UE20CS390A – CAPSTONE PROJECT APPROVAL

Project Title : Monitoring the concentration of air pollutants and its health hazards using Machine Learning models.

Project ID: 102

Project Guide: Prof. Saritha R

Project Team :

Aditi Jain	PES2UG20CS021
Aditya R Shenoy	PES2UG20CS025
Ananya Adiga	PES2UG20CS043
Anirudha Anekal	PES2UG20CS051

Outline

Problem Statement

Concise description of the problem that needs to be addressed

Scope and Feasibility Study

Outlines the goals and objectives of the project and ensures it's viability

Applications/ Use Cases

Specifies the real world scenarios in which our project could be used

Technologies Used

Different technologies we plan on using

Expected Deliverables

Lists the expected Capstone(Phase I & II) deliverables to submit

Project Timeline

Gantt Chart to outline the timeline we propose to follow for our project.

Problem Statement

MONITORING THE CONCENTRATION OF AIR POLLUTANTS AND ITS HEALTH HAZARDS USING MACHINE LEARNING MODELS

Air Pollution

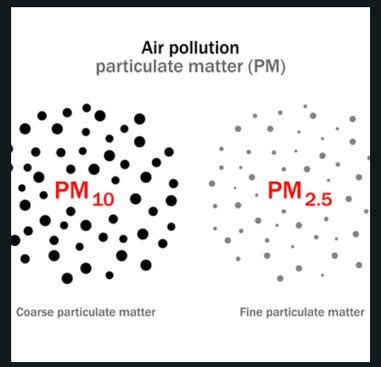
Air pollution is contamination of the indoor or outdoor environment by any chemical, physical or biological agent that modifies the natural characteristics of the atmosphere.

Concerns

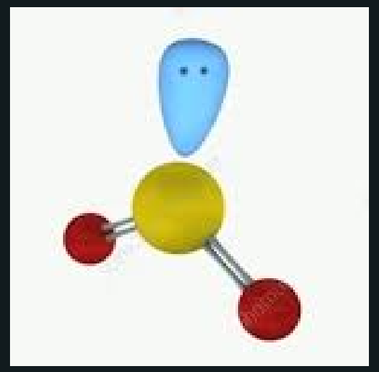
Air pollution has a wide range of negative impacts on the environment and human health like climate change, ozone depletion, acid rain and various harmful diseases.The average deaths per year in India alone is 1.66 million

Major Pollutants

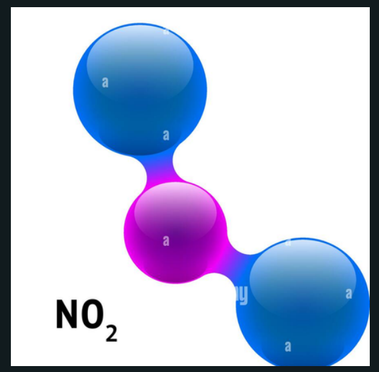
Particulate Matter



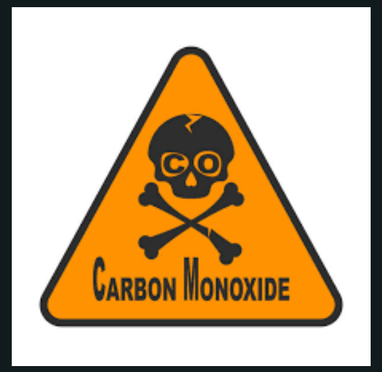
SO_2



NO_2



CO



Cigarette smoke

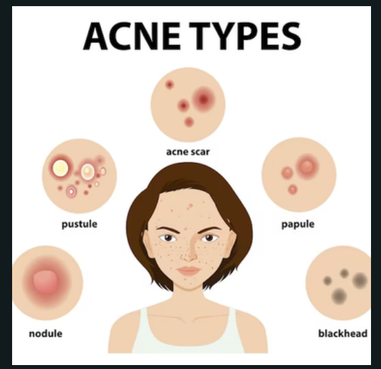
Diseases stemming from Air pollution

Bronchitis

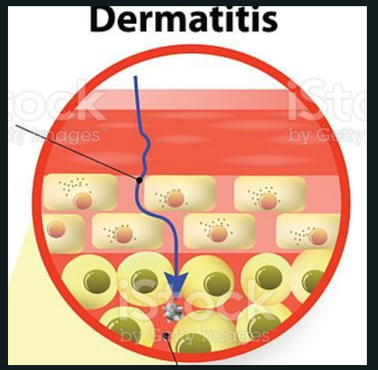
Asthma

Lung Cancer

Acne



Dermatitis



Air Quality Index Scale

Good 0-50	The air is fresh and free from toxins. People are not exposed to any health risk.
Moderate 51-100	Acceptable air quality for healthy adults but mild threat to sensitive individuals.
Sensitive groups 101-150	Inhaling such air can cause slight discomfort and difficulty in breathing.
Unhealthy 151-200	This could be typically problematic for children, pregnant women and the elderly.
Very Unhealthy 201-300	Exposure to air can cause chronic morbidities or even organ impairment.
Hazardous 301-500	Beware! Your Life is in danger. Prolonged exposure can lead to premature death.

Source: Central Pollution Control Board (https://cpcb.nic.in/)

Methodology

We aim to monitor various air pollutants and build a machine learning model to predict the probability of a person contracting various respiratory and skin diseases based on the extent to which they are exposed to harmful pollutants.

We will accomplish this by analyzing the pollution data obtained from various wireless sensors and deploy the trained model on cloud.

Scope

Prognosis using an ML model

We will create and optimize an ML algorithm to predict the probability of a person contracting various conditions or diseases like Bronchitis, asthma, acne and Dermatitis.

Measure pollutant concentration using wireless sensors

We will measure the concentration of various pollutants in the air like PM10, SO2 and NO2 using wireless sensors.

Environment or local atmosphere Analysis

Using our sensors we can generate AQI.
This can be used to understand how polluted our surroundings are.

Deploying the model on a cloud platform

We will upload the data from the wireless sensors to the cloud and pass it through our ML model.

Feasibility Study

Aim

Our project aims to monitor the concentration of pollutants in the air around us and analyze and predict its effects on human health.

Societal Impact

Our project is vital to various sects of society as they seek out information about their surroundings and air quality is a key component in this.

Tech Stack

The technology required to create our project is widely available (various sensors, a cloud platform, collaboration and design tools)

Datasets

There are a large number of publicly accessible datasets for pollution data.
There aren't as many public and accurate datasets for the diseases.

Applications/Use Cases

General Public

The general can be informed about the levels of pollution and the risks associated with it to take preventive measures

Smart home companies

Smart home businesses can incorporate it as an additional layer of detection for current devices

Government

Governments can use this to identify pollution hotspots and create policies and regulations to reduce the harmful emissions and issue public health warnings in high risk areas

City Planning

City planners and private companies can use this information to guide urban planning and development and create communities in a healthier living space

Possible Challenges

Obtaining data

We will need to ask regulating authorities for many of the datasets because they won't be available publicly.

Inaccurate Datasets

We will not be able to verify that the datasets we're utilizing to train our model are authentic.

Energy Consumption

The amount of energy used by IoT technology is significant and it needs to be running constantly.

Data Complexity

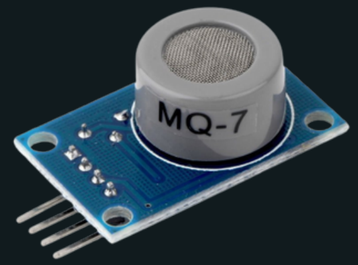
The data produced by IoT devices is often unstructured and provides a limited perspective along with the massive size.

Technology to be used

We will possibly employ the optimising LSTM model, the support vector regression model, the spatial temporal DL model, or a mix of random forest and an adaboost model. Platforms like ThingSpeak could be used to deploy to cloud

Various pollutants- MQ-135

Used for collecting and measuring the concentration of NOx and smoke.



Carbon Monoxide- MQ-7

Used for accurately collecting and measuring the CO concentrations

Cigarette smoke/ Dust Sensor- GP2Y1010AU0F

Used for collecting the concentration of PM and detecting cigarette smoke

AIR QUALITY MONITORING AND DISEASE PREDICTION USING IOT AND MACHINE LEARNING

By: Darshini Rajasekar, Aravind Sekar, Magesh Rajasekar

Case Study

Findings

- In this paper, they measure the concentrations of various pollutants in the atmosphere like SO2, NO2, PM, etc. with wireless sensors.
- The real time data they obtained is passed through a machine learning model and the Air Quality Index (AQI) is calculated
- They then use the AQI to predict the possibilities of various diseases like lung cancer, ventricular hypertrophy, etc.

Concerns

- This paper, as well as a lot of similar papers, solely use the AQI to estimate the likelihood of illness. But ever disease is caused by different pollutants or their combinations. Therefore forecasting the probability solely on AQI is not very helpful and there is a significant chance of errors in prediction.
- They have not cited any verifiable sources for their disease data.
- There is no final conclusion about the correlation between the pollutants and the diseases

EXPECTED DELIVERABLES

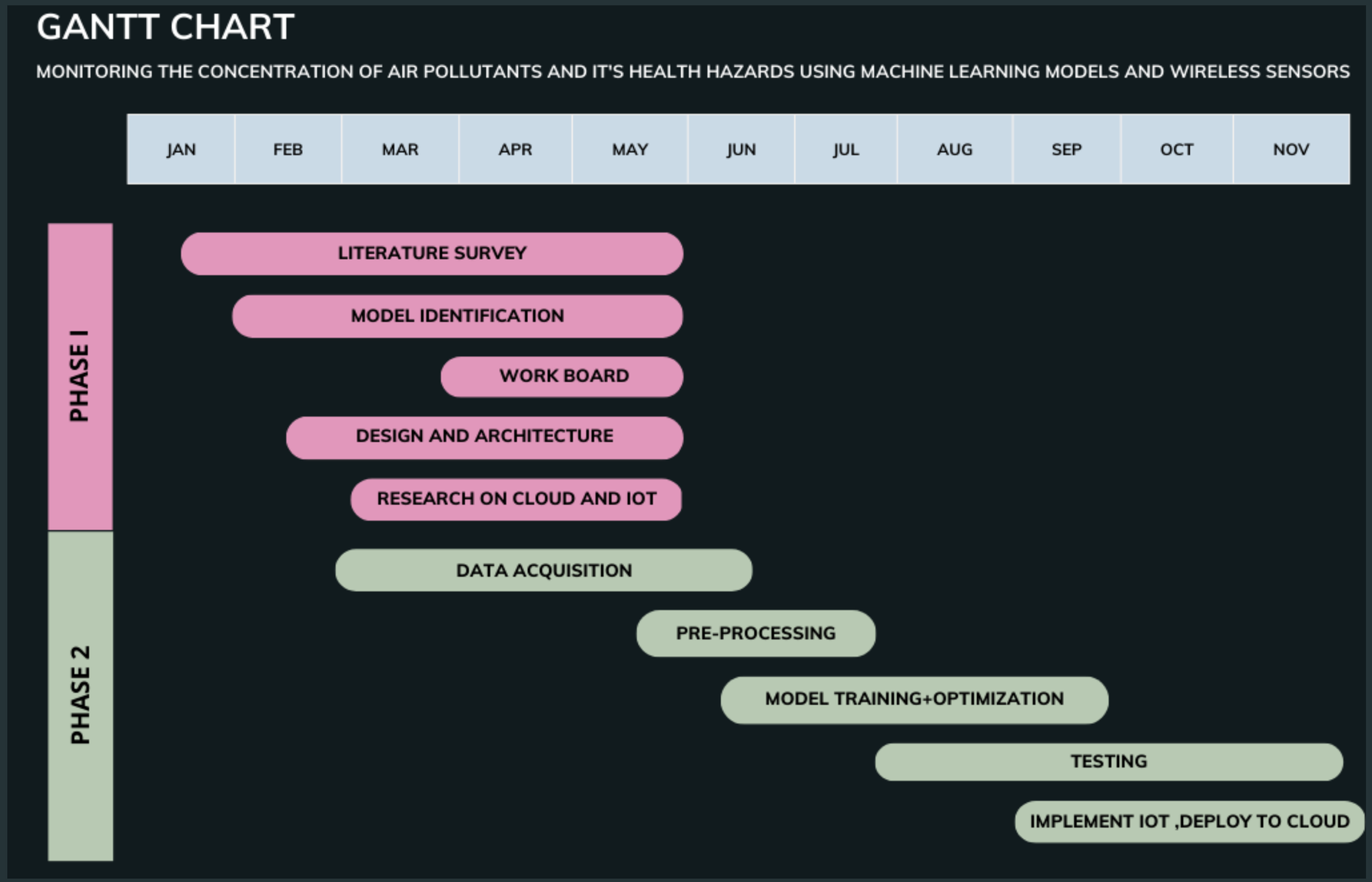
PHASE I

- **Literary survey**
- **Finalize the model and its optimization technique to be used.**
- **Design and architecture .**
- **Tech stack based on a mix of comfort ,usabilities and requirements of the model.**
- **Project plan using a gantt chart or burndown chart**

PHASE II

- Data: generation, collection, optimization etc.
- Optimizing the machine learning model (Theoretical)
- Code implementation (Practical implementation for the model)
- Deploying it on cloud platform
- Multiple prototypes for the sensor

PROJECT TIMELINE



References

[1] Increased risk of emergency department presentations for bronchiolitis in infants exposed to air pollution
By: Elisa Gallo, Silvia Bressan, Simoneta Baraldo
[https://onlinelibrary.wiley.com/doi/full/10.1111/risa.14007]

[2] Deep learning architecture for air quality predictions
By: Xiang Lee, Ling Peng, Yuan Hu
[https://link.springer.com/article/10.1007/s11356-016-7812-9]

[3] Air Quality Prediction using Machine Learning Algorithms –A Review
By: Tanisha Madan, Shrddha Sagar, Deepali Virmani
[https://ieeexplore.ieee.org/document/9362912]

[4] Forecasting Air Pollution Particulate Matter (PM2.5) Using Machine Learning Regression Models
By: Doreswamy, Harishkumar KS, Yogesh KM
[https://www.sciencedirect.com/science/article/pii/S187705092031206O]

[5] Link between environmental air pollution and allergic asthma
By: Quingling Zhang, Zhiming Qiu, Kian Fan Chung
[https://jtd.amegroups.com/article/view/3582/pdf]

Thank You