Assignment

COURSE CODE: ART 203

Submitted to:

Rubayea Ferdows

Senior Lecturer

Department of Computer Science and Engineering

Submitted by:

Afsana Alam Trisha

ID: 19103104

Shima Akter

ID: 19103151

Section: C

Semester: Summer 2022

Date of Submission: 02 July, 2022

Air Pollution Prediction of Dhaka City Using Random Forest Method

Abstract

The aim of this paper is to use random forest method for air pollution prediction of Dhaka city from 2018 to 2021. The datasets of air pollutants like NO2, O3, SO2, CO, PM2.5 and PM10 are collected with daily average concentration. With these datasets the prediction will be calculated and plotted in graphs. There is a good reason to worry about the Air Quality in Dhaka, as Dhaka has ranked the 2nd most air polluted city in the world with the AQI of 164. The aim of this proposal is to use a heterogeneous ensemble of differential evolution with random forest method for air pollution prediction. This is different from existing work as a method is proposing to combine state-of-the-art differential evolution strategies with random forest method instead of focusing on existing single technique.

Introduction

In addition to water and land, the prime resource for sustenance of life is air. Fresh unpolluted air is the basic need of each and every living being. There are different pollutants which are hampering the life on the earth. Air pollution is one the major cause which is affecting life the most. Exposure to air pollution has been associated with morbidity and mortality. Variety of air pollutants are given out into the atmosphere by anthropogenic sources, out of which nitrogen dioxide, sulfur dioxide, ozone, particulate matters, carbon monoxide and are having the significant adverse impact on air quality.

An air pollutant can be produced by human activities or from some natural sources, that effect on human's health and the environment. Solid particles, gases or droplets of liquid can be substance of air pollutant. In this study, six major air pollutants of Dhaka city are predicted. The pollutants predicted in Dhaka are NO_2, CO, SO_2, O_3, PM2.5 and PM10. One pollutant is considered at one time as a target variable and its prediction is made accordingly. And same treatment is made with the rest of the pollutants.

Random forest classifier is the best learning technique. In machine learning and pattern recognition it is very popular and potent technique. The principle of random forests is to build binary sub-trees using the training bootstrap samples coming from the learning sample L and selecting randomly at each node a subset of X. The classification with highest votes among all the trees in the forest opts by decision forest.

Literature Review

In literature, a lot of work is done in the study and analysis of air pollution as well as predicting the future trends. But analyzing and predicting the air pollution of only Dhaka City is not researched yet. As it is not done yet, two main city of India, Delhi and Patna and also two main city of South Korea, Seoul and Incheon air quality papers were reviewed.

In "Assessment of gaseous and respirable suspended Particulate matter (PM10) emission estimates over megacity Delhi: past trends and future scenario (2000-2020) ", Rati Sindhwani has studied the past trends of PM10 emission in Delhi and made prediction about the future till 2020.

In "Development of artificial intelligence based NO2 forecasting models at Taj Mahal, Agra centre for atmospheric sciences", Pramila Goyal has used Artificial Neural Networks for predicting the air pollution at Taj Mahal in Agra.

In "Predicting air quality of Dhaka and Sylhet divisions in Bangladesh: a time series modeling approach", Md Mazharul Islam has used Seasonal Autoregressive Integrated Moving Average (SARIMA) model to forecast weekly air quality of Dhaka and Sylhet divisions.

In "Evolving Differential evolution method with random forest for prediction of Air Pollution", Rubal and dinesh Kumar has worked with differential evolution method with random forest.

In "Vehicular emission inventory of criteria pollutants in Delhi", authors have introduced new grid based mobile source emission inventory using IVE model recognized that vehicular pollution is a main source.

In "Understanding urban vehicular pollution problem vis–a–vis ambient air quality––case study of a megacity (Delhi, India)", authors have studied about growing problem in megacities and they recognized that it all due to vehicular traffic they have also stated the benefits of policies given by government to curb pollution.

In "Forecasting of air quality in Delhi using principal component regression technique", authors have used principal component regression to predict air quality index(AQI). In their study AQI is estimated for 2000-2006 at ITO.

In "Traffic induced emission estimates and trends (2000–2005) in megacity Delhi", authors have developed transport sector emission inventory for Delhi from the period of 2000-2005 to measure vehicular emissions and drew out the

consequences of relevant policy reforms on overall emissions of various air pollutants like CO_2, NO_x, HC, TSP and SO_2 over the years to help in making change for future policies.

In "Long-Term (2003–2019) Air Quality, Climate Variables, and Human Health Consequences in Dhaka, Bangladesh", the authors used Four factors were characterized as estimated sources from PMF modeling.

Author's Name	Paper's Name	Objectives	Methodology	Result	Limitations
Md Riad Sarkar Pavel, Shahid Uz Zaman, Farah Jeba, Md Safiqul Islam and Abdus Salam	Long-Term (2003–2019) Air Quality, Climate Variables, and Human Health Consequences in Dhaka, Bangladesh	-Long-term trends of criteria air pollutants and climate variables were analyzed. -Significant positive trends were observed for the pollutants except ozone.	positive matrix factorization (PMF)	The air pollutant ($PM_{2.5}$, PM_{10}, CO, NO_2, and SO_2) concentrations except O_3 showed significantly increasing trends in this study.	O_3 effects were not analyzed and the impact were not found.
Rubal, Dinesh Kumar	Evolving Differential evolution method with random forest for prediction of Air Pollution	The aim of this paper is to use a heterogeneous ensemble of differential evolution with random forest method for air pollution prediction	Differential evolution method with random forest	The prediction of concentration values of different pollutants is done.	The work with environmental modelling is not done. The AQI is not shown.
Md Mazharul	Predicting air quality of	Air quality in Bangladesh is	Seasonal Autoregressive	Forecast weekly air	Did not show the future of

Islam, Mowshumi Sharmin, Faroque Ahmed	Dhaka and Sylhet divisions in Bangladesh: a time series modeling approach	measured on daily basis using national standards of five pollutants—NO2, CO, Ozone (O3), SO2, and particulate matter (PM2.5 and PM10) and presented as indexed value ranged from 0 to 500 which is further classified from good to extremely unhealthy.	Integrated Moving Average (SARIMA)	quality of Dhaka and Sylhet divisions in Bangladesh.	The graph plot of AQI is not shown. Air pollution of these two cities and how the pollution hampers lives.
Samaher Al-Janabi, Mustafa Mohammad & Ali Al-Sultan	A new method for prediction of air pollution based on intelligent computation	Design an intelligent predictor for the concentrations of air pollutants over the next 2 days.	Deep learning techniques using a recurrent neural network (RNN)	evaluation results are obtained by reading the concentration of each pollutant each hour for at most 30 days then taking the average of the symmetric mean absolute percentage error (SMAPE) for 25 days only.	Accuracy is less.