A Project ETE Review Report

on

Air Pollution Prediction System for Smart City using Data Mining Technique

Submitted in partial fulfillment of the
requirement for the award of the degree of

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE & ENGINEERING

Under The Supervision of

Dr. N. PARTHEEBAN
Project Guide

Submitted By

Prabhat Gupta 19SCSE1010783
SHIVAM 19SCSE1010686

SCHOOL OF COMPUTING SCIENCE AND ENGINEERING DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
GALGOTIAS UNIVERSITY, GREATER NOIDA, INDIA,December 2021

SCHOOL OF COMPUTING SCIENCE AND ENGINEERING
GALGOTIAS UNIVERSITY, GREATER NOIDA

CANDIDATE'S DECLARATION

We hereby certify that the work which is being presented in the project, entitled "Air Pollution Prediction System for Smart City using Data Mining Technique" in partial fulfillment of the requirements for the award of the B.Tech submitted in the School of Computing Science and Engineering of Galgotias University, Greater Noida, is an original work carried out during the period of September,2021 to December, 2021, under the supervision of Dr.N.PARTHEEBAN, Guide Department of Computer Science and Engineering/Computer Application and Information and Science, of School of Computing Science and Engineering , Galgotias University, Greater Noida

The matter presented in project has not been submitted by us for the award of any other degree of this or any other places.

Prabhat Gupta 19SCSE1010783

Shivam 19SCSE1010686

This is to certify that the above statement made by the candidates is correct to the best of my knowledge.

Dr.N.PARTHEEBAN

CERTIFICATE

The Final Thesis/Project/ Dissertation Viva-Voce examination Prabhat Gupta 19SCSE1010783 and Shivam 19SCSE1010686 has been held on Air Pollution Prediction System for Smart City using Data Mining Technique and there work is recommended for the award of B.Tech.

Signature of Examiner(s) **Signature of Supervisor(s)**

Signature of Project Coordinator **Signature of Dean**

Date: 24,December, 2021
Place: Greater Noida

ABSTRACT

Air pollution is one of the major hazards among the environmental pollution. As each living organism needs fresh and good quality air for every second. None of the living things can survive without such air. But because of automobiles, agricultural activities, factories and industries, mining activities, burning of fossil fuels our air is getting polluted. These activities spread sulphur dioxide, nitrogen dioxide, carbon monoxide, particulate matter pollutants in our air which is harmful for all living organism. The air we breathe every moment causes several health issues. So we need a good system that predicts such pollutions and is helpful in better environment. It leads us to look for advance techniques for predicting the air pollution. So here we are predicting air pollution for our smart city using data mining technique. In our model we are using multivariate multistep Time Series data mining technique using random forest algorithm. Our system takes past and current data and applies them to our model to predict air pollution. This model reduces the complexity and improves the effectiveness and practicability and can provide more reliable and accurate decision for environmental protection departments for smart city.

Table of Contents

Title **PageNo.**

Candidates Declaration

Acknowledgement

Abstract

List of Table

List of Figures

Acronyms

Chapter 1. Introduction

1.1 REQUIREMENT OF HALL TICKET

1.2 DISADVANTAGE OF CURRENT SYSTEM

1.3 MERITS OF PROPOSED SYSTEM

Chapter 2 Literature Survey/Project Design

Chapter 3 Functionality/Working of Project

Chapter 4 Results and Discussion

Chapter 5 Conclusion and Future Scope

5.1 Conclusio

5.2 Future Scope

Reference

Publication/Copyright/Product 45

Acronyms

B.Tech.	Bachelor of Technology
M.Tech.	Master of Technology
BCA	Bachelor of Computer Applications
MCA	Master of Computer Applications
B.Sc. (CS)	Bachelor of Science in Computer Science
M.Sc. (CS)	Master of Science in Computer Science
SCSE	School of Computing Science and Engineering

List of Figure

Figure No.	Name Table	Page Number
1.1	World's Most Polluted cities	8
2.1	Knowledge discovery process using data mining	14
2.2	Basic Data Mining Algorithms	18

CHAPTER 1 INTRODUCTION

Air pollution arises when injurious substances, along with particulates and biological molecules, are imported into Earth's atmosphere. It may cause disorder, allergies or death of mankind, it may also cause harm to alternative living organisms such as creatures and food product and may damage the natural or assembled environment. Human activity and natural processes can both generate toxic waste that makes air infection.

A discuss toxin is a fabric in the discussion that can cause opposite impacts on mankind and the surrounding. The meat can be a solid atom, liquid droplets, or Peach States. A toxic waste can be of a usual ancestor or artificial. Breeze impurity is categorized as essential or auxiliary. Essential poisons are commonly issued beginning of a mechanic, specific because dust as of an impulsive explosion, particulate matter, PM2.5 from the climatic experimental condition, carbon paper monoxide (CO) gases starting engine automobile tire out fumes or the sulphur dioxide (SO2) liberated beginning manufacturing plants. Auxiliary poisons are neither breathed straightforwardly or maybe, they shape in the discussion when essential poisons perform or act together. The earth stage breath is an outstanding case of an auxiliary toxin. Outstanding of the boom for mechanical, populace and frequency of activity is developing quickly, individuals are confronting a think condition of for the discussing contamination. In a few nations, their relentless discuss contamination in the metropolis, not as it aggressive the wellbeing of inhabitants other than the plant moreover the climate. Deepest locales, Particulate Matter 2.5 some absorption really on meteorological conditions, implying that weather alter might contain noteworthy impacts on Particulate Matter 2.5. All things considered, so distant small prove to be able to finish respecting the relationships of the PM2.5 towards the weather varies

The ambiguous relationship turns around the troubles too in section the complication for the dependency of a different ingredient of a climatic factor, and to some extent, the active data on the effects of different components (e.g., person variables) about for Particulate Matter2.5 discuss value issue become more over the feasible to the investigation. Recently the analysts of the National Aeronautics and Space Administration addressed tended to that because of the air contamination, the consolidation of the PM2.5 and dihydrogen monoxide steam lying on the Western Pacific compose to create a solid tropical tornado the that loaded with dihydrogen monoxide steam, much all the more, developing the quality of tempests. In this

manner, air contamination reaches towards the one of the fundamental senses that tempests wound up more grounded than previous in North America in current years.

In 2017, 1.2 million people died in India due to pollution, which is 12.5% of total deaths with in our country. Also almost 2 lakh children lost their lives because of air pollution related diseases, which suggests on an average 535 deaths occurred daily. The Union Health Ministry and Indian Council of Medical Research says in our country one out of each eight deaths is attributed to pollution.

Fig1.1:-World's Most Polluted cities

As Figure Shows

World’s most polluted cities are in India. So pollution is one most important hazards among all environmental pollution. As each living organism must needs air for each second. So Air should be fresh and of good quality.None of the living things can survive without such air. These are some reason for our air is getting polluted.

These are some reason for our air is getting polluted:-

1. Combustion of fossil fuels, like coal and oil for electricity and road transport,producing air pollutants like nitrogen and gas.

2. Emissions from industries and factories, releasing great amount of monoxide, hydrocarbon, chemicals and organic compounds into the air.

3. Agricultural activities, thanks to the utilization of pesticides, insecticides, and fertilizers that emit harmful chemicals.

4.Waste production,mostly because of methane generation in landfills.

These activities spread pollutants in our air which is harmful for all of us. Information about those pollutants are given below.

CHAPTER 2 LITRATURE SURVEY

1. Particulate Matter (PM2.5 and PM10):

Living in a city, you've likely walked outside your apartment and noticed a layer of grey haze that prevents you from clearly seeing the landscape miles ahead. That haze appears when there are high concentrations of particulate matter (PM) in the air.PM is a mixture of solids or liquid droplets in the air that are categorized by size:
PM10: Inhalable particles that are less than or equal to 10 micrometers in diameter. Examples include dust, pollen, and mold.

PM2.5: Fine particles that are less than or equal to 2.5 micrometers in diameter. To putthis in perspective,they are about 1/30th of a strand of human hair (too small for the human eye to see).

2. Nitrogen Dioxide (NO2):

Another dangerous form of urban pollution may be a group of gases called nitrogen oxides. because of burning of fuel from road vehicles, cookers and heaters gas is produced during warmth. They react within the air to make particulate (PM) and ozone.

3. Sulphur Dioxide (SO2):

Sulfur dioxide may be a gas. It invisible and incorporates a nasty, sharp smell.About 99%of the dioxide in air comes from human sources. The most source of sulphur dioxide is industrial activity that processes materials that contain sulfur, example. The generation of electricity from gas, oil or coal that contains sulfur. SO2 affects human health. It irritates the nose, throat, and airways to cause coughing, wheezing, shortness of breath, or a good feeling round the chest.

4. Carbon Monoxide (CO):

Carbon monoxide is colorless and odorless, but highly toxic. While often thought of as an inside hazard, it is also a significant outdoor pollution further. Main sources of CO to outdoor air are cars, trucks and other vehicles or machinery that burn fossil fuels.

5.Ozone (O3):

Ground-level ozone (also known as the “bad” ozone) is created by a chemical reaction in the presence of sunlight that forms between man-made VOCs and nitrogen oxides This explains why ozone levels tend to be higher and subsequently more dangerous within the summertime. In rural areas, downwind of urban areas or industrial sites, the highest levels of ozone are mostly found.

Today many health issues are happening due to the air. Our air become more polluted day by day, which is extremelybad for all the living organism. Not only these pollutants but also there are some more parameters which affect the air pollution. Some parameters which affect the air pollution which is explained below.

1. Temperature

   Temperature inversion have a great effect on air quality. Increase of height in atmosphere makes air cooler. However sometimes an upper air layer is warmer than lower air layer. That’s call inversion in atmosphere. And if inversionstays for long time on the ground than pollutants can build up to the higher level. Thatis how temperature affectthe air pollution.

2. Wind direction and speed

   Measurement of wind speed and wind direction is vital in air quality monitoring. It can help identify the situation of the source of the pollution, and also provide a higher overall picture of what’s happening within the air.

3. Air quality of previous day

Previous days quality of air must affect the following day.so here we are using the pollution level of the previous day. So we are able to get the accurate result. So these are the essential of the air pollutants and a few factors that are affect the

pollution. Using of these attribute here we are predicting the pollution. Also now we see the motivation and objective about our project.

DATA MINING FOR PREDICTION

Data mining is a process which is used to turn raw data into useful information. Data mining is the process of gathering information and analyzing it for actionable patterns, which may then be employed in many things. There are many uses of data mining technique.Oneusageofdata mining technique is for prediction. At the end we extract useful information from the data.

Here is basic flow diagram of knowledge discovery from data.

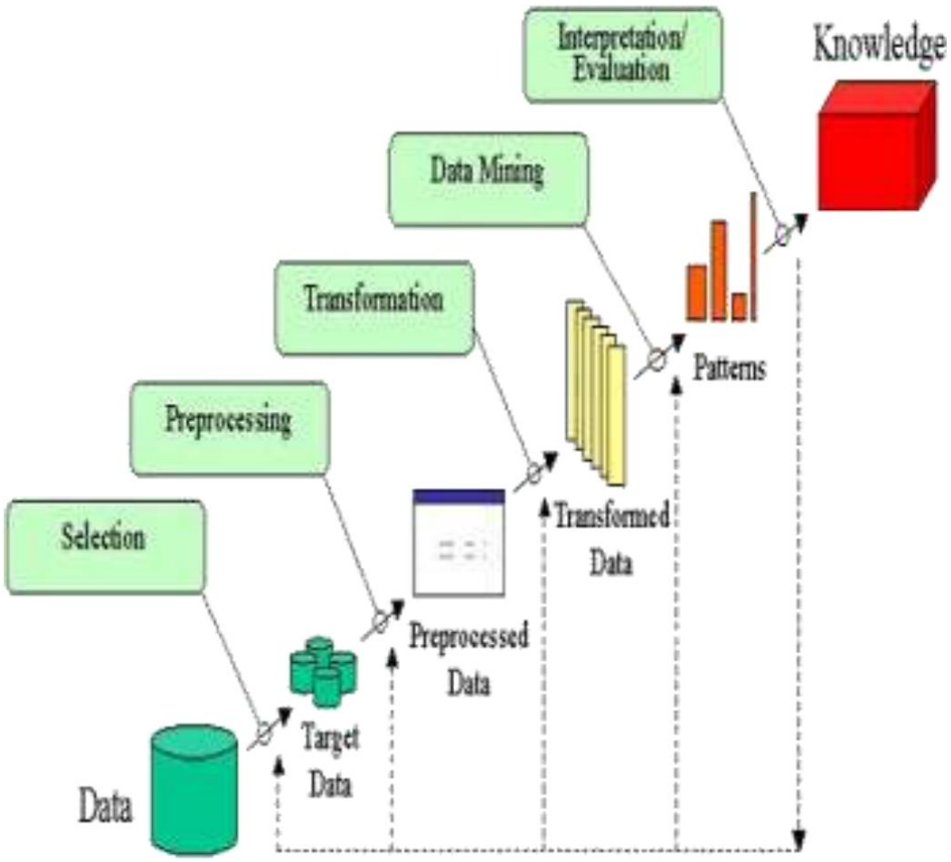

Fig2.1-: knowledge discovery process using data mining

Steps of KDD process:

- Data Cleaning - In this step the noise and inconsistent data is removed.
- Data Integration - In this step multiple data sources are combined.
- Data Selection - In this step relevant to the analysis task are retrieved from the database.
- Data Transformation - In this step data are transformed or consolidated into forms appropriate for mining by performing summary or aggregation operations.
- Data Mining - In this step intelligent methods are applied in order to extract data patterns.
- Pattern Evaluation - In this step, data patterns are evaluated.
- Knowledge Presentation - In this step, knowledge is represented.
- Here are the main types of Data mining algorithms.

Flowchart:-

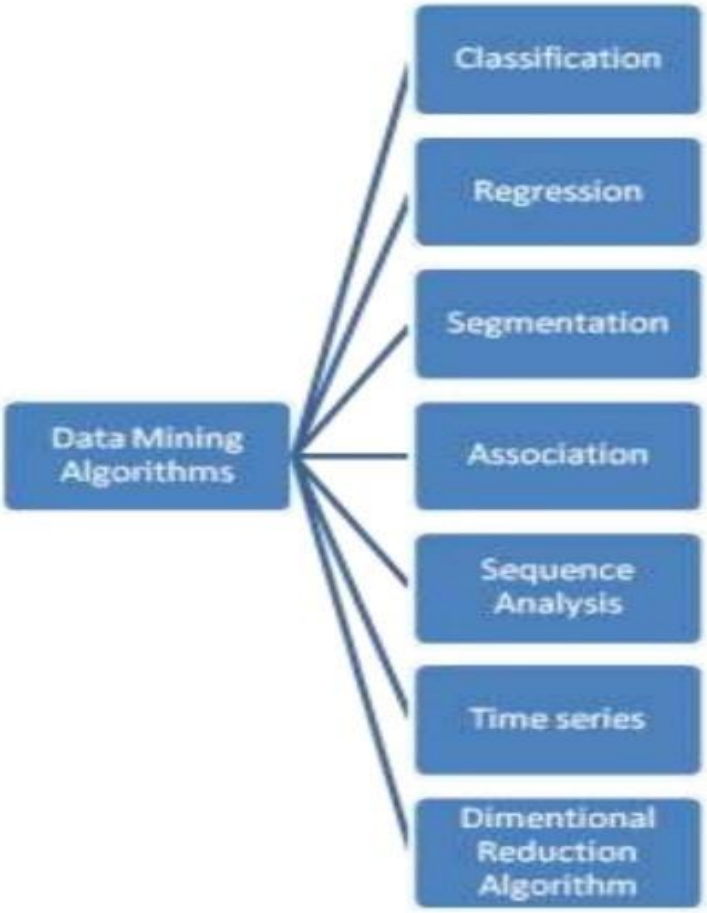

Fig2.2-: Basic Data Mining Algorithms

CHAPTER 3 CONCLUSION

Our system helps in improving the prediction of air pollution in our smart city. Multivariate Multistep Time Series Prediction Using Random For esttechnique improve the performance and reduce the complexity of the air pollution prediction model. Also he rewe are using feature selection technique, which make our prediction even better.

Here we are predicting PM2.5, PM10, SO2,NO2,Ozone, Co pollutants. Also we shown prediction in a bar chart. We give color to the bars, which indicates the quality index of the pollutants. That the pollutant level is good, moderate or bad.

Using this System we are trying tospreadawareness ofthe air pollution in these smart cities, so that everyone canlive healthy and happy life. Also together we will make our India clean and pollution free.

CHAPTER 4 REFERENCE

Harishkumar. K. S , Doreswamy. H "A Survey of Air Pollution Studies Using Data Mining Techniques In Smart City"Conference Paper · May 2018

"Air Pollution Prediction System for Smart city using Data Mining Technique" International Research Journal of Engineering and Technology (IRJET)
Volume: 07 Issue: 06 | June 2020

K. Rajakumari, V. Priyanka "Air Pollution Prediction in Smart Cities by using Machine Learning Techniques" International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075, Volume-9 Issue-5, March 2020

Shweta Taneja, Dr. Nidhi Sharma, Kettun Oberoi, Yash Navoria“Predicting Trends in Air Pollution in Delhi using Data Mining”, IEEE(2016)

Manisha Bisht and K.R. Seeja, “Air Pollution Prediction Using Extreme Learning Machine: A Case Study on Delhi.”, Springer(2018)