Acceleron Aerospace Journal || AAJ.11.2106-2311
Volume 1, Issue 3, pp (47-53)
Research Paper
ISSN- 2583-9942 || DOI: 10.61359/11.2106-2311

FEMTOSAT-Based Air Quality Monitoring: Leveraging Satellite Data and LoRa Communication for Improved AQI Predictions

Tharani R.K*

UG Researcher, Department of Information Technology, Dr. M.G.R. Educational and Research Institute Chennai, India - 600095

ORCID: 0009-0002-0194-1235

Abstract: The exponential growth of the global population has given rise to an alarming surge in air pollution levels, casting profound repercussions on economies, ecosystems, and human well-being. The predominant method for assessing air quality involves the deployment of sensors atop buildings at regular intervals. However, this approach faces significant drawbacks, including heightened energy consumption associated with each building's sensor infrastructure and its inapplicability in sparsely populated rural areas. Addressing these limitations, the utilization of FEMTOSAT technology emerges as a solution, capitalizing on satellite data for autonomous air pollution monitoring, analysis, and mitigation. Amidst the evolving scientific landscape, the integration of Low Power Wide Area Network (LoRa) technology assumes a pivotal role within the Internet of Things (IoT), facilitating long-range data communication with minimal power consumption. In this context, LoRa communication serves as the conduit for transmitting and receiving data from satellites via RF signals. The satellite-derived environmental data, thus collected, serves as the foundation for computing the Air Quality Index (AQI) at specific locations, a critical metric that informs us about air quality conditions, whether pristine or contaminated. The AQI computation factors in various pollutants, including NO2, CO, O3, PM2.5, SO2, and PM10, all of which significantly influence air quality. This study employs a range of machine learning (ML) techniques, including time series analysis, linear regression, Support Vector Machines (SVM), and logistic regression, to predict and forecast AQI values. These models amalgamate AQI data from diverse sources, yielding robust and dependable AQI prediction models. Notably, modern sensor technology simplifies and enhances data collection accuracy. In the realm of environmental data analysis, only ML algorithms can grapple with the complexity of processing vast datasets to generate precise and trustworthy predictions. The incorporation of integrated sensors as payload in the Femto Sat mission epitomizes the mission's objectives. This system is characterized by its cost-effectiveness, lightweight design, durability, redundancy, and user-friendly interface, requiring minimal power consumption for operation.

Table of Contents

1. Introduction

The Air Quality Index (AQI), a key indicator of air quality, depends on various individual factors, including the concentrations of NO2, CO, O3, PM2.5, SO2, and PM10. This research endeavors to predict and forecast AQI using Machine Learning (ML) techniques, including linear regression, Support Vector Machine (SVM), Logistic regression, and time series analysis, among others. To overcome the challenges associated with traditional methods, we leverage FEMTO SAT technology, harnessing satellite data for autonomous air pollution monitoring, analysis, and mitigation.

*UG Researcher, Department of Information Technology, Dr. M.G.R. Educational and Research Institute, Chennai, India. **Contact:** tharanikumaran46@gmail.com
**Received: 09-September-2023 || Revised: 15-September-2023 || Accepted: 20-September-2023 || Published Online: 30-September-2023

Recent advancements in smartphone electronics have significantly facilitated the development of miniaturized satellite components, making this system cost-effective, lightweight, robust, redundant, and power-efficient. In contrast to conventional techniques relying on probability and statistics, ML-based AQI prediction models have demonstrated greater reliability and consistency. The integration of advanced technologies and sensors has streamlined data collection, ensuring accuracy and precision [1-5].

The demands of analyzing extensive environmental data necessitate the efficiency and effectiveness of ML algorithms. We begin by developing a Linear Regression (LR) model, a supervised machine learning approach, to predict AQI. Input features consist of data collected from NO2, Ozone (O3), PM2.5, and SO2 sensors on the satellite, while the optimized AQI serves as the target for training the regression model. Model parameters are validated against new and unseen sensor data.

This project's primary objective is to predict the Air Quality Index using Machine Learning. We compare various learning algorithms, including decision trees, random forests, naive Bayes, and k-nearest neighbors, in our system. Additionally, the utilization of multiple CubeSats in different locations allows for the collection of environmental data, encompassing various air pollutants such as PM10, PM2.5, NO2, SO2, CO, O3, NH3, and Pb, which contribute to AQI level predictions [5-10].

Our mission involves the development of a Femto-CubeSat to collect atmospheric data in an extreme Low Earth orbit (eLEO). Few CubeSat missions operate in this region, making our venture a unique opportunity to gather valuable atmospheric insights from a scientifically rich and underexplored environment.

2. Existing System [5-15]

- The prevailing method for measuring air pollution relies on the installation of sensors fixed atop buildings at specified intervals.
- However, this method suffers from a significant drawback: each building's sensor system consumes additional power, rendering it impractical for deployment in sparsely populated rural areas.
- In today's context, there is an increasing energy expenditure associated with signal transmission.
- Moreover, the utilization of a substantial amount of hardware is required for decoding and monitoring satellite instruments both on the ground and in Earth's orbit. These instruments play a pivotal role in collecting vital data about the composition of our atmosphere.

Instruments both on Earth's surface and aboard orbiting satellites play a crucial role in gathering comprehensive data about the composition of our atmosphere. For instance, NOAA's GOES-R Series satellites continuously monitor particle pollution within our atmosphere. These airborne particles encompass a spectrum, including smoke particles stemming from wildfires, dust and sand particles during storms, pollutants from urban and industrial sources, and volcanic ash. Furthermore, the JPSS satellite series extends its capability to measure ground-level ozone, a key component of air quality assessment.

GOES-R Series satellites offer the advantage of providing particle pollution measurements approximately every five minutes throughout daylight hours. In contrast, JPSS satellites boast a maximum resolution measurement of aerosols across the entire globe, delivering this data once daily. Their capabilities extend to tracking the movement of aerosols across the planet, affording a comprehensive understanding of their dispersion patterns. Additionally, JPSS satellites can gauge carbon monoxide levels, a significant marker of poor air quality resulting from events such as wildfires. Despite these advancements, conventional approaches to air quality assessment have faced significant limitations, primarily stemming from inadequate access to robust longitudinal data.

3. Proposed System

The primary objective of this project is to develop a FEMTOSAT (Environmental Pollution Detection with Integrated Sensors as Payload) for autonomous air pollution monitoring, analysis, and mitigation. The satellite is equipped with a range of low-cost sensors capable of measuring various atmospheric parameters such as Particulate Matter, Ozone, Carbon Monoxide, humidity, temperature, pressure, wind speed, and precipitation at different altitudes within a specific geographic area. The data collected by the satellite serves as a valuable dataset for training and predicting Air Quality Index (AQI) levels in different regions [1-5].

To streamline data processing and reduce the workload on ground stations, cloud computing techniques are integrated to store and monitor the satellite-received data. Machine learning algorithms are employed to efficiently decode the raw satellite signals into digital data, which can then be broadcasted to digital platforms, including websites and apps. The ultimate goal of this project is to interpret the AQI using machine learning models designed to analyze satellite-derived data.

In this research, an ML model is developed to work with existing air pollutant data and historical toxin levels, employing machine learning to forecast future contamination trends. The collected data is stored in an Excel sheet for subsequent analysis. The Femto satellite incorporates a variety of sensors dedicated to gathering information on air pollutants and pollution levels. Multiple Femto satellites may be employed to collect environmental data from diverse locations, encompassing data on air pollutants such as PM10, PM2.5, NO2, SO2, CO, O3, NH3, and Pb, which is then used to estimate AQI levels.

Furthermore, this project explores the feasibility of utilizing cost-effective Femto-satellites in the field of remote sensing, offering students an invaluable opportunity to engage in a genuine engineering project replete with real-world engineering challenges.

4. Architectural Design and Deployment of Femto Sat

Within the FemtoSat, integrated sensors are seamlessly connected to a Microcontroller, alongside the On-Board Computer (OBC) and LoRa Tx transmitter, forming an Embedded IoT system. This system includes the NodeMCU ESP8266, a Wi-Fi module, with the LoRa receiver (Rx) connected to it. Data processing takes place within the NodeMCU, utilizing programming techniques to convert the data into a digital format. The data is then stored in the AWS cloud and further transformed into the required CSV file format. This CSV dataset becomes invaluable for training and predicting Air Quality Index (AQI) levels across various regions, serving as the foundation for model development [12-14].

Upon the launch of the Femto Satellite into extreme Low Earth Orbit (e-LEO), the data can be accessed on-site through a portable ground station. Simultaneously, the data is stored in the ThingSpeak Server, allowing for real-time visualization, analysis, and monitoring.

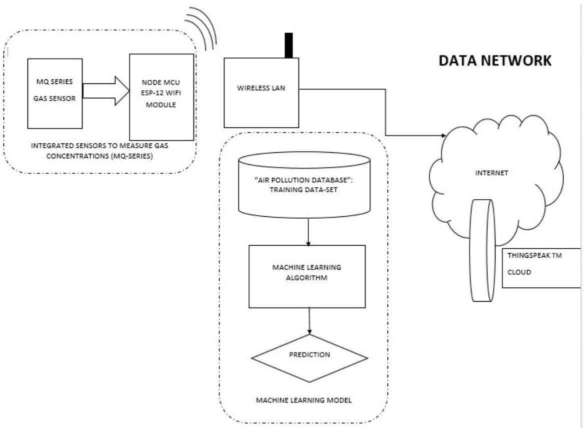


Figure 1: Typical Architectural Design

An IoT analytics solution known as ThingTalk facilitates the aggregation, visualization, and analysis of real-time data streams in the cloud. Data transmitted by devices to ThingTalk is immediately visualized using ThingSpeak. By employing the open-source ThingTalk Internet of Things (IoT) application and API, data can be saved and retrieved from objects via the HTTP protocol, whether over the Internet or a Local Area Network. This approach ensures that data can be remotely accessed from anywhere around the globe. To compute the Air Quality Index (AQI), data for at least three pollutants is required, with one of them being either PM10 or PM2.5. The AQI scale ranges from 0 to 500, with each pollutant having distinct concentration levels and associated health impacts.

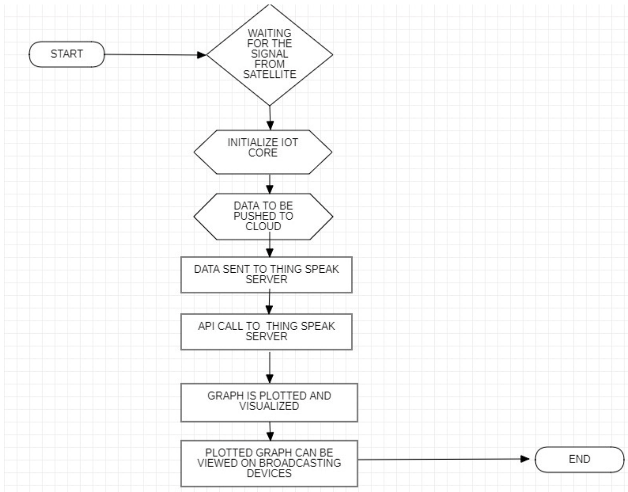


Figure 2: Cloud Flowchart

This project aims to construct a Machine Learning (ML) model for predicting the Air Quality Index (AQI) in pollution monitoring, employing a range of ML techniques such as SVM, Linear Regression, Random Forest, Logistic Regression, KNN, Decision Tree, and Naive Bayes. With this ML model, it becomes possible to predict both current and future air pollution levels in a given area.

Machine learning involves using a set of input variables (x) to forecast an output variable (y), where the input and output variables are interconnected. The ML model is trained using a provided dataset, and subsequently, feature analysis is employed to predict testing data. The ultimate objective is to predict the target column, the AQI Index, using one of the feature columns in the testing data.

The subsequent phase involves data visualization, a critical aspect of applied statistics. Machine learning primarily relies on quantitative justifications, and software analyses are valuable tools for gaining qualitative insights, which can help identify patterns and anomalies in the dataset.

The final steps encompass data pre-processing and validation, which are crucial as raw data from the FemtoSat needs to be refined for analysis. Various machine learning methods, including decision trees, Random Forests, Naive Bayes, Logistic Regression, SVM, and more, are applied to train the dataset and predict the AQI. The model accuracy is assessed by comparing different ML methods, and the prediction model is developed using the ML technique with the highest accuracy **[14-20]**.

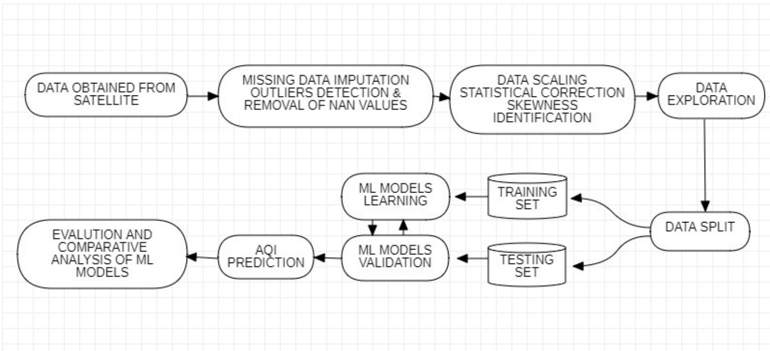
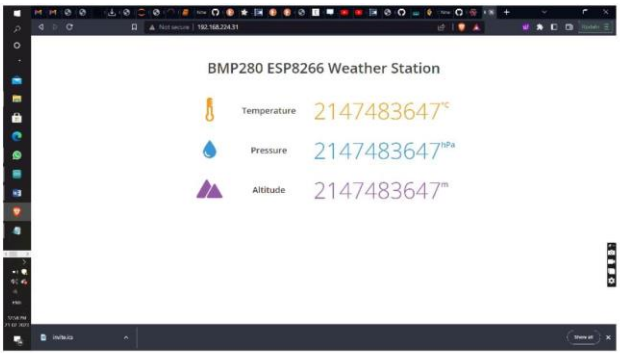


Figure 3: Flowchart of Proposed ML Model

5. Interpreted Results

Figure 4: This figure depicts temperature, pressure, and altitude measurements using the BMP280 sensor, with real-time data updates displayed on a website.



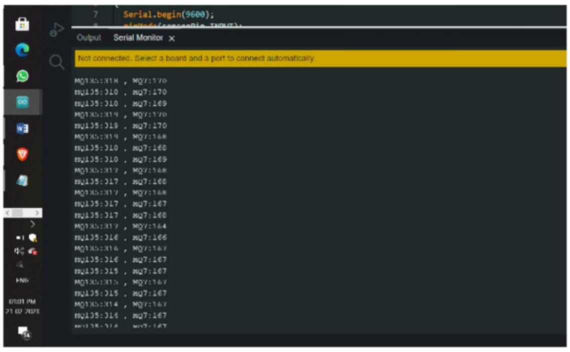
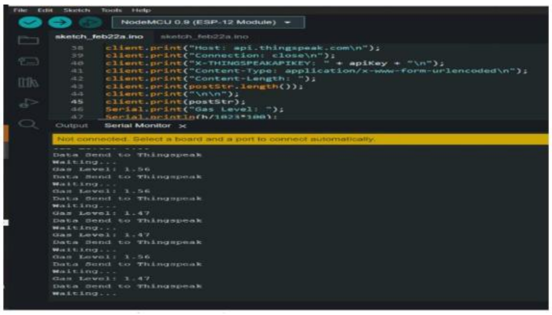


Figure 4 (left) & Figure 5 (right)

Figure 5: Here, sensors like MQ135 and MQ7 are implemented in the Arduino IDE to test sensor functionality, and the output is monitored with timestamps in the serial monitor.

Figure 6: This program monitors gas levels online using an ESP8266 and a Gas MQ135 gas sensor, with data sent to the cloud and monitored on the ThingSpeak server.



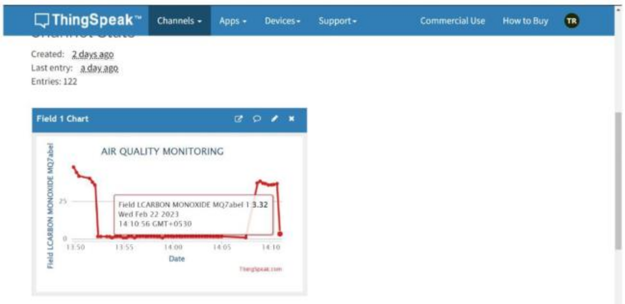


Figure 6 (left) & Figure 7 (right)

Figure 7: ThingSpeak server is utilized to transmit gas percentage measurements over the internet, enabling remote data access from anywhere globally.

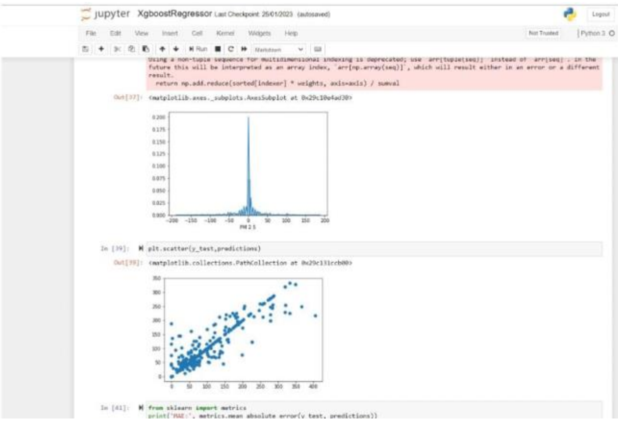


Figure 8

Figure 8: To interpret the pollutant, particulate levels and predict the AQI, various machine learning algorithms are employed, including Decision Tree (DT), k-Nearest Neighbor (k-NN), Support Vector Machine (SVM), Random Forest (RF), and Logistic Regression.

6. Conclusion

Estimating pollution levels is a complex endeavor due to the inherently unstable, dynamic, and variable nature of pollution data, which fluctuates both in location and time. Nevertheless, the profound impact of pollution on both human populations and ecosystems underscores the urgency of developing accurate predictive models. In this study, machine learning (ML) models have been employed to forecast pollutant levels, including NO2, SO2, PM2.5, PM10, and the Air Quality Index (AQI) using publicly accessible data. These models will serve as the foundation for future work, integrating satellite data to enhance predictive accuracy.

The advent of FEMTOSAT technology offers a promising approach to address large-scale air pollutant detection and monitoring. It provides a means to proactively measure pollutants and implement measures to combat air pollution using satellite data. This innovative approach involves deploying 1U femto-satellites equipped with integrated sensors to autonomously monitor air quality at specific locations and detect the presence of harmful pollutants. Our initial experiments have successfully measured concentrations of O3, CO, NH3, and PM2.5. While these findings are promising, further research and refinement are needed to optimize the system's performance.

The collected environmental data from the satellite is instrumental in calculating the Air Quality Index (AQI) for specific locations, highlighting the mission's goal of utilizing integrated sensors as payload. This holistic approach provides a comprehensive understanding of air quality in various regions.

7. Future Work

The solution lies in the deployment of FEMTO SAT technology, equipped with integrated sensors as payload, to monitor pollution levels across diverse locations. The mobility and versatility of these devices enable the monitoring of pollution at varying altitudes and geographic points. The development emphasizes the integration of cost-effective chipset-based modules, including the ATMEGA328P, MQ-7 Gas Sensor, MQ-135 Gas Sensor, and others. This study utilizes machine learning techniques such as Decision Tree (DT), k-Nearest Neighbor (k-NN), Support Vector Machine (SVM), Random Forest (RF), and Logistic Regression to assess pollutant and particle levels and forecast the air quality index. The ML model successfully predicts AQI levels and most pollution components based on both real-time observations and satellite data.

The Femto Sat analysis project's primary objective was to construct a Femto Satellite capable of collecting atmospheric data in an extremely low Earth orbit (eLEO). This initiative facilitates the collection of air data from eLEO, providing researchers with valuable insights from a scientifically rich but underexplored region, as only a limited number of CubeSat missions operate at such low altitudes. This project will continue to evolve, incorporating cutting-edge space science and engineering concepts to further advance our understanding of atmospheric dynamics and environmental monitoring in space.

8. References

[1] Air Quality Prediction and Monitoring using Machine Learning Algorithm based IoT sensor by G.Kalaivani Research Scholar-2021

[2] Smart City Air Quality Prediction using Machine Learning by Rishanti Murugan -2021

[3] Intelligent and Scalable Air Quality Monitoring With 5G Edge by Xiang Su, Xiaoli Liu, Jacky Cao, Petri Pellikka, Yongchun Liu, Pan Hui - March/April 2021

[4] Air Quality Prediction Of Data Log By Machine Learning by Venkat Rao Pasupuleti , Uhasri , Pavan Kalyan , Srikanth , Hari Kiran Reddy-2020

[5] Air Quality Sensing and Reporting System Using IoT by Rohan Kumar Jha- 2020

[6] Learning Algorithm based IoT sensor by G.Kalaivani Research Scholar-2021

[7] Smart City Air Quality Prediction using Machine Learning by Rishanti Murugan -2021

[8] Intelligent and Scalable Air Quality Monitoring With 5G Edge by Xiang Su, Xiaoli Liu, Jacky Cao, Petri Pellikka, Yongchun Liu, Pan Hui - March/April 2021

[9] Air Quality Prediction Of Data Log By Machine Learning by Venkat Rao Pasupuleti , Uhasri , Pavan Kalyan , Srikanth , Hari Kiran Reddy-2020

[10] Air Quality Sensing and Reporting System Using IoT by Rohan Kumar Jha- 2020

[11] Assessment of ambient air quality of a college campus by Komal Daxini , Tejas Turakhia , Rajesh Iyer , Abha Chhabra -2020

[12] Tingwei Wu, DexinQu , Gengxin Zhang , "Research on LoRa Adaptability in the LEO Satellites Internet of Things", IEEE Design and Test, 2016

[13] NASA Goddard Space Flight Center (2013) General Environmental Verification Standard (GEVS) [Online]. Available at: http://edge.rit.edu/content/P16103/public/Systems%20Design%20Documents/P16103_GSFC-STD-7000A.pdf (Accessed: 09 December 2021).

[14] C.-L. Hor, S. J. Watson, and S. Majithia, "Daily load forecasting and maximum demand estimation using arima and garch," in 2006 International Conference on Probabilistic Methods Applied to Power Systems. IEEE, 2006, pp. 1–6.

[15] L. Y. Siew, L. Y. Chin, and P. M. J. Wee, "Arima and integrated arfima models for forecasting air pollution index in shah alam, selangor," Malaysian Journal of Analytical Sciences, vol. 12, no. 1, pp. 257–263, 2008.
[16] J. Zhu, R. Zhang, B. Fu, and R. Jin, "Comparison of arima model and exponential smoothing model on 2014 air quality index in yanqing county, beijing, china," Appl. Comput. Math, vol. 4, pp. 456–461, 2015.
[17] I. Goodfellow, Y. Bengio, and A. Courville, "Machine learning basics," Deep learning, vol. 1, no. 7, pp. 98–164, 2016.
[18] U. Brunelli, V. Piazza, L. Pignato, F. Sorbello, and S. Vitabile, "Three hours ahead prevision of so2 pollutant concentration using an elman neural based forecaster," Building and Environment, vol. 43, no. 3, pp. 304–314, 2008.
[19] R. Sharda and R. Patil, "Neural networks as forecasting experts: an empirical test," in Proceedings of the International Joint Conference on Neural Networks, vol. 2. IEEE, 1990, pp. 491–494.
[20] I. Alon, M. Qi, and R. J. Sadowski, "Forecasting aggregate retail sales:: a comparison of artificial neural networks and traditional methods," Journal of retailing and consumer services, vol. 8, no. 3, pp. 147–156, 2001..

9. Biography

Tharani R.K is an undergraduate researcher in the Department of Information Technology at Dr. M.G.R. Educational and Research Institute in Chennai, India. She is passionate about leveraging space technology and environmental science for innovative solutions. Tharani's recent paper, "FEMTOSAT-Based Air Quality Monitoring," explores the use of satellite data and LoRa communication to improve AQI predictions. Her dedication to cutting-edge research reflects her commitment to addressing environmental challenges with technology. Tharani is poised to make a significant impact in IT and environmental science through her innovative work.

10. Acknowledgement

I wish to express my sincere thanks to my family and mentor for their unwavering support, encouragement, and invaluable guidance.

11. Conflict of Interest

The author have no conflict of interest to report.

12. Funding & Paper Information

No external funding was received to support this study.