast and track the air quality. The hardware configuration consists of an Arduino microprocessor interfaced with DHT11, MQ2, MQ135, and MQ7 sensors and a 16x2 LCD display. The software is made up of the Arduino IDE for microcontroller programming, the Jupyter Notebook for creating and testing models, and Python with libraries like NumPy, Pandas for machine learning and data analysis.

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.53

Volume 4, Issue 1, September 2024

Making sure the LCD is wired appropriately for data display and attaching the sensors to the Arduino with proper cables are the first steps in the hardware assembly. The purpose of the Arduino code is to read, process, and update the LCD display from sensor data. The sensor data is transferred into Jupyter Notebook for pre-processing and analysis in the software component. Then, using this data, a Random Forest Regression model is trained to forecast air quality indicators. The accuracy of the model is assessed and adjusted. In order to provide real-time predictions based on sensor data, the trained model is finally connected with the Arduino setup. Every component of the system is safely built for deployment, and the complete system is tested to guarantee performance and dependability.

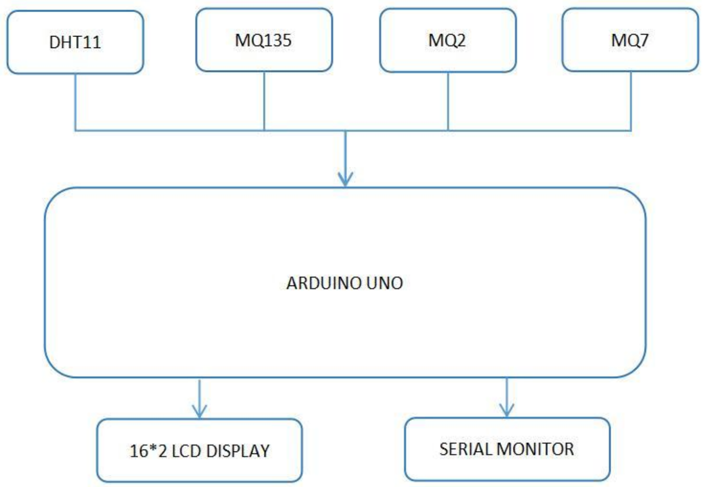

Figure 1: Block diagram of the System

IV. RESULTS AND DISCUSSION

The air quality monitoring system successfully collected real-time environmental data from the sensors, including temperature, humidity, smoke, carbon monoxide, and hazardous gas levels. The Random Forest Regression model, trained using this data, exhibited strong predictive accuracy. These metrics indicate that the model effectively forecasts air quality indices based on sensor inputs. Real-time data displayed on the 16x2 LCD was consistently accurate, validating the system's capability for continuous monitoring.

During testing, the system demonstrated reliable performance across various conditions, proving it to be an effective and low-cost solution for real-time air quality assessment. While the sensors provided good accuracy, occasional fluctuations in sensor readings were noted due to environmental factors, highlighting the importance of regular calibration. Despite these minor challenges, the system's combination of IoT and machine learning technologies offers valuable insights into air quality trends, making it a useful tool for environmental monitoring and public health.

Figure 2: Prototype of the Device

DOI: 10.48175/568

Sl. No.	CO(ppm)	Smoke(ppm)	Sulphur(ppm)
1	184	134	57
2	176	127	62
3	154	135	54
4	187	119	68
5	189	132	55

Figure 3: Results obtained

V. CONCLUSION

The air quality monitoring system created for this project effectively combines cutting-edge machine learning algorithms with reasonably priced hardware components to enable real-time air quality index monitoring and prediction. The system gathers important environmental data, including temperature, humidity, smoke, hazardous gasses, and carbon monoxide levels. It does this by using an Arduino microcontroller interfaced with DHT11, MQ2, MQ135, and MQ7 sensors. For quick feedback, this data is processed and shown on a 16x2 LCD screen. The software component makes effective data analysis and model training possible by utilizing Python and modules like NumPy, Pandas. The Jupyter Notebook was used to train and test the Random Forest Regression model, which showed good accuracy in predicting air quality indices. Throughout the testing phase, the system consistently generated accurate and trustworthy real-time forecasts, demonstrating its use for environmental monitoring. Because of its price and accessibility, the system can be widely employed in environmental research, public health monitoring, and urban planning. However, the study did highlight a few challenges, such as the need for regular sensor calibration and the impact of external ambient variables on sensor performance. Future research may look into other machine learning techniques and employ additional sensors to gather a wider range of pollutants in an effort to increase forecast accuracy. In conclusion, this study demonstrates how machine learning and Internet of Things technology may be combined to create a useful and practical air quality monitoring system. The results promote improved environmental monitoring and public health initiatives with substantial practical value. By laying the groundwork for future advancements in smart environmental monitoring systems, this technology makes deployments more expansive and scalable power supply.

VI. FUTURE SCOPE

In the future, this air quality monitoring system may receive additional sensors to detect a wider range of pollutants, such as particle matter (PM2.5 and PM10). Remote monitoring, historical data storage, and real-time analytics could be facilitated by data transmission to cloud platforms using wireless communication modules like Bluetooth or Wi-Fi. If more complex algorithms, such as neural networks, are incorporated into the machine learning model, prediction accuracy might rise. Real-time warnings and updates on air quality could be provided to consumers through the development of a smartphone application or web application. Automated techniques for sensor calibration could be used to maintain accuracy over time. Moreover, expanding the system's deployment to additional locations would allow for thorough environmental monitoring on a bigger scale.

REFERENCES

[1]. D. Rajasekar, A. Sekar, and M. Rajasekar, “Air Quality Monitoring and disease prediction using IOT and machine learning,” International Journal of Innovative Research in Computer Science & Technology, vol. 8, no. 6, pp. 389–395, Nov. 2020. doi:10.21276/ijircst.2020.8.6.4

[2]. C. C. Goh et al., “Real-time in-vehicle air quality monitoring system using machine learning prediction algorithm,” Sensors, vol. 21, no. 15, p. 4956, Jul. 2021. doi:10.3390/s21154956

[3]. G. K. Kang, the Department of Computer Engineering, San Jose State University, USA, J. Z. Gao, Chiao, S. Lu, and G. Xie, “Air quality prediction: Big data and machine learning approaches,” Int. J. Environ. Sci. Dev., vol. 9, no. 1, pp. 8–16, 2018.

[4]. P. Bhalgat, S. Pitale, and S. Bhoite, “Air quality prediction using machine learning algorithms,” Int. J. Comput. Appl. Technol. Res., vol. 8, no. 9, pp. 367–370, 2019.

[5]. A Mishra, “Air Pollution Monitoring System based on IoT: Forecasting and Predictive Modeling using Machine Learning,” 2018.

[6]. S. Ameer, Comparative Analysis of Machine LearningTechniques for Predicting AirQuality in Smart Cities(2019), vol. 7, 2019.

[7]. S. D. Vito, G. Francia, E. Esposito, S. Ferlito, F. Formisano, E. Massera, Adaptive machine learning strategies for network calibration of IoT smart air quality monitoring devices(2020) Saverio De Vito. Pattern Recognition Letters , 2020