International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-8 Issue-3, September 2019

Novel Framework for Analyzing Air Quality using MatLab

G. Radhika, A. Padmapriya

Abstract*: Nature of the air in city and urban regions is the most significant factor that legitimately impacts the frequency of infections and diminishes the personal satisfaction. Taking suitable choices in an opportune period relies upon the estimation and examination of the parameters of the ongoing air quality checking. On the other hand air contamination is ecological and a social issue. Air contamination is one of the biggest natural wellbeing dangers in the world today. The utilization of multi-parameter air quality observing frameworks makes it conceivable to do an itemized level investigation of real poisons and their sources. These air quality observing frameworks are significant segments in many shrewd city ventures for checking air quality and for controlling the primary poison fixations in urban zones. In this research work a methodology for practical estimation of air quality is proposed. This application has been tried in the city and the estimation was contrasted and the yield information of the neighborhood ecological control expert stations. The results of the performance analysis demonstrate that this methodology can be utilized as an affordable option in contrast to the expert evaluation frameworks. In this research work an investigation on the contamination by traffic framework utilizing dataset with grouping strategy through MatLab.

***Keywords* : Air quality, Clustering, K-Means, ANN, KNN, Accuracy, Floyd Warshall.**

I. INTRODUCTION

Air quality administration is an unpredictable action of incorporated science. There is a requirement for mix of science independent of control. A ten point program is acknowledged and received in for fruitful usage of air quality administration program by the city administration of the European countries.

The ten points are listed below:

(1) Setting the objective for air quality objective with monetarily solid observing component having powerful database on benchmark information, receptor based information and pattern line information for city of concern;

(2) Good quality science based checking instrument with linkage with wellbeing, condition, sway at receptor level, cost examinations including weight to the general public;

(3) Understanding and learning on scientific sciences and barometrical sciences including the data on the nature, sources and effects of air contamination;

(4) Shared comprehension of the issue among significant partners (contamination sources, political pioneers, ecological authorities, open premium gatherings, residents);

(5) Active inclusion of partners and shared pledge to improve air quality;

(6) Shared weight of control crosswise over contributing sources in the city and districts;

(7) Leadership at nearby level, territorial level, national level, to empower support, set reasonable degrees of control, and resolve clashes between partners;

(8) Defined jobs for each degree of government included including characterizing the order;

(9) Strong punishments, clear responsibility, and severe implementation at the national, state, and neighborhood levels, and so on

(10) The city administration has to organize banter on above expressed issues at ordinary interims and at all levels (neighborhood and local).

By doing and placing every one of the issues practically speaking including characterizing neighborhood with deliberate observing and quality information age. It is trusted that the frameworks and motivation behind Continuous Ambient Air Quality will achieve amazing achievement. It might be recollected that the information created are fundamentally implied for open by and large, utilized for gauging, early caution framework including working of certainty among open all in all, controllers and different urban zone the executives association. It is likewise important to comprehend that the achievement of ecological administration is to a great extent subject to the interest of open by and large.

Kids, the older, and individuals with heart or lung ailment, diabetes, minority and low salary networks are especially defenseless against antagonistic wellbeing results from introduction to air contamination, including cardiovascular ailment, asthma and other respiratory ailments, and malignancy. Ongoing proof recommends that air contamination is additionally connected to higher danger of diabetes, chemical imbalance, and lower IQ.

Manuscript published on 30 September 2019
* Correspondence Author
M. Radhika, M. Phil Scholar, Department of Computer Science, Alagappa University, Karaikudi, India.
Email: radhikasolai95@gmail.com
Dr. A. Padmapriya*, Associate Professor, Department of Computer Science, Alagappa University, Karaikudi, India.
Email:drapadmapriyacse@gmail.com, padmapriyaa@alagappauniversity.ac.in

This paper comprises of five sections. The next section will elaborate the related works of the environmental domain especially dealing with quality of air. The section 3 describes the proposed work. The section 4 illustrates the air quality analysis through charts. The section 5 explains the results and discussions. The section 6 concludes the finding.

II. LITERATURE REVIEW

This research work is motivated by the following existing works.

Kok Mun Ng et al [17] developed a MyRIO-LABVIEW based remote air quality monitor. It is designed to monitor the quality of air within buildings. The monitoring is done for three selected locations in Shah Alam. In addition to MyRIO, Arduino Mega and MQ-7 sensors are used for this experimental setup. The CO rate is examined to assess the quality of air.

The authors [18] proposed framework gives minimal effort, low power, smaller and profoundly precise framework for checking the earth with the committed sensors remotely from wherever in this world. An ideal tradeoff among exactness and cost is accomplished by utilizing single board minicomputer Raspberry pi and fitting sensors prompting a well-grounded framework.

Incorporation of IOT with Sensor hubs with the assistance of Raspberry Pi for Air Quality Monitoring gives a successful path than the methodologies that were recently utilized. Sensor web hub is proposed with business gas sensors for distinguishing the gases like CO, CO2 and so forth to screen both indoor and outside air quality.

Sanju Thomas et al [19] developed a hybrid system in package (SiP) acoustic micro sensor. It is used for real time monitoring of particulates. Through the virtual impactor the air is drawn in. The mass of separated PM2.5 particles that are deposited on SMR surface is measured. It is experimented in optimal lab conditions as well as real-time outdoor conditions. The proposed system is efficient in detecting the minute particles. It is a low cost real time monitor the uses sensors based on CMOS-SMR expedients.

Exact assessment and approval of estimations done utilizing minimal effort sensors is significant. Studies have demonstrated that the estimations from the present ease sensors are of fluctuating quality. The sensors, which are expected to quantify air contamination as gas and particles, can for instance be influenced by different climate conditions (wind speed, temperature, and stickiness) or be "cross-touchy" and along these lines unfit to recognize precisely between a few unique toxins. NILU has the important framework to assess minimal effort sensor execution against reference instrumentation in both controlled research center conditions and genuine field conditions.

Shelestov et al. [20] have done a detailed study of various satellite products for monitoring air quality in Kyiv smart city. ERA-PLANET project is aimed to develop a screening model for assessing the quality of air in urban areas. Kyiv is considered for this pilot study. Its ultimate focus it to strengthen the European participation in the Group of Earth observation (GEO). The observations from the Satellites and from the ground are analyzed in the research work.

The authors of [21] proposed a monitoring system named as AirSense. It is an opportunistic crowd sensing based screening system. The collected data are stored in the cloud and the features like individuals pollution footprints, local air quality and index map of air are visualized through smartphones.

Urban street dust monitoring system based on ZigBee was proposed by Hong He et al [22]. The proposed system helps in monitoring the particulate matter on the sides of the roads. It is cost effective and at the same time it has less maintenance. The quality of air in Peru was monitored using three locations in Lima city [23]. By considering uncertainty in the air quality monitoring, it offers better results. The findings of this work will be useful for the city authorities in making decision about the air quality.

III. PROPOSED FRAMEWORK

This framework to analyze the quality of air is shown in the following figure 1. The frame work consists of three phases namely (i) Clustering, (ii) Analysis of the air quality and (iii) Performance analysis of the algorithms used in the framework.

Fig. 1.Workflow of the proposed framework

The phases of the proposed framework will be explained in the following sections.

A. Phase I : CLUSTERING

This is the first phase of the framework. In order to assess the quality of air, the collected air dataset are to be clustered first. The proposed framework employs different algorithms to achieve the same. They are described below.

K Means Clustering

The K means clustering is one among the most commonly used clustering algorithm. Here the labels to the clusters are identified and they are called as centroids. The groups are formed based on the nearness to the centroid. For finding the nearness, Euclidean distance is used here.

Euclidean Distance = $\sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2}$

The clusters thus formed can be refined using subsequent iterations.

Algorithmic steps for k-means clustering

Step 1. The data should be preprocessed, cleaned and transformed for clustering

Step 2. Choose the initial value for K

Step 3. Label the data based on the centroids

Step 4. Calculate the Euclidean distance and assign the nodes to the labels such that the distance from the centroid and the nodes is less

Sept 5. Refinement of clusters by repeating steps 4 and 5.

Artificial Neural Network

Artificial Neural Network is an intelligent computational model. It is usually applied for domains such as bioinformatics, machine learning model, biological networks etc.

The processing is done in the three layers. The input layer is the layer which provides the raw data. The next layer is the hidden layer in which deals with the weights among the connections between the activities. The outputs are given by the output layer.

Algorithmic steps for ANN :

Step1. The data from input layer is given to hidden layer.
Step2. Input values from input layer are used and modified using some weight value and sent to output layer.
Step3. The value is again modified by some weights from connection between hidden and output layer.
Step4. This information is processed and output layer gives final output. Finally, this output is processed by activation function.

Floyd–Warshall's Algorithm

This algorithm will identify the optimal path between any pair of vertices. The distance between the vertices are represented using weights. By contrast to Dijkstra, it has positive as well as negative weight edges.

Steps for Floyd – Warshall's Algorithm:

Step 1. For all vertices assign the length of the shortest path as infinity.
Step 2. Find all the intermediate nodes between the source and destination.
Step 3. Among the choices of intermediate nodes, choose the one having minimum weight.
Step 4. Iterate the above two steps until the path with minimum weight is identified between all pair of vertices

K Nearest Neighbor Method

The KNN algorithm is one of the best and commonly used for clustering and classification. In this method, each sample value fit to the test dataset and it is classified according to the nearest k sample based on the training data. The class numbers values obtained from k sample values, the maximum number is determined from class samples. The distance measurement calculated by using Euclidean distance measure.

Algorithmic steps for KNN :

Step1. Choosing k value: The selection of k value completely based on user idea
step2. Distance measurement: The distance calculation measured by using Euclidean distance method.
step3. Training phase: stores all training dataset A = (ai, bi), i=1... n, here ai training dataset, bi is coherent class and n is number of training dataset.
step4. Testing phase: calculates distance between training data and new feature vector.
step5. The founded distance is arranged in one by one and the minimum k distance value is taken.
Subsequently, the unidentified sample values are selected most relevant to the class from k nearest neighbor algorithm and it is used to find the real value from unidentified sample values.

B. Phase II : DATA ANALYSIS

In this phase the quality of the air is analyzed with the quantity of NO_2 and SO_2 contents in the air molecules. Their emissions rate are presented in this phase

C. Phase III : PERFORMANCE ANALYSIS

The performance of the algorithms is compared in terms of accuracy and time taken.

The next section will elaborate the analysis of air quality with respect to NO_2 and SO_2.

IV. AIR QUALITY ANALYSIS

This section will present the results of air quality analysis. In this research work the analysis is done for two gases namely NO_2 and SO_2.
The presence of the gases, their frequencies, NO_2 and SO_2 expectation as indicated by the sun plate are analyzed. The results are given below.

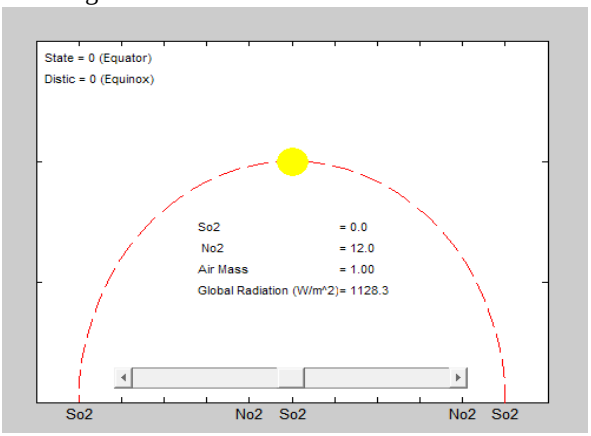

Fig. 2.Examination of both NO_2 and SO_2 gas in a specific locale alongside mass in air and worldwide radiation.

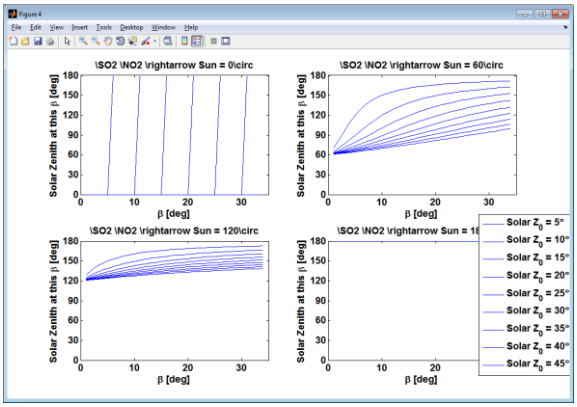

Fig. 3.Frequence of NO2 and SO2

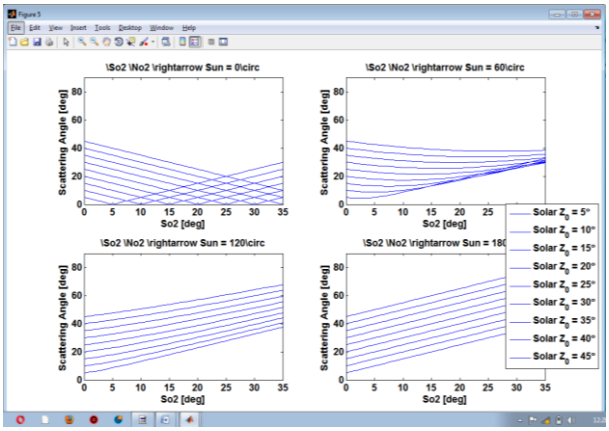

Fig. 4.Representation of Scatering Angle

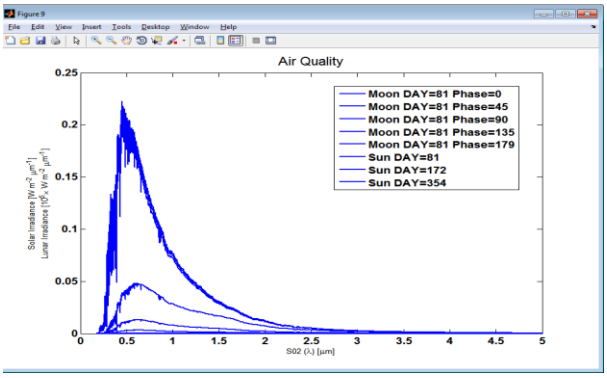

Fig. 5.Accuracy of the Algorithms

V. RESULTS AND DISCUSSION

Air contamination is expanding quickly in the shrewd urban areas and effectively affects human wellbeing. The wellsprings of contamination are many including street traffic, mechanical gases and others. In this examination the proposed framework attempts to locate the most advantageous regions, which are appropriate for leaving, in the keen urban areas by utilizing K-means clustering.

A. Accuracy

The results are tabulated below. The K-Means algorithm offers better accuracy when compared to the others.

Table- I: Name of the Table that justify the values

Algorithm Used	Accuracy
ANN	69%
KNN	74%
Floyd–Warshall's Algorithm	86%
K - Means	92%

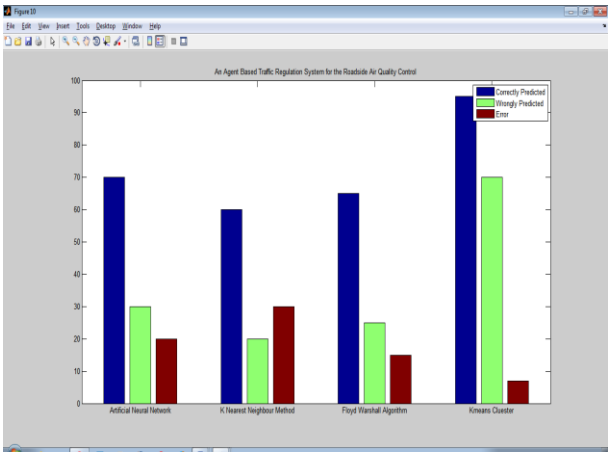

Fig. 6.Accuracy of the Clustering Algorithms

B. Time Taken

The time taken for execution of the four algorithms used in the proposed framework is tabulated below.

Table- II: Name of the Table that justify the values

Algorithm Used	Time taken (in seconds)
ANN	900ms
KNN	819ms
Floyd–Warshall's Algorithm	800ms
K - Means	778ms

The time taken for execution of K-Means is minimum when compared to the others

C. Time Complexity

The time complexity of the four algorithms used in the proposed framework is given below.

1) The time complexity of the Floyd-Warshall algorithm is $O(n^3)$
2) KNN time complexity is O(nd).
3) The k-means algorithm is known to have a time complexity of $O(n^2)$
4) The Artificial Neural Network algorithm is having O(n^3) complexity for n vertices.

VI. CONCLUSION

In this research work, the quality of the air is analysed using clustering algorithms. Here three different algorithms and Floyd Warshall algorithm are employed to cluster and assess the quality of air. Among the algorithms the K-Means clusters offers better accuracy. The impact of SO_2 and NO_2 levels in the air is clearly depicted through the different data analysis charts. The complexity of the algorithms used as well as their accuracy are analysed and presented. The K-Means algorithm gives higher exactness pace of 92% when compared to the other algorithms. The accuracy can be improved by using advanced wireless sensors for measuring air quality.

ACKNOWLEDGMENT

This article has been written with the financial support of RUSA – Phase 2.0 grant sanctioned vide Letter No.F.24-51/ 2014-U, Policy (TNMulti-Gen), Dept. of Edn. Govt. of India, Dt. 09.10.2018

REFERENCES

1. World Health Organisation, “Health effects of Particulate Matter,” 2013.
2. Hajjam, A., Wilson, J. C., & Pourkamali, S. (2011). Individual air-borne particle mass measurement using high-frequency micromechanical resonators. *IEEE Sensors Journal*, 11(11), 2883-2890.
3. Johnston, M. L., Kymissis, I., & Shepard, K. L. (2009, June). An array of monolithic FBAR-CMOS oscillators for mass-sensing applictions. In TRANSDUCERS 2009-2009 *International Solid-State Sensors, Actuators and Microsystems Conference* (pp. 1626-1629). IEEE.
4. EPCOS, “Design-guide for the SAW oscillator,” 2008.
5. S. V. Elardo, ‘‘‘Air pollution: Health-related problems V/S increase of transportation activities in an urban area,’’’ in 2017IEEE 9th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM), 2017, pp. 1-7.
6. L. L. Chen, J. Xu, Q. Zhang, Q. H. Wang, Y. Q. Xue, and C. R. Ren, ‘‘Evaluating impact of air pollution on different diseases in Shenzhen, China,’’ IBM J. Res. Dev., vol. 61, no. 6, p. 2:1-2:9, Nov. 2017.
7. A. Delgado and I. Romero, ‘‘Social impact assessment on a hydrocarbon proyect using triangular whitenization weight functions,’’ in 2016 IEEE Congreso Argentino de Ciencias de la Informática y Desarrollos de Investigación (CACIDI), 2016, pp. 1--6.
8. S. Liu and Y. Lin, Grey Systems: Theory and Applications. Berlin: Springer, 2010.
9. Vrajesh Sharma; Dr. Manju Bala,” A Credits Based Scheduling Algorithm with K-means Clustering” 2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC). DOI: 10.1109/ICSCCC.2018.8703201.

10. Shi Na, Liu Xumin, Guan Yong," Research on k-means Clustering Algorithm: An Improved k-means Clustering Algorithm", 2010 Third International Symposium on Intelligent Information Technology and Security Informatics. DOI: 10.1109/IITSI.2010.74.
11. Nuwan Ganganath, Chi-Tsun Cheng, Chi K. Tse," Data Clustering with Cluster Size Constraints Using a Modified K-Means Algorithm", 2014 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery. DOI: 10.1109/CyberC.2014.36.
12. K.Rajeswari, Omkar, Acharya, Mayur, Sharma, Mahesh, Kopnar, Kiran, Karandikar, "Improvement in k-Means Clustering Algorithm Using Data Clustering", 2015 International Conference on Computing Communication Control and Automation. DOI: 10.1109/ICCUBEA.2015.205.
13. Mingwei Leng, Xiaoyun Chen, Longjie Li," K-means clustering algorithm based on semi-supervised learning", 2008.
14. Liu Zhenpeng, Guo Ding, Zhang Xizhong, Wang Xu, Zhu Xianchao," An Improved Semi-supervised K-Means Algorithm Based on Information Gain", 2014 IEEE 17th International Conference on Computational Science and Engineering. DOI: 10.1109/CSE.2014.358.
15. Archana Singh Amity, Avantika Yadav, Ajay Rana," K-means with Three different Distance Metrics",2013 International Journal of Computer Applications.
16. Shruti Kapil, Meenu Chawla," Performance evaluation of K-means clustering algorithm with various distance metrics", 2016 IEEE 1st International Conference on Power Electronics, Intelligent Control, and Energy Systems (ICPEICES). DOI: 10.1109/ICPEICES.2016.7853264.
17. Kok Mun Ng, Muhamad Azri Haziq Mohd Suhaimi, Adizul Ahmad, N.A. Razak. "Remote Air Quality Monitoring System by Using MyRIO-LabVIEW", 2018 9th IEEE Control and System Graduate Research Colloquium (ICSGRC), 2018
18. Kumar, S., & Jasuja, A. (2017). Air quality monitoring system based on IoT using Raspberry Pi. *2017 International Conference on Computing, Communication and Automation (ICCCA)*, 1341-1346.
19. S. Thomas, M. Cole, F. H. Villa-Lopez, J. W. Gardner, J. Peters and J. Theunis, "A low-cost acoustic microsensor based system in package for air quality monitoring," *2016 IEEE SENSORS*, Orlando, FL, 2016, pp. 1-3.
20. A. Shelestov *et al*., "Air Quality Monitoring in Urban Areas Using in-Situ and Satellite Data Within Era-Planet Project," *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, Valencia, 2018, pp. 1668-1671.
21. Dutta, J., Chowdhury, C., Roy, S., Middya, A.I., & Gazi, F. (2017). Towards Smart City: Sensing Air Quality in City based on Opportunistic Crowd-sensing. *ICDCN*.
22. H. He, Y. Gao and Z. Zhang, "The urban road dust monitoring system based on ZigBee," *2016 Chinese Control and Decision Conference (CCDC)*, Yinchuan, 2016, pp. 1793-1796.
23. A. Delgado, P. Montellanos and J. Llave, "Air quality level assessment in Lima city using the grey clustering method," *2018 IEEE International Conference on Automation/XXIII Congress of the Chilean Association of Automatic Control (ICA-ACCA)*, Concepcion, 2018, pp. 1-4.

AUTHORS PROFILE

G. Radhika is currently pursuing M. Phil in the Department of Computer Science, Alagappa University, Karaikudi – 630003, Tamilnadu, India. Her area of interest is Data Mining.

Dr. A. Padmapriya is working as an Associate Professor in Computer Science of Alagappa University, Karaikudi, Tamilnadu, India. Her areas of interest include Communication Networks, Data mining and Information security. She has published many articles in reputed National and International journals.

.