Air Quality Prediction using Machine Learning

Team Number : 15

Team Members:

19114006: Alankrit Anand
19114010: Anurag Bansal
19114022: Bhupendra Malakar
19114044: Kavish Lodha
19114090: Gaurav Wasnik

Aim of Project:

The aim of the project is to train the computer for predicting the air quality index(AQI) along with other parameters based on images of the sky around ourselves.

Goals of the Project:

1. It will help common people have knowledge about their environment's air quality which can help them to take appropriate precautions while going out.
2. It can help decrease the load of calculating Air Quality, which first involves calculating the concentrations of pollutants.

WorkFlow of the Project:

Basically the project is divided into 3 parts:

1. First part involves creating a Dataset for training the model upon, which will be a map between images of places and the corresponding AQI of that place when the photo was clicked.
    a. The dataset could be collected either manually by going around and clicking photos around ourselves and recording the corresponding AQI with apps like Breathe, AirVisual etc.

    b. The second way could be to search out newspapers on the internet and try to find out images of places with a considerable portion of sky, and map it to the AQI of that place on that specific date. Sites such as https://aqicn.org/ can be used here for getting aqi data.

    c. The third way could be to create an app and distribute it in the IITR community and ask members to take pictures around themselves and fill the required AQI details in the app's form which finally will be stored in backend services like Firebase.

We need around 5000-6000 images for effectively training the model.

2. The second part after collecting the dataset is to train the model for predicting the Air Quality Index.
    This again involves three steps:
1. First we need to extract only the sky part from the images and use those images instead in later steps.
2. Then we need to use the images gathered in earlier steps for extracting useful information or parameters out of them.
3. Then we have to train a model which will use the information extracted earlier and map it to the Air Quality Index.

3. The third part is optional, where we can design a website or an app as a deliverable product where people can upload the images and get the AQI displayed.

Packages / Languages / Environment :

Development Environment:

1. VS Code
2. Android Studio
3. Ubuntu

Languages:

1. Dart
2. Python
3. HTML/CSS

Packages/Frameworks:

1. Tensorflow / Pytorch / Scikit-Learn Machine Learning Libraries.
2. Keras API for designing neural networks faster.
3. Flutter Framework for designing the App for collecting dataset.
4. Firebase API for storing data collected from the App.
5. Django / Flask Framework for developing the website as a final product.

Limitations:

Collecting the dataset is a limitation to the project because to make people in the IITR community involved in gathering data is uncertain and based on the load of work we already have from our curriculum this semester, it would be difficult for us to gather 5000 images.