

Machine Learning Based Predictive Model for Closed Loop Air Filtering System

P Bhaskar¹, Farooq Sunar Mahammad^{1*}, A Hemanth Kumar², D Raj Kumar³, S M Abdul Khadar⁴, P Moin Khan⁵, P Veer Sekhar Reedy⁶

Department of Computer Science Engineering, Santhiram Engineering College, Nandyal

Email: bhaskar.cse@srecnandyal.edu.in

ABSTRACT

An automaton is showcased here to assess the high-quality of air with two elements: the quantity of dirt debris present inside the air and the temperature of the room. These two pieces of information are taken and a simple pinnacle-down mathematical version with a conditional clause is employed for assessing the satisfaction of air in a room. The mathematical model will consist of an easy equation. Destiny studies may be accomplished by optimising the mathematical model to enhance the accuracy of assessing the air quality.

Keywords: Linear Regression, Automaton, Prediction, Machine learning

1. INTRODUCTION

The maximum essential and existential want for human beings is smooth air to respire. With the appearance of industries and vehicles, the high-quality of air has appreciably long gone down. Air pollution is a combination of strong debris and gases. Within the air, vehicles ejecting chemical compounds from industries' dirt powder and mould spores may be suspended as elements. Ozone and gasoline are a prime part of air corruption in towns, whilst the upper region of the atmosphere forms air pollutants also referred to as smog. A few air pollutants are toxic except significant carbon dioxide (CO₂) emissions substantial quantities of carbon monoxide co hydrocarbon (HC) nitrogen oxide (NO_x) and suspended particulate matter (SPM) and different air pollutants are emitted from motor cars within the ecosystem inflicting. Severe environmental and health influences had to grow to be a want to put air purifiers in buildings now.

2. LITERATURE SURVEY

Primarily based on the inference from our reference base paper we got here to realize that “the urban environmental issues are contributed by several elements, together with greenhouse gasoline emissions and air high-quality. Material combustion from motors, production flora, fossil gas intake, and plenty of other business activities are the number one assets of atmospheric pollution. Those noxious implications purpose hundreds of premature deaths, and millions of years of existence misplaced, and are a contributor to weather exchange, frequently in growing international locations. Therefore, pollution forecast and simulation earlier than occurring are the principle awareness of this paper. It facilitates dealing with pollution stages, and infrastructure set up of towns and underpins various environmental guidelines. Awareness of the characteristics of the atmospheric quality, along with the mathematical foundation for assessing pollution, lets in for constructing the pollutants evaluation models required to recognize air high-quality and controls. In this paper, an effective version for simulating and studying atmospheric high-quality the use of cellular automata is proposed.”[1]

Based totally on the overview of the literature we suggest an efficient software program model with the use of a linear regression algorithm.

3. WORKING

Proposed system:

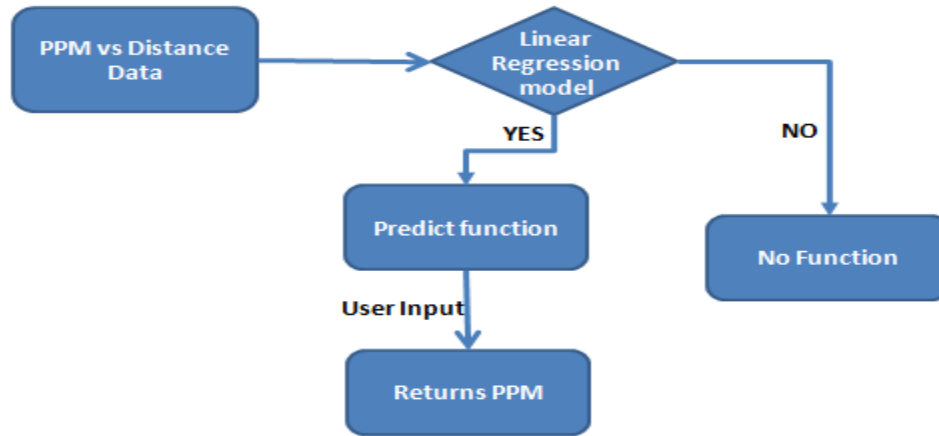


Figure 1 Block Diagram

In this project, the distance from the main road to the house and the air particulate dust measured in PPM are taken as two parameters to identify the proportionality between them. Linear regression is employed and the inverse proportionality between the distance from the main road and the dust particulate measure in PPM. Linear Regression is a regression machine learning algorithm. In this project, two-dimensional Linear Regression is employed. That is, one input parameter and one output parameter is taken and the given dataset is plotted in a 2D graph. The output model of Linear regression is the best fit line, which is the line with ' $y = mx + c$ ' equation passing closest to all the given points of the dataset. The model of this machine learning algorithm with just one input and one output parameter can be described with the value of the slope ' m ' and the value of the y-intercept ' c '. This algorithm maps the proportionality, be it direct or inverse, between the input and the output parameter.

Linear Regression Model:

“This method showcases the courting among the impartial linear variable (X-axis) and the structured variable (Y-axis). It is applied to get the value based on the value of another number. If there's a solo enter number (x), it's uncomplicated regression. However, if there are a couple of entering digits, it is multiple regression. This method version offers a sloped straight line describing the connection inside the numbers.” [2]

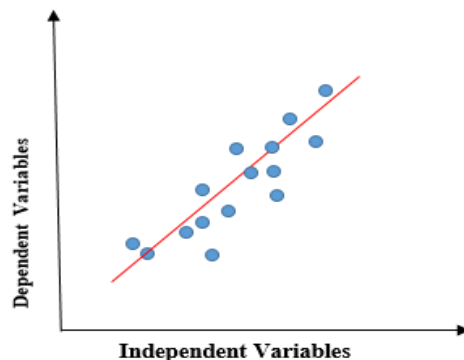


Figure 2 Linear Regression graph

“The above graph provides linear courting among the structured variables and free variables. Whilst the value of x (free variable) will increase, the value of y (structured variable) is likewise increasing. The red line is referred to as

the high-quality match straight line. primarily based on the given statistics factors, we attempt to plan a line that models the points the satisfactory.”[2]

4. ARTIFICIAL INTELLIGENCE

The algorithmic glide provided in this paper falls underneath the field of computing Intelligence and machine studying. supplied under is a definition of these technologies.

“Robotic or artificial intelligence (AI) is intelligence tested via machines, unlike the natural intelligence displayed with the aid of humans and animals, which entails cognizance and emotionality. The difference between the previous and the latter classes is frequently found via the acronym selected. Robust AI is commonly labelled as artificial general intelligence (AGI) whilst trying to emulate 'natural' intelligence was referred to as artificial biological intelligence (ABI). Main AI textbooks outline the sphere because of the study of sensible retailers: any device that perceives its surroundings and takes moves that maximise its risk of achieving its desires. Colloquially, the term artificial intelligence is regularly used to explain machines that mimic cognitive features that people partner with the human thoughts, which includes mastering and trouble fixing.”[3]

“As technologies end up increasingly more successful, duties taken into consideration to require intelligence are frequently eliminated from the definition of AI, an occurrence called the AI impact. A quip in Tesler's rule state AI is anything that hasn't been executed yet. As an instance, optical individual recognition is often excluded from matters taken into consideration to be AI, having emerged as a habitual technology. Modern-day system talents typically categorised as AI encompass efficiently expertise in human speaking abilities, competing at the highest degree in strategic recreation structures (consisting of chess and move), and additionally imperfect-records video games like poker, self-using vehicles, smart routing in content material transport networks, and military simulations.”[3]

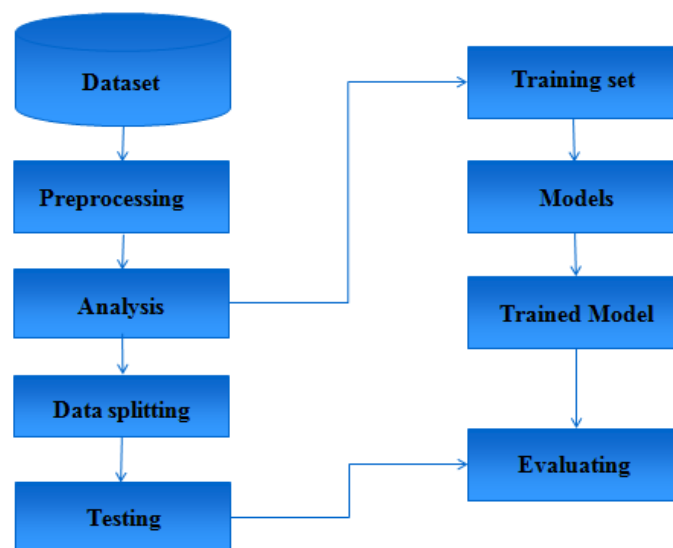


Figure 3 Flow Chart

“AI studies have been divided into subfields that frequently fail to talk with each different. those sub-fields are primarily based on technical issues, along with precise desires (e.g. robotics or system studying), the usage of unique equipment (good judgment or synthetic neural networks), or deep philosophical variations. Sub-fields have additionally been based totally on social elements (particular establishments or the work of specific researchers).”[3]

“Within the twenty-first century, AI techniques have skilled a resurgence following concurrent advances in pc strength, big quantities of records, and theoretical expertise; and AI strategies have to end up an essential part of the technology enterprise, assisting to clear up many difficult troubles in computer science, software engineering and operations studies.”[3]

5. MACHINE LEARNING

“Machine learning(ML) involves the research of computing algorithms that improve themselves over the duration as a consequence of expertise plus information. It’s considered an element of computer intelligence. Artificial learning algorithms create a framework relying on retraining information to generate forecasts or judgments despite having to get specifically configured to accomplish it. Machine intelligence algorithms were utilised across a vast range of applications, including healthcare, spam filtration, audio identification, and desktop sight, where developing algorithms to perform those required jobs is complicated or impossible.”[4]

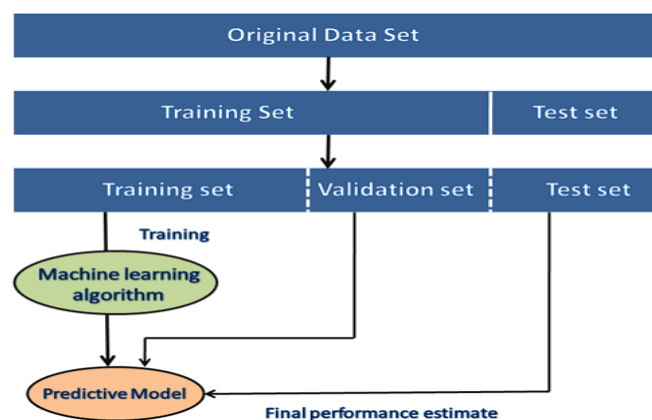


Figure 4 Machine learning dataset

“Machine intelligence and computing analytics are intimately linked, which concentrates upon utilising computers to make forecasts; nevertheless, nearly most computer intelligence involves statistical analysis. The discipline of machine intelligence benefits from computational efficiency research since it provides tools, concepts, and applicable fields. Data extraction is a similar branch of research that focuses on uncontrolled learning for interactive data processing. Computer intelligence is also known as anticipatory analytics when intelligence is used to solve commercial challenges.”[4]

5. INDUSTRY 4.0

“This same Fourth Economic Transformation (4IR aka Industrial 4.0) involves the continued digitization of old commercial and international operations, utilising new intelligent systems, according to several websites. Regarding increasing robotics, massive device connectivity (M2M) plus the network interconnected objects (IoT) are combined enhanced personality and interaction, as well as the development of intelligent technologies which can assess and resolve problems without any requirement of living thing interaction.”[5]

Automation under Industry 4.0 has a particular schema or pattern at its outset. Presented below is how automation in the mass production industry as well as consumer level products are built in today’s technological era.

The schema presented in Figure 5 has a lot of other components involved but the generic outline of it stands justifiable for all kinds of automation today.

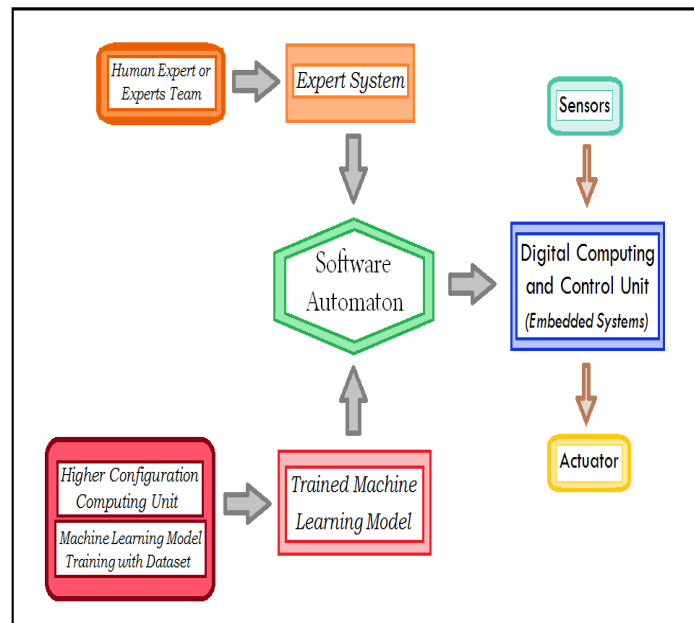


Figure 5 Schema of Automation

The software automaton of the conventional automation model, which is the status quo, was built by a human expert or a team of human experts till now. With the advent of machine learning technology, the software automaton was not fully directly designed by human experts. The human experts build the machine learning software and give the real-world data set as training information. Within the formats of a numerical framework, computer intelligence software finds statistical connections between the intake and outcome variables of the database. This mathematical model can be downloaded as a working software module to other electronic computing devices. This mathematical model is referred to as the ‘trained machine learning module’. The software automaton of all the current digital embedded devices is a mathematical model that gives a numerical output for a numerical input based on arithmetic and logical conditions. This software automaton, as explained above can be either directly developed by a set of human experts by means of setting the boundary conditions themselves based on observation and requirement or can be downloaded as an executable module from machine learning training systems that are trained with the relevant dataset. In whatever way the software automaton is developed, it can be loaded onto the relevant embedded computing module that can be used for either sensor-based closed-loop automation or open-loop automation.

The technological components of Industry 4.0 includes IoT, augmented reality, virtual reality, cloud computing, 3D printing, big data analytics, networking, data security, human-machine interaction and others. IoT is a very effective way to collect real world data. Sensors integrated with data acquisition and transmission systems can be placed anywhere and the collected data can be pre-processed if required and used as datasets to train machine learning models.

Cloud computing is employed for the optimised utilisation of computing resources. There are many third-party vendors like Google and Amazon which are very reliable in terms of data security and speed of computation. These services offer companies and organisations a cheap and reliable way to leverage digital intelligence as well as machine learning to their advantage.

Analytics on big datasets is the set of technological components involved with collecting, collating and managing large quantities of data for analytics and decision making. When so much data is involved, especially with third party service providers, data security plays an important role.

One of the paramount concerns about Industry 4.0 is the unemployment it can create due to powerful automations. The field of human-machine interactions and co-working has been a very developing field now to mitigate the above mentioned problem.

6. RESULTS AND DISCUSSION

Distance from Main Road	PPM Count of Air Dust
21	908
34	621
45	543
56	501
63	432
67	400
70	342
75	299
76	206
79	200
84	199
100	203
106	197
112	154
121	123

Figure 6 Input Sample Dataset

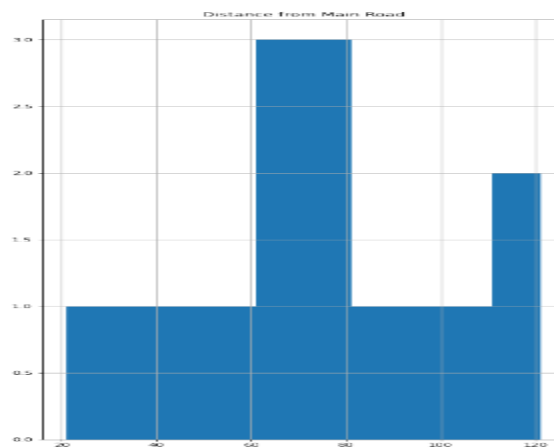


Figure 7 Distance from main road

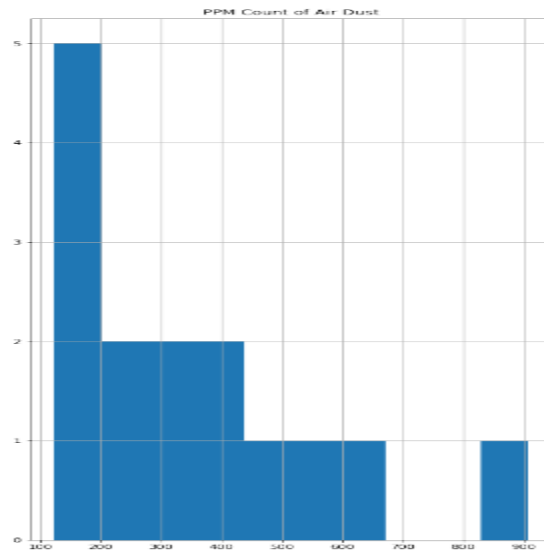


Figure 8 PPM count of air dust

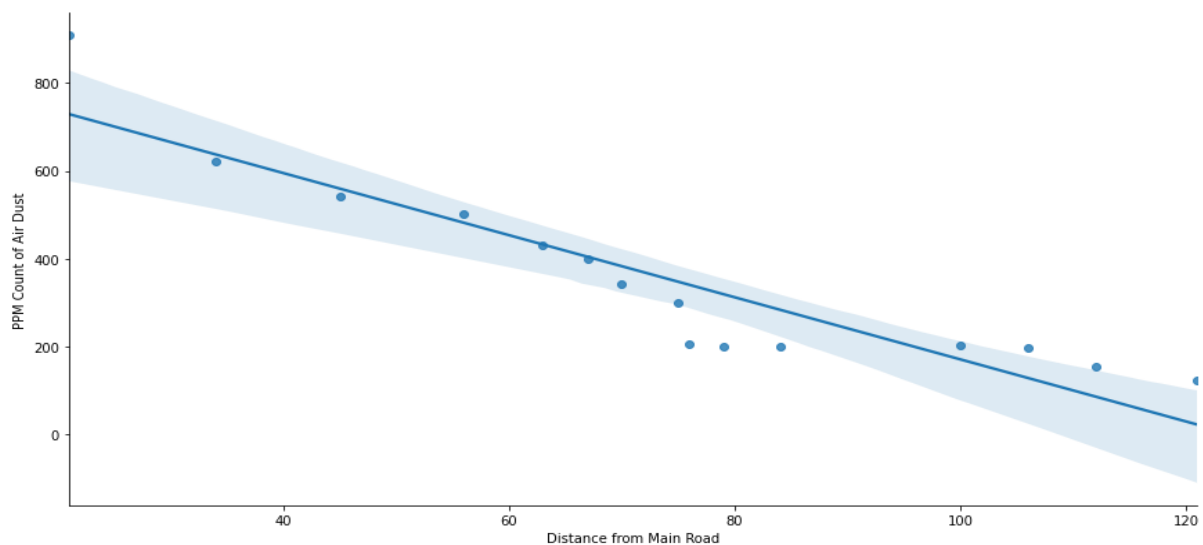


Figure 9 Output graph

This machine learning based predictive model is implemented employing Python programming language. The relevant library files have been included for execution of the code. The dataset was given as input and the predictive model obtained. The prediction model was tested and the results were satisfactory. The output graph of the machine learning model has been presented above. In this project, a Linear Regression algorithm is employed. The two parameters are the distance from the main road and PPM Count of air dust. The Linear Regression model maps the inverse proportionality between the two above mentioned parameters.

7. CONCLUSION AND FUTURE WORK

The software automaton presented here can be employed in all air purification systems that are integrated with IoT modules. The PPM measured in the air filter and the position of the unit away from the main road will either validate the trained machine learning model presented here or it will debate from it. If it deviates, it means that new parameters are required to make the model more effective. Automation under Industry 4.0 has a particular schema

or pattern at its outset. The software automaton of the conventional automation model, which is the status quo, was built by a human expert or a team of human experts till now. With the advent of machine learning technology, the software automaton was not fully directly designed by human experts. The human experts build the machine learning software and give the real-world data set as training information. The machine-learning software program identifies the sample between the input and the output parameters of the dataset in the shape of a mathematical version. This mathematical model can be downloaded as a working software module to other electronic computing devices.

REFERENCES

- [1] Hiep Xuan Huynh, Phuoc Thanh Luu, Huong Hoang Luong and Nghia Duong Trung, “Analysis of Atmospheric Quality based on Cellular Automata Simulation”, February 2020, Conference: ICIIT 2020: 2020 5th International Conference on Intelligent Information Technology
- [2] <https://www.analyticsvidhya.com/blog/2021/06/linear-regression-in-machine-learning/>
- [3] https://en.wikipedia.org/wiki/Artificial_intelligence
- [4] https://en.wikipedia.org/wiki/Machine_learning
- [5] https://en.wikipedia.org/wiki/Fourth_Industrial_Revolution