International Journal of E-Health and Medical Communications
Volume 12 • Issue 2 • March-April 2021

Prediction of Environmental Pollution Using Hybrid PSO-K-Means Approach

Manish Mahajan, Graphic Era University (Deemed), India
https://orcid.org/0000-0002-5670-650X

Santosh Kumar, Graphic Era University (Deemed), India
https://orcid.org/0000-0002-1008-0804

Bhasker Pant, Graphic Era University (Deemed), India
https://orcid.org/0000-0002-6070-5104

ABSTRACT

Air pollution is increasing day by day, decreasing the world economy, degrading the quality of life, and resulting in a major productivity loss. At present, this is one of the most critical problems. It has a significant impact on human health and ecosystem. Reliable air quality prediction can reduce the impact it has on the nearby population and ecosystem; hence, improving air quality prediction is the prime objective for the society. The air quality data collected from sensors usually contains deviant values called outliers which have a significant detrimental effect on the quality of prediction and need to be detected and eliminated prior to decision making. The effectiveness of the outlier detection method and the clustering methods in turn depends on the effective and efficient choice of parameters like initial centroids and number of clusters, etc. The authors have explored the hybrid approach combining k-means clustering optimized with particle swarm optimization (PSO) to optimize the cluster formation, thereby improving the efficiency of the prediction of the environmental pollution.

KEYWORDS

Air Quality, Clustering, Data Mining, K-Means, Outliers, Particle Swarm Optimization

1. INTRODUCTION

With the evolution of the economy and society everywhere on the planet, the world is experiencing increased concentrations of air pollutants. Air quality has a direct bearing on how people live and breathe. Currently, the environmental downside is the most severe problem that features a major influence on human health and ecosystem. Air pollution is increasing day by day, adversely affecting the world economy, degrading the quality of life and resulting in a major productivity loss. At present, this is one of the most critical problem. It has a significant impact on human health and ecosystem. Recently there have been scenarios where the air pollution has surged to significant levels and had a severe detrimental effect on human health. The Amazon forest fires, severe air quality degradation in Delhi, India and the fires in the Australian forests are some of the biggest air pollution hazards

DOI: 10.4018/IJEHMC.2021030104

in recent times. Various efforts are placed by government towards the management of pollution, and much success has been obtained within the same (Gulia, Shiva Nagendra, Khare, & Khanna, 2015). Human health problem is one of the necessary consequences of air pollution, particularly in urban areas. The global warming from phylogeny greenhouse gas emissions may be a long run consequence of air. Correct air quality prediction can cut back the effect of a pollution peak on the encircling population and ecosystem, hence rising air quality prediction is a very important goal for society(Bellinger, Mohomed Jabbar, Zaïane, & Osornio-Vargas, 2017). Most recent air quality prediction uses effortless methods viz box models, Gaussian models and linear statistical models. All of the above models are quite simple to implement and enable for the quick calculation of predictions (Moltchanov et al., 2015). Nevertheless, they normally don't describe interactivity & non-linear relationships that command the transfer and nature of adulterants in air. With these provocations, machine learning approaches like outlier detection have become favoured in air quality prediction and other environmental related areas (D. Zhu, Cai, Yang, & Zhou, 2018).

There have been proposed a number of ways to analyse such time series data in order to develop prediction models. Any model that can be developed for the purpose of air quality or air pollution component prediction can have a significant impact on the real-world applications. One major aspect to air pollution prediction is that the results vary significantly based on the geographical location, season, and other factors. For example, if air pollution in Delhi, India is compared in different seasons, the major pollutants will be different. in the stubble burning season, the main culprit would be Carbon Dioxide or Carbon Monoxide whereas in the other seasons, vehicular emissions would be the main culprits. Data mining approaches help us to analyse the air quality data available in order to predict the various determiners of air pollution. There are many data mining methods that can be used to develop models for the purpose. The most used approach in such application areas that has shown a significant performance in temporal data is Clustering. Clustering is the data analysis procedure that is used to inspect and interpret the vast collection of data. Another major setback faced by the data mining methods being employed for the purpose is that the air quality data collected from sensors usually contains deviant values called outliers. Outliers are the data values that do not conform to the general trend of the time series. These outliers have a significant detrimental effect on the quality of prediction and need to be detected and eliminated prior to decision making. Outlier detection has been used for ages to find and, where possible, eliminate ambiguous data points. Outliers emerge because of mechanical defects, changes in system behaviour, fraudulent nature, human mistakes, or instrumental mistakes (Agarwal C.C.., 2017). Identifying outliers can determine glitches before they will intensify with the probable disastrous results and results in finding out the errors and reduce their contaminating impact on the complete set of data. This data is then as such purified for further processing. Outlier detection has found applicability in a wide variety of areas like frauds-detection, intrusions-detection for cyber-security, insurance etc. (Chandola, Banerjee, & Kumar, 2009). The performance of the outlier detection methodology and subsequently the efficiency of the air quality prediction using the outlier detection as a component depends to a large extent on the underlying data mining algorithm (Nazari & Kang, 2018). Several attempts have been made to improve the performance of the data mining algorithms with varying levels of success. One of the most commonly used technique to improve the data mining algorithm is to use nature inspired meta-heuristics like Particle Swarm Optimization, Ant colony Optimization etc to optimize the underlying data mining algorithm (Kant & Mahajan, 2019). This study proposes to use particle swarm optimization to improve the performance of the k-means clustering algorithm in order to increase the efficiency of the air quality prediction.

The remaining part of the paper is organised as: Section 2 gives a brief overview as well as the recent developments of the techniques used in the proposed methodology, viz., Clustering and Particle Swarm Optimization; section 3 outlines the experimental setup used and the results obtained; section 4 discusses the conclusions drawn from the experimental results and the future scope and work required.

2. BACKGROUND AND RELATED WORK

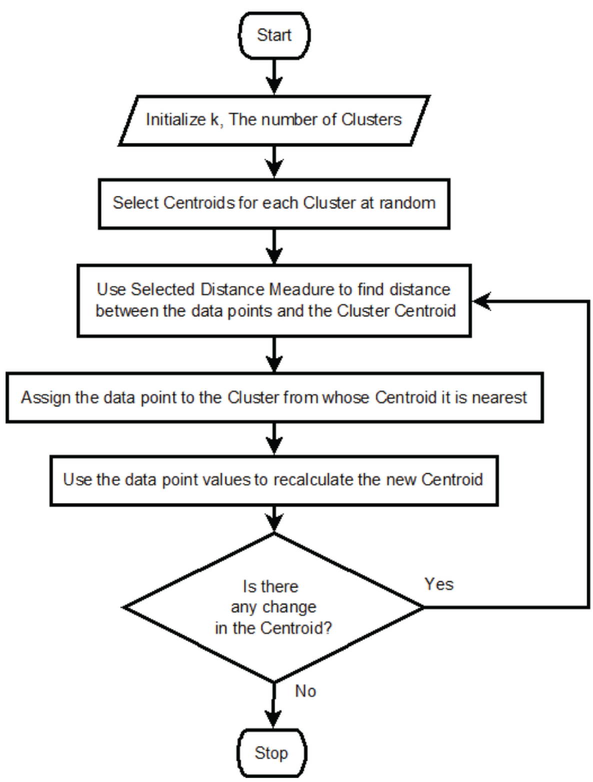
2.1 Clustering

Clustering is one of the significant data analysis procedure which is utilized to investigate and comprehend the huge collection of data. Clustering has demonstrated its probability in different fields like bioinformatics, pattern acknowledgement, picture processing, medical mining and some more. The purpose of the clustering is to organize objects into clusters dependent on the estimations of their qualities (Oktar & Turkan, 2018). Recently, numerous analysts have a noteworthy enthusiasm in developing clustering algorithms. Clustering has a clear disadvantage in that it does not have early information knowledge about the given dataset. Also, the decision of input parameters such as the number of clusters, number of closest neighbours, halting criteria and other factors in these algorithms make the clustering increasingly challengeable (Berkhin, 2006). These algorithms also experience the ill effects of unsuitable precision when the dataset contains clusters with various complex shapes, densities, and anomalies. Furthermore, nature-based algorithms have also been developed to find the exact solution of clustering problems. These algorithms provide better quality results in comparison to traditional algorithms (Pacheco, Gonçalves, Ströele, & Soares, 2018).

2.1.1 K-Means Clustering

K-means Algorithm is one of the broadly utilized unsupervised clustering techniques. This methodology utilizes the distance measures to discover clusters in a given dataset as shown in Figure 1. It starts with the identification of random centroids for a predetermined specific number

Figure 1. K-means Clustering Algorithm



of clusters, and subsequently relegates the various datapoints into the most suitable cluster based on their distance from the chosen cluster centres (Kaur, 2017). The distance measures commonly used include Euclidean distance, Manhattan distance and so forth. The data points which does not seem to belong to any of the clusters may represent the anomalies for that dataset (Sharma, Goel, & Kaur, 2013). The process of re-evaluating the distance measures and reassigning the data points to the various centroids is carried out iteratively.

It was shown in (Aggarwal & Singh, 2018) that although being easy to implement and having fast computation speed, the k-means algorithm has some shortcomings. The efficiency of the k-means clustering algorithm depends to a very large extent on the initially chosen number of clusters and centroids. If not selected properly, it may cause the k-means to get stuck up in the local optima and defeats the purpose of the decision making algorithm employing k-means as a clustering tool(Sieranoja & Fränti, 2018). So, if a centroid is initialized to be a "far-off" point, it might just end up with no points associated with it and at the same time more than one clusters might end up linked with a single centroid. Similarly, more than one centroid might be initialized into the same cluster resulting in poor clustering. A number of studies have been made on the comparative evaluation of the k-means algorithm so as to find the impact of the determining factors on its performance (Steinley & Brusco, 2007). To overcome the dependability of K-means on the initial centroid choices, K-means++ algorithm was proposed. This algorithm ensures a smarter initialization of the centroids and improves the quality of the clustering. Apart from initialization, the rest of the algorithm is the same as the standard K-means algorithm. K-means++ thus involves a smarter initialization of the centroids in the K-means algorithm (Jain, 2010) (Arthur & Vassilvitskii, 2007). The modified algorithm tries to pick up centroids which are far away from one another. This increases the chances of initially picking up centroids that lie in different clusters. Also, since centroids are picked up from the data points, each centroid has some data points associated with it at the end.

2.2 Dimensionality Reduction – PCA

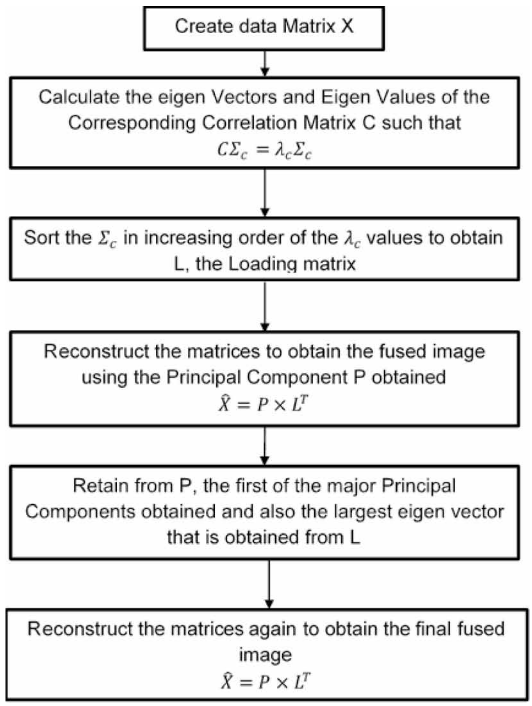
Data that is used for making decisions by using data mining often has a large number of dimensions, in extreme cases the number of dimensions or attributes may even outnumber the number of data instances. Reducing the dimensionality of such data becomes a necessity before it can be subjected to further analysis using data mining methods. This becomes even more important in cases where the dimensions are measured using different scales. There are a number of ways that are used for the purpose of dimensionality reduction, a commonly used one being Principal Component Analysis (PCA) (Abdi & Williams, 2010) (Lever, Krzywinski, & Altman, 2017).

Figure 2 shows the basic PCA approach. PCA is a measurable methodology that uses an orthogonal transformation to change over a lot of perceptions of potentially related factors into a set of values of directly uncorrelated variables called principal components. It is a technique of summarizing data. It is basically used to reduce the dimensions of a large dataset. It is also used to ensure whether the variables are statically independent of each other (Jacques & Preda, 2014).

2.3 Particle Swarm Optimization

Particle Swarm Optimization is an algorithm that takes its motivation from the natural behaviour of swarms of birds or fishes. PSO starts with initializing a simulated swarm of birds in the given solution space, where each individual bird, representing the individual solutions, is called a particle (Kennedy & Eberhart, 1995). PSO then uses a strategy employing both local and global searches to find the optimized solution to the given problem. The movement of the particles in the solution space is specified with the help of a change in their position and velocity which is calculated iteratively as a vector sum (Bansal, 2019). The vector sum for a particle 'p' is calculated using its current position, inertial value keeping in mind its current position and velocity, the global best position and velocity and the personal best value of the position and velocity of the particle itself. The value thus obtained

Figure 2. Principal Component Analysis (PCA)



is evaluated at each step using the cost function chosen (Chopard & Tomassini, 2018). The velocity of the particle in the n*th* iteration is calculated as:

$$v_p^{n+1} = initwt \times v_p^{(n)} + \left(coeff_1 \times rand_1 \times \left(localbest_p - x_p^{(n)}\right)\right)$$
$$+ \left(coeff_2 \times rand_2 \times \left(globalbest_p - x_p^{(n)}\right)\right)$$

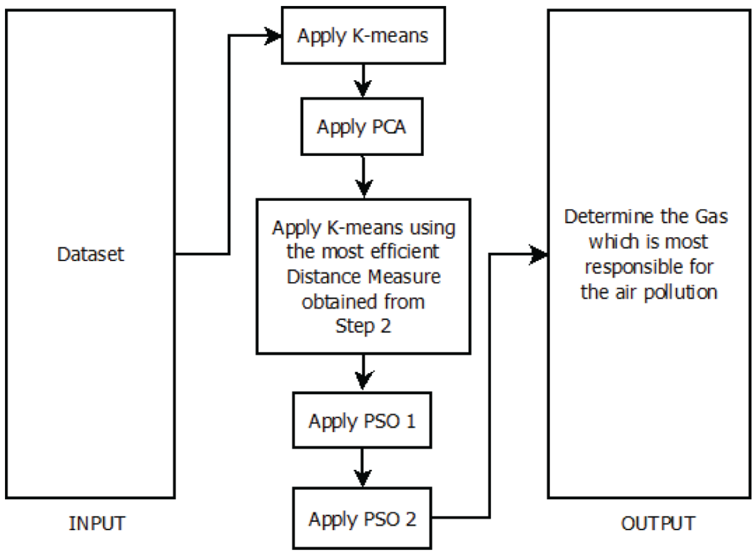
- x denotes the position of the particle;
- rand1 and rand2 are two random numbers in the range [0,1];
- coeff1 is the cognitive coefficient and coeff2 is the social coefficient;
- Localbest is the best position of the particle; and
- globalbest is the best position of the swarm.

It has been observed that using the Particle Swarm Optimization algorithm, improves the efficiency of the data mining algorithm and also enhances the outlier detection and elimination capabilities, thereby making the decision making process more effective.

3. EXPERIMENT ANALYSIS AND RESULTS

The proposed method uses two variants of the Particle Swarm Optimizers PSO1 and PSO2 to optimize the K-means algorithm being used on the Air Quality dataset for the purpose of predicting the gas which is the major cause of air pollution. The proposed methodology has been tested using the UCI air quality dataset (De Vito, Piga, Martinotto, & Di Francia, 2009).

Figure 3. Architecture of the proposed model



The proposed methodology has been implemented in the form of an architecture as shown in Figure 3.

The experimental setup used to test the methodology included:

- Cognitive coefficient value of 0.5
- Social Coefficient Value of 0.3
- Initial Weight for particles as 0.9
- Number of clusters = 3
- Silhouette being used as the Effectiveness measure
- PCA as the dimensionality reduction method
- A number of distance measures were used as the cost functions with the intention of using the most effective of them for the final evaluation of the method

The results obtained at the various stages of the proposed method are described in Table 1.

First a comparison was made among the various distance measures so as to identify the best suited one for the application area. It has been observed that the application area plays an important role in the effectiveness of the distance function and conversely the proper choice of the distance measure plays an important role in the effectiveness and efficiency of the results obtained.

It was clear from the results that for the Air Quality and Pollutant prediction task, the Euclidian Distance Measure was having the best performance.

Next, PCA was used based on the correlation matrix shown in Figure 4 in order to reduce the dimensional complexity of the task.

In the next step, the simple K-means clustering algorithm and the K-means++ algorithm was applied on the dataset whose dimensionality had been reduced. Three (3) clusters have been used in the proposed approach. The data points that don't exist near any clusters are the anomalies in the dataset. Figure 5 and Figure 6 show the clusters formed by using these approaches. The obvious difference arising when using the two approaches yield the results that have been tabulated next ad help to select the best approach to use.

The performance of the two approaches has been compared in Table 2.

Table 1. Comparison of various distance measures

Distance Measure	MSE Loss	RMSE Loss
Euclidian Distance	293.9933911027594	0.17725580589292042
Spearman Distance	426.7923290000836	0.2135698721887785
Manhattan Distance	300.5555022789991	0.1792231172418168
Pearson Distance	334.2810173952336	0.18901120153408418
Chebyshev Distance	454.2843047552455	0.22034111942243143

Figure 4. Dimensional matrix of the Air Quality dataset

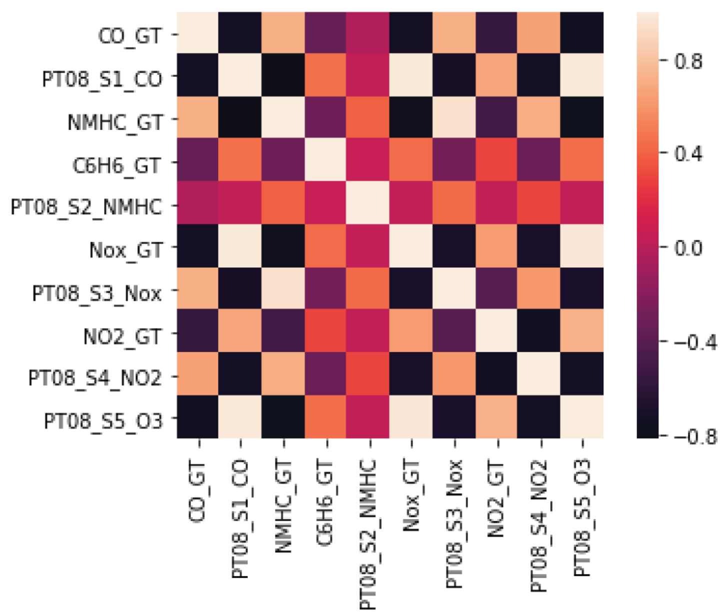


Figure 5. Cluster formation using the simple K-means algorithm

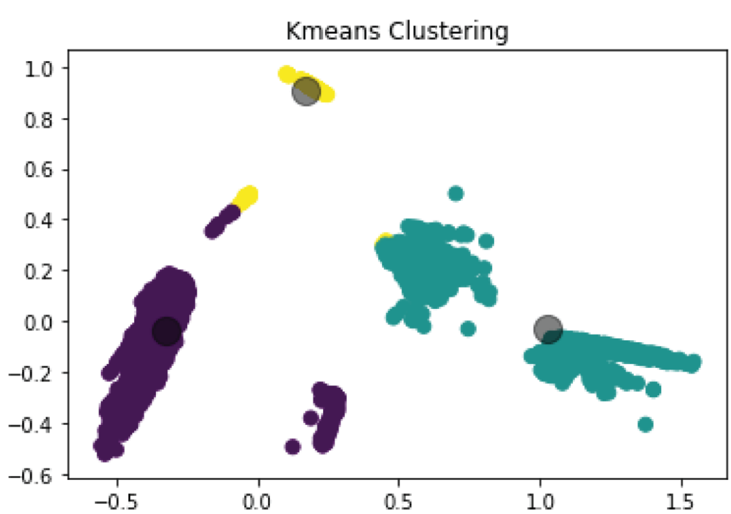


Figure 6. Cluster formation using the K-means++ algorithm

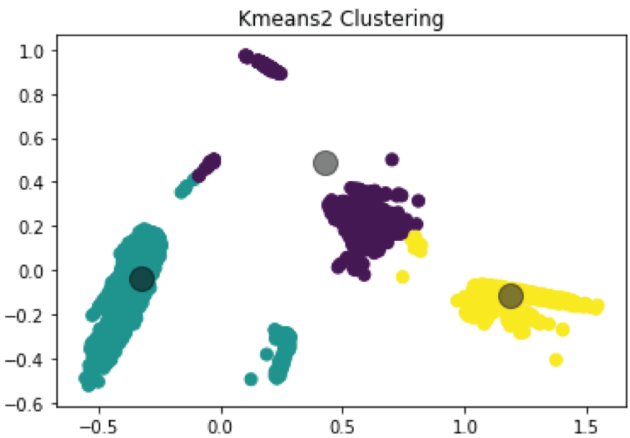


Table 2. Comparative performance of the two K-means implementations

Parameter	Simple K-Means	K-means++
Silhouette	0.8158111150597671	0.8158111150597671
SSE	369.409262582403	295.40050339463886
Quantization	9.710605103745937	9.223337549823496

It was clear from the results, especially, the SSE values that the K-means++ algorithm was better performing for the proposed problem area, so it was used for further processing.

Next the two implementations of the PSO algorithms were used to optimize the clustering process. Figure 7 and Figure 8 shows the cluster formation as obtained after application of the two Particle Swarm implementations. Once again there are marked differences in the visualizations of the clusters on using the two approaches, differences that are supported by the results that have been tabulated.

The two approaches were compared based on the final global best score obtained on using the two implementations. The results are shown in Table 3.

Figure 7. Cluster formation using the Hybrid K-Means++ PSO1 implementation

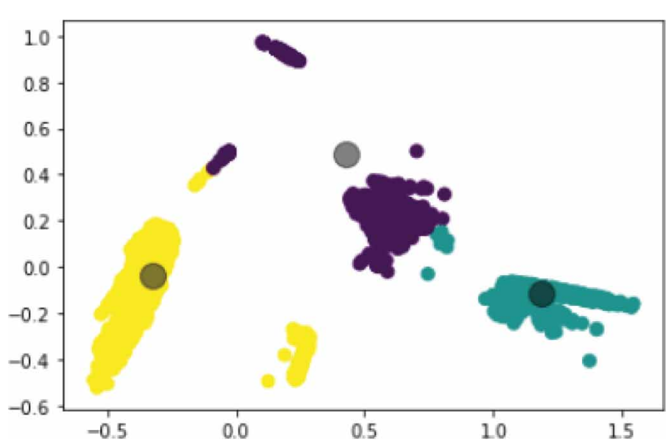


Figure 8. Cluster formation using the Hybrid K-Means++ PSO2 implementation

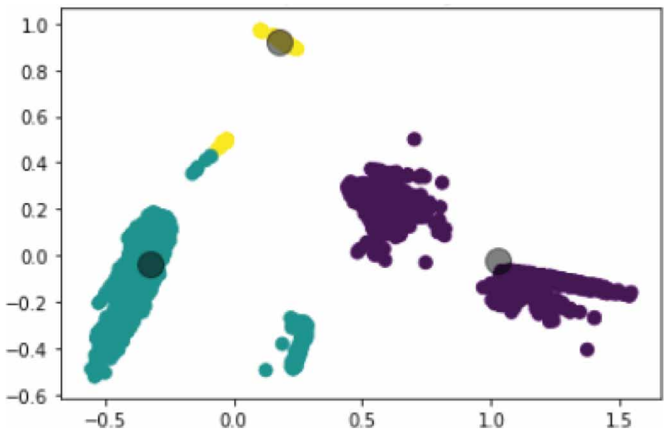


Table 3. Comparative performance of the two PSO implementations

Parameter	PSO1	PSO2
Initial global best score	9.71060510374594	9.223337549823492
Finish with gbest score	9.660661668493398579	9.223314483500631766
Silhouette	0.8162848761193342	0.816868941175955
SSE	369.71096041257806	295.40050339463835
Quantization	9.660661668493399	9.22333754982349

From the above results, it can be concluded that PSO 1 was performing better than PSO 2 as is indicated by the fact that the value of gbest score of PSO 1 less as compared to that of PSO 2.

The outlier values were identified and eliminated from the different components of the air quality data series in order to have a better prediction and to increase the effectiveness and efficiency of decision making. Next the concentration levels of the various gases as obtained after the processing was checked in order to find the component responsible for the maximum pollution. The results were tabulated as a chart as shown in Figure 9.

The graph shows the concentration level of the various gases present in the air. So, from this graph it can be concluded that the gas PT08_S5_O3 is the major cause of air pollution as it is present in the atmosphere in the highest concentration thus depleting the quality of the air. The relative performances of the various methods have been presented in Table 4.

These results provide a comparative analysis of the approaches used in the proposed methodology. The results help to compare the various approaches and to select the optimal approach towards the problem domain.

4. CONCLUSION AND FUTURE WORK

The proposed methodology was intended to find out the component which is most responsible for the air pollution. The effort was to make this decision in a more effective manner to make the decision making more efficient. The use of Particle Swarm Optimization to optimize the clustering process, makes it possible to analyse the results in a focussed manner and helps in this regard. Two implementations of the K-means clustering algorithms were compared and the K-means++ algorithm

Figure 9. Concentrations of the various gases in air as obtained after outlier removal

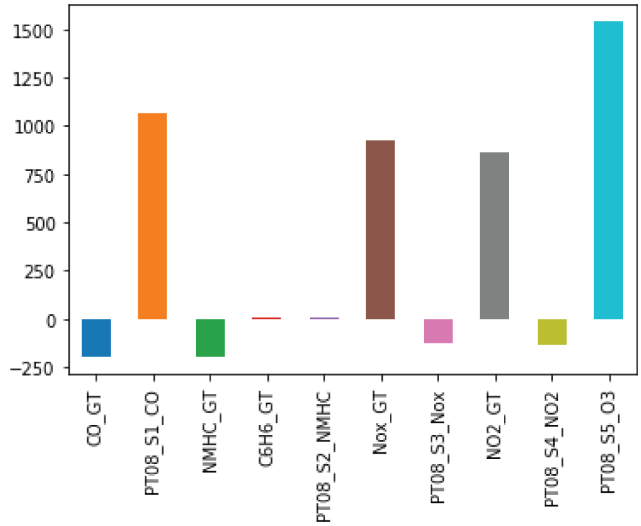


Table 4. Comparison of performances

Parameter	K-Means++	PSO	Hybrid PSO
SSE_Mean	22.732944	366.77343	317.70808
SSE_stdev	2.866618	90.04187	33.963804
Silhouette_mean	0.478871	0.808124	0.816796
Silhouette_stdev	0.018372	0.016836	0.000337
Quantization_mean	0.715906	9.799435	9.350725
Quantization_stdev	0.005868	0.656476	0.200679

was found to be more efficient for the prediction purpose. The choice of the distance measure to be used for the clustering was also based on an empirical comparison and here the Euclidian distance measure was found to be the better performer. The RMSE values were chosen as an indicator of the performance of the clustering approach and were used to compare the different approaches. Further two implementations of the Particle Swarm Optimization algorithm, *viz.* the canonical PSO and the Hybrid PSO were used to optimize the K-means++ algorithm being used for clustering. The performances of the swarm implementations were compared based on the global best scores and Hybrid PSO was found to be better suited for the task. The outliers thus identified were eliminated and then the clusters thus obtained were used to predict the gas component which was contributing the most towards the air pollution. The removal of outliers from the dataset before a decision could be made results in a better and efficient prediction model. This will help the authorities focus their efforts in a much better manner to combat the menace of air pollution more effectively.

The approach has been tested on the UCI Air quality dataset and needs to be tested on more such datasets. The results that are expected from applying the proposed methodology to other datasets may vary because in different regions, climates, time periods there may be different primary components of the air pollution. Moreover, the approach can be extrapolated to other categories of tasks as well, wherever decision making based on clustering is required.

REFERENCES

Abdi, H., & Williams, L. J. (2010). Principal Component Analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, *2*(4), 433–459. doi:10.1002/wics.101

Aggarwal, C. C. (2017). Outlier Analysis. Artificial Intelligence Review. doi:10.1007/978-3-319-47578-3

Aggarwal, S., & Singh, P. (2019). Cuckoo, Bat and Krill Herd based k-means++ clustering algorithms. *Cluster Computing*, *22*(S6), 14169–14180. doi:10.1007/s10586-018-2262-4

Arthur, D., & Vassilvitskii, S. (2007). K-Means++: The Advantages of Careful Seeding. *Proc. of the Annu. ACM-SIAM Symp. on Discrete Algorithms*, *8*, 1027-1035. doi:10.1145/1283383.1283494

Bansal, J. C. (2019). Particle Swarm Optimization. In J. Bansal, P. Singh, & N. Pal (Eds.), *Evolutionary and Swarm Intelligence Algorithms. Studies in Computational Intelligence* (Vol. 779). Springer. doi:10.1007/978-3-319-91341-4_2

Bellinger, Zaïane, Vargas, & Jabbar. (2017). A systematic review of data mining and machine learning for air pollution epidemiology. *BMC Public Health, 17*, 907. 10.1186/s12889-017-4914-3

Berkhin, P. (2006). A Survey of Clustering Data Mining Techniques. In J. Kogan, C. Nicholas, & M. Teboulle (Eds.), *Grouping Multidimensional Data* (pp. 25–71). Springer., doi:10.1007/3-540-28349-8_2

Chandola, V., Banerjee, A., & Kumar, V. (2009, July). Anomaly detection: A survey. *ACM Computing Surveys*, *41*(3), 15. Advance online publication. doi:10.1145/1541880.1541882

Chopard, B., & Tomassini, M. (2018). Particle Swarm Optimization. In *An Introduction to Metaheuristics for Optimization. Natural Computing Series* (pp. 97–102). Springer. doi:10.1007/978-3-319-93073-2_6

De Vito, S., Piga, M., Martinotto, L., & Di Francia, G. (2009). CO, NO2 and NOx urban pollution monitoring with on-field calibrated electronic nose by automatic bayesian regularization. *Sensors and Actuators. B, Chemical*, *143*(1), 182–191. doi:10.1016/j.snb.2009.08.041

Gulia, S., Shiva Nagendra, S. M., Khare, M., & Khanna, I. (2015). Urban air quality management-A review. *Atmospheric Pollution Research*, *6*(2), 286–304. doi:10.5094/APR.2015.033

Jacques, J., & Preda, C. (2014). Model-based clustering for multivariate functional data. *Computational Statistics & Data Analysis*, *71*(March), 92–106. doi:10.1016/j.csda.2012.12.004

Jain, A. K. (2008). Data Clustering: 50 Years Beyond K-means. In W. Daelemans, B. Goethals, & K. Morik (Eds.), Lecture Notes in Computer Science: Vol. 5211. *Machine Learning and Knowledge Discovery in Databases. ECML PKDD 2008* (pp. 3–4). Springer. doi:10.1016/j.patrec.2009.09.011

Kant, N., & Mahajan, M. (2019). Time-series outlier detection using enhanced k-means in combination with PSO algorithm. In Lecture Notes in Electrical Engineering (Vol. 478, pp. 363–373). doi:10.1007/978-981-13-1642-5_33

Kaur, P. (2017). Outlier Detection Using Kmeans and Fuzzy Min Max Neural Network in Network Data. *Proceedings - 2016 8th International Conference on Computational Intelligence and Communication Networks,* 693-696. doi:10.1109/CICN.2016.142

Kennedy, J., & Eberhart, R. (1995). Particle swarm optimization. *Proceedings of ICNN'95 - International Conference on Neural Networks*, 1942-1948. doi:10.1109/ICNN.1995.488968

Lever, J., Krzywinski, M., & Altman, N. (2017). Points of Significance: Principal component analysis. *Nature Methods*, *14*(7), 641–642. doi:10.1038/nmeth.4346

Moltchanov, S., Levy, I., Etzion, Y., Lerner, U., Broday, D. M., & Fishbain, B. (2015). On the feasibility of measuring urban air pollution by wireless distributed sensor networks. *The Science of the Total Environment*, *502C*, 537–547. doi:10.1016/j.scitotenv.2014.09.059 PMID:25300018

Nazari, Z., & Kang, D. (2018). Evaluation of Multivariate Outlier Detection Methods with Benchmark Medical Datasets. *International Journal Of Computer Science And Network Security*, *18*(4), 36–43.

Oktar, Y., & Turkan, M. (2018). A review of sparsity-based clustering methods. *Signal Processing*, *148*, 20–30. doi:10.1016/j.sigpro.2018.02.010

Pacheco, T. M., Gonçalves, L. B., Ströele, V., & Soares, S. S. R. F. (2018). An Ant Colony Optimization for Automatic Data Clustering Problem. *2018 IEEE Congress on Evolutionary Computation, CEC 2018 - Proceedings*, 1-8. doi:10.1109/CEC.2018.8477806

Sharma, S., Goel, M., & Kaur, P. (2013). Performance Comparison of Various Robust Data Clustering Algorithms. *International Journal of Intelligent Systems and Applications*, *7*(7), 63–71. doi:10.5815/ijisa.2013.07.09

Sieranoja, S., & Fränti, P. (2018). Random projection for k-means clustering. Lecture Notes in Computer Science, 10841, 680-689. doi:10.1007/978-3-319-91253-0_63

Steinley, D., & Brusco, M. J. (2007). Initializing K-means batch clustering: A critical evaluation of several techniques. *Journal of Classification*, *24*(1), 99–121. doi:10.1007/s00357-007-0003-0

Zhu, D., Cai, C., Yang, T., & Zhou, X. (2018). A Machine Learning Approach for Air Quality Prediction: Model Regularization and Optimization. *Big Data and Cognitive Computing*, *2018*(2), 5. doi:10.3390/bdcc2010005

Manish Mahajan received his B.E. (Electronics Telecommn.) from Marathwada University in 1992 and M.tech in 2005. He is a member of IEEE, ACM, SCRS, IAENG, and has published 14 research papers in National and International Journals/Conferences in the fields of Computer Security, Software Testing and Data Mining. He has over 20 years of experience in teaching/research of UG(B.Tech) and PG(M.Tech) level courses as a Lecturer/ Assistant Professor/Associate Professor in organizations like ABES IT, Ghaziabad, IPEC, Ghaziabad, ABES EC Ghaziabad and Graphic Era University, Dehradun to name a few. He is currently associated with Graphic Era University as Associate Professor and is Heading the of Department of Information Technology. His research interests include Data Mining, Outlier Analysis, Software Testing, Cyber and Information Security.

Santosh Kumar received Ph.D. from IIT Roorkee (India) in 2012, M. Tech. (CSE) from Aligarh Muslim University, Aligarh (India) in 2007 and B.E. (IT) from C.C.S.University, Meerut (India) in 2003. He has more than 13 years of experience in teaching/research of UG (B. Tech.) and PG (M.Tech.) level courses as a Lecturer/Assistant Professor/ Associate Professor in various academic /research organizations. He has supervised 01 Ph.D. Thesis and 22 M.Tech Thesis and presently mentoring 06 Ph.D. students (singly and jointly. He has also completed a consultancy project titled "MANET Architecture Design for Tactical Radios" of DRDO, Dehradun in between 2009-2011. He is an active reviewer board member in various national/International Journals and Conferences. He has memberships of ACM (Senior Member), IEEE, IAENG, ACEEE, ISOC (USA) and contributed more than 52 research papers in National and International Journals/conferences in the field of Wireless Communication Networks, Mobile Computing and Grid Computing and software Engineering. Currently holding position of Associate professor in the Graphic Era Deemed to be University, Dehradun (India). His research interest includes Wireless Networks, MANET, WSN, IoT, AI, and Software Engineering.

Bhasker Pant is currently working as Dean Research and Development and Associate Professor in Department of Computer Science and Engineering. He is Ph.D. in Machine Learning and Bioinformatics from MANIT, Bhopal. Has more than 15 years of experience in Research and Academics. He has till now guided as Supervisor 3 Ph.D. candidates (Awarded).and 5 candidates are in advance state of work. He has also guided 28 MTech. Students for dissertation. He has also supervised 2 foreign students for internship. Dr.Bhasker Pant has more than 70 research publication in National and international Journals. He has also chaired a session in Robust Classification & Predictive Modelling for classification held at Huangshi, China.