

# Enhancing Linear Regression Models to Predict Air Pollution using Internet of Things(IoT)

Mary Subaja Christo<sup>1</sup>, Uma Priyadarsini<sup>2</sup>, Hridya Venugopal<sup>3</sup>

<sup>1</sup>Associate Professor, <sup>2</sup>Assistant Professor(SG), <sup>3</sup>Assissant Professor,  
<sup>1,2,3</sup>Department of CSE, Saveetha School of Engg.

Saveetha Institute of Medical and Technical Sciences,Chennai,Tamilnadu

<sup>1</sup>marysubajachristo.sse@saveetha.com, <sup>2</sup>umapriyadarsini@gmail.com, <sup>3</sup>hridyavenugopal.sse@saveetha.com,

## Article Info

Volume 83

Page Number: 3238-3241

Publication Issue:

May - June 2020

## Abstract

Air contamination has critical impact on the centralization of constituents in the climate prompting impacts like a worldwide temperature alteration and corrosive downpours. To keep a calculated distance beginning such lopsided characteristics of the nature, an air contamination checking is most extreme significant. Right now answer for observing air contamination levels in condition is proposed and we foresee Air Pollution. In light of the past put away records, we can anticipate the future contamination level by utilizing direct relapse rule. The framework uses city transports to gather toxin gases, for example Smoke. The information displays the toxin levels then their neighborhood air value gauges. Here effectively plan android web server which displyas the continuous air contamination information/air poisons on site page which can be open from anyplace on the planet along these lines, here have structured circuit which makes restorative move on the expansion of air contamination on the specific edge esteem. The execution is tried for two parameters like smoke concerning the typical levels which give a power over the contamination observing to create the earth brilliant.

## Article History

Article Received: 19 August 2019

Revised: 27 November 2019

Accepted: 29 January 2020

Publication: 12 May 2020

**Keywords:** *IoT, Raspberry pi, Android Studio, Linear Regression, Prediction.*

## 1. Introduction

Air Pollution rose in numerous pieces of the world because of touchy modern development. Street transport is likewise one of the significant givers of air contamination, which add to environmental change that has hazardous household and worldwide outcomes. Present advancements in innovation predominantly center around controlling and observing of various exercises. These are progressively developing to arrive at the human needs. A gainful checking framework is required to screen and evaluate the conditions in the event of beating the bolstered degree of parameters (e.g., smoke). How much condition gets affected is noted and implied to the customer through the compact application.

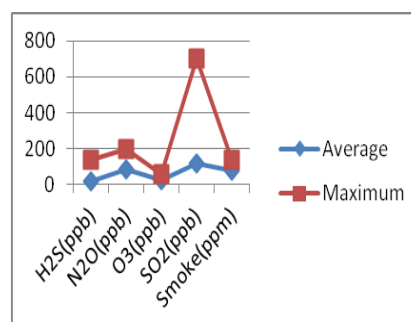


Figure 1: Different Pollution gases in atmosphere

Checking framework relies upon the sort of the information assembled by the sensor gadgets. At first the sensor gadgets are sent in condition to identify the

parameters (e.g., smoke) while the information obtaining, calculation and checking activity. Sensor gadgets are set at various areas to gather the information to anticipate the conduct of a specific region of intrigue. The primary point of this paper is to structure and actualize a proficient checking framework through which the necessary parameters are observed remotely utilizing web and the information assembled from the sensors are put away for future reference.

## 2. Existing System

These days various sullyng checking system are arranged by considering different common parameters. Existing structure form is obtainable in figure 3 which is used Zigbee [2] method remote antenna frameworks to screen objective and normal setting through an enormous number of utilization in dissimilar field. In this sensor centers clearly talked by touching centers passed on top of the article of attention which avoid the usage of multifaceted coordinating estimation anyway close by figurings are immaterial.

RFID [3] is a techniques for taking care of and recouping facts from side to side electromagnetic broadcast to a RFID impeccable composed path. This is normally worn to name then tract things inside markets with manufactories. This structures include 2 guideline sections: names then perusers. This label has a distinctive verification (Identifier) quantity and a recollection to supplies further information, for instance, producer, thing type, and normal components, for instance, temperature, tenacity, etc.

The peruser can scrutinize just as form data to marks by methods for remote transmissions. In an ordinary Radio Frequency Identifier purpose, marks added otherwise surrounded addicted to things with the aim of requiring ID or following. RFID marks can be gathered into three critical classes by their ability source: dynamic names, idle names, and semi idle (semi-dynamic) names.

Flexible [4] phones or propelled cells which are accessed with circuits be second-hand for influence in public with how convenient advancement must be used for regular guaranteeing, distinguishing and to affect without a second to save information to make improvements and exercises environmental pleasant. PDA sensors be sent then used lying on built-up domains intended for checking then it be arranged keen on noteworthy classes, participatory identifying where customer is honestly included and clever 2 distinguishing anywhere customer is excluded, yet its requirement joins influence and motionless in sequence dealing with

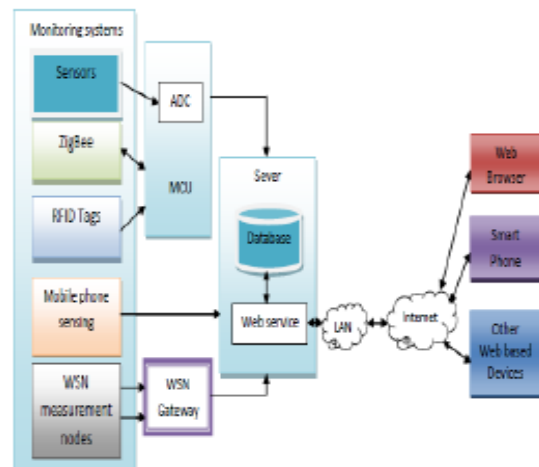


Figure 2: Existing System Model

A Wireless Sensor Network [5] comprise of numerous reasonable remote sensors, which are equipped for gathering, putting away, preparing ecological data, and speaking with neighboring hubs. Previously, sensors are associated by wirelines.

## 3. Related Work

A bit of the assessment works finished for checking the tainting parameter inside a meticulous district of excitement used creation the earth insightful around there, different frameworks and systems which were used previously. Insightful surroundings monitor using IEEE 802.11 standards frameworks [1] – This effort they be essentially revolve around the creation the municipality condition stylish, by passing on remote method organizes within wherever every one through the capital and touching open carrying structure transports and cars.

Beside gettin to each and every one the antenna frameworks, biological practices be accumulated because a spouting statistics stand to perceive the regular surroundings. In this framework give the watching statistics starting fixed center point sent in metropolis to the adaptable centers on top of open carrying transports and vehicles. Toward a Green Campus with the Internet of Things – The Application of Lab Management this investigation work gets "Web of Things" and executes an idea of imperativeness saving by proper organization of PCs and constrained air frameworks. The structure and the model of the system is explained in [5]. At this time the substance of IOT Things are PCs and atmosphere control frameworks.

Wireless Sensor Network, Intrenet of Things – based on their extension of Homes and their things[6] – In This module fundamentally hopes in develop and make strong, beneficial, versatile, traditionalist, ceaseless and pragmatic wellbeing sensor frameworks for insightful building model. In this antenna and actuator center points reliant lying remote frameworks organization progressions are sent into the home condition. These centers make steady statistics associated to the thing

procedure and improvement within the residence. Additional loosens up the adroit residence organization towards clever structures and structure the arrangement issues linked towards the sharp structure condition.

#### 4. Proposed System

The anticipated gadget is used to checking smoke range within the air towards create the earth wise. In this system replica is appeared in Figure 3 increasingly versatile and flexible to screen the ecological values.

Main objective is creating air contamination observing is to improve personal satisfaction by utilizing innovation to improve the effectiveness of administrations and address inhabitants' issues. Consider a territory that is being overviewed for evaluating how much the region is influenced by contamination. The constituents of air alongside its extent are determined and on the off chance that it higher than typical, at that point the authorities are insinuated about it. At that point the individuals are cleared to a sheltered spot.

The MQ-3 is a combustible gas and burn antenna distinguishes the groupings of flammable chat noticeable all around, yields its perusing similar to a simple power. The sensor can gauge convergences of combustible gas of Three Hundred to Ten Thousand ppm. This gas antenna is touchy to LPG, I-butane, propane, methane, liquor, Hydrogen and gas. They are utilized in gas spillage distinguishing supplies in family and industry and in compact gas indicator.

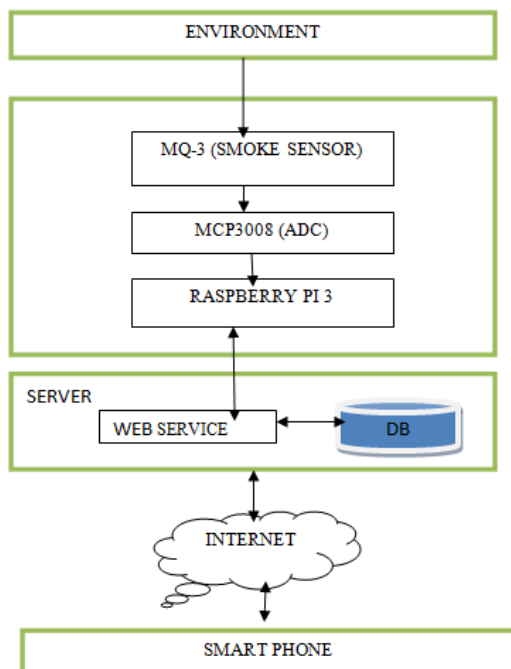


Figure 3: Proposed System Model

Since MQ-3 yields its perusing as a simple voltage yet raspberry pi just procedure advanced qualities. So we have to change over simple qualities to relating computerized values before going to raspberry pi. For this

change, MCP3008 (simple to computerized convertor) is utilized. Raspberry Pi is a minimal effort, fundamental PC that was initially planned to assist prod with intriguing in processing among school-matured kids. The Raspberry Pi is contained on a solitary circuit board and highlights ports for:

- HDMI
- USB 2.0
- Composite video
- Analog sound
- Power
- Internet
- SD Card

Raspberry pi process the digital values of smoke levels and compares it with actual level of smoke level is same as in the impression. If it goes not in the limit the normal level, then it will be dejected to the user via smart phones. In addition, the values of smokes store in the memory device for the future reference through internet service

**MQ-3 Methodology:**The MQ-3 which is combustible gas and burn feeler detect the concentration of explosive gas in the sky This application helps to display the real time Particular Matter values in the air (PM2.5 levels). This system provide some machine learning methods to find and expect the PM2.5 levels depend on a statistics set developing of special condition in a main capital. The responsibilities of proposed systems are

- (i) Detects the level of Matter Values on given impressive
- (ii) Predicts the level of Matter Value for a particular date

From Fig 4 shows the LR is used to notice whether the facts sets are stained or not and with help of auto regression to expect prospect standards of Pollution dependes on the earlier effluence values. The main goal is to predict air pollution level in City with help of existing data sets

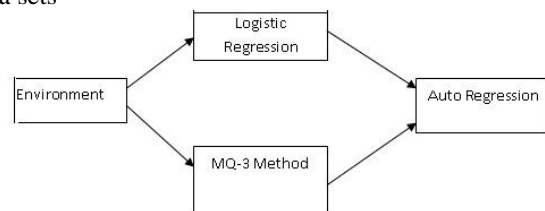


Figure 4: Prediction System Model

Two main stages are there in the system

1. Education Stage: Existing data used to train the system
2. Testing Stage: Based on the input and output the systems are used to tested the data

Finally the accuracy is calculated and the output shows the high levels of accuracy. Datas are used to train the model and test the model

## 5. Results and Discussion

As shown in the fig.5 when Checked with other machine learning replica applied on the data values, Logistic Regression go with the best for this system with the mean accuracy and standard deviation accuracy to be 0.98859 and 0.000512 respectively. For this reason, Logistic regression can be used to clearly classify and differentiate the PM2.5 value produced based on the given sample impressive conditions to be infected or not.

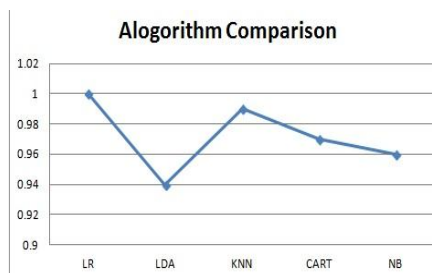


Figure 5: Comparison of Algorithm

Autoregression applied on time arrangement informational collection to anticipate the PM2.5 esteem 7 days preceding the present date, delivered the Mean Squared Error(MSE) to be 27.00. MSE can be diminished by diminishing the distinction between the present date and the date on which the estimation of PM2.5 is to be anticipated

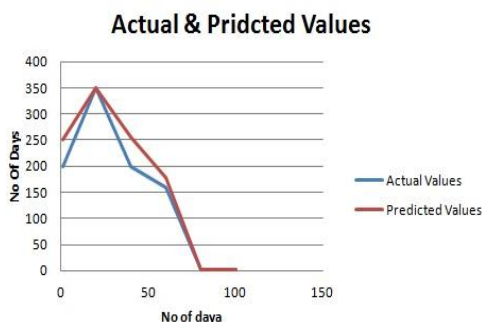


Figure 6: Actual and Predicted Method

As in the above fig. 6 the chart demonstrates the real values in blue shading and the anticipated qualities by red. The diagram was plotted for 100 qualities that were part into test information.

## 6. Conclusion

The framework uses city transports to gather toxin gases, for example smoke. This information displays the toxin level and neighborhood air value gauges. Here effectively plan android web server which displays the continuous air contamination information/air poisons on site page which can be open from anyplace on the planet along these lines, here have structured circuit which makes restorative move on the expansion of air contamination on the specific edge esteem.

## References

- [1] Platt, Nathan, and Dennis DeRiggi. "Comparative investigation of Source Term Estimation algorithms using FUSION field trial 2007 data: linear regression analysis." *International Journal of Environment and Pollution* 48, no. 1-4 (2012): 13-21.
- [2] Alexeeff, S.E., Carroll, R.J. and Coull, B., 2016. Spatial measurement error and correction by spatial SIMEX in linear regression models when using predicted air pollution exposures. *Biostatistics*, 17(2), pp.377-389.
- [3] Rosenlund, M., Forastiere, F., Stafoggia, M., Porta, D., Perucci, M., Ranzi, A., Nussio, F. and Perucci, C.A., 2008. Comparison of regression models with land-use and emissions data to predict the spatial distribution of traffic-related air pollution in Rome. *Journal of exposure science & environmental epidemiology*, 18(2), pp.192-199.
- [4] Pantazopoulou, A., Katsouyanni, K., Kourekremastinou, J. and Trichopoulos, D., 1995. Short-term effects of air pollution on hospital emergency outpatient visits and admissions in the greater Athens, Greece area. *Environmental research*, 69(1), pp.31-36.
- [5] Magari, S.R., Hauser, R., Schwartz, J., Williams, P.L., Smith, T.J. and Christiani, D.C., 2001. Association of heart rate variability with occupational and environmental exposure to particulate air pollution. *Circulation*, 104(9), pp.986-991.
- [6] Ruidavets, Jean-Bernard, Maxime Cournot, Sylvie Cassadou, Michel Giroux, Mariam Meybeck, and Jean Ferrières. "Ozone air pollution is associated with acute myocardial infarction." *Circulation* 111, no. 5 (2005): 563-569.
- [7] Pandey, Gaurav, Bin Zhang, and Le Jian. "Predicting sub-micron air pollution indicators: a machine learning approach." ; *Environmental Science: Processes & Impacts* 15.5 (2013): 996-1005
- [8] Dixian Zhu, Changjie Cai, Tianbao Yang and Xun Zhou: A Machine Learning Approach for Air Quality Prediction: Model Regularization and Optimization. *Big data and cognitive computing* [2018].