A PRELIMINARY REPORT ON

# Urban Air Quality Forecasting & Visualization Dashboard

SUBMITTED TO THE VISHWAKARMA INSTITUTE OF INFORMATION TECHNOLOGY, PUNE

IN THE PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE AWARD OF THE DEGREE

OF

**BACHELOR OF TECHNOLOGY (COMPUTER ENGINEERING)**

SUBMITTED BY

**RAHUL TALELE** **EXAM SEAT NO. 22010379**

**DEPARTMENT OF COMPUTER ENGINEERING**

**BRACT'S**

**VISHWAKARMA INSTITUTE OF INFORMATION TECHNOLOGY, PUNE**

**2022 -2023**

SURVEY NO. 3/4, KONDHWA (BUDRUK), PUNE – 411048, MAHARASHTRA (INDIA).

# CERTIFICATE

This is to certify that the project report entitles

## Urban Air Quality Forecasting & Visualization Dashboard

Submitted by

**RAHUL TALELE** **EXAM SEAT NO .22010379**

is a bonafide student of this institute and the work has been carried out by him under the supervision of **Dr. Namrata Wasatkar** and it is approved for the partial fulfillment of the requirement of VISHWAKARMA INSTITUTE OF INFORMATION TECHNOLOGY, for the award of the degree of **Bachelor of Technology** (Computer Engineering).

**(Dr. Namrata Wasatkar)**
Guide
Department of Computer Engineering

**(Dr. S. R. Sakhare)**
Head
Department of Computer Engineering

**(Dr. Vivek Deshpande)**
Director,
BRACT's Vishwakarma Institute of Information Technology, Pune-48

Place: Pune

Date:

# ACKNOWLEDGEMENT

It gives us great pleasure to present the preliminary project report on "Urban Air Quality Forecasting & Visualization Dashboard." I would like to express our heartfelt gratitude to all those who have contributed to the successful completion of this project.

First and foremost, I would like to extend my sincere appreciation to my project supervisor, Dr. Namrata Wasatkar, for her invaluable guidance, continuous support, and insightful suggestions. Her expertise and encouragement have been instrumental in shaping the direction of my project. I am truly grateful for her unwavering commitment and dedication.

I would also like to acknowledge the support and valuable suggestions provided by Prof. S. R. Sakhare, Head of the Computer Engineering Department at Vishwakarma Institute of Information Technology. His wisdom and expertise have greatly contributed to the development of my project.

**Rahul Talele**

# ABSTRACT

In the face of escalating urbanization, air quality is an integral concern, critically influencing both human health and environmental well-being. As urban areas expand, the imperative for precise, real-time air quality information intensifies. This research introduces the Urban Air Quality Forecasting & Visualization Dashboard, a pioneering platform tailored for Pune city's distinct needs.

Capitalizing on machine learning's capabilities, we merged multiple forecasting models, such as ARIMA, LSTM, Facebook Prophet, and TBTAS, to project the Air Quality Index (AQI). Our outcomes elucidate the layered challenges of AQI prediction, especially with the unexpected performance quirks of LSTM and SARIMA's recognized limitations for certain datasets.

Nevertheless, through consistent model refinement, our dashboard proficiently dispenses real-time AQI insights, further enhanced by discerning visualizations via Tableau. In amalgamating technology with environmental science, our ambition is to furnish city planners, health agencies, and residents with vital insights, nurturing a forward-thinking paradigm in urban air quality management.

# TABLE OF CONTENTS

# CHAPTER 1: INTRODUCTION

## 1.1 OVERVIEW

The Urban Air Quality Forecasting & Visualization Dashboard is a groundbreaking initiative aimed at addressing the escalating challenges posed by urban air pollution, particularly within the city limits of Pune. The relentless pace of urbanization has ushered in unprecedented growth but concurrently heightened concerns about the quality of air that urban populations breathe. Recognizing the critical intersection of public health, environmental sustainability, and city planning, this project aspires to deliver a comprehensive solution that not only forecasts air quality but visualizes the data in an accessible manner.

At the core of this initiative is the integration of advanced machine learning algorithms, including ARIMA, LSTM, Facebook Prophet, and TBTAS. These models are meticulously trained on historical air quality data, allowing them to discern intricate patterns and trends that influence air quality variations. The selection of multiple models reflects a commitment to robustness and adaptability, acknowledging the diverse and dynamic nature of urban air quality data.

One of the key challenges in this endeavor is the accurate prediction of the Air Quality Index (AQI). While traditional models like ARIMA lay the foundation for time series forecasting, unexpected outcomes from models such as LSTM prompt a deeper exploration into the intricate dynamics of air quality fluctuations. This iterative process of model refinement underscores the commitment to precision in AQI forecasting.

The visualization aspect of the project is materialized through the renowned data visualization platform, Tableau. Leveraging the power and flexibility of Tableau, the dashboard transforms complex air quality data into intuitive visualizations. These visualizations are designed not only for city planners and health agencies but also for residents like Rahul, ensuring that the insights derived from the data are accessible to a diverse audience.

Beyond the technological facets, the project envisions a broader impact on urban living. By empowering stakeholders with timely and accurate air quality information, the dashboard aims to influence proactive decision-making in city planning, enhance public health awareness, and encourage collective efforts towards a cleaner, healthier urban environment. The Urban Air Quality Forecasting & Visualization Dashboard stands as a testament to the potential of data-driven solutions in mitigating the challenges posed by urbanization on air quality.

## 1.2 MOTIVATION

The genesis of the Urban Air Quality Forecasting & Visualization Dashboard is rooted in a deep-seated concern for the escalating impact of air pollution on urban life. As cities worldwide witness unprecedented growth, the resultant increase in vehicular emissions, industrial activities, and other anthropogenic factors contributes significantly to the deterioration of air quality. The motivation for this project stems from the urgent need to address the consequences of poor air quality on public health, environmental sustainability, and the overall quality of urban living.

The World Health Organization's identification of air pollution as the single largest environmental health risk underscores the criticality of proactive measures. Rising pollution levels are associated with a spectrum of health issues, from respiratory ailments to cardiovascular diseases. Recognizing this nexus between air quality and public health, the project seeks to empower individuals, city planners, and health agencies with a tool that not only predicts air quality but transforms data into actionable insights.

The motivation for integrating advanced machine learning algorithms lies in the inadequacies of traditional air quality forecasting methods. The dynamic and nonlinear nature of urban air quality data demands adaptive models capable of discerning intricate patterns. The employment of ARIMA, LSTM, Facebook Prophet, and TBTAS reflects a commitment to staying at the forefront of forecasting technologies, ensuring the accuracy and reliability of predictions.

Furthermore, the project's motivation extends beyond technical advancements to encompass societal impact. By providing a user-friendly dashboard accessible to all, the project aims to raise public awareness about the implications of air quality on health and lifestyle. The visualization component, facilitated by Tableau, serves as a bridge between complex data and actionable insights, making the information comprehensible and relevant for residents, city planners, and health practitioners alike.

In essence, the motivation behind the Urban Air Quality Forecasting & Visualization Dashboard lies in a collective responsibility to safeguard the well-being of urban populations. By leveraging technology, data, and innovation, the project aspires to instigate positive change, fostering a healthier and more sustainable urban environment.

### 1.3 PROBLEM DEFINITION AND OBJECTIVES

**Problem Definition:**

The Urban Air Quality Forecasting & Visualization Dashboard addresses the multifaceted challenges posed by urban air pollution; a critical issue exacerbated by rapid urbanization. The core problem lies in the inadequacy of traditional air quality monitoring and forecasting methods to provide timely, accurate, and accessible information to residents, city planners, and health agencies. Existing approaches often fall short in capturing the dynamic and nonlinear nature of urban air quality data, hindering proactive decision-making and public awareness.

**Objectives:**

**1. Accurate Air Quality Forecasting:**

Develop and integrate advanced machine learning algorithms, including ARIMA, LSTM, Facebook Prophet, and TBTAS, to enhance the accuracy of Air Quality Index (AQI) forecasting.

Evaluate and refine these models iteratively, addressing the challenges posed by the nuanced patterns inherent in urban air quality data.

**2. Real-Time Visualization:**

Utilize Tableau as the visualization platform to transform complex air quality data into intuitive, real-time visualizations. Ensure the dashboard is accessible to a diverse audience, providing actionable insights for residents, city planners, and health agencies.

**3. Public Awareness and Engagement:**

Foster public awareness about the impact of air quality on health and lifestyle. Engage residents in the monitoring process, empowering them to make informed decisions based on the provided air quality insights.

**4. Decision Support for City Planning:**

Provide city planners with a tool that aids in making informed decisions regarding urban development, traffic management, and environmental policies.
Enable a proactive approach to mitigate the impact of urbanization on air quality through data-driven strategies.

**5. Continuous Improvement:**

Establish a feedback loop for continuous improvement, incorporating user feedback, new data sources, and advancements in machine learning techniques. Ensure the dashboard remains adaptive and responsive to the evolving dynamics of urban air quality.

The overarching objective is to create a holistic solution that not only addresses the immediate challenges of air quality forecasting but also contributes to a broader cultural shift towards environmental consciousness and sustainable urban living.

## 1.4 PROJECT SCOPE AND LIMITATIONS

### 1.4.1 Project Scope:

The Urban Air Quality Forecasting & Visualization Dashboard aims to create a comprehensive and user-centric solution for addressing urban air pollution challenges within the geographic scope of Pune city. The project encompasses the following key aspects:

**1. Geographic Focus:**

The primary focus is on providing air quality forecasts and visualizations for specific locations within Pune city, catering to the unique environmental dynamics of the urban landscape.

**2. Machine Learning Integration:**

The integration of machine learning models, including ARIMA, LSTM, Facebook Prophet, and TBTAS, enhances the accuracy and reliability of Air Quality Index (AQI) predictions.

**3. Real-Time Visualization:**

Leveraging Tableau, the project offers real-time visualizations of air quality data, making complex information accessible to residents, city planners, and health agencies.

**4. Public Engagement:**

The project encourages public engagement by raising awareness about air quality issues, fostering informed decision-making among residents.

**5. Decision Support for City Planning:**

City planners can utilize the dashboard as a decision support tool for urban development, traffic management, and environmental policies.

**6. Continuous Improvement:**

The project establishes mechanisms for continuous improvement, incorporating user feedback and staying adaptable to emerging technologies and data sources.

### 1.4.2 Limitations:

While the project endeavors to create a robust solution, certain limitations are acknowledged:

**1. Data Availability:**

The accuracy of forecasts is contingent on the availability and quality of historical air quality data. Limitations in data sources may impact the precision of predictions.

**2. Generalization to Other Cities:**

The machine learning models, and forecasting algorithms are specifically tuned for the characteristics of Pune city. Generalizing the approach to other cities may require re-calibration and additional data.

**3. External Factors:**

The accuracy of forecasts may be influenced by external factors not explicitly considered in the models, such as sudden industrial emissions, natural disasters, or extreme weather events.

**4. Public Engagement Challenges:**

Achieving widespread public engagement poses challenges related to communication strategies, diverse user demographics, and varying levels of environmental literacy.

**5. Technology Adoption:**

The effectiveness of the dashboard relies on technology adoption by relevant stakeholders. Encouraging consistent usage among residents and city planners may require targeted awareness campaigns.

**6. Resource Constraints:**

Resource constraints, both in terms of computational capabilities and data processing capacities, may impact the scalability of the project to handle increasing data volumes.

Despite these limitations, the project represents a significant step towards addressing urban air quality challenges, with a focus on Pune city, and lays the foundation for future enhancements and adaptations.

## 1.5 METHODOLOGIES OF PROBLEM SOLVING

The Urban Air Quality Forecasting & Visualization Dashboard employs a systematic approach to address the intricate challenges posed by urban air pollution. The methodologies of problem-solving encompass a blend of innovative techniques, technological integrations, and an iterative refinement process.

**1. Understanding Data Dynamics:**

- **Data Exploration:** Commencing with a thorough exploration of historical air quality data, the project aims to understand the dynamic patterns, trends, and anomalies inherent in urban air quality fluctuations.

- **Descriptive Analytics:** Utilizing descriptive analytics, the project identifies key statistical measures, distributions, and correlations within the dataset to inform subsequent forecasting models.

**2. Machine Learning Integration:**

- **Model Selection:** The integration of machine learning models, including ARIMA, LSTM, Facebook Prophet, and TBTAS, involves a meticulous selection process based on their historical effectiveness and adaptability to the nonlinear characteristics of urban air quality data.

- **Training and Evaluation:** Models are trained on historical data, and their accuracy is rigorously evaluated, with a keen focus on iteratively refining algorithms to improve predictive capabilities.

3. **Visualization through Tableau:**

- **Data Transformation:** Leveraging Tableau, the project transforms complex air quality data into visually intuitive representations, facilitating a deeper understanding of trends, outliers, and patterns.

- **Interactive Dashboards:** The visualization methodology prioritizes the creation of interactive dashboards accessible to a diverse audience, ensuring that insights are comprehensible and actionable.

4. **Public Engagement and Awareness:**

- **Communication Strategies:** To foster public engagement, the project employs effective communication strategies, utilizing user-friendly language and visuals to convey the implications of air quality on health and lifestyle.

- **Educational Initiatives:** Public awareness initiatives educate residents about the importance of air quality, empowering them to interpret and act upon the insights provided by the dashboard.

**5. Iterative Development and Continuous Improvement:**

- **Feedback Loops:** The project establishes feedback loops to gather user input, ensuring that the dashboard remains responsive to the evolving needs of residents, city planners, and health agencies.

- **Adaptive Model Refinement:** Continuous improvement involves adapting machine learning models to changing urban dynamics, incorporating new data sources, and staying abreast of advancements in forecasting methodologies.

**6. Collaborative Decision Support:**

- **City Planner Involvement:** The dashboard serves as a collaborative decision support tool for city planners, enabling them to make informed decisions regarding urban development, traffic management, and environmental policies.

In essence, the methodologies employed in solving the complex challenges of urban air quality integrate data-driven insights, machine learning prowess, and user-centric design to create a comprehensive solution with the potential for positive societal impact.

# CHAPTER 2: LITERATURE SURVEY

Urban air quality has emerged as a paramount concern in recent decades, affecting millions worldwide. With the rapid urbanization of cities, there's an increasing need for timely and accurate information regarding air pollution and its potential implications [8]. Poor air quality has been intrinsically linked to a plethora of health issues, from respiratory ailments to cardiovascular diseases, posing a significant threat to urban populations. Furthermore, air quality fluctuations can influence various socio-economic factors, including healthcare costs, workforce productivity, and overall quality of life.

Given the dynamic nature of urban environments, traditional methods of monitoring and predicting air quality often fall short in terms of real-time accuracy and accessibility [8]. However, with the convergence of advanced technologies, vast data repositories, and enhanced computational prowess, a shift is evident. Innovative platforms are emerging, employing data-driven methodologies to forecast air quality metrics with remarkable precision. These platforms not only predict but also visually represent complex data, rendering it comprehensible to diverse stakeholders ranging from policymakers and health agencies to the public.

Our project, the Urban Air Quality Forecasting & Visualization Dashboard, is rooted in this paradigm shift. Harnessing machine learning algorithms, it endeavors to forecast air quality indices for Pune city, culminating in an interactive dashboard accessible to all. This literature survey delves into the evolution of air quality monitoring, the advent and potential of machine learning in forecasting, and the profound impact of effective data visualization. By understanding the tapestry of research and innovation in this domain, we aim to contextualize our project's significance and its contribution to a healthier, informed urban future.

The challenges posed by urban air pollution are not a recent phenomenon. As early as the Industrial Revolution, cities like London and Pittsburgh witnessed severe air pollution events, dramatically impacting the environment and public health. Often referred to as 'pea-soupers' or smog, these events were the consequence of unchecked industrial emissions combined with adverse weather conditions [9].

In the initial stages, monitoring air quality was a rudimentary process. Observational methods, based primarily on visual cues such as visibility range and the color of emissions, were the norm [10]. Instruments like wet and dry bulb thermometers provided basic insights into humidity and temperature, factors that heavily influenced air quality.

By the mid-20th century, the detrimental health effects of poor air quality became glaringly evident. Catastrophic events, such as the 1952 Great Smog of London, which resulted in thousands of premature deaths, accelerated the need for systematic air quality monitoring [11]. This urgency catalyzed the development of the first generation of air quality monitoring instruments. Devices like the smoke shade comparator and early sulfur dioxide detectors came into existence [14].

Towards the latter part of the century, advancements in sensor technologies and analytical techniques enabled more precise and comprehensive air quality measurements. Stations equipped with an array of instruments began to monitor multiple pollutants, from particulate matter (PM) and nitrogen oxides (NOx) to volatile organic compounds (VOCs). These stations became the precursors to modern air quality monitoring networks.

However, while technology progressed, several challenges remained. Traditional stations were often sparsely located, leading to data gaps. Moreover, the data, while accurate, was not always timely, rendering real-time response to pollution events difficult [12].

In parallel, the scientific community began to understand the nuanced relationship between various pollutants and their synergistic effects on health. This realization underscored the need for integrated and real-time monitoring solutions that could capture the multifaceted nature of urban air quality.

It's against this historical backdrop that projects like the Urban Air Quality Forecasting & Visualization Dashboard emerge. By leveraging modern technologies and vast data sources, such initiatives aim to fill the gaps left by traditional methods, offering real-time, comprehensive insights into urban air quality.

**1. Evolution and Integration:**

Machine learning (ML) offers a suite of algorithms that can learn patterns from data without being explicitly programmed. Over the past two decades, the environmental sciences community has started harnessing these powerful techniques. Air quality prediction, with its intricate multi- factor influences ranging from weather patterns to urbanization, presents a particularly challenging problem that traditional statistical models sometimes struggle with. This led to the exploration and eventual integration of ML techniques into air quality research [18].

**2. Key Advantages of ML in AQI Forecasting:**

- **Complexity Handling:** ML models, especially deep learning architectures like LSTM, can handle the multi- dimensional nature of air quality data, considering various factors simultaneously.

- **Adaptability:** Unlike traditional models, ML algorithms can adapt to new data, refining their predictions as more information becomes available [18].

- **High Accuracy:** Many ML models, when correctly trained and fine-tuned, can achieve superior accuracy levels, especially when forecasting short to medium-term AQI levels.

**3. Common ML Techniques in AQI Prediction:**

- **Time Series Forecasting:** Techniques like ARIMA and Facebook Prophet are used to predict future values based on past data. While ARIMA captures linear relationships, Prophet offers flexibility with daily observations and different time scales [13].

- **Neural Networks:** LSTMs, a type of recurrent neural network, are particularly suited for sequence data, like time series AQI values, capturing long-term dependencies in the data.

- **Ensemble Methods:** Combining multiple models can help enhance accuracy and reduce overfitting. Techniques like Random Forests or Gradient Boosted Trees can be employed to capture non-linear relationships in AQI data.

**4. Challenges Faced:**

- **Data Quality:** The effectiveness of ML largely depends on the quality of data. Inaccurate or incomplete data can lead to misleading predictions.

- **Overfitting:** Given the complexity of some ML models, there's a risk of overfitting where the model becomes too tailored to the training data, reducing its generalization capability [20].

- **Interpretability:** ML models, especially deep neural networks, can sometimes act as 'black boxes', making it challenging to understand the reasoning behind predictions.

**5. Broader Implications:**

The integration of ML into AQI forecasting doesn't just offer improved prediction accuracy. It provides a more holistic understanding of various factors influencing air quality. For city planners, health agencies, and residents, ML-powered forecasts offer actionable insights, enabling preemptive measures in the face of deteriorating air quality.

Moreover, as computational capacities grow and data becomes more abundant, the synergy between machine learning and air quality research will likely deepen, opening avenues for further innovation and refinement in AQI forecasting methodologies [15].

The accurate forecasting of Air Quality Index (AQI) is a pivotal component of the Urban Air Quality Forecasting & Visualization Dashboard. To achieve this, the project leverages multiple machine learning models, each tailored to capture distinct patterns and trends within the air quality data. The selection of these models is based on their historical effectiveness in forecasting environmental data and their applicability to non-seasonal urban air quality datasets.

The key machine learning models utilized for AQI prediction include:

**ARIMA (AutoRegressive Integrated Moving Average):**

ARIMA is a time series forecasting model that captures linear relationships and trends within time series data. While ARIMA has been a cornerstone in time series forecasting, it's found limited applicability in this project. The urban air quality data often exhibits complex patterns that go beyond linear relationships [15].

**LSTM (Long Short-Term Memory):**

LSTM is a type of recurrent neural network (RNN) designed for processing sequences of data. In the context of urban air quality, LSTM was one of the models tested, but it yielded unexpected results. A negative accuracy score indicated that the model was not effectively forecasting the data [16].

**Facebook Prophet:**

Facebook Prophet is an open-source forecasting tool designed for time series data with daily observations that display patterns on different time scales. While Facebook Prophet is known for its versatility in handling various time series datasets, its performance in this project was carefully examined [16].

**TBTAS (Theta Beta Time Series Analysis):**

TBTAS is an advanced time series forecasting method that considers multiple factors, including seasonality and trends. TBTAS is among the models used for urban air quality forecasting, taking into account the specific characteristics of the dataset [19].

The forecasting process entails training these models with historical air quality data, enabling them to identify patterns and relationships. However, it's essential to note that not all models proved equally effective. For instance, the SARIMA (Seasonal AutoRegressive Integrated Moving Average) model was deemed unsuitable for analysis due to its inherent limitation in dealing with non-seasonal datasets. Likewise, LSTM's inability to provide accurate forecasts underscored the challenges in predicting the nuanced and variable nature of urban air quality [17].

The project's objective is to continuously assess and refine these models to ensure the highest level of accuracy in forecasting AQI. This commitment to improving forecasting methodologies underscores the dynamic and evolving nature of the Urban Air Quality Forecasting & Visualization Dashboard.

# CHAPTER 3: SOFTWARE REQUIREMENTS SPECIFICATIONS

## 3.1 Product Perspective:

The Urban Air Quality Forecasting & Visualization Dashboard is envisioned as a standalone web-based application that fits into a broader ecosystem focused on urban development and health. While it operates autonomously, it exists within the context of addressing urban challenges by harnessing data-driven insights to assist various stakeholders.

The system is designed to fill a significant gap in urban planning and health advisories. In many urban areas, while data about air quality might be available, it often lacks actionable insights. Our dashboard bridges this gap by not just presenting the data, but by providing forecasts and trends which can be crucial for planning and decision-making.

*This dashboard serves multiple purposes:*

- **Decision-making tool for City Planners:** Offering a lens into the future air quality can guide infrastructural and environmental decisions.

- **Health Advisory for Agencies:** With forecasted data, health agencies can issue timely advisories to residents, potentially saving lives and improving health outcomes.

- **Planning Aid for Residents:** Residents can plan their activities, such as outdoor exercises or events, based on the forecasted air quality.

While it begins its journey focused on Pune city, the underlying architecture is designed with scalability in mind, suggesting that the system is adaptable and can be expanded to include other urban regions in the future.

### 3.1.1 System Interfaces

**AQODP Data Interface:**

The primary system interface will be the connection to the AQODP, which serves as the data backbone for the dashboard. This connection is crucial for fetching the required historical AQI data and is typically established through an API or data feed provided by AQODP. It ensures real-time data retrieval and updates.

### 3.1.2 User Interfaces

**Dashboard Interface:**

The primary user interface is the Tableau-based dashboard where users will interact with the displayed data. This dynamic and intuitive interface will provide current AQI values, forecasts, and historical trends. It is designed to be user-friendly, with features like zoom, pan, and location-specific data retrieval to ensure a seamless experience for all users, whether they're city planners or local residents.

### 3.1.3 Hardware Interfaces

While the dashboard is primarily a web-based application and doesn't interact directly with specific hardware components, its design considers various devices:

- **Desktops & Laptops:** The dashboard will be optimized for larger screens, ensuring data visualizations are clear and detailed.

- **Mobile Devices & Tablets:** With the increasing use of mobile devices for accessing web-based applications, the dashboard will be responsive, adapting to different screen sizes and ensuring a consistent user experience across devices.

### 3.1.4 Software Interfaces

**Tableau:**

The visualization component of the system relies heavily on Tableau. This software provides the necessary tools to create dynamic and interactive dashboards, transforming raw data into understandable visual formats.

**Machine Learning Libraries:**

For the forecasting functionality, the system will make use of popular machine learning libraries and frameworks. Libraries such as TensorFlow, Scikit-learn, or PyTorch may be employed to develop, train, and deploy AQI forecasting models.

## 3.2 Product Function:

**Data Retrieval and Integration:**

- **Description:** The system will connect to AQODP to access and retrieve historical AQI data. This function ensures the dashboard has the most up-to-date information available, providing an accurate backdrop for both current conditions and forecasts.

- **Benefits:** Provides users with a consistent stream of accurate and real-time data, making the dashboard a reliable tool for decision-making.

**AQI Forecasting:**

- **Description:** Leveraging machine learning algorithms, the dashboard will predict future AQI values. Multiple algorithms will be tested and compared for accuracy, and the most precise model will be employed for final forecasting.

- **Benefits:** Gives users a forward-looking perspective, allowing them to anticipate changes in air quality and plan accordingly. This can be particularly beneficial for health advisories and city planners.

**Interactive Visualization:**

- **Description:** Data, both current and forecasted, will be visualized using Tableau. Users will be able to interact with this data, exploring different locations within Pune, zooming into specific timeframes, and viewing historical trends.

- **Benefits:** A dynamic visualization interface enables users to glean insights quickly, making the data more accessible and understandable, regardless of the user's technical expertise.

**Model Accuracy Comparison:**

- **Description:** The system will not just forecast AQI values but will also assess the accuracy of the different forecasting models employed. This ensures that the system evolves and improves its predictions over time.

- **Benefits:** Ensures the highest level of accuracy in AQI forecasting, boosting user confidence in the predictions provided by the dashboard.

**Location-Specific Insights:**

- **Description:** While the dashboard covers Pune city, it allows users to drill down into specific locations or zones within the city. This granularity offers more localized AQI information and forecasts.

- **Benefits:** Allows for more tailored decision-making, as users can get insights specific to their locality or areas of interest.

### 3.3 User Characteristics - Rahul:

**Profile Name:** Rahul

**Age:** 32

**Occupation:** Software Engineer at a multinational corporation

**Residence:** Koregaon Park, Pune

**Educational Background:** Master’s in computer science from a reputed institution

**Technical Proficiency:** Advanced (Familiar with technology tools, software applications, and has basic coding knowledge)

**Primary Device for Access:** Laptop during work hours; Smartphone during evenings and weekends

***Primary Goals with the Dashboard:***

- **Personal Health:** Rahul is health-conscious and likes to jog early in the morning. He wants to ensure that he's jogging when the AQI is safe.

- **Family Safety:** With a young daughter at home, Rahul wants to ensure the environment is safe for her, especially regarding outdoor activities and playtime.

- **Relocation Considerations:** Rahul is considering buying a new house and wants to ensure he chooses a location with consistently good air quality.

- **Technical Curiosity:** Given his background, Rahul is also interested in the underlying technology and might delve deeper into the forecasting algorithms and data sources.

- **Frequency of Use:** Rahul is expected to check the dashboard daily, mostly in the early mornings before his jog and occasionally during the day when making weekend plans or considering outdoor activities with his family.

- **Preferred Mode of Interaction:** Rahul prefers visual data representations (graphs, charts, maps) but is also keen on understanding the numeric values and trends. He would likely utilize both the interactive visualization features and detailed data metrics provided by the dashboard.

### 3.4 Constraints:

- **Data Reliability:** The system's accuracy and forecasting capabilities are inherently tied to the quality of the data sourced from AQODP. If AQODP provides inaccurate or delayed data, it could impact the dashboard's reliability.

- **Data Access Limitations:** There might be limits imposed by the AQODP in terms of how frequently data can be accessed or the volume of data that can be retrieved in a given timeframe.

- **Unpredictable External Factors:** While the system can forecast based on historical data, sudden and unpredictable events (e.g., industrial accidents, unexpected weather patterns) can influence air quality and may not be accurately predicted by the system.

- **Technological Dependencies:** The system's performance depends on the continuous functionality of external software like Tableau and the chosen ML libraries. Any changes or disruptions in these technologies could impact system functionality.

- **Scalability Challenges:** While the system is designed with scalability in mind, rapid expansion to cover additional cities or regions may require significant architectural changes or adjustments.

- **Internet Dependency:** Being a web-based dashboard, it requires a stable internet connection for data retrieval and updates. Users might experience difficulties accessing the dashboard in areas with poor internet connectivity.

- **Maintenance and Updates:** Machine learning models might need periodic retraining to maintain forecasting accuracy. Ensuring this happens seamlessly, without disrupting user experience, is a constraint.

- **Security Concerns:** Ensuring the secure transmission and storage of data, especially if user-specific data or preferences are stored, is crucial. This demands continuous monitoring and updates to adhere to the best security practices.

### 3.5 Assumptions and Dependencies

***Assumptions:***

- **Data Availability:** AQODP will consistently provide historical AQI data for Pune, ensuring the dashboard always has access to up-to-date information.

- **Technological Consistency:** The technologies, like Tableau and chosen ML libraries, will remain available, supported, and updated throughout the lifecycle of the dashboard.

- **User Internet Access:** Most of the target users will have access to a stable internet connection, allowing them to use the web-based dashboard without interruptions.

- **Data Consistency:** AQI data for Pune remains consistent in format and structure, ensuring the dashboard's data retrieval and processing mechanisms do not frequently need adjustments.

- **Server Uptime:** The server hosting the dashboard will maintain high uptime, ensuring consistent availability to users.

- **Forecasting Models:** Machine learning models will remain relevant and accurate for AQI prediction for a considerable period post their last training.

***Dependencies:***

- **AQODP Reliability:** The dashboard is heavily dependent on AQODP for data. Any changes, interruptions, or discontinuations in their service will directly impact the dashboard.

- **Software Updates:** Periodic updates and patches for software tools and libraries, such as Tableau and ML frameworks, are required to ensure security and optimal performance.

- **External Data Sources:** If additional data sources, like weather forecasts or pollution event databases, are integrated in the future, the system will become dependent on the reliability and accuracy of these sources.

- **Cloud Service Providers:** If the dashboard utilizes cloud services for hosting or data storage, it will depend on the reliability and performance of the chosen cloud service provider.

- **Regulatory Changes:** Any regulatory changes regarding air quality data sharing, data privacy, or related areas could impact the system's operations or features.

- **External APIs:** If other APIs (e.g., for map visualization or additional data sources) are integrated, the dashboard will depend on the consistent performance and availability of these APIs.

## 3.6 Specific Requirements

### 3.6.1 Functional requirements:

**Data Retrieval and Integration:**

- The system must connect to AQODP via an API or suitable data interface, ensuring a continuous and updated stream of air quality data.

- The system should automate data retrieval, ensuring real-time accuracy and the most recent AQI data is always available.

- Mechanisms to check the validity and consistency of the incoming data are crucial, ensuring data integrity and preventing potential errors.

**AQI Forecasting:**

- The system should employ multiple machine learning algorithms, providing redundancy and an opportunity to choose the best performer.

- Periodic evaluation of these algorithms against actual data ensures the models remain relevant and adapt to changing data patterns.

- The most accurate algorithm should be automatically selected for forecasting, ensuring the highest reliability at any given time.

- Confidence intervals or accuracy metrics should be provided alongside forecasts, offering users insights into prediction variability.

**Interactive Visualization:**

- A dynamic dashboard displaying real-time AQI data, historical trends, and forecasts allows users to glean insights at a glance.

- Users should be able to select specific locations within Pune for a more granular view of AQI data.

- Features like zoom, pan, and other interactive elements should be incorporated, enhancing navigation and exploration.

- Notable AQI changes or significant events should be highlighted, allowing users to quickly identify and respond to potential concerns.

**Model Accuracy Comparison:**

- Continuous evaluation of forecasting model accuracy ensures the best predictions are offered to users.

- Models should be automatically retrained using the latest data, ensuring they remain updated and accurate over time.

- A historical log of model performances aids in tracking improvements and ensuring accountability.

**User Preferences and Customization:**

- The dashboard should be customizable, offering users a tailored experience based on their specific interests and needs.
- Users should have the option to set up alerts or notifications based on specific AQI thresholds.
- Storing user preferences ensures a consistent and familiar experience during each visit.

**Help and Documentation:**

- A help section or tutorial ensures that new users can quickly become familiar with the dashboard's features.
- Tooltips and quick reference guides should be available, helping users interpret and understand the displayed data.

### 3.6.2 Non-Functional Requirements:

**Usability:**

- The dashboard should have an intuitive interface, ensuring users of varying technical backgrounds can navigate and comprehend the data without difficulty.
- A user-friendly help section, tutorials, and tooltips should be available to guide users through the platform's features.
- The system should be designed to minimize the number of steps required to accomplish key tasks, promoting a smooth user experience.

**Performance:**

- Data retrieval and updates from AQODP should occur promptly, ensuring the most recent information is always displayed.
- Forecasting models should execute in real-time or near real-time to offer immediate insights based on the latest data.
- The dashboard should load quickly, with minimal lag, to provide an efficient user experience, even during peak usage times.

**Reliability:**

- The system should have a high uptime, ensuring users can access the dashboard whenever needed.
- Regular backups of data and system configurations should be conducted to allow for quick recovery in case of any failures.
- Forecasting models should be robust, providing consistent and reliable predictions even when faced with anomalous data.

**Security:**

- Data transfers, especially from AQODP, should be encrypted using protocols like HTTPS to ensure data integrity and confidentiality.
- If user accounts or preferences are implemented, user data should be securely stored, and best practices for password storage and authentication should be followed.

- The system should be regularly audited and updated to address potential security vulnerabilities.

**Scalability:**

- The system should be designed to handle an increasing amount of data as more historical AQI records are added.

- The architecture should allow for the potential inclusion of additional cities or regions in the future without major reconfigurations.

- The platform should be capable of accommodating a growing user base without degradation in performance.

**Maintainability:**

- The codebase and system architecture should be modular and well-documented, allowing for easy updates, modifications, or fixes.

- Regular system health checks and performance monitoring should be in place to identify and address potential issues proactively.

- A clear update and patching mechanism should be established to ensure the system stays current with technological advancements and best practices.

**Accessibility:**

- The dashboard should be designed with accessibility in mind, ensuring it's usable by individuals with disabilities. This includes considerations for screen readers, high-contrast modes, and keyboard navigation.

- The platform should be optimized for various devices, from desktops to mobile phones, ensuring a consistent experience across devices.

# CHAPTER 4: Software Design Specification

## 4.1 USE CASE DIAGRAM

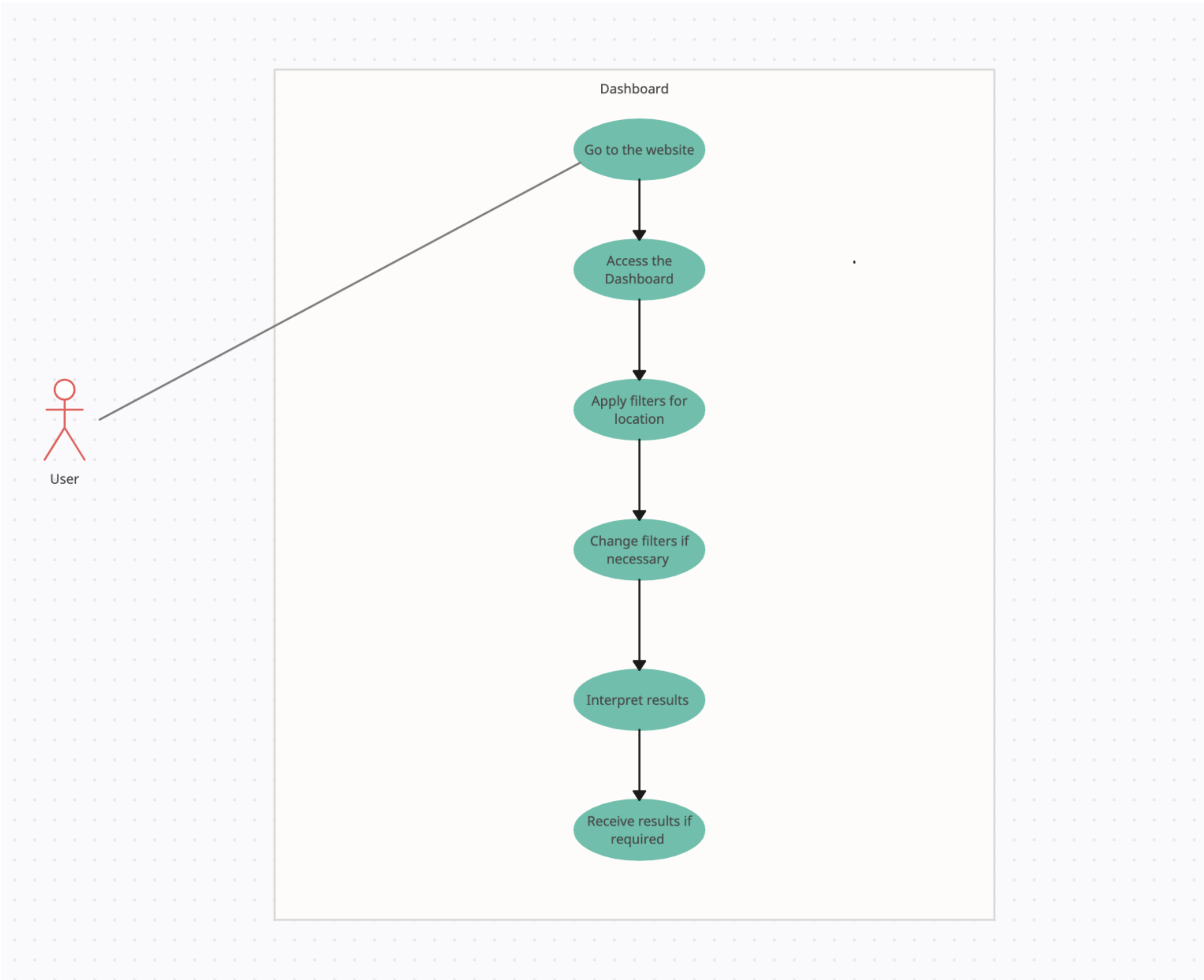



Figure 4.1 Use Case Diagram

# CHAPTER 5: PROJECT IMPLEMENTATION

## 1. Data Collection and Preprocessing:

The project's implementation initiates with the acquisition of historical air quality data from the AQODP (Air Quality Open Data Platform) for selective locations within Pune city. This data, enriched with various meteorological and environmental parameters, serves as the foundation for machine learning model training. A rigorous preprocessing phase follows, encompassing data cleaning, handling missing values, and normalization to ensure the quality and uniformity of the dataset.

## 2. Machine Learning Model Integration:

The core of the project lies in the integration of diverse machine learning models, each tailored to capture specific nuances within urban air quality data. The models, including ARIMA, LSTM, Facebook Prophet, and TBTAS, undergo a comprehensive training process. Historical data is partitioned into training and validation sets, allowing models to discern patterns and relationships. The performance of each model is evaluated against benchmark metrics, and iterative refinement cycles are employed to enhance predictive accuracy.

## 3. Visualization with Tableau:

Leveraging the powerful visualization capabilities of Tableau, the project transforms complex datasets into interactive and comprehensible dashboards. Real-time visualizations of air quality metrics are designed with a user-centric approach, ensuring accessibility for residents, city planners, and health agencies. Visual representations, such as time series graphs, heat maps, and geographical overlays, provide diverse insights into air quality trends.

4. **Public Engagement and Communication:**

To foster public engagement, the implementation includes targeted communication strategies. The project employs user-friendly language, infographics, and educational initiatives to raise awareness about the importance of air quality. Interactive elements within the dashboard encourage residents to explore and understand the presented data, empowering them to make informed decisions regarding outdoor activities and lifestyle choices.

5. **Continuous Improvement Mechanisms:**

Implementation is not a static process but a dynamic one that incorporates continuous improvement mechanisms. Feedback loops are established to gather user input, ensuring that the dashboard remains responsive to user needs. Machine learning models are adaptively refined, incorporating new data sources and advancements in forecasting methodologies, to enhance their predictive capabilities over time.

6. **Decision Support for City Planners:**

The implementation extends to providing city planners with a valuable decision support tool. The dashboard equips planners with insights into the air quality implications of urban development, traffic patterns, and environmental policies. This collaborative aspect ensures that the project contributes to informed and sustainable city planning practices.

In conclusion, the project's implementation is characterized by a meticulous approach to data handling, robust machine learning integration, intuitive visualization, public engagement strategies, and a commitment to continuous improvement. These elements collectively contribute to the realization of an impactful Urban Air Quality Forecasting & Visualization Dashboard for Pune city.

# CHAPTER 6: SOFTWARE TESTING

For the Urban Air Quality Forecasting & Visualization Dashboard, a thorough testing regimen is critical to ensure data accuracy, system reliability, and optimal user experience. Various forms of testing will be employed:

1. **Unit Testing:**

- **Description:** This type of testing involves verifying the smallest parts of the application in isolation (like functions or methods) to ensure they work as intended.
- **Relevance:** Given the critical nature of forecasting and data retrieval functionalities, unit tests will be pivotal in identifying and rectifying potential issues at the granular code level.

2. **Integration Testing:**

- **Description:** Integration tests are conducted to ensure that different components of the software work well together.
- **Relevance:** This is crucial for our dashboard, especially when integrating AQODP data with our forecasting algorithms and subsequently with the visualization components in Tableau.

3. **System Testing:**

- **Description:** Here, the entire system is tested to ensure it meets the specified requirements and functions correctly as a complete entity.

- **Relevance:** This ensures that the entire dashboard, from data retrieval to visualization, functions seamlessly, providing accurate and timely AQI information.

4. **Usability Testing:**

- **Description:** Usability tests assess the software from an end-user's perspective, ensuring it is user-friendly and intuitive.

- **Relevance:** As our dashboard targets a range of users, from city planners to residents like Rahul, it's vital to ensure that users can easily navigate and extract value from the tool.

5. **Performance Testing:**

- **Description:** This testing assesses the software's robustness, speed, and overall efficiency.

- **Relevance:** Given the real-time requirements of AQI data and the potential volume of users, it's essential to ensure the dashboard remains responsive and doesn't experience lags or crashes.

6. **Security Testing:**

- **Description:** Security tests identify vulnerabilities in the software, ensuring data protection and unauthorized access prevention.

- **Relevance:** As our system fetches data online and might store user preferences or other data in the future, it's crucial to safeguard against potential security threats.

7. **Acceptance Testing:**

   - **Description:** This phase checks if the software meets the business requirements and if it's ready for deployment and use.

   - **Relevance:** Before launching the dashboard to the public, acceptance testing ensures that it aligns with the project's initial objectives and stakeholder expectations.

8. **Regression Testing:**

   - **Description:** After updates or changes, regression tests ensure that previously developed and tested software still operates after a change.

   - **Relevance:** Given that our system will undergo periodic updates, especially in the forecasting algorithms, it's essential to ensure that new additions don't disrupt existing functionalities.

Each of these testing phases will employ both manual testing techniques and automated testing tools, ensuring thorough coverage and efficient issue identification. The testing process aims to ensure that the dashboard delivers accurate AQI forecasts and offers a robust and user-friendly experience

# CHAPTER 7: RESULTS

1. **Katraj- Pune**

| **Model Used** | ARIMA | FACEBOOK PROPHET | TBTAS |
|---|---|---|---|
| **Accuracy *(in percentage)*** | 89.11 | 82.09 | 53.85 |

*Table No. 1: Forecast Results- Katraj, Pune*

2. **Pashan - Pune**

| **Model Used** | ARIMA | FACEBOOK PROPHET | TBTAS |
|---|---|---|---|
| **Accuracy *(in percentage)*** | 96.67 | 63.35 | 26.04 |

*Table No. 2: Forecast Results- Pashan, Pune*

3. **ShivajiNagar - Pune**

| **Model Used** | ARIMA | FACEBOOK PROPHET | TBTAS |
|---|---|---|---|
| **Accuracy *(in percentage)*** | 77.88 | 82.09 | - |

*Table No. 3: Forecast Results- ShivajiNagar, Pune*

4. **Karve Road - Pune**

| **Model Used** | ARIMA | FACEBOOK PROPHET | TBTAS |
|---|---|---|---|
| **Accuracy *(in percentage)*** | 73.23 | - | 63.22 |

*Table No. 4: Forecast Results- Karve Road, Pune*

**5. Nigdi - Pune**

| **Model Used** | ARIMA | FACEBOOK PROPHET | TBTAS |
|---|---|---|---|
| **Accuracy** ***(in percentage)*** | 83.19 | 30.80 | 26.04 |

*Table No. 5: Forecast Results- Nigdi, Pune*

**6. Bhosari - Pune**

| **Model Used** | ARIMA | FACEBOOK PROPHET | TBTAS |
|---|---|---|---|
| **Accuracy** ***(in percentage)*** | 83.67 | - | 77.42 |

*Table No. 6: Forecast Results- Bhosari, Pune*

**7. Lohegaon - Pune**

| **Model Used** | ARIMA | FACEBOOK PROPHET | TBTAS |
|---|---|---|---|
| **Accuracy** ***(in percentage)*** | 87.76 | - | 77.42 |

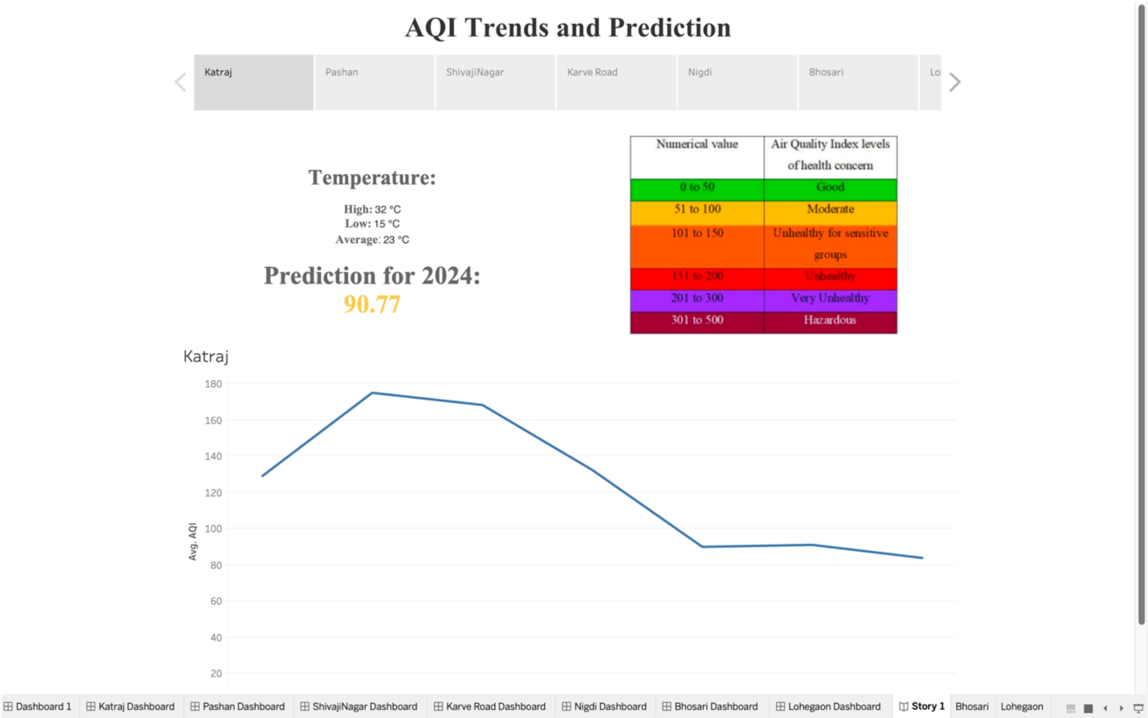
*Table No. 7: Forecast Results- Lohegaon, Pune*

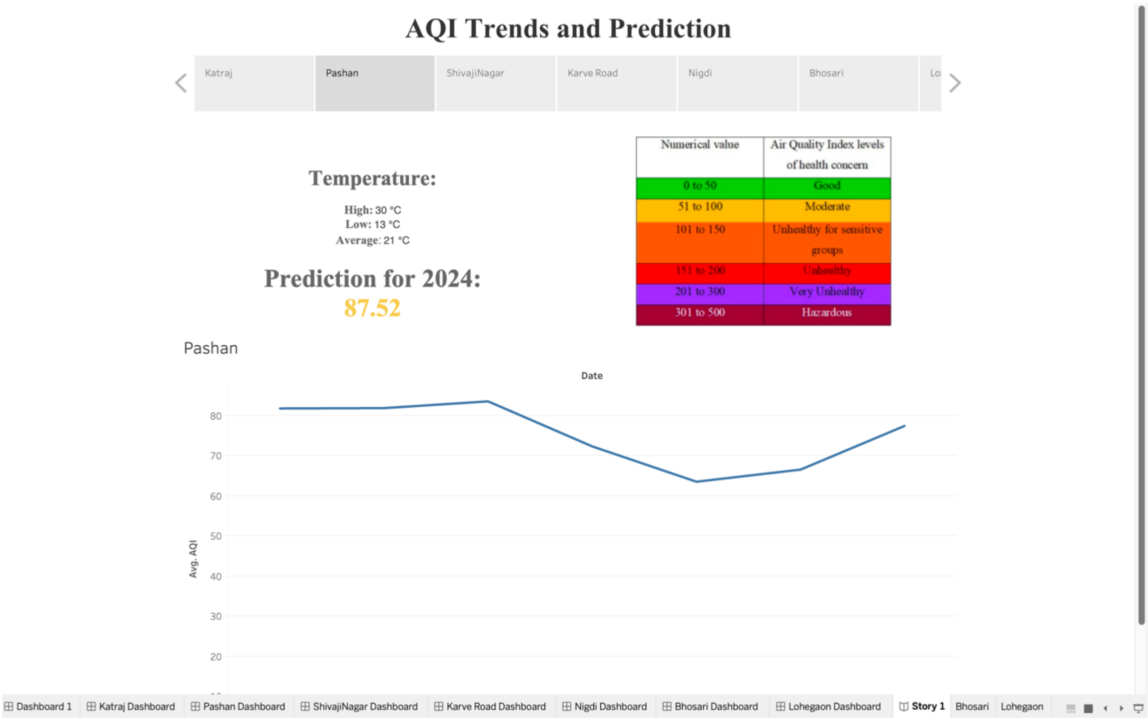
**Predicted Results:**

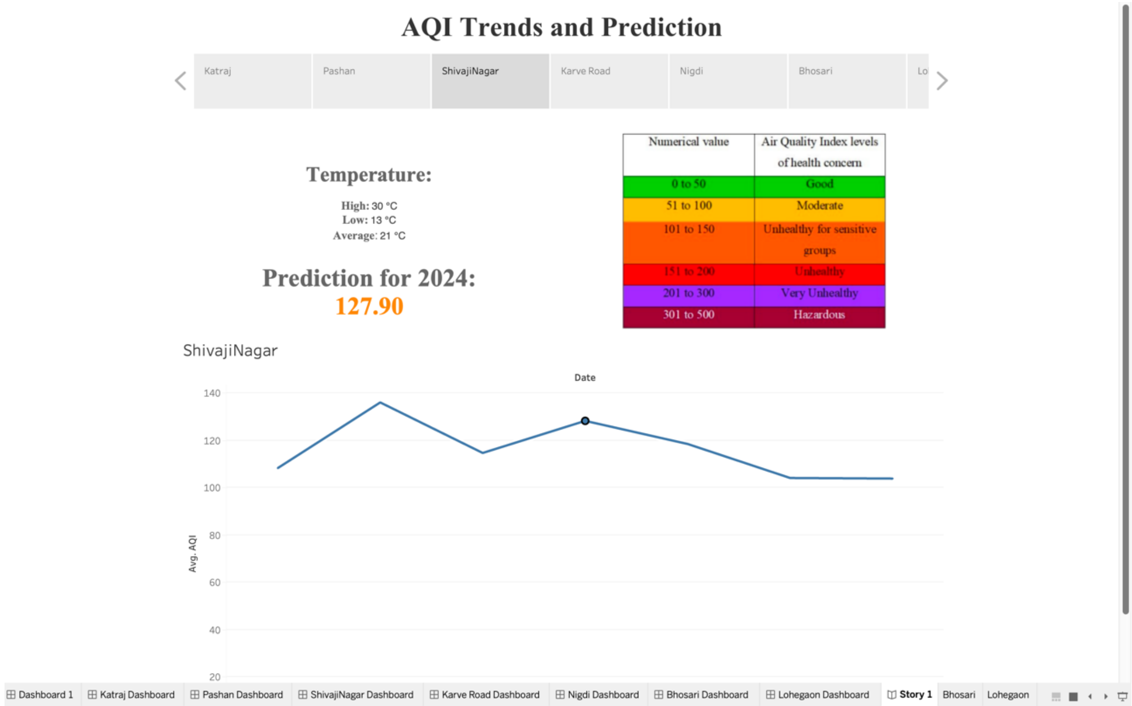
| Sr. No. | Region (Pune District) | Average AQI predicted |
|---|---|---|
| 1 | Katraj | 90.77 |
| 2 | Pashan | 87.52 |
| 3 | ShivajiNagar | 127.90 |
| 4 | Karve Road | 93.76 |
| 5 | Nigdi | 94.45 |
| 6 | Bhosari | 114.16 |
| 7 | Lohegaon | 110.13 |

*Table No. 8: Predicted Results*

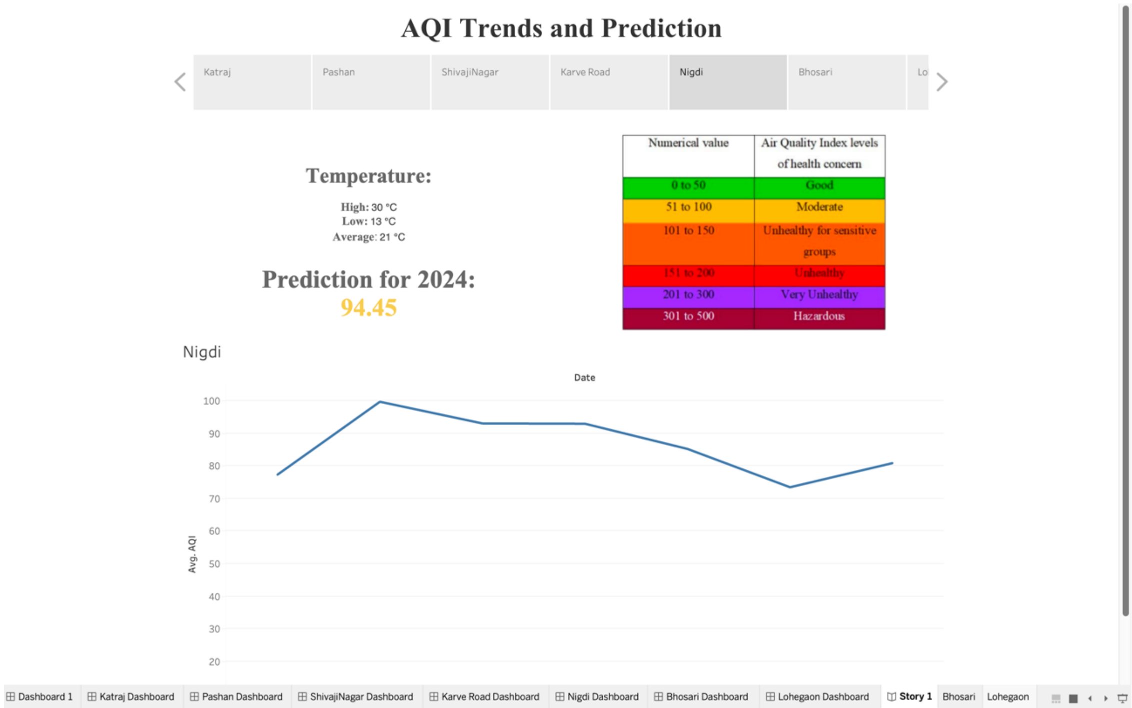
## 7.1 Screenshots:





AQI Trends and Prediction
Katraj
Pashan
ShivajiNagar
Karve Road
Nigdi
Bhosari
Temperature:
High: 30 °C
Low: 13 °C
Average: 21 °C
Prediction for 2024:
127.90
Numerical value
Air Quality Index levels of health concern
0 to 50
Good
51 to 100
Moderate
101 to 150
Unhealthy for sensitive groups
151 to 200
Unhealthy
201 to 300
Very Unhealthy
301 to 500
Hazardous
ShivajiNagar
Date
Avg. AQI
140
120
100
80
60
40
20
Dashboard 1
Katraj Dashboard
Pashan Dashboard
ShivajiNagar Dashboard
Karve Road Dashboard
Nigdi Dashboard
Bhosari Dashboard
Lohegaon Dashboard
Story 1
Bhosari
Lohegaon

AQI Trends and Prediction
Katraj
Pashan
ShivajiNagar
Karve Road
Nigdi
Bhosari
Temperature:
High: 30 °C
Low: 13 °C
Average: 21 °C
Prediction for 2024:
94.45
Numerical value
Air Quality Index levels of health concern
0 to 50
Good
51 to 100
Moderate
101 to 150
Unhealthy for sensitive groups
151 to 200
Unhealthy
201 to 300
Very Unhealthy
301 to 500
Hazardous
Nigdi
Date
Avg. AQI
100
90
80
70
60
50
40
30
20
Dashboard 1
Katraj Dashboard
Pashan Dashboard
ShivajiNagar Dashboard
Karve Road Dashboard
Nigdi Dashboard
Bhosari Dashboard
Lohegaon Dashboard
Story 1
Bhosari
Lohegaon

# CHAPTER 9: CONCLUSION

## 9.1 CONCLUSION

AQI is a widely recognized environmental metric that quantifies the quality of the air we breathe. Lower AQI values signify healthier air quality, while higher values indicate increased levels of pollutants and a potential health threat. With the threshold of 100 often being considered as a differentiation point between satisfactory and moderate air quality, several regions like ShivajiNagar, Bhosari, and Lohegaon are predicted to cross this benchmark, highlighting potential environmental concerns.

From the perspective of homeowners, these predictions can be invaluable. The data suggests that regions like ShivajiNagar, Bhosari, and Lohegaon might experience diminished air quality, which can influence property values, desirability for residential purposes, and, most importantly, the health of the residents. Residents in these areas might need to invest in air purifiers or take precautions to minimize outdoor activities during peak pollution periods.

For environmentalists, these predictions underscore the urgency for intervention. The elevated AQI levels in regions like ShivajiNagar demand immediate attention, possibly through strategies like increasing green spaces, regulating local sources of pollution, or implementing stricter emissions guidelines. Such predictions offer a roadmap for future environmental strategies, emphasizing areas that need priority attention.

In conclusion, the forthcoming AQI predictions for Pune's regions serve as an environmental compass, guiding both homeowners and environmentalists. While homeowners can leverage this data to make informed decisions about their residences and health precautions, environmentalists are provided with clear markers to direct their efforts and champion for sustainable urban living. This data underscores the symbiotic relationship between urban growth and environmental consciousness and underscores the importance of proactive strategies to ensure a healthier tomorrow.

## 9.2 FUTURE WORK

1. **Enhanced Machine Learning Models:**

   Future iterations of the project will focus on exploring and integrating more advanced machine learning models to further improve the accuracy and robustness of air quality forecasts. This includes investigating ensemble methods and hybrid models that combine the strengths of different algorithms.

2. **Incorporation of External Factors:**

   To enhance the predictive capabilities of the dashboard, future work will involve incorporating additional external factors that may influence air quality, such as industrial activities, construction projects, and traffic congestion. This expanded dataset will contribute to a more comprehensive understanding of urban air quality dynamics.

3. **Integration of Sensor Networks:**

   The project envisions incorporating real-time data from sensor networks strategically placed across the city. Integrating this high-frequency data will provide a more granular view of air quality variations and enable the dashboard to respond more swiftly to sudden changes or events.

4. **User Feedback Integration:**

   Establishing a structured mechanism for collecting and incorporating user feedback will be a priority in future work. This will involve refining the dashboard interface based on user experiences, preferences, and specific needs, ensuring that it remains user-friendly and effective.

5. **Predictive Analytics for Health Impact:**

   Future iterations of the project could explore the integration of predictive analytics models to assess the potential health impacts of varying air quality conditions. This could involve correlating air quality data with health statistics to provide residents with personalized health recommendations based on forecasted conditions.

6. **Expansion to Other Cities:**

   With a solid foundation in Pune city, the future vision includes expanding the Urban Air Quality Forecasting & Visualization Dashboard to other urban centers. This expansion will involve adapting the machine learning models and visualization strategies to the unique characteristics of different cities.

7. **Collaboration with Environmental Agencies:**

   Collaborative efforts with environmental agencies and research institutions could be a focal point in the future. This involves sharing insights, data, and methodologies to contribute to broader research on urban air quality and to align the project with evolving environmental standards.

8. **Integration with Smart City Initiatives:**

   Future work will explore closer integration with smart city initiatives, leveraging technologies like the Internet of Things (IoT) for real-time data collection and decision-making. This could enhance the overall smart city ecosystem and contribute to more sustainable urban development.

9. **Climate Change Impact Assessment:**

As climate change continues to influence weather patterns, future work may involve assessing the impact of these changes on urban air quality. This forward-looking approach can contribute to proactive measures in mitigating the effects of climate-induced air quality variations.

## 9.3 APPLICATIONS

1. **City Planning and Infrastructure Development:**

The Urban Air Quality Forecasting & Visualization Dashboard serves as a pivotal tool for city planners, providing insights into the impact of urban development and infrastructure projects on air quality. By integrating air quality forecasts, city planners can make informed decisions to create sustainable and environmentally conscious urban spaces.

2. **Health Advisory and Public Awareness:**

Residents benefit from the dashboard's real-time air quality visualizations and forecasts, empowering them to make informed decisions about outdoor activities and health precautions. The dashboard acts as a health advisory tool, raising public awareness about the correlation between air quality and well-being.

3. **Environmental Policy Formulation:**

Environmental agencies and policymakers can leverage the dashboard to formulate data-driven environmental policies. The insights provided contribute to evidence-based decision-making, enabling the creation of regulations and initiatives aimed at improving air quality.

4. **Traffic Management Strategies:**

With the integration of air quality forecasts, the dashboard aids in the formulation of effective traffic management strategies. By understanding the correlation between vehicular emissions and air quality, traffic authorities can implement measures to alleviate congestion and reduce pollution.

5. **Urban Resilience and Emergency Response:**

The dashboard enhances urban resilience by providing a proactive tool for emergency response planning. In the event of sudden air quality deterioration due to natural disasters or industrial incidents, emergency response teams can use the dashboard to implement timely and effective measures.

6. **Research and Academic Studies:**

The project contributes to academic and research endeavors focused on urban air quality. Researchers can utilize the dataset and methodologies for in-depth studies, furthering the understanding of air quality dynamics and the effectiveness of forecasting models.

## REFERENCES

1. World Health Organization. (2018). Ambient (outdoor) air quality and health. WHO.
2. U.S. Environmental Protection Agency. (2020). Air Quality Index (AQI) Basics. EPA.
3. Beelen, R., Raaschou-Nielsen, O., Stafoggia, M., et al. (2014). Effects of long-term exposure to air pollution on natural-cause mortality: an analysis of 22 European cohorts within the multicentre ESCAPE project. The Lancet.
4. Mitchell, T. (1997). Machine Learning. McGraw Hill.
5. Balling, R.J., & Cerveny, R.S. (1998). Weekly cycles of air pollutants, precipitation and tropical cyclones in the coastal NW Atlantic region. Nature.
6. Chauhan, A., & Singh, R.P. (2020). Decline in PM2.5 concentrations over major cities around the world associated with COVID-19. Environmental Research.
7. Tableau Public. (n.d.). Tableau Public Gallery. Tableau Software.
8. World Health Organization. (2016). Ambient air pollution: A global assessment of exposure and burden of disease. WHO.
9. Dockery, D.W., Pope III, C.A., Xu, X., et al. (1993). An association between air pollution and mortality in six U.S. cities. The New England Journal of Medicine.
10. "Air Pollution". (n.d.). National Institute of Environmental Health Sciences. Retrieved from [https://www.niehs.nih.gov/health/topics/ agents/air-pollution/index.cfm].
11. Bell, M.L., Davis, D.L., & Fletcher, T. (2004). A retrospective assessment of mortality from the London smog episode of 1952: the role of influenza and pollution. Environmental Health Perspectives.
12. "Real-time Air Quality Index for more than 60 countries in the world". (n.d.). AQICN. Retrieved from [https://aqicn.org/].
13. Géron, A. (2019). Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow. O'Reilly Media, Inc.
14. "The History of Air Quality Monitoring". (n.d.). Aeroqual. Retrieved from [https://www.aeroqual.com/history-air-quality-monitoring].
15. Brundson, C., & Comber, L. (2015). An Introduction to R for Spatial Analysis and Mapping.
16. "Machine Learning: What it is and why it matters". (n.d.). SAS. Retrieved from [https://www.sas.com/en_us/insights/analytics/ machine-learning.html].

17. "Time Series Forecasting with ARIMA in Python 3". (2018). DigitalOcean. Retrieved from [https://www.digitalocean.com/ community/tutorials/a-guide-to-time-series-forecasting-with-arima-in- python-3].

18. "Machine Learning with Python: Regression Analysis". (2019). RealPython. Retrieved from [https://realpython.com/linear-regression- in-python/].

19. Chollet, F. (2017). Deep Learning with Python. Manning Publications.

20. IBM "What is machine learning?" [https://www.ibm.com/topics/ machine-learning]

BRACT's
Vishwakarma Institute of Information Technology
Department of Computer Engineering.

**B.Tech Project Formal Technical Review 1: Evaluation Report (AY 2022-23: Sem 2)**

Project Group No.: 3 Date: 11/09/2023

Project Title: **Urban Air Quality Forecasting & Visualization Dashboard**

Category: In-house

| Sr. No. | Student Name | Gr No. | Rno | Phone no. |
|---|---|---|---|---|
| 1 | Rahul Gokul Talele | 22010379 | 422072 | 9552589611 |

Checklist:

1. Purpose of the project known (Yes/No): Yes
2. Scope of the system defined (yes/No): Yes
3. Are the domain keywords identified by students and does it support the title? Yes
4. Does the title of the project explain Domain and Algorithm?( Not technology)? Yes
5. Is project plan prepared (Yes/No): Yes
6. Is Literature survey file maintained(Yes/No): Need to be work on it.
7. Remarks:

Name of Internal Guide: Dr. N. N. Wasatkar. Expert Name ~~and~~ Sign: Prof. Y. V. Dongre.
Signature with date: 07/10/2023.

(If title is not feasible, guide and expert reserve right to reject it and suggest suitable title)

BRACT's
Vishwakarma Institute of Information Technology
Department of Computer Engineering.

**Formal Technical Review Report 2**
**(B. Tech. Project)**

Date: 19/10/2023

Project Group No.: **3**

Project Title: **Urban Air Quality Forecasting & Visualization Dashboard**

Comments on the previous review report:

Checklist:

1. SRS in IEEE format prepared (Yes/No): yes
2. Log Report is up to date (Yes/No): yes
3. Is vision Document prepared: -
4. Is mathematical modeling done (yes/No): ys
5. Cost and Effort estimation done (yes/No): -
6. Use Case Diagrams Prepared (yes/No): yes
7. State Transition Diagrams prepared (yes/No)
8. Is selection of current technology, programming Language, tools for project implementation is discussed and presented?(Yes/No): yes
9. Is sample code of Modules designed (yes/No): ye
10. Is project implementation as per UML diagram? (yes/No): yes
11. Is the input output table prepared?( Results in the form of Table/Graph, images etc? (Yes/No): yes
12. Is Conference paper published and its status(published or accepted or Communicated)?(Yes/No):

13. Marking scheme of CIE(100 marks) and Oral (50 marks) explained by Respective Guide and Understood same (Yes/No): yes

14. Remarks:________________________________________________

**Internal Guide Name:**

Dr. Namrata N Wasatkar

**Sign:**

**Expert Name:**

Prof. Y. V. Dongre

**Sign:**

BRACT's
Vishwakarma Institute of Information Technology
Department of Computer Engineering.

**Formal Technical Review Report 3 (B. Tech Projects)**

Project Group No.: 3 Date: 28/11/2023

Project Title: **Urban Air Quality Forecasting & Visualization Dashboard**

Comments on previous review report and necessary actions taken (if)

—

| Sr.No. | Student Name | Phone No. |
|---|---|---|
| 1 | Rahul Gokul Talele | 9552589611 |

Checklist:

1. Is result of implemented modules meets the actual problem statement (Yes/No): yes
2. Is Project Vision and scope maps with the outcome(Yes/No): —
3. Have the modules implemented as per functional and non functional requirement (yes/No): yes
4. Are the implemented modules tested using specific testing tools(yes/No) yes
5. Have the implemented modules critically analyzed?(yes/No) yes
6. Is Human Computer Interaction(GUI) and Design appropriate?(yes/No): yes
7. Is Coding standards applied? (yes/No) yes
8. Are the modules technically understood? (Yes/No) yes
9. Are the implemented modules tested well?(yes/No) yes

10. No of modules planned and implemented (e.g 4/5): 7

11. Paper presentation status Published

Remark: Excellent/Good/Average/Below Average/Not Satisfactory)

Name of Guide: **Dr. Namrata N Wasatkar**

Sign:

Name of Expert: **Prof. Y. V. Dongre**

Sign:

# CERTIFICATE OF PRESENTATION

This is to certify that **RAHUL GOKUL TALELE** of Vishwakarma Institute of Information Technology, has presented the paper titled “Predictive Modeling for Urban Air Quality: A Machine Learning Comparison Study” at **1st International Conference on Deep Learning, IoT, Drone Technology, Smart Cities & Application (ICDIDSA-23)** held on 29-30 November 2023 and organized by the Department of Electronics and Telecommunication Engineering, Sandip Institute of Technology and Research Centre, Nashik, Maharashtra, India.

| Ankur Saxena | Dr. Gayatri M. Phade | Dr. M.M. Patil |
|---|---|---|
| Conference Coordinator | Conference Chair | Convenor |
| SITRC | HOD, SITRC | Principal, SITRC |