

***A Mini Project Synopsis on***  
**Air Quality Index Prediction**

**T.E. - I.T Engineering**

**Submitted By**

**Yugandhar Ghatge            19104046**

**Mayuri Patil                19104012**

**Abhijeet Mishra            19104019**

**Under The Guidance Of**

**Prof. Neha Deshmukh**



**DEPARTMENT OF INFORMATION TECHNOLOGY**

**A.P.SHAH INSTITUTE OF TECHNOLOGY**

**G.B. Road, Kasarvadavali, Thane (W), Mumbai-400615**

**UNIVERSITY OF MUMBAI**

**Academic year : 2021-22**

## CERTIFICATE

This to certify that the Mini Project report on Stock Prediction has been submitted by **Yugandhar Ghatge (19104046), Mayuri Patil (19104012) and Abhijeet Mishra (19104019)** who are a Bonafede students of A. P. Shah Institute of Technology, Thane, Mumbai, as a partial fulfilment of the requirement for the degree in **Information Technology**, during the academic year **2021-2022** in the satisfactory manner as per the curriculum laid down by University of Mumbai.

Ms. Neha Deshmukh  
Guide

Prof. Kiran Deshpande  
Head Department of Information Technology

Dr. Uttam D.Kolekar  
Principal

External Examiner(s)

- 1.
- 2.

Place: A. P. Shah Institute of Technology, Thane  
Date:

## TABLE OF CONTENTS

1.	Introduction.....	1
1.1.	Purpose.....	1
1.2.	Objectives.....	1
1.3.	Scope.....	2
2.	Problem Definition .....	3
3.	Proposed System .....	4
	3.1. Features and Functionality .....	4
4.	Project Outcomes .....	7
5.	Software Requirements.....	8
6.	Project Design .....	9
7.	Project Scheduling.....	15
8.	Screenshot of Application.....	16
9.	Conclusion .....	17

References

Acknowledgement

# **Chapter 1**

## **Introduction**

Air pollution and its prevention are constant scientific challenges during last decades. However, they still remain huge global problems. Affecting human's respiratory and cardiovascular system, they are cause for increased mortality and increased risk for diseases for the population. Many efforts from both local and state government are done in order to understand and predict air quality index aiming improved public health. This paper is one scientific contribution towards this challenge.

As the largest growing industrial nation, India is producing record amount of pollutants specifically Co<sub>2</sub>, pm<sub>2.5</sub> etc and other harmful aerial contaminants. Air quality of a particular state or a country is a measure on the effect of pollutants on the respected regions, as per the Indian air quality standard pollutants are indexed in terms of their scale, these air quality indexes indicates the levels of major pollutants on the atmosphere. There are various atmospheric gases which causes pollution on our environment. Each pollution has individual index and scales at different levels. The major pollutants Such as (no<sub>2</sub>, so<sub>2</sub>, rspm, spm) indexes AQI is acquired, with this individual AQI, the data can be categorized based on the limits. We collected the data from the Indian government database, which contains pollutant concentration occurring at various places across India. We start by calculating the individual index of the pollutant for every available datapoints and find their respective AQI for the region. We have designed a model to predict the air quality index of every available data points in the dataset, our model is capable of forecasting the air quality of India in any given area. By predicting the air quality index, we can backtrack the major pollution causing pollutant and the location affected seriously by the pollutant across India. With this forecasting model, various knowledge about the data are extracted using various techniques to obtain heavily affected regions on a particular region(cluster). This gives more information and knowledge about the cause and seniority of the pollutants.

## **1.1 Purpose**

The purpose of the Air Quality Index (AQI) is to help you understand what local air quality means to your health. To make it easier to understand, the AQI is divided into six (6) categories. Each category corresponds to a different level of health concern. There are six (6) levels of health concern.

1. Good : The AQI value for your community is between 0 and 50. Air quality is considered satisfactory, and air pollution poses little or no risk.
2. Moderate: The AQI for your community is between 51 and 100. Air quality is acceptable; however, for some pollutants there may be a moderate health concern for a very small number of people. For example, people who are unusually sensitive to ozone may experience respiratory symptoms.
3. Unhealthy for Sensitive Groups: When AQI values are between 101 and 150, members of sensitive groups may experience health effects. This means they are likely to be affected at lower levels than the general public. For example, people with lung disease are at greater risk from exposure to ozone, while people with either lung disease or heart disease are at greater risk from exposure to particle pollution. The general public is not likely to be affected when the AQI is in this range.
4. Unhealthy: Everyone may begin to experience health effects when AQI values are between 151 and 200. Members of sensitive groups may experience more serious health effects.
5. Very Unhealthy: Values between 201 and 300 trigger a health alert, meaning everyone may experience serious health effects.
6. Hazardous: Values over 300 trigger health warnings of emergency conditions. The entire population is more likely to be affected.

## 1.2 Objective

1. The AQI is an index for reporting daily air quality.
2. To compare air quality conditions at different locations/cities.
3. The purpose of the AQI is to help people know how the local air quality impacts their health.
4. It also helps in identifying faulty standards and inadequate monitoring programs.
5. AQI helps in analysing the change in air quality (Improvement or degradation).
6. AQI informs the public about environmental conditions. It is especially useful for people suffering from illness aggravated or caused by air pollution.
7. Different areas have different levels of air quality at different times so, it is important for us to monitor what is happening.
8. The AQI focuses on health effects you may experience within a few hours or days after breathing unhealthy air.

### **1.3. Scope**

- Air Quality Index Prediction is designed to secure the aerial life of humanity. The air quality index is an index for reporting air quality on daily basis.
- Machine learning algorithms make it easy to predict the PM 2.5 value and with simple steps, AQI of any desired city can be obtained.
- The aim is to investigate machine learning based technique for air quality forecasting using AQI index by predicting results in best accuracy.
- This system proves beneficial for cities as well as environment departments to take early actions.

## **Chapter 2**

### **Problem Definition**

- Air pollution is considered to occur whenever harmful or excessive quantities of defined substances such as gases, particulates, and biological molecules are introduced into the atmosphere.
- These excessive emissions have obvious consequences, causing diseases and death of populations and other living organisms and impairing crops.

- The most common air pollutants are known as the criteria pollutants, which correspond to the most widespread health threats, e.g., CO, SO<sub>2</sub>, lead, ground-level ozone (O<sub>3</sub>), NO<sub>2</sub>, and PM.
- The levels of these pollutants are measured by the Environmental Protection Agency (EPA), which controls overall air quality.
- Scientific research has demonstrated a correlation between short-term exposure to this kind of pollutants and many health problems, like limited ability to respond to increased oxygen demands when exercising (especially for people with heart conditions), airway inflammation in healthy people and increased respiratory symptoms for people with asthma, respiratory emergencies particularly for children and the elderly, and so on.



## **Chapter 3**

### **Proposed System**

#### **Features and Functionalities :**

Prediction : AQI is calculated based on chemical pollutant quantity. By using machine learning, we can predict the AQI. The air quality index is an index for reporting air quality on a daily basis. The main part of the system is where we are trying to create a resource and give the predictability to most accurate data. Based on the given city input of the user, the model predicts the Air Quality Index for that city.

#### **Environment :**

According to the PM 2.5 value, Air Quality Index range can be figured out as unhealthy or healthy. Formerly developed to make information available about the health effects of the five most common air pollutants, and how to avoid those effects.

#### **Ease :**

It is a user friendly application and time & location doesn't really matter.

## **Chapter 4**

### **Project Outcomes**

- Air Quality Index Prediction is capable of successfully predicting the upcoming AQI of any particular region.
- This system will help in minimizing the amounts of pollutants in air.
- The higher the AQI value, the greater the level of air pollution and the greater the health concerns.
- Improving air quality can deliver substantial health benefits; reducing air pollution levels means reducing premature deaths and diseases from stroke, heart disease, lung cancer, and both chronic and acute respiratory diseases, including asthma.
- This system not only proves beneficial to the environmental organizations but also to normal civilians so as they get a right to know their city's AQI.
- AQI can be obtained anytime and anywhere as it is a system application.

## Chapter 5

### Software Requirements

Air Quality Index prediction system is a user friendly web application requiring an operating environment with below functionalities :

#### **Softwares** :-

Python : Python is the main language used with specific ML algorithm.

Web Browsers – The application can run on any of the browsers such as Chrome, Firefox, etc.

IDE - Pycharm

#### **Modules** (Main) :-

Pickle - Python pickle module is used for serializing and de-serializing a Python object structure. Any object in Python can be pickled so that it can be saved on disk. What pickle does is that it “serializes” the object first before writing it to file.

Seaborn - Seaborn is a library for making statistical graphics in Python. It builds on top of matplotlib and integrates closely with pandas data structures. Seaborn helps you explore and understand your data.

#### **Hardware** :-

Operating System – Windows, Mac or Linux

RAM – Min. 500 MB or more

Hard Disk – Min. 2 GB of space

## **Chapter 6**

### **Project Design**

- Air Quality Index prediction is a simple web application which predicts the air quality of a city based on its previous analysis.
- It is user friendly and a user just needs to access the portal. The foremost page provides user an option to select which city's AQI he/she wants to predict.
- Along with that, AQI prediction system provides a user guide for all the non-technical people about the basic AQI indexes.
- Accordingly, the PM 2.5 value of that city gets generated and users can then get an idea about the current AQI.
- This system fetches real time data from openweathermap to get live climatic updates and accordingly displays the output.
- The AQI totally depends on what is the PM 2.5 value of that city. For example, if the PM 2.5 value of any city is 31.24 then it's AQI would be unhealthy.
- As it is a web based application, users can get access to it easily and thereby look forward to the AQI result.

## Chapter 7

### Project Scheduling

Sr. no.	Name	Duration	Functionality
1.	Yugandhar Ghatge	Feb - Mar	Collection of different datasets/API
		Mar - Apr	Integrating all the modules and framework
2.	Mayuri Patil	Feb - Mar	Analysing previously implemented projects
		Mar - Apr	Training of dataset by Randomforest
3.	Abhijeet Mishra	Feb - Mar	Finalising the algorithm and tech stack
		Mar - Apr	API implementing with testing

## Screenshot of Application

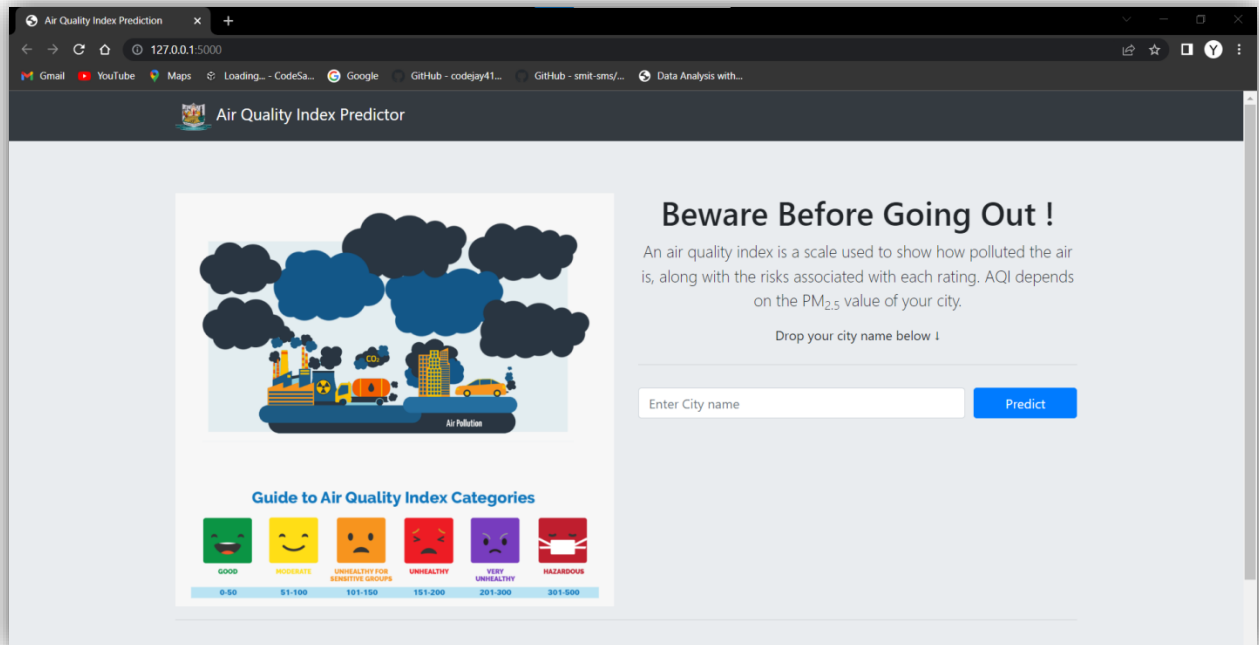


Fig 1 : User Interface

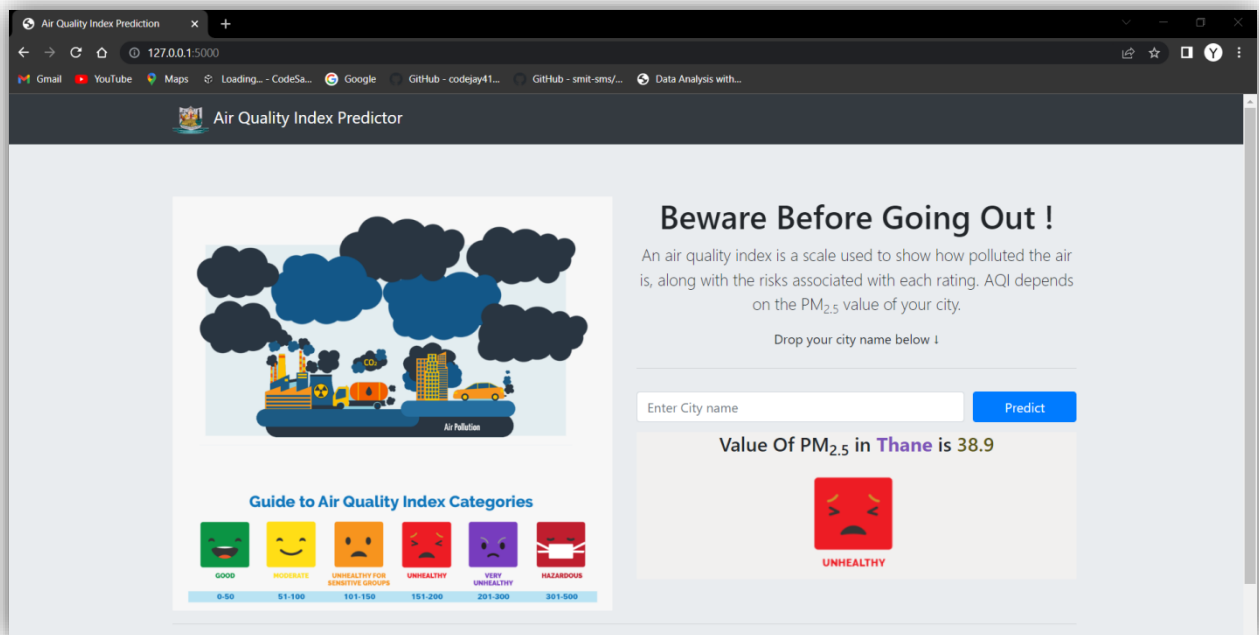


Fig 2 : AQI Prediction

## **Chapter 9**

### **Conclusion**

- According to the AQI, various environmental effects can be studied including social health.
- With this model we can forecast the AQI and alert the respected region of the country also it a progressive learning model it is capable of tracing back to the particular location needed attention provided the time series data of every possible region needed attention.
- The ability to predict air quality enables the government and other concerned organizations to take necessary steps to shield the most vulnerable, from being exposed to the air with hazardous quality.
- The exploratory data analysis and feature engineering methods implemented for the prediction models revealed interesting correlations between weather and pollution data.
- This system is completely user friendly and can be accessed anywhere and anytime.

### **References**

- <https://openweathermap.org/api>
- <https://www.bootstrapcdn.com/>
- Machine Learning With Random Forests And Decision Trees: A Visual Guide For Beginners by Scott Hartshorn
- <https://cpcb.nic.in/>

## **ACKNOWLEDGEMENT**

This project would not have come to fruition without the invaluable help of our guide **Prof. Neha Deshmukh**. Expressing gratitude towards our HoD, **Prof. Kiran Deshpande**, and the Department of Information Technology for providing us with the opportunity as well as the support required to pursue this project. We would also like to thank our teacher Ms. Neha Deshmukh who gave us her valuable suggestions and ideas when we were in need of them. We would also like to thank our peers for their helpful suggestions.