iJETRM
International Journal of Engineering Technology Research & Management

AGENT BASED FRAMEWORK FOR THE ANALYSIS OF AIR POLLUTANT LEVEL USING DATA MINING TECHNIQUES IN CHENNAI CITY

Dr. K Shyamala [*1], G.Fathima [2]

[*1] Associate Professor, PG & Research Department of Computer Science, Dr. Ambedkar Government Arts College, Chennai, Tamilnadu, India

[2] Research Scholar, PG & Research Department of Computer Science, Dr. Ambedkar Government Arts College, Chennai, Tamilnadu, India

ABSTRACT

An air pollutant is a material in the air that can have adverse effects on human's beings and the ecosystem. The substance can be solid particles, liquid droplets, or gases. Chennai is one of the fast-progressing metropolitan cities which are battling against air pollution. It suffers from many issues like Vehicle exhaust fumes, Fossil Fuel-Based power plants, exhaust from Industrial plants and factories, construction and agricultural activities, natural causes and household activities. In this paper Manali, Alandur and Velachery have been selected for air pollution analysis in Chennai city and it explores Particular Matter ($PM_{2.5}$), Sulpher diOxide (SO_2), Ozone (O_3), Carnon Monoxide (CO) and Nitrogen diOxide (NO_2) pollutant levels and compare these pollution levels in Chennai city by using Bubble sort Algorithm and categorize the value by using Categorize agent and Finally sending alerts to the Pollution Control Board when pollution level raises. Recent Reports shown that Air Quality in Chennai worsens almost equal to Delhi, so it is more important to analyze the pollutant level in developing cities like Chennai. To attain these objectives, concepts of Data mining and Agent Based Framework is used. It helps the Government to take initiatives in making policies and control air pollution.

Keywords: Air pollution, Data mining, Chennai city, Pollutant, $PM_{2.5}$, SO_2, NO_2, CO, O_3, Agent based framework, Bubble sort Algorithm

INTRODUCTION

Air pollution refers to the release of harmful substances in the air that cause danger to human beings Air pollution is one of the major environmental problems especially in Chennai because of the growth of urbanization, industrialization, and the growth of the population. According to many research study, it is found that, over one lakh deaths in 29 Indian cities may be attributed to the rising PM 2.5 levels [1].

Data mining is playing an important role in air pollution. The main aim of Data mining is used to extract useful patterns from a large set of databases. In this paper, Particular Matter ($PM_{2.5}$), Sulpher diOxide (SO_2), Ozone (O_3), Carnon Monoxide (CO) and Nitrogen diOxide (NO_2) day wise pollutant levels are collected from the certain region like Manali, Alandur and Velachery in Chennai city. The data has been collected from the website of Chennai Real Time Air Quality Index (http://aqicn.org/city/chennai//) [2]. The pollutant components emission level for the month January to August, 2020 is considered for this analysis.

LITERATURE REVIEW

Mihaela Oprea, Mădălina Cărbureanu, Elia Georgiana Dragomir [3] has described the collaborative intelligence in agents and it uses the C45 algorithm to extract the patterns from the large datasets. It has been implemented in Zeus and See5 Data mining. The AirQMAS system (Collaborative Multi-Agent System for Air Quality Analysis) will separate the air quality levels into strong pollutants and weak pollutants and send reports to the supervisor agent. It provides the air quality index and general report. The air pollutants that were analyzed are CO, SO_2, PM_{10}, H_2S, NO, NO_2, O_3 from 6 stations and it also uses the meteorological stations

iJETRM
International Journal of Engineering Technology Research & Management

Sabri Ghazi, Julie Dugdale, and Tarek Khadir [4] have described the human decision in the impacts of air pollution. Most of the paper describes the ecological and social parameters, but it introduces the human decision making. It used the two-stage air pollution modeling method: a GPD dispersion model and an ANN forecasting model. Cooperation strategy is also adopted in managing Air quality. The reward/penalty strategies also are used for effective air quality analysis. The main aim of the paper is the software agents are playing N-Prisoner Dilemma game in which they have two decisions either to increase or decrease the emission. It also gives a brief explanation of the cooperation strategies. The simulation used PM_{10}, NOx, and SO_X data collected hourly from the city of Annaba and it also included climatic parameters like wind speed, humidity, and temperature. It has used two years of data for processing.2003 dataset is used for training the ANN and 2004 datasets are used for testing.

Fazziki, Abdelaziz & Benslimane, Djamal & Sadiq, Abderrahmane & Ouarzazi, Jamal & Sadgal, Mohammed [5] have described the multi-agent framework for on-road air quality monitoring and to generate recommendations and regulations for the road users it has used the Artificial Neural Network Model for the prediction of air quality and used the Dijkstra algorithm for finding the least polluted path in the road network. And it has been implemented using a Hadoop based framework and Prometheus Development tool for the modeling of the agent-based framework. The pollutant parameters such as PM_{10}, NOx, CO, SO_X, and Ozone data are collected from the sensor in the city of Marrakech. And it also used meteorological parameters like temperature, Wind speed direction, precipitation, and pressure and used GPS(Global Positioning System) for storing the current location and temperature of the Vehicles.

G.Fathima and Dr. K Shyamala [6] have described the analysis of air pollutant level in Chennai city using the Data mining technique.. The main air pollutants like Nitrogen diOxide, Sulpher diOxide, and Particulate Matter are collected from Tamilnadu Pollution Control Board website. The five different regions like Anna Nagar, Kilpauk, Adyar, Nungambakkam, and T Nagar in Chennai city are selected for the observations. It has categorized the values into six scale values (i.e.,) Good, Moderate, Very Good, Excellent, Medium, Bad, and very bad. Finally, it states the higher region in the Chennai city. The system is implement using the Data mining tool called WEKA. From the observation, it is found that Annan agar is polluted when compared to all the five region in Chennai city and second more polluted region is Nungambakkam
.
Elia Georgiana Dragomir. Mihaela Oprea [7] have developed a Multi-agent system for monitoring urban air pollution due to power plants operating in the Ploiesti town area. It uses the Artificial intelligence and Distributed simulation. Based on the real-time values, the air quality index is calculated. The main air pollutants like Carbon Monoxide, Nitrogen Monoxide, Nitrogen diOxide, Sulpher diOxide and particulate Matter are collected in three key places in Ploiesti city. This is mainly used for the decision factors, when the pollution level increases. It has been implemented in Zeus

Ioannis N. Athanasiadis and Pericles A. Mitkas [8] have explained an intelligent environmental monitoring system, that uses software agents to analyze the air pollution from different sensors and used Data mining to add extra intelligence into the system and finally give an alarm when the pollution level exceed the limit level to appropriate users. The C4.5 algorithm is used for validating incoming measurements, estimating the erroneous ones and the system was implemented using WEKA, JADE and protege 2000 OntologyEditor.
Sabri Ghazi, Julie Dugdale, and Tarek Khadir [9] have described the multi-agent-based simulation approach and it was integrated with dispersion and prediction models to simulate an air pollution crisis. Including uncontrolled sources of pollutants is a key element for making the simulation more realistic. The simulation used PM_{10}, NOx and SO_X data collected hourly from the city of Annaba and it also included climatic parameters like wind speed, humidity, and temperature. It also uses Cooperation strategies to reduce the effect of pollutant concentration.

iJETRM
International Journal of Engineering Technology Research & Management

METHODS AND MATERIALS

As mentioned in the Literature review Agent based Framework helps in the process of decision making and Data mining is used to discover patterns from the large dataset. In the proposed framework, a Preprocessor agent is used to find out the erroneous data and Simple-K Means Clustering is used to detect patterns from the datasets and Report agent is used to convert the days into month for better prediction of results and used for decision making. Then, Sorting Agent is used to find out which parameter is highest among the different parameter by using Bubble sorting algorithm. Categorize Agent is a way of showing changes in the amount of pollution in the air. Finally, an Alert agent is used to inform Higher officials of Pollution Board of Manali, Alandur and Velachery when pollution level rises. The Proposed Agent based framework using Data mining process is shown in Figure 1.

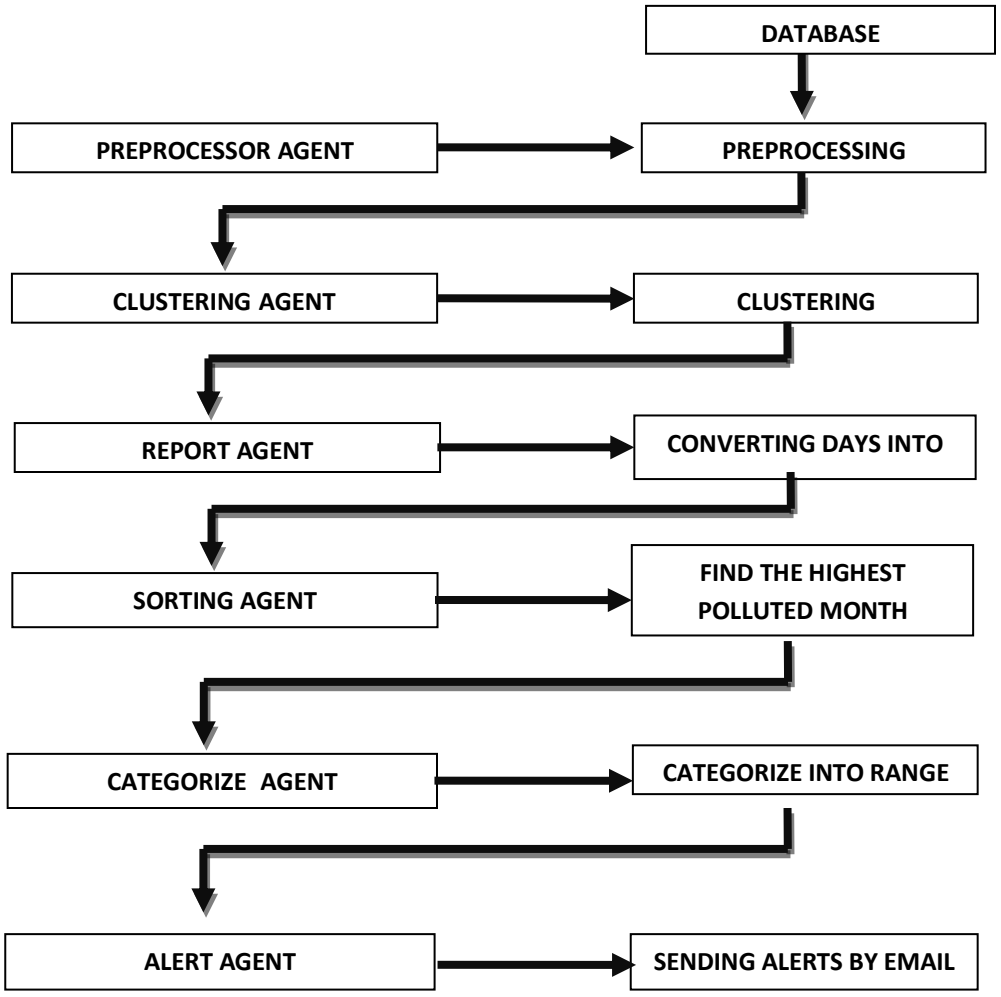

Figure 1 Proposed Agent based Framework

iJETRM
International Journal of Engineering Technology Research & Management

A. PREPROCESSOR AGENT

Data cleaning and transformation are methods used to remove outliers and standardize the data so that they take a form that can be easily used to create a model. The database sometimes contains unwanted data, and it leads to wrong analysis, so it is used to remove the unwanted data. And it is important to add the missing values. In the preprocessor, all the data will be preprocessed and fill the missing values by taking the mean ,mode or median [10].

B. CLUSTERING AGENT

Clustering is the process of partitioning the data (objects) into the same class. The data in one class is more similar to each other than to those in other clusters. The process of partitioning data objects into subclasses is called clusters [11] Simple K- means clustering algorithm is used to find groups in the data, with the number of groups represented by the k variable. K means is an iterative clustering algorithm that aims to find local maxima in each iteration.

C. REPORT AGENT

A report is a document that presents information in an organized format for a specific purpose [12]. A report is the formatted result of database queries and contains useful data for decision-making and analysis. It is often used form making recommendations for future action. It is used to convert the days into month for better prediction of results.

D.SORTING AGENT

Sorting is any process of arranging items systematically with distinct meaning and arranging items in a sequence ordered by using some criterion [13]. It is used to find out which parameter is highest among the different parameter. Bubble sort algorithm is used for sorting the values to state the highest polluted region in the Chennai city among the 3 different regions. Bubble sorting algorithm is comparison based algorithm in which each pair of adjacent elements is compared and the elements are swapped if they are not in ordered [14].

E.CATEGORIZE AGENT

Categorizing air pollution level is used to communicate to the public how much air surround is polluted [15]. Instead of prediction in numerical scales, Categorizing in Air Quality Index levels help the people to take care about the surroundings and maintain their health. Categorize is a way of showing changes in the amount of pollution in the air. Categorize Agent is used to categorizing the pollutant values as Very-low, Low, Moderate, High, Very-high, and Hazardous.

F.ALERT AGENT

An Alert agent is used to inform Higher officials of Pollution Control Board, when pollution level rises. So that it helps the people to know about pollution levels in the air and advises vulnerable groups such as the elderly, children, and people with lung disease to take precautions.

RESULTS AND DISCUSSION

Comparison results of $PM_{2.5}$, NO_2, SO_2, CO and O_3 in the following three areas like Manali, Alandur and Velachery in Chennai city has been displayed in the Figure 2, Figure 3, Figure 4, Figure 5 and Figure 6 respectively.

iJETRM
International Journal of Engineering Technology Research & Management

0-1500 µg/m3 = VERYLOW 1501-2000 µg/m3 = LOW 2001-2500 µg/m=MODERATE

2501-3000 µg/m3 = VERYHIGH ABOVE 3000 µg/m3 = HAZARDOUS

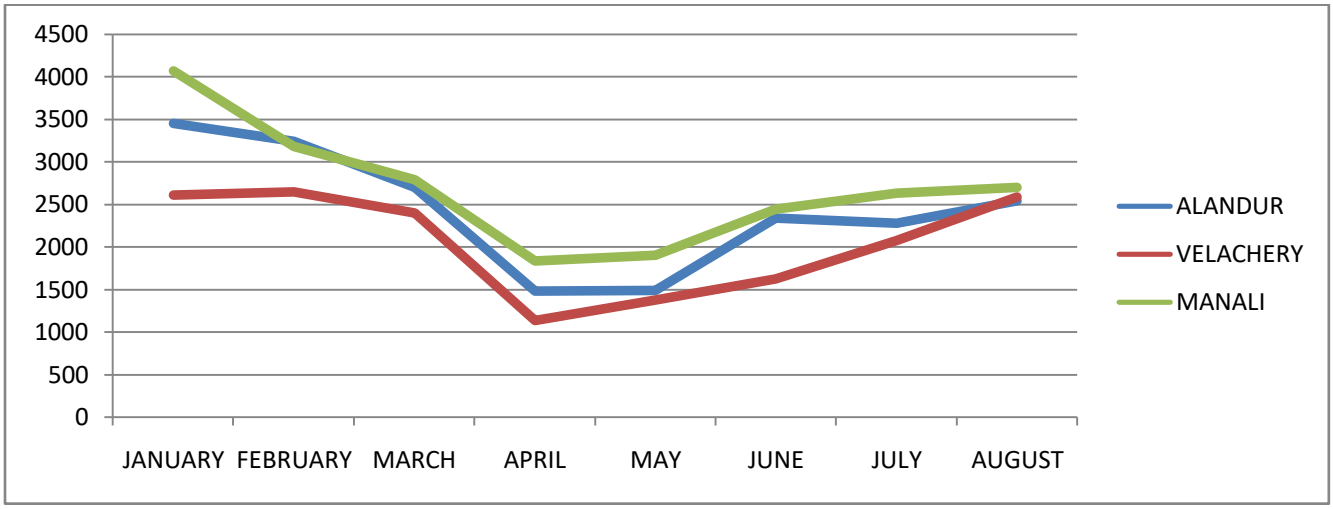

Figure 2 $PM_{2.5}$ Comparison levels in Manali, Alandur and Velachery areas in Chennai city.

0-100 µg/m3 = VERYLOW 101-150 µg/m3 = LOW 151-200 µg/m3=MODERATE

201-250 µg/m3 = VERYHIGH ABOVE 300 µg/m3 =HAZARDOUS

\\

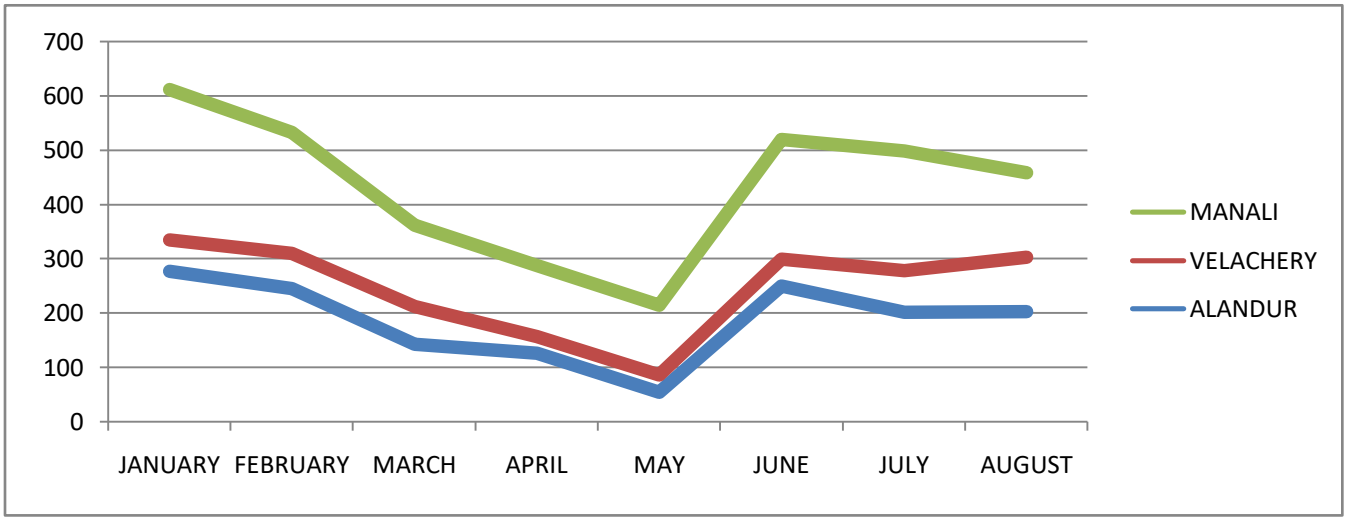

Figure 3 NO_2 Comparison levels in Manali, Alandur and Velachery areas in Chennai city.

iJETRM
International Journal of Engineering Technology Research & Management

0-100 µg/m3 = VERYLOW 101-150 µg/m3 = LOW 151-200 µg/m3= MODERATE

201-250 µg/m3 = VERYHIGH ABOVE 300 µg/m3 = HAZARDOUS

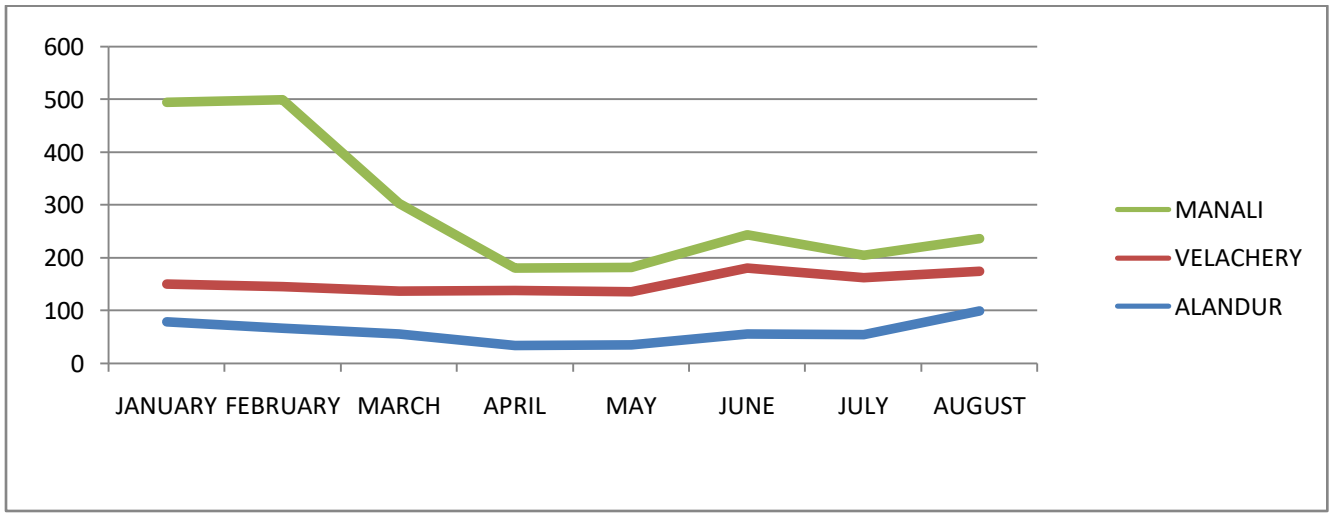

Figure 4 SO_2 **Comparison levels in Manali, Alandur and Velachery areas in Chennai city.**

0-100 µg/m3 = VERYLOW 101-150 µg/m3 = LOW 151-200 µg/m3 =MODERATE

201-250 µg/m3 = VERYHIGH ABOVE 300 µg/m3 = HAZARDOUS

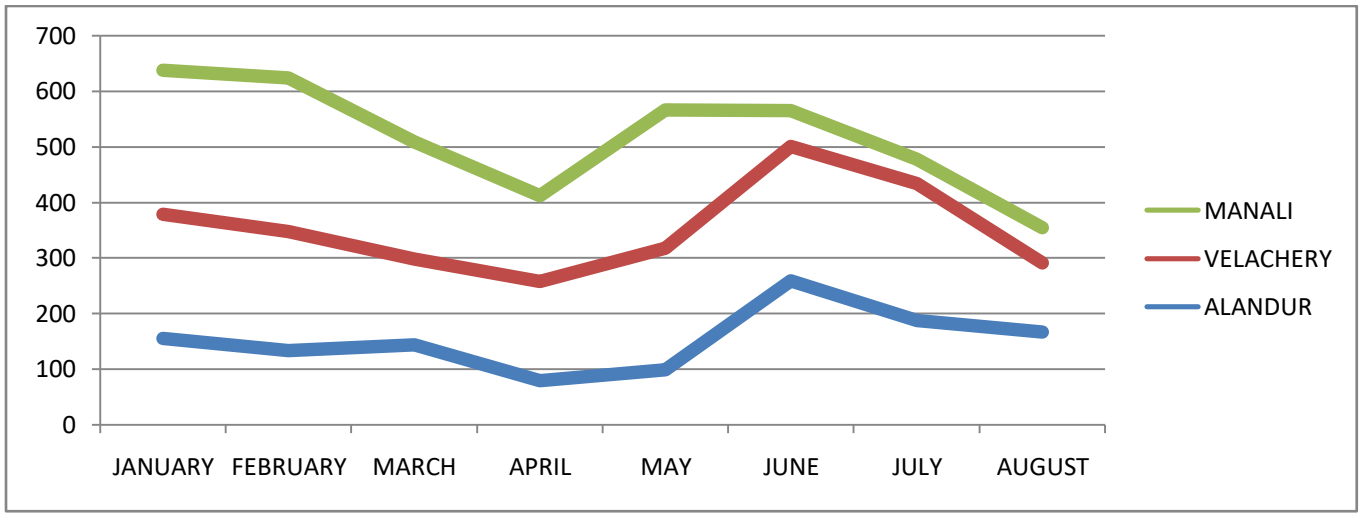

Figure 5 CO **Comparison levels in Manali, Alandur and Velachery areas in Chennai city.**

iJETRM
International Journal of Engineering Technology Research & Management

0-300 µg/m3 = VERYLOW 301-400 µg/m3 = LOW 401-500 µg/m3 = MODERATE

501-600 µg/m3 = VERYHIGH ABOVE 600 µg/m3 = HAZARDOUS

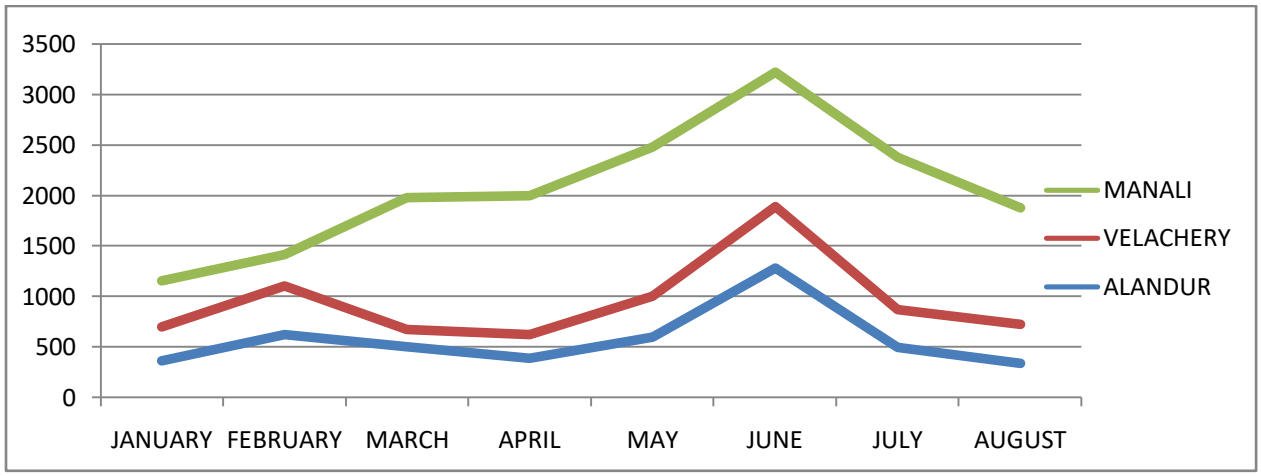

Figure 6 O_3 **Comparison levels in Manali, Alandur and Velachery areas in Chennai city.**

1. AIR POLLUTION IN MANALI

Manali is an industrial area and it is located in the North of Chennai city. It has the main hub of many manufacturing industries like Petroleum industries, Fertilizer industries, Foam industries, Rubber Factory, Chemical industries and Pesticides industries [16]. From the Observation, it is found that Manali is more polluted when compared to all the other regions in Chennai city. It is more pollutant during January because of fire accident on the Chennai Petroleum Cooperation Limited (CPCL) [17]. A health camp conducted in October 2017 by Huma Lung Foundation and Community Environment Monitoring among 200 participants residing close to the Vallur Thermal Power plant of Kuruvimedu in Ennore found five common afflictions: dry cough, body pain, loss of appetite, breathlessness and skin allergies [18]. And the second polluted month is February because, there was a massive fire broke out in the Chemical industry in Madavaram, it causes high pollutants of $PM_{2.5}$. As per the "Time of India" newspaper its stated that North Chennai is clock-in pollution levels equivalent to Delhi.

2.AIR POLLUTION IN VELACHERY

Velachery is a residential area and it is the central hub to many IT companies [19].At the distance of 5 kilometer it has Chennai International Airport**.** It is the shorter way to reach the Observation, it is third major polluted region in Chennai. During March, it makes many IT people to work on home and even make many people to lost their jobs due to COVID-19.It made cancellation of Visa to many people, who is working in many foreign countries.. Air quality index of Central Pollution Control Board's real-time monitors shows that ozone was the prominent pollutant in Velachery and it is more prominent during the Sundays and the Ozone at the ground level is high at the Evening [20].

iJETRM
International Journal of Engineering Technology Research & Management

3. AIR POLLUTION IN ALANDUR

Alandur is an urban region and it is located in the Guindy division. It has a landmark of Katipara junction. The pollution as been highest during January and it has been gradually decreased because of Covid19 spread. Pollution is more prominent during February, Because of the hot winds. Many Migrant people working in the industries of Alandur region forced to travel on foot to their own land due to Lockdown during Covid19 [21] . It has many welding stations, manufacturing factories are the major reason for the release of NO_2 and it has many damaged roads which makes the travel uneasy for people. And it is more polluted because of vehicular emission. If the pollutant becomes high it results in worsened cough and wheezing and leads to many health effects like asthma and bronchitis.

CONCLUSION

Due to the rapid urbanization and industrialization, it makes more pollutants in Chennai particularly in Alandur ,Velachery and Manali. It is found that the pollution rapidly decreases in the month of March due to Covid -19 Lockdown, and gradually increases in the following month due to Lockdown relaxation. It mainly helps in the reduction of air pollutant levels and it paves the way for many animals and birds wandering on our doorsteps. According to the survey of Pollution Control Board, it is found that up to the third lockdown. the pollution level decrease from 61.5 percentage to 29.5 percentage and after relaxation it has been increased up to 34.3 percentage. It also helps the Government to take initiatives to reduce the pollutant level and to keep it under control. From the observation, it is found that Manali is more polluted when compared to other region in Chennai city.

FUTURE ENHANCEMENT

This study is further extended by using real time sensors for prediction of air quality values. This work is also extended to other region in Chennai, so we can discover alternative ways for a cleaner and greener environment.

REFERENCES

[1] https://economictimes.indiatimes.com/news/politics-and-nation/over-1lakh-deaths-in-29-cities-du to-airpollution
[2] http://aqicn.org/city/chennai//
[3]Mihaela Oprea, Mădălina Cărbureanu, Elia Georgiana Dragomir, " *AirQMAS: A Collaborative Multi agent System for Air Quality Analysis* ", 2012.
[4]Sabri Ghazi, Julie Dugdale and Tarek Khadir, " *Modelling Air Pollution Crises Using Multi-agent Simulation* " , 49th Hawaii International Conference on System Sciences (HICSS), 2016.
[5]Fazziki, Abdelaziz & Benslimane, Djamal & Sadiq, Abderrahmane & Ouarzazi, Jamal & Sadgal, Mohammed., " *Agent Based Traffic Regulation System for the Roadside Air Quality Control* ",IEEE Access, 2017.
[6]G.Fathima and Dr. K Shyamala, " *Analysis of air pollutant level using Data mining in Chennai city* ",2020, International Journal of Engineering Technology Research & Management.
[7]Elia Georgiana Dragomir. Mihaela Oprea, " A Multi-Agent System for Power Plants Air Pollution Monitoring ", 2nd IFAC Workshop on Convergence of Information Technologies and Control Methods
[8]Ioannis N Athanasiadis and Pericles A , " *An agent-based intelligent environmental monitoring system* ", Management of Environmental Quality, An International Journal vol.15, 2004.
[9]Sabri, Ghazi & Dugdale, Julie & Khadir, Mohamed ,. " *MASAP: Multi-Agent Simulation of Air Pollution "*., (2018). 10.1007/978-3-319-94580-4_27.
[10]https://www.tutorialspoint.com/weka/ weka_preprocessing_data.html
[11]https://www.analyticsvidhya.com/blog/2016/11/an-introduction-to-clustering-and-different-methods-of-clustering/
[12]https://en.wikipedia.org/wiki/Report
[13]https://en.wikipedia.org/wiki/Sorting//
[14]https://www.tutorialspoint.com/data_structures_algorithms/bubblesortalgorithm.htm
[15]https://en.wikipedia.org/wiki/Category:Air_pollution

iJETRM
International Journal of Engineering Technology Research & Management

[16]https://en.m.wikipedia.org/wiki/Manali,_Chennai//
[17]https://www.thehindu.com/news/cities/chennai/fire-at-distillation-unit-of-cpcl-in-manali/article30738127.ece
[18]https://chennai.citizenmatters.in/industry-emissions-air-pollution-health-manali-ennore-coal-ash-8222
[19]https://en.m.wikipedia.org/wiki/Velachery
[20]https://timesofindia.indiatimes.com/city/chennai/ozone-pollution-high-in-velachery-and-manali/articleshow/72100454.cms
[21]https://www.thenewsminute.com/article/migrant-labourers-chennai-take-streets-demand-transport-go-home-123810