Particulate Matter on Human Health and their Feasibility Study Using Machine Learning Algorithms

Prathik Anandhan[1], Uma Kaliappan[2]* and Anuradha Jageadeesan[3]

[1]School of Information Technology and Engineering, VIT University, Vellore, Tamil Nadu, India
[2]School of Advanced Sciences, VIT University, Vellore, Tamil Nadu, India1
[3]School of Computing Science and Engineering, VIT University, Vellore, Tamil Nadu, India

_____

ABSTRACT

The safety levels of NAAQS pollution standards been set but the measures should be relative to factors such as climate, population and population. Therefore it is essential to reset the levels for the particular location, in order to create the safer environment. This paper aims in determining the pollution level using particulate matters. Naive Bayes and Relevance Vector Machine algorithms are applied to the sample particulate matters studied at various locations across India. In particular Ranipet region. The steps must be taken not only to reduce pollution but also to eradicate it through careful methods of safer waste disposal in all the sectors including vehicular and industrial emissions. This result in the aggravating the pollution. The experimental work conducted on particulate matters such as nitrogen dioxide, sulphur dioxide levels at different test sites indicates that there is an urgency to set modified NAAQS safety limits in order to revert the declining health rate of atmosphere in India.

Keywords: Naive bayes; Relevance vector Machine algorithms; NAAQS; Pollution

_____

INTRODUCTION

The healthy atmosphere is decided by the quality of water and air taking only into the account that the country has stepped into industrial revolution without considering environment, and also caring that how much the environment gets polluted and how much it affect the health conditions of the people due to pollution is quite not acceptable by any citizen of the country. People living in any country not demand reduce industrial revolution otherwise it indirectly affecting economic, employment problem and other social related issues, but they demand to use safer ways to discharge harmful gases released by industrial sectors and reduction of vehicular pollution. Since last ten years various epidemiological studies were demonstrated the consortium between air pollutants related issues, daily deaths, admission to the hospital and visiting rooms that are under emergency and so on. In particularly PM (particulate matter) have been commonly associated to long as well as short-term human health effects. Multinational organisations and governments around the world are trying to minimise discharge and exposure of humans to PM_{10} (particulate having aerodynamic width less than 10 μm) by taking several legal measures [1-3]. The European Union was adopted PM_{10}air quality standards [4]. In 2012 Europe, daily limit were crossing in traffic by 22%, urban backdrop by 27 %, industrial by 17 % and most importantly rural sites by 7 % [5]. The current alarming issue related to greenhouse effect, global warming, arises as the industrial revolution has started [6]. Chlorofluorocarbon (CFC)-11, 12, 13 when CFC entered into air this increase the temperature of environment. The carbon dioxide (CO_2) level statistics however not disclosed to public by governments agencies one must move towards personal calculations towards the measurements of the CO_2 and other pollutant concentrations gases in air. The rise in the atmospheric concentration of carbon dioxide (CO2) and other harmful gases air is widely considered as the main contributor to the complete changes in air quality and global warming effect. Likewise$PM_{10,}$ $NO_{2,}$ andSO_2these are considered some of harmful gases by NAAQS have to be kept in check to keep alive healthy atmosphere. The capacity to minimise pollution present in trees and due to deforestation the air quality, soil fertility get affected in large. There are two types data analysis used to [7], prediction and classification were two different types of data analysis [8] which are used for selecting the models for explaining usefulness predict future data trends or data classes. Grouping the categorical labels, called "classes", while prediction models are used for predict the

_____

continuous valued variables. Typically, both prediction and classification data analyses is a dual process. In which initial step describes the prediction or classification algorithm is applied on the available data and a decision model is removed. The principle component analysis has been carried out for the percentage contribution of elements of PM_{10} [9]. The NAQI value ranking higher at traffic, tannery site [10]. Element are released to the atmosphere from both anthropogenic and natural sources [11]. Uncontrolled use of fossil fuels in industries and transport sectors [12] Air pollution is the major source of pollution [13]. Higher level of NO_2, CO, and hydrocarbons [14]. Fossil based technology is the primary sources in India [15]. The mined decision model encapsulates the command lying in the data in a form regression neural network decision tree support vector machine. This stage focuses on testing the potential of the decision model be close to data not used in training. In this respect both statistical prediction, classification algorithms been applied for function verge in several, yet diverse domains. To well known the terms "classification" and "prediction".

Need for setting strict NAAQS limits on industrialized countries
Air is the source of almost all diseases and has a capacity to infect a living organism as it breathes. The environment polluted to a sizable extent which is approximately a point less than that of the NAAQS safety level and labelling that particular region is strictly following NAAQS safety standards cannot be fully acceptable. Take an insect as an example; first inject a small dosage of insecticide which does not cause sudden death of the insect at the point of time, however after a few repeated dosages the insect dies. By seeing this example to establish that just because we do not suffer from imminent death or disease any imminent side effects cannot say that we live in a healthy environment. Due to air pollution even plants or trees where produce healthy daily commodities for man tend to lose a considerable portion of their health interest. This situation continues as long as harsh anti-pollution rules were set on all the industries functioning in the country. On taking vehicular pollution into account one has to realize that substitute to fuel must be introduced all over the world.
The country around the world apprises petroleum is the second biggest advantage of the country after gold so that no country is ready to introduce full-fledged effort on bringing down usage of petroleum. Battery a powered car was introduced long ago, but still it is not come into public use. If consider power shortage occur in case if launch battery powered vehicles then one must know that cost of producing power is well below the cost of producing petroleum. Environment will continue to heat up as continue the use of fuels. The stern NAAQS safety limits must be set on industrialized countries for bringing down pollution.

EXPERIMENTAL SECTION

Sample study
India has diversity of industries vary from small scale to large scale industries and industries shape the backbone of Indian economy. Petroleum products, machineries, chemicals, steel iron, automobiles etc., were some of the major goods produced in India. Waste generated by industries were not treated properly, if it get mix up with air aggravate contaminations in that particular location. Few research works has done on the Vellore district Tamil Nadu, India. Regarding careless disposal of industrial waste and the research was done on a particular place called Ranipet where more number of leather industries located and the residents accuse that there was nasty smell in air surrounding the places of the leather industries also paint manufacturing industries. The plant in Ranipet manufacture chemicals such as sodium chromate, tanning powder, basic chromium sulphate, chromium salts were used in the industry almost 15 lakh ton of strong squanders gathered more than two many years of plant operation are stacked in an open spot on the premises. This result in polluting the ground level water. These contaminations affect the health, resources and livelihood of thousands.

RESULTS AND DISCUSSION

Using machine learning algorithms Bayes, RVM to classify the dataset
The study has been extended to selected part of India and among 246 test sites which monitor pollution statistics in India, the year-wise datasets were collected from the CPCB. This composed of ambient air quality, concentration levels of SO_2, NO_2, and PM10 in air. Machine learning algorithms were used to grade them to predict which pollutant is present in large quantity and side effect caused by pollutants. The dataset yield the final concentration value of SO_2, NO_2, PM_{10} pollutants with respect to the year. The calculations were based on various sources like the number of industries present in that particular area, their share among the pollution, dust, vehicular emission levels, etc.

$$y = f(x|\mu, \alpha) = \left(\frac{1}{\alpha\sqrt{2\pi}}\right)e^{\xi}$$

_____

NO$_2$ (Nitrogen Dioxide)
Nitrogen dioxide is one of the dangerous gases which have to be kept constant check and the high level of NO$_2$ in air because breathing problems irritation in eye. Bayes algorithm was used to classify the on which range the concentration of the gas is in excess. The results were shown below graph 1.
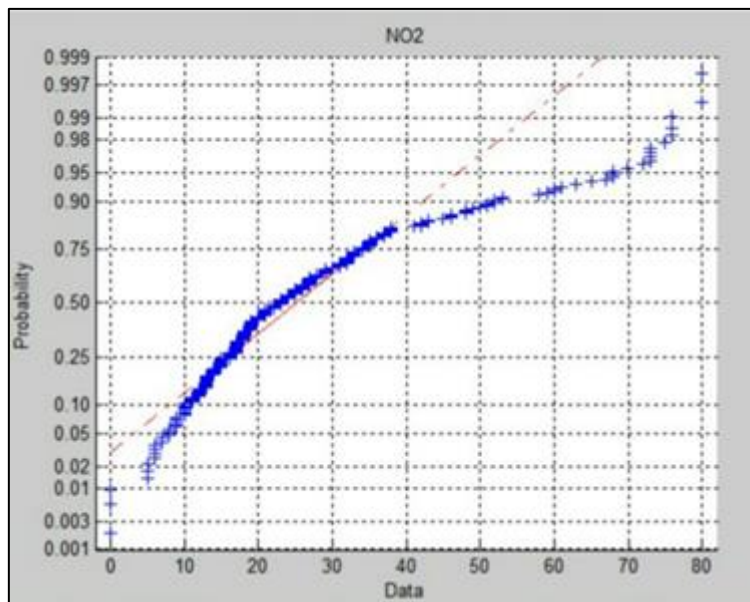
Graph 1: Level of NO$_2$ gas in air

SO$_2$ (Sulphur Dioxide)
Burning of coal and oil inside a fuel causes the emission of sulphur dioxide gas. More quantity of the gases air causes breathing and heart related problems. Acidification of lakes, reduced visibility, and corrosion of buildings and metallic structure are also some of the worst side effects of excess concentrationof sulphur dioxide if air. Results were summarised given the graph 2.
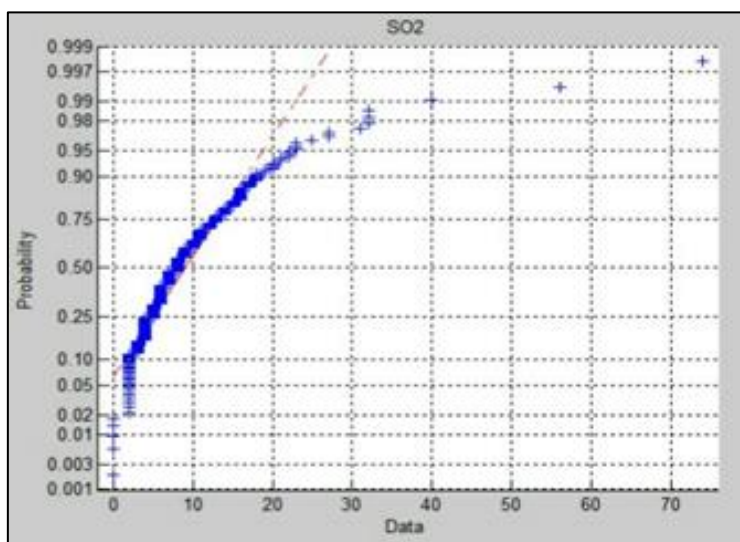

Graph 2: Level of NO$_2$ gas in air

PM$_{10}$ (Particulate Matter)
The particulate matter mainly a blend of minute particles has the capacity to reach the lower regions of the respiratory tracks. Causes lung related problems in human and other respiratory problems particularly asthma is major adverse effect of high particulate matter in air were some of its health impacts. Results were shown in the graph 3.

_____
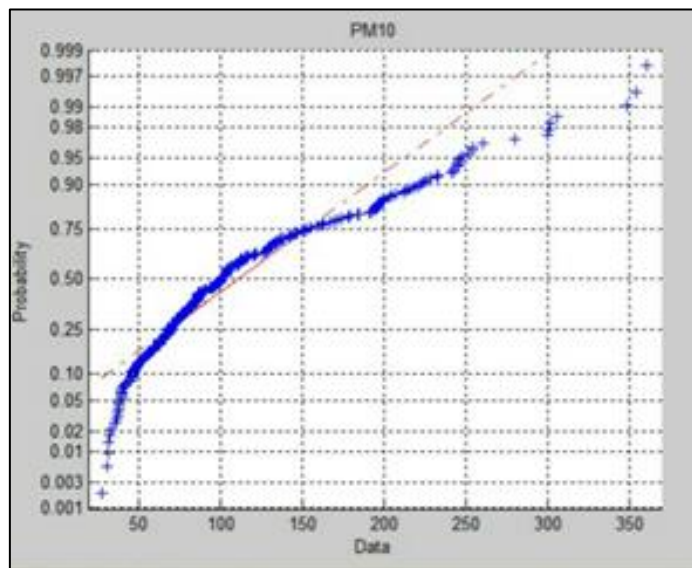
Graph 3: Level of particulate matter in air

Total concentration level (graph 4)
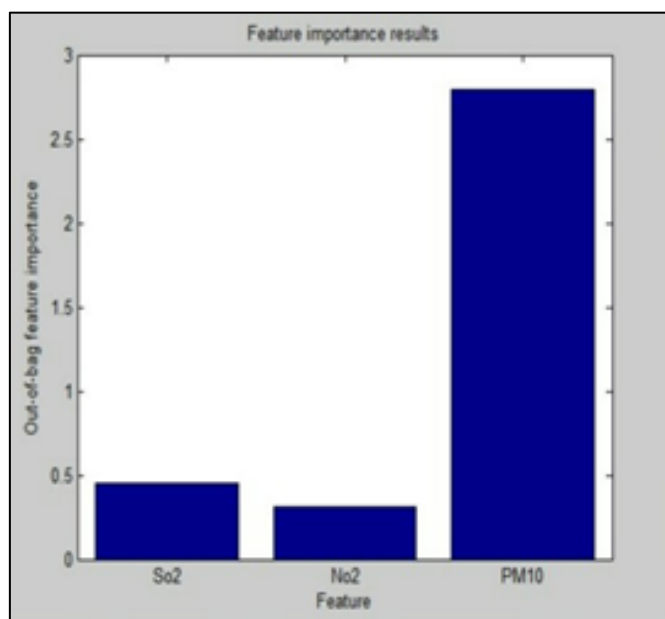
Graph 4: Level of total concentration level in air

Proposed algorithm to determine pollution level in India
The suggested algorithm states the newer way of estimating how much quantity of pollution were released in air by different sources. By taking into the account only the total number of sources emitting pollutants and quantity of pollutants emitted in part per millions (ppm) and also the total quantity of the pollution treated with respect to particular area we able to estimate the approximate amount of pollutant level concentration in air at the particular point of time. If the government fix the pollution level restrictions based on this value and also in addition to the limitations on total quantity of individual pollutant generated one can more accurate and more efficient way of determining pollution levels of several other sources contributing to pollution should also be included in the scenario in accordance to India like open let drains, pollution due to tannery, quantity of untreated garbage in roads etc. The more analyses of the sources of pollution, greater extent can be accurately predict the pollution index.

_____

CONCLUSION

Encompassing air contamination has greater significant and assorted wellbeing impacts, in cowhide tannery industries and its surrounding areas, The present levels of chromium, lead, H_2S, SO_2, NO_2 furthermore, particulates in babies, children' sand grown-ups are among the most liable to be influenced particularly, those in who were living in around cowhide tannery particulates remain dangerous in numerous parts of the Ranipet, Vellore district of Tamil Nadu, where the present National Ambient Air Quality Standards may not ensure the public satisfactorily. There is a need to convincing to advance on endeavours to guarantee the clean air for all. Concoction vaporous contaminations levels have a tendency to be most extreme toward the evening and it might be likely to diminish at night kids and grown-up's introduction by draw in strenuous open air work prior in the day. The individuals who were having higher convergence of chromium levels in blood and urine were get extreme type of bronchial asthma and there were no neurological disarranged.

REFERENCES

[1] I Barmpadimos; C Hueglin; J Keller; SP Henne, ASH. *Atmos Chem Phys,* **2011**, *11*(4), 1813-1835.
[2] CA Pope. *Inhal Toxicol*, **2007**, 19(1), 33-38.
[3] CA Pope; DW Dockery, *J Air Waste Manage*, **2006**, 56(6), 709–742.
[4] EC Directive 2008/50/EC of the European Parliament and of the Council on ambient air quality and cleaner air for Europe, **2008**.
[5] EEA, Air quality in Europe – **2014** report. Tech. report No. 5/2014.
[6] http://www.global-greenhouse-warming.com/deforestation
[7] V Jayathi; R Krishnamurthy, Curr Sci, **2016**, 90(3), 403-413.
[8] J Han; M Kambar, Data mining-Concept and technique Morgan Kaufmann publishers Wyman street, 225, USA.
[9] TV Lakshmi Kumar; P Ajay; K Aruna; JS Sudarsan, *Int Chem Tech Res*., **2015**,7(3), 1414-1420.
[10] V Parthipa; SA Raja, *Int Chem Tech Res*., **2015**, 8(6), 45-55.
[11] SR kumar; Rajkumar, *Int Chem Tech Res*., **2015**, 8(7), 373-382.
[12] Kistan; SS Ansari, *Int Chem Tech Res.,* **2015**, 8(4), 1916-1922.
[13] NVE Mariappan; P Mohana; C Murugesan, *Carbon*, **2015**, 5(1), 172-179.
[14] G AbdElAziz, *Int Chem Tech Res*., **2015**, 8(6), 468-482.
[15] G Murali, *Int Chem Tech Res*., **2015,** 8(1), 30-36.