IJITCE

ISSN 2347- 3657

International Journal of Information Technology & Computer Engineering

Email : ijitce.editor@gmail.com or editor@ijitce.com

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 13, Issue 1, 2025

A MACHINE LEARNING APPROACH TO DETECT ASSOCIATIONS BETWEEN AIR QUALITY AND ASTHMA IN URBAN ENVIRONMENTS

[1]Vinitha B, [2]Navika B, [3]Thahzeeba M, [4]Kavisri K, [5]Ajuma B

[1,2,3,4,5]UG Student, Department of Computer Science

[1,2] KSR College of Arts and science for Women, KSR Kalvi Nagar, Thiruchengode -637215, Namakkal (Dt).

*Corresponding Author

E-mail : vinithapavi.17@gmail.com

ABSTRACT

Traffic and power generation constitute the primary sources of urban air pollution. The notion that outdoor air pollution can aggravate pre-existing asthma is substantiated by a substantial body of evidence accumulated over several decades, with numerous studies indicating a potential role in the onset of new asthma cases as well. This paper examines the impacts of particulate matter (PM), gaseous pollutants (ozone, nitrogen dioxide, and sulfur dioxide), and air pollution from mixed traffic sources. We concentrate on clinical studies, encompassing both epidemiological and experimental research, published within the last five years. From a mechanistic standpoint, air pollutants likely induce oxidative damage to the airways, resulting in inflammation, remodeling, and an elevated risk of sensitization. While numerous pollutants have been associated with the onset of asthma, the robustness of the evidence varies. We also examine clinical implications, policy concerns, and research deficiencies pertinent to air pollution and asthma.

Keywords: Air pollution, Asthma prediction, Supervised learning, Light Gradient Boosting Model.

1. INTRODUCTION

Outdoor air pollution contributed more than 3% of the annual disability-adjusted life years lost in the 2010 Global Burden of Disease comparative risk assessment, a notable increase since the previous estimate was made in 2000.1 Previous assessments of global disease burden attributed to air pollution were restricted to urban areas or by coarse spatial resolution of concentration estimates.2 In a study of ten European cities, 14% of the cases of incident asthma in children and 15% of all exacerbations of childhood asthma were attributed to exposure to pollutants related to road traffic.3 Urbanisation is an important contributor to asthma and this contribution might be partly attributed to increased outdoor air pollution (figure 1).4–6 Because many urban centres in the developing world are undergoing rapid population growth accompanied by increased outdoor air pollution, the global burden of asthma is likely to increase. In this context, it is notable that the populations of China, India, and Southeast Asia are equal to the rest of the world combined. In view of the burden of asthma attributed to outdoor air pollution, a better understanding of why asthmatic individuals are susceptible to this exposure should enable the design of effective preventive strategies. The idea that air pollution can cause exacerbations of preexisting asthma is supported by an evidence base that has been accumulating for several decades,7–10 but evidence has emerged that suggests air pollution might cause new-onset asthma as well.11–21 Not all studies support a causal link between air pollution and asthma, and a recent meta- analysis22 of cross-sectional studies that compared communities with different levels of pollution showed no effect of long-term exposure to pollution on asthma prevalence. Although outdoor air pollution almost always occurs as a mixture, air quality is regulated by most jurisdictions in terms of its individual components. Such regulation has meant that experimental studies of humans and animals have been focused on individual pollutants. Because epidemiological studies inherently involve exposure to mixtures of pollutants, substantial efforts are usually made to try to identify the individual effects of pollutants, which often obscures the health effect of the mixture as a whole. With increasing attention to traffic-related air pollution (TRAP) as the exposure variable of interest, a shift has occurred away from a focus on individual components of the pollution mixture. In this Series paper, we will attempt to discuss the effects of several gaseous pollutants (ozone, nitrogen dioxide, and sulphur dioxide), the independent effects of various forms of PM, and then focus on the effects of TRAP as a mixture. We concentrate on studies published in the past 5 years that report results relevant to both exacerbation and onset of asthma. We focus primarily, although not exclusively, on epidemiological and experimental clinical studies. Controlled exposure studies in human beings are restricted by small sample size and an inability to study the potentially most susceptible subgroups (eg, children and adults with severe asthma) and the effects of chronic exposure. Epidemiological studies are restricted by imprecise methods of both exposure and asthma outcome assessment and often inadequate data about potentially confounding variables. Although the potential effect of indoor air pollution on asthma is an important concern, especially in developing countries where much domestic cooking is done with solid

fuels, it is outside the scope of this review.

2. LITERATURE SURVEY

2.1 Asthma So Affected by Exposure to Air Pollution

Why are individuals with asthma so affected by exposure to air pollution? At high concentrations, such as those noted in megacities in India and China, air pollutants might have direct irritant and inflammatory effects on airway neuroreceptors and epithelium, but such levels of exposure rarely occur in North America or Europe. At the lower concentrations that are more typical in high-income countries, other mechanisms are probably in operation. Specific pollutants can induce airway inflammation (eg, ozone, nitrogen dioxide, and PM <2•5 μm in diameter [PM2•5])23–28 and airway hyper-responsiveness (ozone and nitrogen dioxide),23,29 two characteristic features of asthma. In addition, oxidative stress (a feature of severe asthma) has been associated with pollutant exposures (ozone, nitrogen dioxide, and PM2•5).30–32 Therefore, exposure to these pollutants is unsurprisingly associated with exacerbations and possibly even the onset of asthma. The mechanisms by which pollutants induce these effects are not completely clear.

2.2 Air Pollution Might Contribute

A framework for how air pollution might contribute to the development and exacerbation of asthma proposed by the UK's Committee on the Medical Effects of Air Pollutants identified four main mechanisms: oxidative stress and damage, airway remodelling, inflammatory pathways and immunological responses, and enhancement of respiratory sensitisation to aeroallergens (figure 2).33 Variation in the genes that regulate these mechanisms could confer increased susceptibility to development of new-onset asthma or exacerbations of existing disease with exposure to air pollution.

2.3 Traffic-Related Air Pollution

Because the pollutants of interest, including TRAP, can cause oxidative stress, the ability of antioxidant defenses to handle the increased load of reactive oxygen species generated in the lungs after exposure is an important determinant of risk for subsequent adverse effects. Specific polymorphisms in antioxidant enzyme genes, such as glutathione S-transferase genes, GSTM1 and GSTP1, can modify risk of asthmatic responses to pollutants34,35 and these variants (GSTM1 null and GSTP1 Ile105Val) might also interact with a tumour necrosis factor (TNF) promoter variant (G-308A) that affects expression of TNF and hence the early inflammatory response.36 Additionally, neonatal rats are more prone to oxidative stress from PM exposure at least in part due to relative deficiency of nuclear factor-like 2 (Nrf2).37 Proinflammatory effects of oxidative stress are mediated by the redox-sensitive MAP kinase and nuclear factor-κB cascades that are responsible for the

expression of cytokines, chemokines, and adhesion molecules, and reduced antioxidant capacity in the airways can result in altered expression after pollutant exposure.38 Other pathways through which oxidising pollutants might affect severity of asthma involve control of immune responses. TRAP, specifically ambient polycyclic aromatic hydrocarbons and diesel-exhaust particles, affect regulatory T cell (Treg) function through an epigenetic mechanism.39,40 Hypermethylation of CpG islands in Foxp3 associated with chronic exposure to polycyclic aromatic hydrocarbons39 or diesel-exhaust particles40 leads to suppression of Treg function and increased asthma severity as assessed by symptoms and lung function. Hypermethylation of interferon γ in effector T cells, contributing to a shift towards a Th2 response, has also been associated with exposure to air pollution.41Studies in animals and in vitro42,43 suggest that exposure to PM results in allergic inflammation with Th2 and Th17 phenotypic differentiation, with a specific role for environmentally persistent free radicals and polycyclic aromatic hydrocarbon fractions of PM in this differentiation. In addition, exposure to diesel-exhaust particles is associated with increased serum interleukin 17 and increased symptoms in children with allergic asthma; a parallel study44 that used a murine model of allergic airway inflammation showed that combined exposure to diesel-exhaust particles and antigen from a house dust mite induced a mixed Th2/Th17 response.

2.4 A Potential Enhancing Effect of Pollutant Exposure On Responses

A potential enhancing effect of pollutant exposure on responses to inhaled allergen has been studied in both animals and man, with evidence for such an effect on lung function and inflammatory responses to ozone, nitrogen dioxide, sulphur dioxide, and diesel-exhaust particles.45–48 Several mechanisms through which air pollutants could enhance sensitisation to aeroallergens have been proposed and include increased deposition of allergen in the airways due to carriage by particles, increased epithelial permeability due to oxidative injury, increased antigenicity of proteins from chemical modification, and a direct adjuvant effect (including for diesel-exhaust particles in human beings).49In summary, air pollutants might cause oxidative injury to the airways that leads to inflammation and remodelling, which in a genetically predisposed individual could result in clinical asthma. One predisposing factor might be atopy, and air pollutants could increase the risk of sensitization and the responses to inhaled allergen in individuals with asthma.

3. PROPOSED SYSTEM

LightGBM is a gradient boosting framework that uses tree-based learning algorithms. It is designed to be distributed and efficient with the following advantages:

- Faster training speed and higher efficiency.
- Lower memory usage.
- Better accuracy.
- Support of parallel and GPU learning.

- Capable of handling large-scale data.

At present, decision tree-based machine learning algorithms dominate Kaggle competitions. The winning solutions in these competitions have adopted an algorithm called XGBoost. A couple of years ago, Microsoft announced its gradient boosting framework LightGBM. Nowadays, it steals the spotlight in gradient boosting machines. Kagglers start to use LightGBM more than XGBoost. LightGBM is 6 times faster than XGBoost. The size of the dataset is increasing rapidly. It is become very difficult for traditional data science algorithms to give accurate results. Light GBM is prefixed as Light because of its high speed. Light GBM can handle the large size of data and takes lower memory to run. Another reason why Light GBM is so popular is because it focuses on accuracy of results. LGBM also supports GPU learning and thus data scientists are widely using LGBM for data science application development. It is not advisable to use LGBM on small datasets. Light GBM is sensitive to overfitting and can easily overfit small data.

LightGBM intuition

LightGBM is a gradient boosting framework that uses tree-based learning algorithm. LightGBM documentation states that –

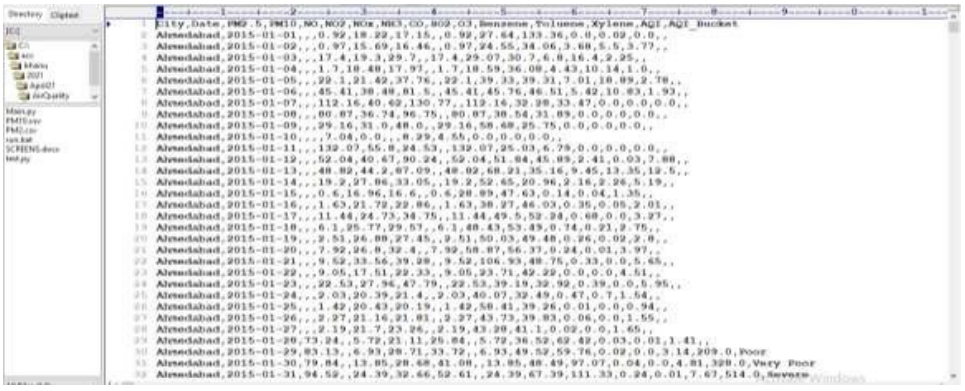
- LightGBM grows tree vertically while other tree-based learning algorithms grow trees horizontally. It means that LightGBM grows tree leaf-wise while other algorithms grow level- wise. It will choose the leaf with max delta loss to grow. When growing the same leaf, leaf- wise algorithm can reduce more loss than a level-wise algorithm.
- So, we need to understand the distinction between leaf-wise tree growth and level-wise tree growth.

Important points about tree-growth

- If we grow the full tree, best-first (leaf-wise) and depth-first (level-wise) will result in the same tree. The difference is in the order in which the tree is expanded. Since we don't normally grow trees to their full depth, order matters.
- Application of early stopping criteria and pruning methods can result in very different trees. Because leaf-wise chooses splits based on their contribution to the global loss and not just the loss along a particular branch, it often (not always) will learn lower-error trees "faster" than level-wise.
- For a small number of nodes, leaf-wise will probably out-perform level-wise. As we add more nodes, without stopping or pruning they will converge to the same performance because they will literally build the same tree eventually.

4. RESULTS AND DISCUSSION

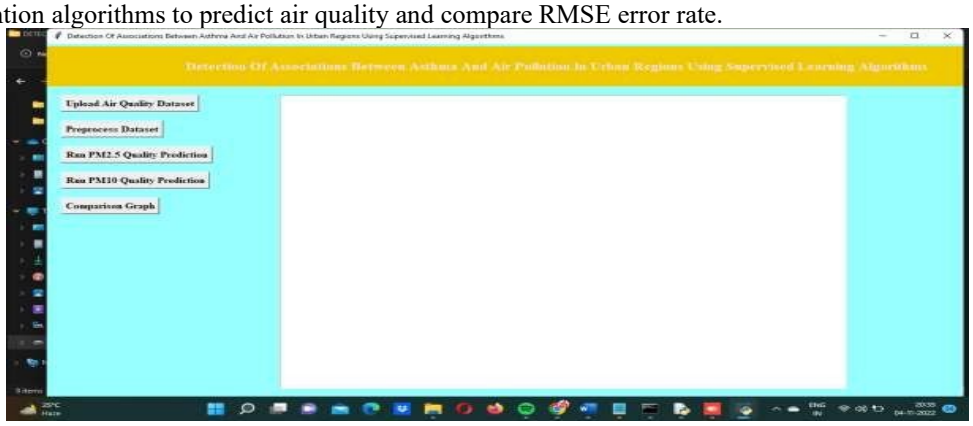
This project employed various machine learning algorithms to predict air quality as this air quality has a heavy effect on human health. Here, LightGBM algorithm is used as a proposed ML classifier which can predict air quality efficiently from high-dimensional large-scale data. RMSE is used as a quality metric for comparing with other algorithms such as ADABOOST, GBDT (Gradient Boosting Decision Tree), XGBoost and DNN etc.. The lower the RMSE error rate the better is the prediction model and in all algorithms Light GBM is giving less error rate. Here, Air Quality dataset from KAGGLE website is used the dataset and below screen showing details of dataset. In this paper we are using PM2.5 and PM10 quality values to predict air quality.



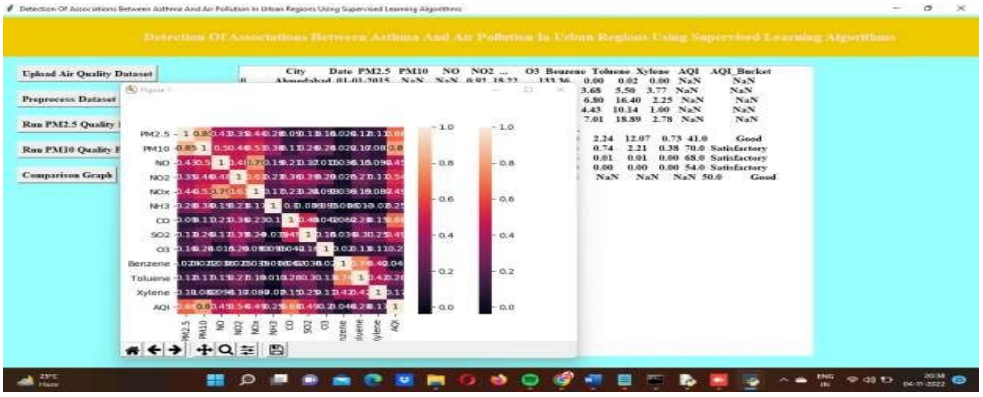
In the above dataset screen we have city name, date and other air quality values and we will use above dataset with

International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 13, Issue 1, 2025

above mention algorithms to predict air quality and compare RMSE error rate.



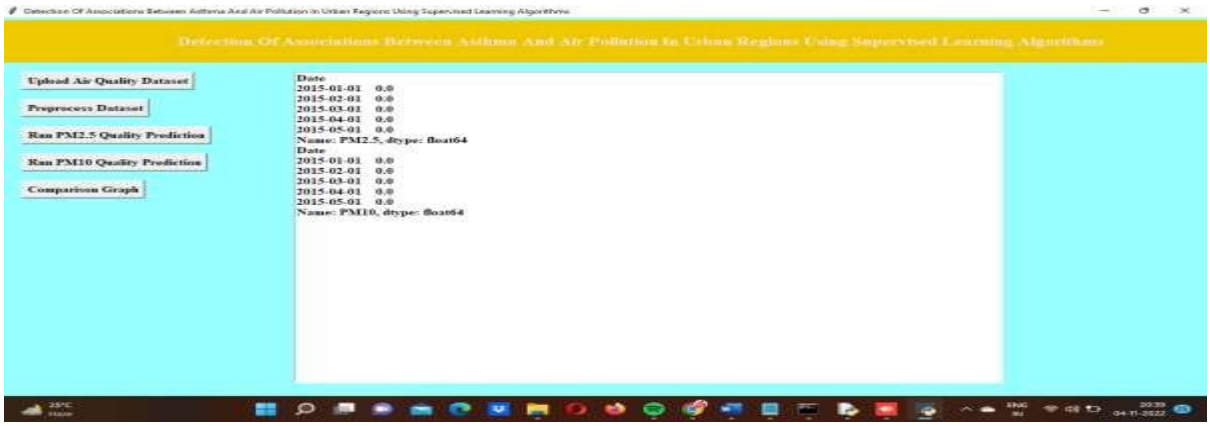
In above screen click on 'Upload Air Quality Dataset' button to upload dataset and to get below screen

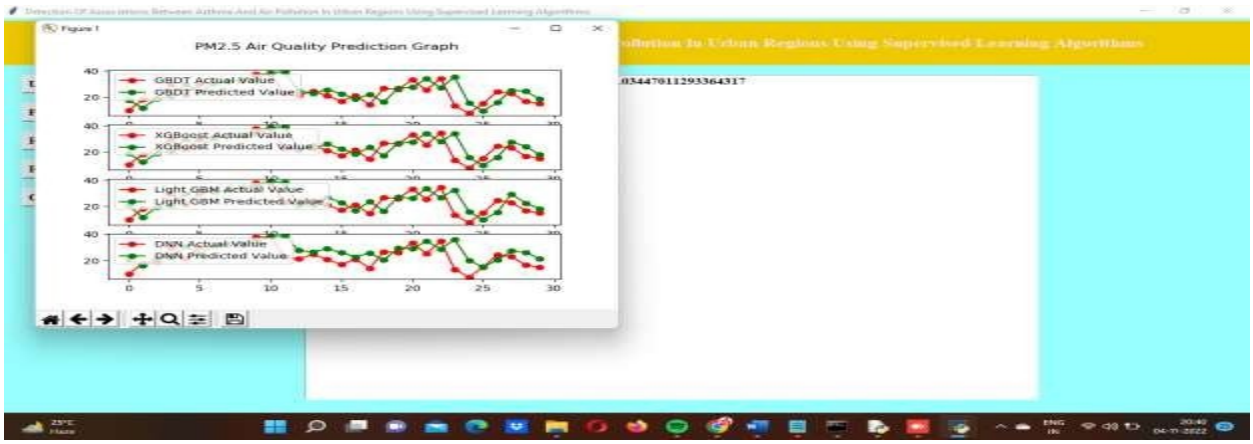


In above screen dataset loaded and we can see dataset values in text area and in graph we can see column name from dataset and its correlation feature importance values in graph boxes. Now close above graph and click on 'Preprocess Dataset' button to remove missing values and then calculate air pollution rate date wise in dataset like below screen.

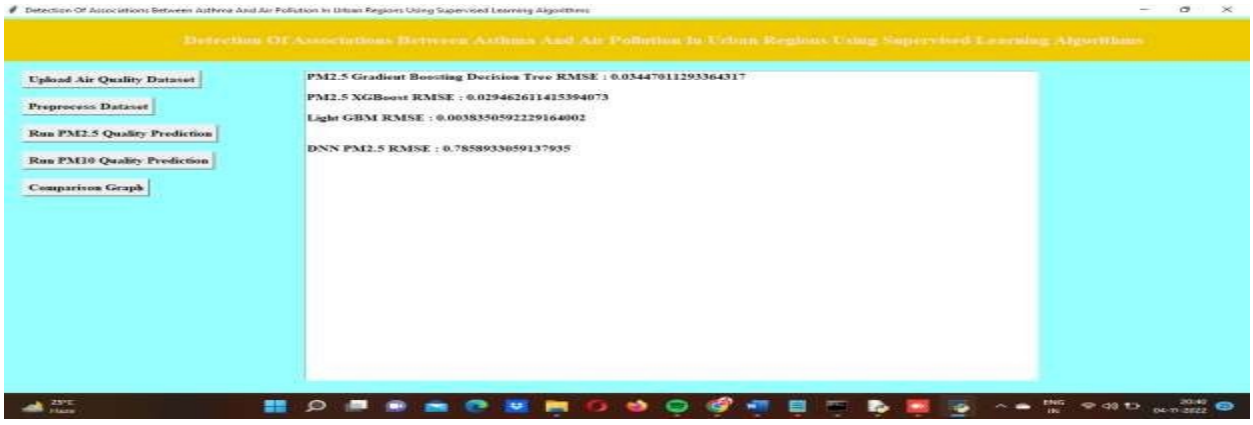
In above screen for few days I am printing pollution rate and now all missing values are removed and now click on 'Run PM2.5 Prediction' button to predict air quality and for testing purpose we are using test values from dataset and then calculate RMSE error rate between original and prediction values.

International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 13, Issue 1, 2025



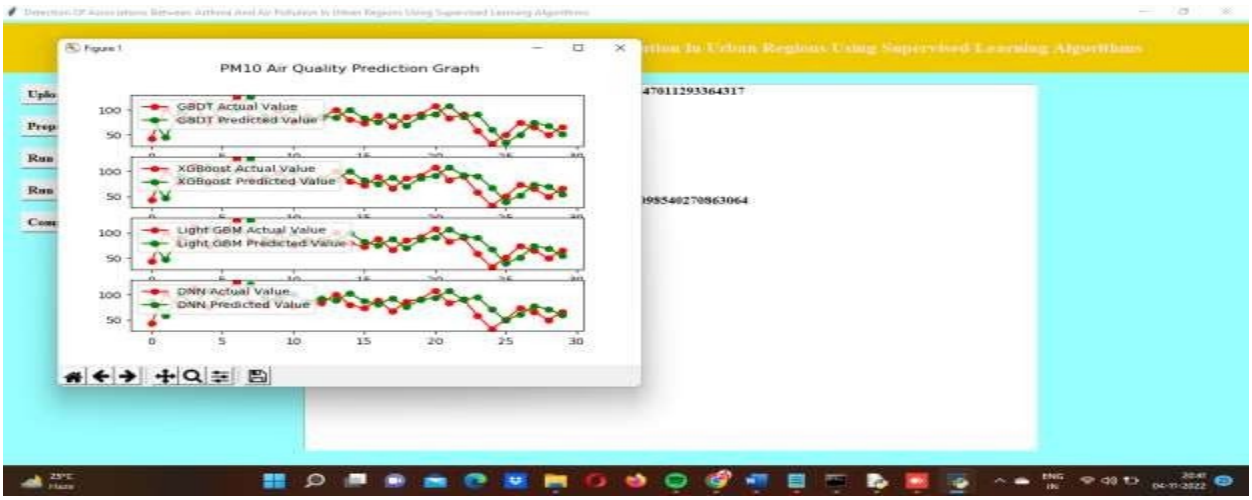


In above screen we can see RMSE error rate for all 4 algorithms and in all algorithms Light GBM got less RMSE error rate. Below is the predicted and original values from all 4 algorithm

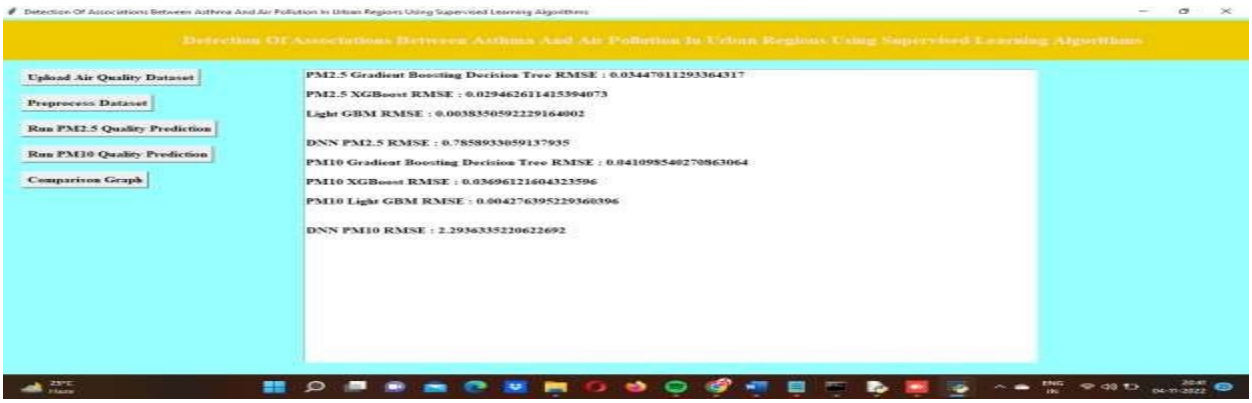


In above screen we can see RMSE error rate for all 4 algorithms and in all algorithms Light GBM got less RMSE error rate. Below is the predicted and original values from all 4 algorithms

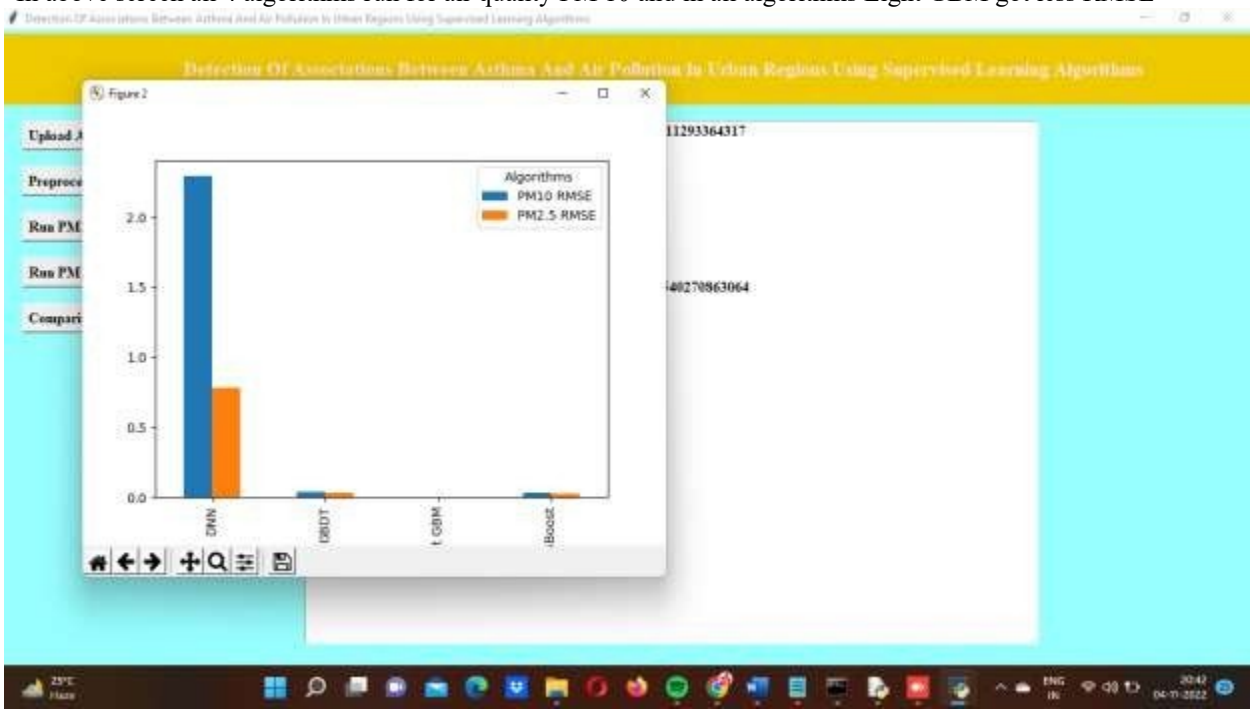
International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 13, Issue 1, 2025



In above graph red line represents original air quality and green line represents predicted air quality and I am displaying prediction graph for all 4 algorithms and from all 4 Light GBM is closed and now closed above graph and then click on 'Run PM10 Quality Prediction' button to predict air quality for PM10 using all 4 algorithms



In above screen all 4 algorithms run for air quality PM 10 and in all algorithms Light GBM got less RMSE



error and below is the prediction graph for all 4 algorithms

In above graph x-axis represents algorithm names and y-axis represents RMSE error rate and for both PM10 and PM2.5 air quality Light GBM got less RMSE error

5. CONCLUSION

A substantial body of research on the effects of air pollution on asthma has been published in the past 5 years, adding to the body of knowledge that has accumulated over several decades. Presently, short- term exposures to ozone, nitrogen dioxide, sulphur dioxide, PM2•5, and TRAP is thought to increase the risk of exacerbations of asthma symptoms. Increasing amounts of evidence also suggest that long- term exposures to air pollution, especially TRAP and its surrogate, nitrogen dioxide, can contribute to new-onset asthma in both children and adults. Much more about the mechanisms that are involved with exacerbations induced by pollution and onset of asthma needs to be understood, but oxidative stress and immune dysregulation are probably both involved. Young children with asthma, especially those growing up in poor neighborhoods, are at increased risk of adverse effects from exposures to air pollution. Unravelling which components of the traffic pollution mixture are responsible for asthma exacerbations and onset is a substantial challenge. Improved air quality to prevent exacerbations and new cases of asthma will require strong governmental efforts to move economies in both developed and developing countries away from combustion of fossil fuels for transportation and energy production; this approach is also needed to mitigate climate change.

REFERENCES

[1] Lim SS, Vos T, Flaxman AD, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet. 2012; 380:2224–60. [PubMed: 23245609]

[2] Brauer M, Amann M, Burnett RT, et al. Exposure assessment for estimation of the global burden of disease attributable to outdoor air pollution. Environ Sci Technol. 2012; 46:652–60. [PubMed: 22148428] 3. Perez L, Declercq C, Iniguez C, et al. Chronic burden of near-roadway traffic pollution in 10 European cities (APHEKOM network). Eur Respir J. 2013; 42:594–605.[PubMed: 23520318]

[3] Wong GWK, Chow CM. Childhood asthma epidemiology: insights from comparative studies of rural and urban populations. PediatrPulmonol. 2008; 43:107–16. [PubMed: 18092349]

[4] Robinson CL, Baumann LM, Romero K, et al. Effect of urbanisation on asthma, allergy and airways inflammation in a developing country setting. Thorax. 2011; 66:1051–57. [PubMed: 21730351]

[5] Brunekreef B, Stewart AW, Anderson HR, Lai CKW, Strachan DP, Pearce N. Self-reported truck traffic on the street of residence and symptoms of asthma and allergic disease: a global relationship in ISAAC phase 3. Environ Health Perspect. 2009; 117:1791–98. [PubMed: 20049134]

[6] Friedman MS, Powell KE, Hutwagner L, Graham LM, Teague WG. Impact of changes in transportation and commuting behaviors during the 1996 Summer Olympic Games in Atlanta on air quality and childhood asthma. JAMA. 2001; 285:897–905. [PubMed: 11180733]

[7] O'Connor GT, Neas L, Vaughn B, et al. Acute respiratory health effects of air pollution on children with asthma in US inner cities. J Allergy Clin Immunol. 2008; 121:1133–39. [PubMed: 18405952]

[8] Weinmayr G, Romeo E, De Sario M, Weiland SK, Forastiere F. Short-term effects of PM10 and NO2 on respiratory health among children with asthma or asthma-like symptoms: a systematic review and meta-analysis. Environ Health Perspect. 2010; 118:449–57. [PubMed: 20064785]

[9] Health Effects Institute. Traffic-related air pollution: a critical review of the literature on emissions, exposure, and health effects. Boston, MA: Health Effects Institute; 2010. Panel on the Health Effects of Traffic-Related Air Pollution.

[10] Dong G-H, Chen T, Liu M-M, et al. Gender differences and effect of air pollution on asthma in children with and without allergic predisposition: northeast Chinese children health study. PLoS One. 2011; 6:e22470. [PubMed: 21811617]

[11] Nishimura KK, Galanter JM, Roth LA, et al. Early-life air pollution and asthma risk in minority children. The GALA II and SAGE II studies. Am J Respir Crit Care Med. 2013; 188:309–18. [PubMed: 23750510]

[12] Jacquemin B, Sunyer J, Forsberg B, et al. Home outdoor NO2 and new-onset of self-reported asthma in adults. Epidemiology. 2009; 20:119–26. [PubMed: 18923331]

[13] Jerrett M, Shankardass K, Berhane K, et al. Traffic-related air pollution and asthma onset in children: a prospective cohort study with individual exposure measurement. Environ Health Perspect. 2008; 116:1433–38. [PubMed: 18941591]

[14] McConnell R, Islam T, Shankardass K, et al. Childhood incident asthma and traffic-related air pollution at home and school. Environ Health Perspect. 2010; 118:1021–26. [PubMed: 20371422]

[15] Zhou C, Baiz N, Zhang T, Banerjee S, Annesi-Maesano I. Modifiable exposures to air pollutants related to asthma phenotypes in the first year of life in children of the EDEN mother- child cohort study. BMC Public

Health. 2013; 13:506. [PubMed: 23705590]
[16] Carlsten C, Dybuncio A, Becker A, Chan-Yeung M, Brauer M. Traffic-related air pollution and incident asthma in a high-risk birth cohort. Occup Environ Med. 2011; 68:291–95. [PubMed: 20952556]
[17] Clark NA, Demers PA, Karr CJ, et al. Effect of early life exposure to air pollution on development of childhood asthma. Environ Health Perspect. 2010; 118:284–90. [PubMed: 20123607]
[18] Kunzli N, Bridevaux P-O, Liu L-JS, et al. Traffic-related air pollution correlates with adult- onset asthma among never-smokers. Thorax. 2009; 64:664–70. [PubMed: 19359271]
[19] Modig L, Toren K, Janson C, Jarvholm B, Forsberg B. Vehicle exhaust outside the home and onset of asthma among adults. Eur Respir J. 2009; 33:1261–67. [PubMed: 19251785]
[20] Anderson HR, Favarato G, Atkinson RW. Long-term exposure to air pollution and the incidence of asthma: meta-analysis of cohort studies. Air Qual Atmos Health. 2013; 6:47–56.
[21] Anderson HR, Favarato G, Atkinson RW. Long-term exposure to outdoor air pollution and the prevalence of asthma: meta-analysis of multi-community prevalence studies. Air Qual Atmos Health. 2013; 6:57–68.
[22] Seltzer J, Bigby BG, Stulbarg M, et al. O3-induced change in bronchial reactivity to methacholine and airway inflammation in humans. J Appl Physiol. 1986; 60:1321–26. [PubMed: 3084448]
[23] Aris RM, Christian D, Hearne PQ, Kerr K, Finkbeiner WE, Balmes JR. Ozone-induced airway inflammation in human subjects as determined by airway lavage and biopsy. Am Rev Respir Dis. 1993; 148:1363–72. [PubMed: 8239177]
[24] Solomon C, Christian DL, Chen LL, et al. Effect of serial-day exposure to nitrogen dioxide on airway and blood leukocytes and lymphocyte subsets. Eur Respir J. 2000; 15:922–28. [PubMed: 10853860]
[25] Frampton MW, Boscia J, Roberts NJJ, et al. Nitrogen dioxide exposure: effects on airway and blood cells. Am J Physiol Lung Cell Mol Physiol. 2002; 282:L155–65. [PubMed: 11741827]
[26] Dales R, Wheeler A, Mahmud M, et al. The influence of living near roadways on spirometry and exhaled nitric oxide in elementary schoolchildren. Environ Health Perspect. 2008; 116:1423–27. [PubMed: 18941589]
[27] Delfino RJ, Staimer N, Gillen D, et al. Personal and ambient air pollution is associated with increased exhaled nitric oxide in children with asthma. Environ Health Perspect. 2006; 114:1736– 43. [PubMed: 17107861]
[28] Poynter ME, Persinger RL, Irvin CG, et al. Nitrogen dioxide enhances allergic airway inflammation and hyperresponsiveness in the mouse. Am J Physiol Lung Cell Mol Physiol. 2006; 290:L144–52. [PubMed: 16085673]
[29] Liu L, Poon R, Chen L, et al. Acute effects of air pollution on pulmonary function, airway inflammation, and oxidative stress in asthmatic children. Environ Health Perspect. 2009; 117:668– 74. [PubMed: 19440509]
[30] Patel MM, Chillrud SN, Deepti KC, Ross JM, Kinney PL. Traffic-related air pollutants and exhaled markers of airway inflammation and oxidative stress in New York City adolescents. Environ Res. 2013; 121:71–78. [PubMed: 23177171]
[31] Neophytou AM, Hart JE, Cavallari JM, et al. Traffic-related exposures and biomarkers of

systemic inflammation, endothelial activation and oxidative stress: a panel study in the US trucking industry. Environ Health. 2013; 12:105. [PubMed: 24314116]
[32] Gowers AM, Cullinan P, Ayres JG, et al. Does outdoor air pollution induce new cases of asthma? Biological plausibility and evidence; a review. Respirology. 2012; 17:887–98. [PubMed: 22672711]
[33] Gilliland FD, Li Y-F, Saxon A, Diaz-Sanchez D. Effect of glutathione-S-transferase M1 and P1 genotypes on xenobiotic enhancement of allergic responses: randomised, placebo-controlled crossover study. Lancet. 2004; 363:119–25. [PubMed: 14726165]
[34] Islam T, Berhane K, McConnell R, et al. Glutathione-S-transferase (GST) P1, GSTM1, exercise, ozone and asthma incidence in school children. Thorax. 2009; 64:197–202. [PubMed: 18988661]
[35] Li Y-F, Gauderman WJ, Avol E, Dubeau L, Gilliland FD. Associations of tumor necrosis factor G-308A with childhood asthma and wheezing. Am J Respir Crit Care Med. 2006; 173:970–76. [PubMed: 16456144]
[36] Chan JKW, Charrier JG, Kodani SD, et al. Combustion-derived flame generated ultrafine soot generates reactive oxygen species and activates Nrf2 antioxidants differently in neonatal and adult rat lungs. Part FibreToxicol. 2013; 10:34. [PubMed: 23902943]
[37] Li N, Xia T, Nel AE. The role of oxidative stress in ambient particulate matter-induced lung diseases and its implications in the toxicity of engineered nanoparticles. Free Radic Biol Med. 2008; 44:1689–99. [PubMed: 18313407]
[38] Nadeau K, McDonald-Hyman C, Noth EM, et al. Ambient air pollution impairs regulatory T- cell function in asthma. J Allergy Clin Immunol. 2010; 126:845–52. [PubMed: 20920773]
[39] Brunst KJ, Leung Y-K, Ryan PH, et al. Forkhead box protein 3 (FOXP3) hypermethylation is associated with diesel exhaust exposure and risk for childhood asthma. J Allergy Clin Immunol. 2013; 131:592–94. [PubMed: 23260754]