AIR QUALITY ANALYSIS

Abstract---

We forecast the air quality of India by using machine learning to predict the air quality index of a given area. Air quality index of India is a standard measure used to indicate the pollutant (s02, n02, rspm, spm. etc.) levels over a period. We developed a model to predict the air quality index based on historical data of previous years and predicting over a particular upcoming year as aGradient decent boosted multivariable regression problem. we improve the efficiency of the model by applying cost Estimation for our predictive Problem. Our model will be capable for successfully predicting the air quality index of a total county or any state or any bounded region provided with the historical data of pollutant concentration. In our model by implementing the proposed parameterreducing formulations, we achieved better performance than the standard regression models. our model has 96% accuracy on predicting the current available dataset on predicting the air quality index of whole India, also we use AHP MCDM technique to find of order of preference by similarity to ideal solution.

Keywords—AQI,dataset,preprocessing,

outliers, BVA,prediction

I.INTRODUCTION

As the largest growing industrial nation. India is producing record amount of pollutants specifically c02, pm2.5 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 (n02, s02, 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 give more information and knowledge about the cause and seniority of the pollutants.