

School of Engineering and Applied Science (SEAS), Ahmedabad University

B.Tech (CSE Semester VI)/M.Tech/PhD:
Machine Learning (CSE 523)

Project Abstract Submission #1

Submission Deadline: January 31, 2020 (11:59 PM)

- **Group No.:** *S_ECC6*
- **Project Area:** Environment and climate change
- **Project Title:** AQI (Air Quality Index) Prediction using ML techniques
- **Name of the group members :**
 1. Kalagee Anjaria (AU1741052)
 2. Nidhi Golakiya (AU1741062)
 3. Rajvee Kadchha (AU1741064)

Abstract

Air pollution is rising day by day which is a serious issue. In cities like Delhi , Ghaziabad, Noida it is a prime issue. AQI (Air Quality Index) is basically an index for measuring air pollution and various pollutants like PM_{2.5}, NO_2 , O_3 , CO etc [2]. PM_{2.5} is a very harmful pollutant, it is for all particles having radius less than 2.5. These particle do not settle down and stay longer in air which if inhaled is harmful. There are six levels for AQI. Low value of AQI means the air is safe to be inhaled. Pollution is increasing due to urbanisation, more people are living in urban areas which leads to more number of vehicles, more traffic and more harmful gasses released into the environment.

Predicting AQI has become almost a necessity in the highly polluted cities, so that people can take appropriate precautions regarding their healths. For that, accurate measurements of AQI is needed. There are many existing techniques for predicting AQI (Air Quality Index). But using an appropriate technique which gives us accurate results in less time is a must. In our base article, four regression techniques have been used for predicting AQI, which include:- Decision Tree regression, Random Forest regression, Multi-Layer Perceptron regression and Gradient Boosting regression. All these four techniques are compared to determine which gives lesser error. And for comparing the errors between actual and predicted data, the evaluation parameters used are:- Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) [1].

In our project, we are going to predict the AQI for an Indian city. We'll acquire the dataset from an official government website for datasets. Then, we'll clean the dataset, because it is not necessary that all the values are available for all the parameters given. For calculating the AQI, first we have to calculate the AQI for all the parameters taken, and the maximum AQI out of all these parameters is considered to be the AQI for that city. Then for predicting the future AQI of that city, we may use either any technique of linear regression or any of the four techniques mentioned in the base article. Thereafter, we'll apply dimensionality reduction on our dataset, to reduce the complexity by reducing the number of random variables. For this, we'll use the dimensionality reduction technique most suitable to our dataset.

References

- [1] Ameer, Saba Shah, Munam Khan, Abid Song, Houbing Maple, Carsten Islam, Saif Asghar, Muhammad. (2019). Comparative analysis of machine learning techniques for predicting air quality in smart cities. *IEEE Access. PP. 1-1*.
- [2] Srivastava, Chavi Singh, Shyamli Singh, Amit. (2018). Estimation of Air Pollution in Delhi Using Machine Learning Techniques. 304-309. 10.1109/GUCON.2018.8675022.