

Tamper Proof Air Quality Management System using Blockchain

Vaneeta M^{1*}, Deepa S R², Sangeetha V³, Kamalakshi Naganna⁴, Kruthika S Vasisht⁵, Ashwini J⁶, Nikitha M⁷,
Srividya H.R⁸

Department of Artificial Intelligence and Machine Learning, K. S. Institute of Technology, Bangalore, India^{1*}

Department of Computer Science and Design, K. S. Institute of Technology, Bangalore, India²

Department of Computer Science and Engineering, Ramaiah Institute of Technology, Bangalore, India³

Department of Computer Science and Engineering, Sapthagiri College of Engineering, Bangalore, India⁴

Department of Computer Science and Engineering, K S Institute of Technology, Bangalore, India⁵

Department of Development, Gupshup Technologies, Bangalore, India⁶

Department of Development, TVAST IT Solutions, Bangalore, India⁷

Department of Development, Accolite Digital, Bangalore, India⁸

Abstract—One of the most important concerns facing urban regions across the world is air pollution. As a result, it's critical to monitor pollution levels and notify the public on the state of the air. An indicator called the Air Quality Index (AQI) does this by mapping the concentration of different contaminants into a single number. Because the examination of pollutant data is frequently opaque to outsiders, poor environmental control judgments may result. This has led to a need for a tamper-proof pollution management system for use by authorities, like the state and central pollution boards. To address these challenges, a model using machine learning algorithms to predict the air quality and store that information in the blockchain is proposed. Machine learning algorithms are used to categorize the air quality, and blockchain technology guarantees a permanent, tamper-proof record of all air quality data. Such a system might address the persistent issues with data dependability, immutability and trust in pollution control. The execution time of two main operations on blockchain are measured. The execution time of the put block is measured as 10 ms and the get block gets executed in 1 ms that fetches data from the blockchain ledger.

Keywords—Air pollution; air quality index; machine learning; blockchain technology

I. INTRODUCTION

Air quality refers to how well the air is suited for breathing by people, animals, and plants. An average healthy person breathes approximately 14,000 liters of air each day. As a result, poor air quality may have an effect on the quality of life for both the present and future generations by hurting human well-being, the environment, the economy, and urban sustainability.

AQI Category (Range)	PM ₁₀ (24hr)	PM _{2.5} (24hr)	NO ₂ (24hr)	O ₃ (8hr)	CO (8hr)	SO ₂ (24hr)	NH ₃ (24hr)
Good (0-50)	0-50	0-30	0-40	0-50	0-1.0	0-40	0-200
Satisfactory (51-100)	51-100	31-60	41-80	51-100	1.1-2.0	41-80	201-400
Moderately polluted (101-200)	101-250	61-90	81-180	101-168	2.1-10	81-380	401-800
Poor (201-300)	251-350	91-120	181-280	169-208	10-17	381-800	801-1200
Very poor (301-400)	351-430	121-250	281-400	209-748	17-34	801-1600	1200-1800
Severe (401-500)	430+	250+	400+	748+	34+	1600+	1800+

Fig. 1. Air quality index (AQI) category range

The government keeps an eye on the air quality in various locations to determine the pollution level and to ensure that pollutant levels are within acceptable limits for human health. Air quality agencies can better plan how and when they will take action to safeguard the public's health by identifying how much pollution is present in a given location. Fig. 1 shows the AQI category range of the major pollutants.

The current technique for tracking industrial pollution is centralized, with a lack of openness and the possibility of data falsification. As a result, a consistent and tamper-proof mechanism must be utilized, such as secure software with data encryption and simultaneous data transfer directly to the regulator. Blockchain delivers Distributed ledger technology (DLT), which possesses the potential to solve many of the present system's open issues. Blockchain nodes are a network of multiple storage and computing devices that replicate data over a highly available and fault-tolerant infrastructure. Thus, blockchain facilitates the operation of a distributed database that is transparent and tamper-resistant. There is a need to design and develop an application using machine learning to predict Air quality category and store it on the blockchain that ensures it is tamper-proof and secure. The proposed system has three modules namely machine learning model, Blockchain network and Client application

The machine learning model is trained using industrial air pollution data. Supervised learning algorithms such as random forest classifier, decision tree classifier and Naive Bayes are used to predict the air quality index and the quality range of the given input data. The design of the ML model has these phases. The dataset comprises pollutant concentration information from over 15 industrial areas across India. This data set has around 37-40 pollutants, but the seven most appropriate pollutants are considered. The dataset is cleaned and partitioned into training and testing data. On comparing the results, best results were obtained from decision tree classifier with an accuracy of 99.6%.

The next module is the Blockchain network. The chaincode contains the ML model deployed in it. Once the client supplies the data to the blockchain, the chaincode that has the ML

model will start executing. Once the air quality category has been predicted, it is sent as a transaction to the users of blockchain for endorsement according to the endorsement policy decided prior. The users of this system are the members of state and central pollution control boards. The transaction is then ordered and validated by the respective chain code. If the transaction is validated, it is stored on the ledger.

S. Benedict [1] proposed a blockchain enabled IoT cloud implementation to tackle the existing issues of security hazards and performance inefficiencies in smart cities. It particularly highlights the implementation of chaincodes for air quality monitoring systems in Smart Cities. The proposed architecture named as IoT enabled Blockchain for Air Quality Monitoring System (IB-AQMS) is illustrated using experiments.

Abu Buker [2] proposed an indoor AQI monitoring system to predict the AQI through the Neural Network Algorithm and Block-chain. The Indoor Air Quality system consists of sensors such as temperature, humidity, Carbon Di Oxide, Particulate Matter, Carbon Mono Oxide, and LPG. The Neural Network decision-making model is used to predict the AQI. The suggested IoT-based smart block-chain technology plays a vital role by imparting scalability, privacy, and reliability.

The significant objectives of proposed work are:

- To design and experiment with Random Forest classifier, Decision tree classifier and Linear regression algorithms to predict air quality category and consider the one with best accuracy.
- To implement blockchain-based solutions to resolve the ongoing issues with data dependability in pollution monitoring.
- To ensure a permanent, immutable record of all air quality data of industries.
- To develop GUI for the user to maintain an immutable record of all air quality information.

The organization of the paper is as follows: Section II discusses about the background literatures used for modelling AQI index and use of blockchain based technology for storing data. Section III elaborates the proposed solution and discusses about detailed implementation followed by discussion of result in Section IV. Finally, the conclusive remarks are provided in Section V and future scope is presented in Section VI.

II. BACKGROUND

The work [3] by S. Mahanta investigates the efficacy of different existing prediction models in forecasting AQI values based on input values. According to Dyuthi Sanjeev's article [4], the AQI is calculated based on pollutants or attributes that have the greatest impact on air pollution. The Random Forest model is the most efficient, according to the research, with a score of 99.4 percent accuracy. Timothy M's research [5] proposes a method for evaluating air quality by building prediction models that link sensor data to an air quality score. Aditya C R [6] uses logistic regression to determine if a data sample is contaminated, as well as auto regression to predict future PM2.5 values based on present PM2.5 data. The study's

purpose [7] is to examine a range of existing prediction models to see how effective they are at predicting data from the study area.

Yue-Shan Changa [8] offers an ALSTM (Aggregated Long Short-Term Memory Model) that combines regional air-quality monitoring stations, industrial zone stations, and external emission source stations. Mahmoud Reza Delavar's [9] study provides a novel method for predicting air pollution in urban regions based on both stationary and non-stationary sources, using machine learning and statistical approaches.

This study [10] will be using Data Mining and Machine Learning models in this research project, to forecast the AQI and classify the AQI into buckets labeled as Good, Satisfactory, Moderate, Poor, Very Poor and Severe. Regression models are used to predict AQI. In order to predict AQI bucket, KNN (K Nearest Neighbors) algorithm with repeat CV classification is used. Station-level data from Indian cities was used to accurately classify and forecast AQI Labels. KNN and repeat CV classification performed best in terms of accuracy. M. Lücking et al. [11] offer a software design for a pollution monitoring system (PMS) based on distributed ledger technology and the long-range protocol, which is described in this paper. It provides adaptable, traceable, and energy-efficient monitoring. Multiple unresolved issues in the functioning of pollution monitoring systems, such as storing data that is invalid or susceptible to tampering, are addressed by distributed ledger technology in a Hyperledger Fabric blockchain.

One of the prime components of blockchain is cryptography for providing confidentiality and authentication using efficient keys. In regard to this author Vaneeta in paper [12] proposes multi-tier framework by including a superior authentication scheme using enhanced public key encryption and digital signature. Sina Rafati Niya's paper [13] proposes an automated approach for measuring, monitoring, and storing air and water quality in factories, lakes, and other sites, based on an IoT and Blockchain-based system.

The proposed system in paper [14] collects real-time air pollution data from industrial locations using 5G wireless IoT sensors and transmits encrypted blockchain data to the index measurement service and cloud via a periodic blockchain transaction. This device enables real-time pollution monitoring in industrial settings and also protects data from tampering. The distributed messaging protocol and blockchain's encryption technologies increase the efficiency of data processing and exchange, while maintaining data integrity. The main objective of this research study [15] is to give an overview of technologies such as Artificial Intelligence, Blockchain and Internet of Things (IoT) and their current applications in the fields of public healthcare and the environment.

Air pollution has been a source of great concern for a long time, but it has come to the attention of stakeholders only recently. The Air Act of 1981 was the piece of legislation that established the requirement for air quality monitoring and opened the door to the monitoring techniques employed in India under the CPCB's oversight. This study [16] suggests that we adopt a thorough approach to manage air pollution. In this

paper [17] development of less expensive, simple-to-use, portable air pollution monitoring sensors, which deliver high-time resolution data in almost real-time and makes access to environmental data convenient is discussed. A variety of air contaminants can already be monitored by sensor devices, and new technologies are constantly being developed.

Regulating and monitoring pollution emissions becomes increasingly essential for battling the illness. This research [18] proposes an Internet of Things (IoT) based system that uses low sensors to monitor pollutants. It is developing a hardware layer of device that is capable of measuring concentrations of pollutants by means of three sensors, respectively, MQ-131, PMSA003 and MICS-6814. The given research study [19] proposes use of Internet of Things (IoT) sensors to periodically collect air quality information such as pollutant concentration and transmit the same over Low power wide area (LPWA) network. IoT cloud is used to process and analyze the data. The participatory urban sensing architecture for PM2.5 monitoring described in this research [20] has more than 2500 devices operating in Taiwan and 29 other nations. The framework's open system design, which is built on the ideas of open hardware, open source software, and open data, is its defining feature.

The paper [21] introduces CNN-ILSTM, an Air Quality Index prediction model based on Convolution Neural Networks (CNN) and Improved Long Short-Term Memory (ILSTM). The experimental data set includes air quality data from 00:00 on April 4, 2019 to 23:00 on June 30, 2021 in Shijiazhuang, Hebei Province, China. Air pollution is affecting public health and causing a slew of health-related issues, resulting in a significant medical bill each year. Taking air quality information into account, the study [22] offers a safe path when the air quality index is poor to reduce the impact on human health. Dijkstra's method is used to discover the safest path between source and destination. This study also points out research gaps in various studies on similar grounds. The first completely expandable blockchain architecture for supporting distributed applications is called Hyperledger Fabric. Additionally, Fabric is the first blockchain platform to support distributed applications created in common, all-purpose programming languages, independent of a native coin as a system backend.

As indicated in the study [23], Fabric proposes a revolutionary architecture that is evocative of middleware-replicated databases and isolates transaction execution from consensus while enabling policy-based endorsement. Author Sangeetha in paper [24] proposed that Security issues are one of the core problems in mobile adhoc networks owing to the decentralized architecture. The proposed system introduces a new scheme that acts as multi-layer security under two different stages and enhances security in MANET's by modelling the different interactions among a malicious node and with a number of legitimate nodes.

Based on two publicly accessible datasets, this paper [25] regression models using support vector regression (SVR) and random forest regression (RFR) to predict the Air Quality Index (AQI) in Beijing and the nitrogen oxides (NOX) content in an Italian city. The performance of the regression models was assessed using the root-mean-square error (RMSE), correlation coefficient (r), and coefficient of determination (R²). SVR based model predicts AQI better whereas RFR based model predicts NOX concentration better. According to study in this paper [26], most of the research on Air quality uses Machine learning techniques and Big data analytics on data collected by IoT sensors. Aim of this study is to evaluate such techniques on air quality forecasting. Based on the observations made, study suggests the need for more research and development in real time air quality monitoring. To suffice the needs of future cities, an integrated air quality monitor with hybrid machine learning models can be developed that address impacts of dynamic quality on various atmospheric levels.

Research paper [27] proposes a real-time IoT based system for air quality monitoring. Study uses models such as CNN-LSTM-BOA (Convolution Neural Networks-Long Short Term Memory-Bayesian Optimization Algorithm) and other baseline models such as SVM (Support Vector Machine), ANN (Artificial Neural Network), Ensemble model. The LSTM model proved to be good for prediction. Results suggest that the proposed models perform better than baseline models. As part of further research, this study also suggests using statistical criteria, AI algorithms can evaluate performance and compare the results to publicly available data sets. Proposed technique in paper [28] used IoT sensors and Artificial Intelligence techniques that are said to reduce implementation costs to $\frac{2}{3}$ since before. The proposed model includes wireless sensor nodes to measure gas concentrations (MQ series) which are connected by IEEE 802.11 Wireless LAN AP to an IoT cloud that stores and maintains data in turn connected to a machine learning model responsible for predicting air quality levels. Model uses ARIMA prediction technology which stands for differential autoregressive moving normal model.

Research [29] proposes a predictive model using a multilayer perceptron, support vector regression and linear regression to predict future condition of air quality in a vehicle, based on data collected from sensors. Performance of these models can be evaluated using Root mean square error, coefficient of determination (R²), Mean Squared Error and Mean Absolute Error. For data collected, the support vector regression model had the highest performance in terms of R² and had a lower error rate.

III. PROPOSED SOLUTION

The suggested system's mechanism is depicted in Fig. 2. It demonstrates how the entire technique is broken down into five system implementation components.

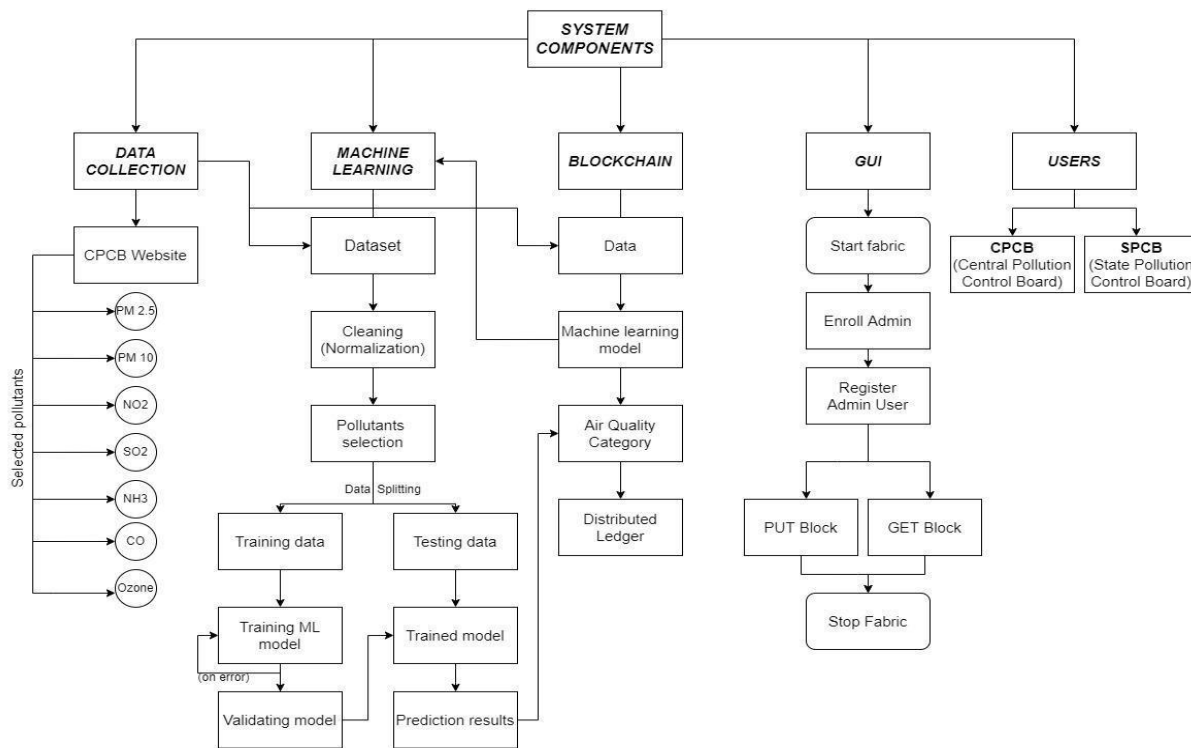


Fig. 2. System components

A. Data Collection

- Web scraping is used to collect data for the seven pollutants directly from the official CPCB website. This scraped data is given as input into the machine learning model deployed as the chaincode on the blockchain network. Chaincode then gives the AQI category as Good, Satisfactory, Moderately polluted, Poor, Very Poor and severe. The values of all seven pollutants and AQI categories are then stored on the distributed ledger.
- The web scraping process is divided into three steps. The first step is to set up to and from dates. For these consecutive dates 24hr-data is received. Second, is to pull the data for the setup dates in JSON format and finally, to parse this data and store this in a table form. The data is collected from an industrial area named Peenya in the city of Bangalore in Karnataka, India which is one of the biggest industrial areas in Asia.
- The data set is obtained from the Central Pollution Control Board's official website (CPCB). The website has around 15 stations in industrial areas all over India. The data is from Jan 1st 2020 till April 23rd 2022 for seven pollutants. Data is being considered on a 24-hour basis for five pollutants (PM2.5, PM10, NO2, SO2, NH3) and eight hourly basis for two pollutants (CO and ozone). The dataset has around 12600 rows.

B. Machine Learning Model

- The above air pollution data set is divided 80 percent for the machine learning model's training and 20 percent for its testing. Supervised learning methods-Linear regression, Random Forest classifier, and

Decision Tree classifier were the three machine learning techniques used to predict the air quality category of the given input data. The core nodes of the decision tree classifier represent dataset properties, the branches represent decision rules, and each leaf node represents the classification outcome. A 99.56 percent accuracy rate was generated by the decision tree classifier.

- Random forest algorithm is considered as an ensemble learning algorithm. The algorithm's core concept is to construct short and weak decision-trees with few attributes, in parallel, and then merge the trees to generate a single, powerful learning model by taking the majority vote or by just taking the average. An accuracy score of 99.06 % was observed for this model.
- Linear regression, a supervised learning algorithm, is the most common regression model which is used to determine how the independent variable(s) and the dependent variable are related. It is employed to ascertain how the value of the independent variable affects the value of the dependent variable. Finding the best fit line is essential when using linear regression since it reduces the error between the actual and projected values. A sloping straight line is used to represent the connection between the variables. The line that fits the data the best is the one with the least error or inaccuracy. An accuracy score of 91.79 % was observed for this model. Comparing the three learning algorithms. Decision tree classifier produced the most accurate results, with a 99.56 percent accuracy.

C. Blockchain Network

A blockchain is a type of distributed ledger or database that securely and impenetrably keeps a chain of data in the form of blocks in chronological order. The chain of blocks, also known as a ledger, is continuously expanding, thus new blocks are added to the end of the ledger. Each new block retains a reference to the content of the preceding block using a hash value. The distributed ledger material is secured using the public key encryption process, which also ensures consistency, irreversibility, and non-reputability. The block's immutability, anonymity, and compactness are guaranteed by the use of a cryptographic one-way hash function, such as SHA256. In a peer-to-peer (P2P) network, the ledger and its contents are copied and synced among several peers, forming a distributed ledger. There are three basic categories of blockchains: consortium blockchains, private permissioned blockchains, and public permissionless blockchains. All blockchain data is open to and visible to the general public since the permissionless blockchain type stresses the public component. The Bitcoin and Ethereum blockchains are two examples of such a blockchain. A private blockchain, on the other hand, permits only selected nodes to join the network, making it appear to be a type of distributed but nonetheless centralized network. Combining the two, the consortium blockchain only allows a predetermined set of nodes to take part in the distributed consensus process.

The proposed work implements private blockchain using Hyperledger Fabric. Hyperledger Fabric is a distributed ledger platform that is open-source and enterprise-ready. It features extensive privacy controls that ensure that only the information you want shared with "permissioned" (known) network participants is shared.

The operation of fault-tolerant distributed ledgers is principally due to distributed ledger technology (DLT). Each distributed ledger node keeps a local copy of the data, and new data is added to the ledger in the form of transactions. New transactions are validated using digital signatures and saved in node memory, which is then passed on to other DLT nodes in the network. Eventually, approved transactions are either directly put to the ledger or recorded in a block and then added to the ledger. Most DLT consensus systems (like Kafka or Raft) or even Byzantine fault-tolerant ones are crash-tolerant (e.g., Nakamoto consensus). Crash fault tolerance is the capacity of a consensus mechanism to reach consensus across all validating nodes notwithstanding the (temporary) unavailability of nodes.

Comparing private-permissioned distributed ledgers versus public-permissionless distributed ledgers, the former frequently provides greater flexibility (maintainability), faster speed (rapid transaction confirmation and maximum throughput), and a high level of transparency.

The users of our blockchain network are the State Pollution Control Boards (SPCB), Central Pollution Control Board organization (CPCB) and an orderer organization. All the three organizations form a private "subnet" of communication called channel as shown in Fig. 3.

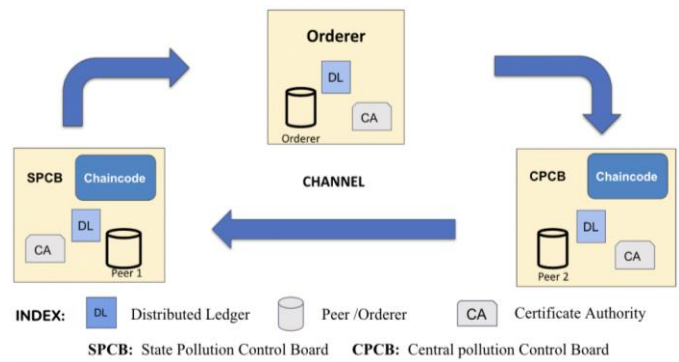


Fig. 3. Blockchain architecture

Chaincode (smart contracts) in Hyperledger Fabric are small programs written in Go, JavaScript/TypeScript or Java that contain the business logic to be executed as transactions on the blockchain. The chaincode has methods to store data received from machine learning models and query the ledger.

To digitally sign the response and endorse the transaction, the endorsement system chaincode is employed. After the transaction is ordered, a validation system chaincode compares the endorsements in the transaction to the endorsement policy defined for the chaincode. If the policy on endorsement isn't followed, then that transaction is marked invalid. Once the transaction has been endorsed, it is ordered and validated by respective peers and chaincode. If the transaction is validated successfully, it is stored on the ledger and every peer will maintain a copy of the same. Otherwise it is rendered unsuccessful.

Every transaction should be endorsed by either peers of SPCB or CPCB organizations, which is mentioned in the endorsement policy. Orderer nodes sequence groups of approved transactions into blocks and bundle them. These blocks are added to the blockchain. The Orderer then distributes blocks to all peers associated with it. Every peer validates the distributed block separately, maintaining consistency, to ensure the block is endorsed by the peers of the right organization and follows the endorsement policy. To prevent altering the ledger's state, all invalidated blocks are appended to an immutable block that was produced by the orderer and was designated invalid by the peer.

For a transaction to occur between the client application and the blockchain network, the client application primarily needs to have a certificate saying it can interact with the blockchain network and should have the necessary information of the network. The steps in transaction are as shown in Fig. 4.

1) The client application enrolls a user in order to get a valid certificate required to communicate with the Hyperledger Fabric blockchain network. Next the client application calls one or more peers to discover the network and since Hyperledger fabric is a private permission blockchain, users get to access only a part of the network. This step is not to be performed before every transaction. Once the required certificate along with topology of the required part of the blockchain network is obtained by the client application, the actual transaction process can be started.

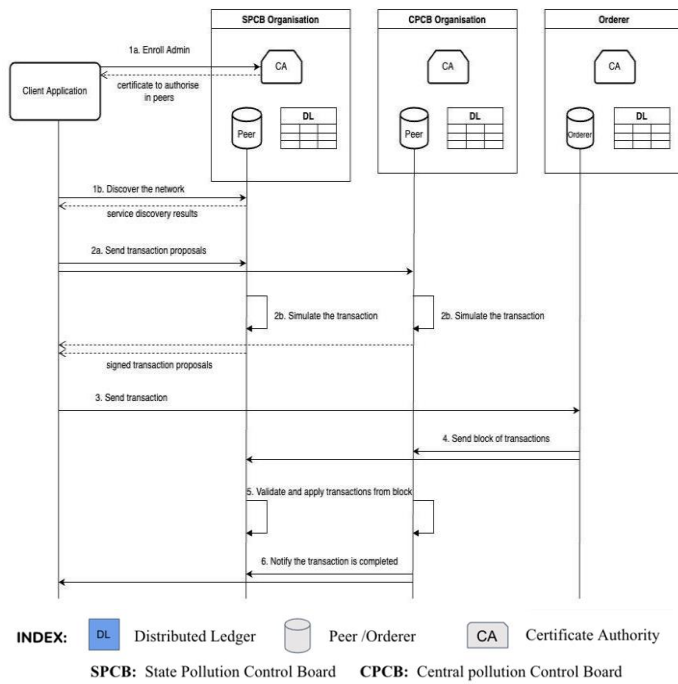


Fig. 4. Transaction proposal

2) Transaction proposals are sent to peers by the client application. These transactions should satisfy the endorsement policy, hence peers simulate the transaction by calling the smart contract which determines what has to be read and what has to be written to the world state based on their copy of the ledger if the transaction succeeds. This information along with the digital signatures from the peers is returned to the client application.

3) The client application next sends the transaction (which contains the simulation results and peer signatures) to the ordering service.

4) The ordering service creates a block once it has collected, validated, and ordered the appropriate number of transactions. The block is then transmitted to the channel's lead peers, who pass it along to the other peers.

5) The transactions are validated and applied by each peer that receives the block. The world state databases are updated with the transaction read/write sets as well as the blockchain copies on the peers are appended with new transactions.

6) The client application is expected to wait until relevant peers notify them that the transaction has been successfully completed. This notification indicates that it actually was appended to the blockchain network on a given peer.

Fig. 5 presents the flowchart of GUI system. The user application is for SPCB and CPCB. Once they access the website they can click on "START NETWORK" to start the network setup. Once the network is up and running, one can enroll the admin by clicking "ENROLL ADMIN" and once this is done, the admin user can be registered through "REGISTER USER".

D. GUI

The major functionality of the client is to collect values of pollutants on a 24 hr basis and transfer these to the blockchain along with the details of the concerned industrial area. The user inputs the "to" and "from" date after clicking the "PUT BLOCK" button as shown in Fig. 5 for which the required data is to be extracted. Once the date is entered, the concentration values of each pollutant for that particular "to"-date is fetched and displayed with the help of an API designed by us. The pollution data is then sent to the ML model with the help of another API. The model then calculates the AQI and determines the AQI category. This predicted category is then stored in the blockchain ledger along with the prefetched pollution data. Further, the ledger can be accessed through the GUI by clicking on the "GET BLOCK" button and the contents are displayed as shown in Fig. 5.

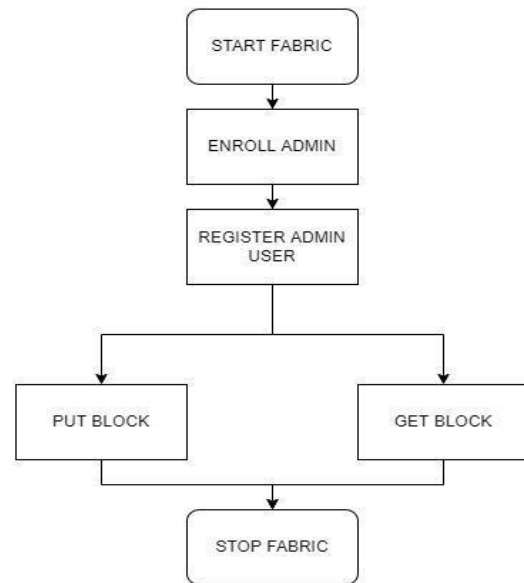


Fig. 5. GUI of the system

E. Users

The users of our system are the State pollution control board (SPCB) and the Central pollution control board (CPCB).

The SPCB is a council that analyzes, supervises, and partakes in an inquiry. The Board has a team of experts and a testing facility to evaluate the quality of various samples taken from industrial areas' soil, water, and air samples. It operates in accordance with the guidelines that the government occasionally sets. The CPCB advises the government and SPCB on issues pertaining to the implementation and enforcement of the Air, Water, and Environmental Acts.

IV. SYSTEM IMPLEMENTATION

Steps in setting up blockchain Network:

1) *Generate certificates using cryptogen tool:* All entities must be first recognized and granted permission before entering a consortium network under a permissioned blockchain. With the help of a bin/cryptogen tool provided by Hyperledger Fabric, crypto material is created. The

cryptographic components of the Test Network are created using a configuration file, and the finished product is stored in the directory structure. Along with these and docker compose files consortium network can be started.

2) *Generating orderer genesis block*: A "genesis block" is the initial block of a freshly established channel and first block of the "orderer system channel".

3) *Generating channel configuration transaction 'channel.tx'*: The shared configuration for a Hyperledger Fabric blockchain network is stored in a set of configuration transactions, one per channel.

4) *Generating anchor peer update for Org1MSP and Org2MSP*: Peers that are present outside of the organization are known as anchor peers. Any communication between organizations needs anchor peers.

Build your first network (BYFN) end-to-end test

Channel name: mychannel

Creating a channel...

Channel 'mychannel' created

This will launch all of the containers, and then drive a complete end-to-end application scenario.

5) *Having all peers join the channel...*

- a) peer1.org1 joined channel 'mychannel'
- b) peer0.org2 joined channel 'mychannel'
- c) peer1.org2 joined channel 'mychannel'

Administrators can conduct channel-related actions on a peer, such as joining the peer to a channel, using the peer channel command.

6) *Updating anchor peers for org1...*

a) Anchor peers updated for org 'Org1MSP' on channel 'mychannel'

Updating anchor peers for org2...

b) Anchor peers updated for org 'Org2MSP' on channel 'mychannel'

7) *All GOOD, BYFN execution completed*

- Installing smart contract on peer0.org1.example.com
- Installing smart contract on peer0.org2.example.com
- Instantiating smart contract on mychannel
- Waiting for instantiation request to be committed ...
- Submitting initLedger transaction to smart contract on mychannel
- The transaction is sent to the two peers with the chaincode installed (peer0.org1.example.com and peer0.org2.example.com) so that chaincode is built before receiving the following requests

8) *Total setup execution time: 181 secs ...*

V. RESULTS AND DISCUSSION

The below Fig. 6 shows the User Interface with start network, Enroll admin and register user button options. Once the start network option is selected the blockchain network is established between two organizations of CPCB and SPCB. During start network steps 1 to 8 get executed. After the network is up the admin is enrolled by entering user name and password. Once the admin is logged the user can register by entering username and password.

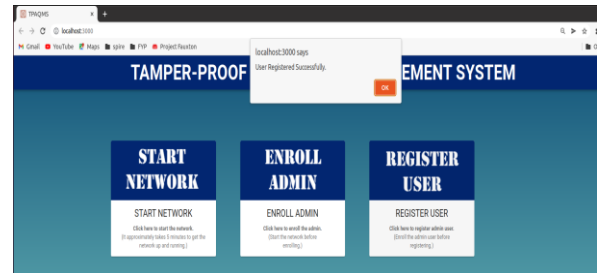


Fig. 6. User Interface

The Fig. 6 shows the User Interface for inserting the from and to dates for which the data must be fetched from the CPCB website. After fetching data of pollutants from CPCB website the Machine Learning model for calculation of AQI index is executed and its air quality category is identified. Put Block option places the transaction consisting of pollutant values along with AQI category on the blockchain ledger. Get block option retrieves block from ledger. The blockchain network can be stopped by selecting Stop Network option (see Fig. 7).

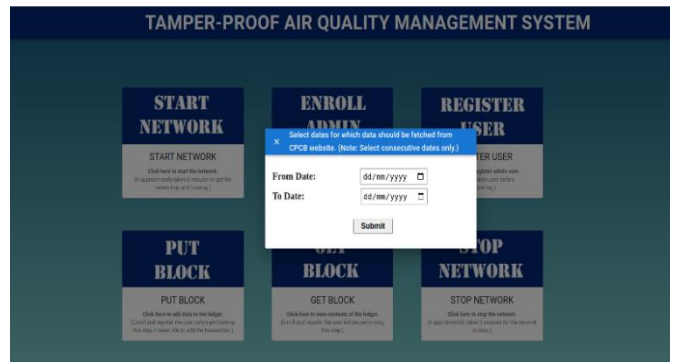


Fig. 7. Input dates to get pollutants

Fig. 8 shows the ledger data maintained by blockchain network accessible only to two organizations. It stores date with pollutant values and AQI category.

Date	PM2.5	PM10	NO2	SO2	NO	CO	Count	AQI Category
10 Jun 2022	37.51	93.24	8.99	16.93	36.86	0.57	2812	Moderately Polluted
15 Jun 2022	31.44	67.55	10.2	17.17	33.31	0.52	2937	Moderately Polluted
19 Jun 2022	27.9	72.35	13.97	9.05	24.59	0.5	414	Satisfactory

Fig. 8. Ledger data

The performance parameter considered is time required for setting the blockchain network. The graph in Fig. 9 represents the average execution time of certain transactions in blockchain network.

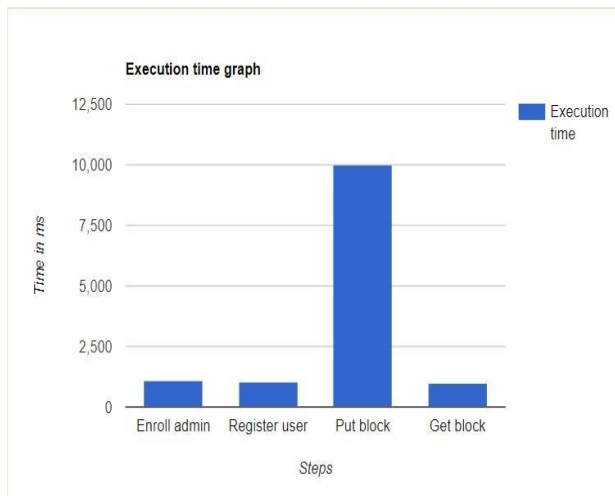


Fig. 9. Execution time

The configuration of a computer system considered is 64bit OS, 12 GB RAM, i7 core processor, 1.99 GHz. For a transaction to occur in a private blockchain network such as Hyperledger fabric, the client application needs to enroll and register the user whose execution times are 1000ms. The execution time of the put block is 10,000ms which includes fetching pollution data blockchain ledger. Lastly, the get block gets executed in 1000ms that fetches data from the blockchain ledger from the CPCB website that is sent to the ML model which determines the AQI category that is to be stored in the blockchain ledger.

The proposed system uses Hyperledger Fabric blockchain network. It is a private network. Hyperledger Fabric is private, permissioned network and does not use currencies. Proposed system is a simple attempt to store transaction securely on ledger. Ethereum blockchain network used in [30] is private or public without any permissions for users and even uses currency called as ether. As a solution for verifying the accuracy of sensitive pollution data, offers a blockchain management system. Data provided by an air quality monitoring network with high geographical and temporal resolution could be traced back in time. Ethereum blockchain supports preserving data on the average densities of zones regarded to be the city's key areas.

VI. CONCLUSION

The present system for tracking pollutants emitted in industrial areas is centralized, lacks complete transparency and is highly susceptible to data tampering. Proposed tamper-proof air quality management platform combines a machine learning model to predict the AQI category which is then stored on Blockchain. Prediction accuracy has improved using the Decision tree algorithm which is a machine learning algorithm that gives us an accuracy of 99.56%.

It is discovered that a blockchain-based solution can address data dependability issues in pollution monitoring while

also providing a permanent, tamper-proof record of industrial air quality data. As a result, industrial area-specific air quality data may be provided in a credible and transparent manner, allowing industries and the government to take the required steps to minimize pollution.

VI. FUTURE SCOPE

The proposed tamper proof air quality management system is limited to collecting pollutant data from CPCB website calculating AQI category and storing on blockchain ledger of two organizations. The system can be enhanced by incorporating Internet of Things for collection for pollutant values. IoT device with sensors for different pollutants can be installed and real time data can be collected, calculate the AQI category and store the transaction on blockchain ledger.

REFERENCES

- [1] S. Benedict, P. Rumaise and J. Kaur, "IoT Blockchain Solution for Air Quality Monitoring in SmartCities," 2019 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS), Goa, India, 2019, pp. 1-6, doi: 10.1109/ANTS47819.2019.9118148.
- [2] Abu Buker Siddique, Rafaqat Kazmi, Habib Ullah Khan, Sikandar Ali, Ali Samad & Gulraiz Javaid (2022) An Intelligent and Secure Air Quality Monitoring System Using Neural Network Algorithm and Blockchain, IETE Journal of Research, DOI: 10.1080/03772063.2022.2052984
- [3] S. Mahanta, T. Ramakrishnu, R. R. Jha and N. Tailor, "Urban Air Quality Prediction Using Regression Analysis," TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON), 2019, pp. 1118-1123, doi: 10.1109/TENCON.2019.8929517.
- [4] Dyuthi Sanjeev, 2021, Implementation of Machine Learning Algorithms for Analysis and Prediction of Air Quality, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) Volume 10, Issue 03 (March 2021)
- [5] T. M. Amado and J. C. Dela Cruz, "Development of Machine Learning-based Predictive Models for Air Quality Monitoring and Characterization," TENCON 2018 - 2018 IEEE Region 10 Conference, 2018, pp. 0668-0672, doi: 10.1109/TENCON.2018.8650518.
- [6] C R, Aditya & Deshmukh, Chandana & K, Nayana & Gandhi, Praveen & astu, Vidyav. (2018). Detection and Prediction of Air Pollution using Machine Learning Models. International Journal of Engineering Trends and Technology. 59. 204-207. 10.14445/22315381/IJETT-V59P238.
- [7] Stojov, Vladimir & Koteli, Nikola & Lameski, Petre & Zdravetski, Eftim. (2018). Application of machine learning and time-series analysis for air pollution prediction.
- [8] Yue-Shan Chang, Hsin-Ta Chiao, Satheesh Abimannan, Yo-Ping Huang, Yi-Ting Tsai, Kuan-Ming Lin, An LSTM-based aggregated model for air pollution forecasting, Atmospheric Pollution Research, Volume 11, Issue 8, 2020.
- [9] Delavar, M.R.; Gholami, A.; Shiran, G.R.; Rashidi, Y.; Nakhaeizadeh, G.R.; Fedra, K.; Hatefi Afshar, S. A Novel Method for Improving Air Pollution Prediction Based on Machine Learning Approaches: A Case Study Applied to the Capital City of Tehran. ISPRS Int. J. Geo-Inf. 2019, 8, 99.
- [10] Londhe, Mayuresh. "Data mining and machine learning approach for air quality index prediction." International Journal of Engineering and Applied Physics 1.2 (2021): 136-153.
- [11] M. Lücking et al., "The Merits of a Decentralized Pollution-Monitoring System Based on Distributed Ledger Technology," in IEEE Access, vol. 8, pp. 189365-189381, 2020, doi: 10.1109/ACCESS.2020.3028430.
- [12] Vaneta M, S Swapna Kumar, "Multi-tier Framework for Optimizing Pairwise Key Predistribution in Sensory Applications", International Journal of Innovative Technology and Exploring Engineering, 2019, 2278-3075, Volume-9, Issue-2.

- [13] S. R. Niya, S. S. Jha, T. Bocek and B. Stiller, "Design and implementation of an automated and decentralized pollution monitoring system with blockchains, smart contracts, and LoRaWAN," NOMS 2018 - 2018 IEEE/IFIP Network Operations and Management Symposium, 2018, pp. 1-4, doi: 10.1109/NOMS.2018.8406329.
- [14] Yohan Han, Byungjun Park, Jongpil Jeong, A Novel Architecture of Air Pollution Measurement Platform Using 5G and Blockchain for Industrial IoT Applications, *Procedia Computer Science*, Volume 155, 2019, Pages 728-733, ISSN 1877-0509 <https://doi.org/10.1016/j.procs.2019.08.105>.
- [15] M. Bublitz, F.; Oetomo, A.; S. Sahu, K.; Kuang, A.; X. Fadrigue, L.; E. Velmovitsky, P.; M. Nobrega, R.; P. Morita, P. Disruptive Technologies for Environment and Health Research: An Overview of Artificial Intelligence, Blockchain, and Internet of Things. *Int. J. Environ. Res. Public Health* 2019, 16, 3847. <https://doi.org/10.3390/ijerph16203847>
- [16] Sunil Gulia, Nidhi Shukla, Lavanya Padhi, Parthaa Bosu, S.K. Goyal, Rakesh Kumar, Evolution of air pollution management policies and related research in India, *Environmental Challenges*, Volume 6, 2022, 100431, ISSN 2667-0100, <https://doi.org/10.1016/j.envc.2021.100431>. (<https://www.sciencedirect.com/science/article/pii/S2667010021004054>)
- [17] The Changing Paradigm of Air Pollution Monitoring Emily G. Snyder, Timothy H. Watkins, Paul A. Solomon, Eben D. Thoma, Ronald W. Williams, Gayle S. W. Hagler, David Shelow, David A. Hindin, Vasu J. Kilaru, and Peter W. Preuss *Environmental Science & Technology* 2013 47 (20), 11369-11377 DOI: 10.1021/es4022602
- [18] H. P. L. d. Medeiros and G. Girão, "An IoT-based Air Quality Monitoring Platform," 2020 IEEE International Smart Cities Conference (ISC2), 2020, pp. 1-6, doi: 10.1109/ISC251055.2020.9239070
- [19] K. Zheng, S. Zhao, Z. Yang, X. Xiong and W. Xiang, "Design and Implementation of LPWA-Based Air Quality Monitoring System," in *IEEE Access*, vol. 4, pp. 3238-3245, 2016, doi: 10.1109/ACCESS.2016.2582153.
- [20] L. -J. Chen et al., "An Open Framework for Participatory PM2.5 Monitoring in Smart Cities," in *IEEE Access*, vol. 5, pp. 14441-14454, 2017, doi: 10.1109/ACCESS.2017.2723919.
- [21] Wang J, Li X, Jin L, Li J, Sun Q, Wang H. An air quality index prediction model based on CNN-ILSTM. *Sci Rep.* 2022 May 19;12(1):8373. doi: 10.1038/s41598-022-12355-6. PMID: 35589914; PMCID: PMC9120089.
- [22] Shriram, Pranav and Malladi, Srinivas, A Study and Analysis of Air Quality Index and Related Health Impact on Public Health (January 18, 2021). ICICNIS 2020, Available at SSRN: <https://ssrn.com/abstract=3768477>
- [23] Androulaki, E., "Hyperledger Fabric: A Distributed Operating System for Permissioned Blockchains"
- [24] Sangeetha V, Vaneeta M, S Swapna Kumar, Piyush Kumar Pareek, Sunanda Dixit, "Efficient Intrusion detection of malicious node using Bayesian Hybrid Detection in MANET", CCRDA 2020, IOP Conf. Series: Materials Science and Engineering 1022 (2021) 012077, IOP Publishing, doi:10.1088/1757-899X/1022/1/012077,2021.
- [25] Liu, H.; Li, Q.; Yu, D.; Gu, Y. Air Quality Index and Air Pollutant Concentration Prediction Based on Machine Learning Algorithms. *Appl. Sci.* 2019, 9, 4069. <https://doi.org/10.3390/app9194069>
- [26] Kaur, Gaganjot & Gao, Jerry & Chiao, Sen & Lu, Shengqiang & Xie, Gang. (2018). Air Quality Prediction: Big Data and Machine Learning Approaches. *International Journal of Environmental Science and Development*. 9. 8-16. 10.18178/ijesd.2018.9.1.1066.
- [27] Kataria, A., Puri, V.: AI- and IoT-based hybrid model for air quality prediction in a smart city with network assistance. *IET Netw.* 11(6), 221– 233 (2022). <https://doi.org/10.1049/ntw2.12053>
- [28] Sambana, Bosubabu. (2020). An Artificial Intelligence based Air Pollution Prediction and Monitoring System using Internet of Things. *Xi'an Dianzi Keji Daxue Xuebao/Journal of Xidian University*. 14. 1634-1644. 10.37896/jxu14.7/187.
- [29] Goh, C.C.; Kamarudin, L.M.; Zakaria, A.; Nishizaki, H.; Ramli, N.; Mao, X.; Syed Zakaria, S.M.M.; Kanagaraj, E.; Abdull Sukor, A.S.; Elham, M.F. Real-Time In-Vehicle Air Quality Monitoring System Using Machine Learning Prediction Algorithm. *Sensors* 2021, 21, 4956. <https://doi.org/10.3390/s21154956>
- [30] Sofia, D.; Lotrecchiano, N.; Truccillo, P.; Giuliano, A.; Terrone, L. Novel Air Pollution Measurement System Based on Ethereum Blockchain. *J. Sens. Actuator Netw.* 2020, 9, 49.