

# Reliable E-Nose System using the Improved Optimization Technique based ANN



Jambi Ratna Raja Kumar, Rahul K. Pandey, Biplab K. Sarkar

**Abstract:** (Since from last decade, there is a growing interest in a system that detects the pollutant gases and other environmental information is called Electronic Nose (E-Nose) networks. The gases such as methanol, Liquid Petroleum Gases, ammonia, etc. are harmful for human beings; therefore, such frailness required detecting automatically as well as safety alarm promoted in a specific field. The critical challenges of the E-nose system are efficient to detect with minimum error and overhead. In this paper, we targeted to design the optimized machine learning-based algorithm to detect and alert the pollutant gases, Humidity, O<sub>2</sub> Level, and Air Temperature in the real-time datasets. We initiated E-nose design using Artificial Neural Network (ANN). Using essential ANN leads to poor accuracy and error rates, as they failed to select the best solutions during the training process. Thus, we next use the Particle Swarm Optimization (PSO) based ANN called ANN-PSO to improve the accuracy rate and error performances for E-Nose systems. Finally, the proposed Improved Optimization Technique based ANN (IOT-ANN) machine learning model designed and evaluated in current this research work. The IOT-ANN it is based on a bio-inspired algorithm to achieve reliable training during the E-Nose prediction.

**Index Terms:** E-nose system, pollutant gases, humidity, artificial intelligence, prediction, artificial neural network.

## I. INTRODUCTION

In current techniques, the prediction procedures using ML is used current recognizable applications. Machine Learning is latest and current trend in analysis for research and mainly focused on finding the models and the other similarity in data. Data processing is the Task of finding patterns in extensive database involving procedures at the intersection of machine learning, analytics systems. One of the important steps in ML in those methods is applied to extract data patterns. Also called known as an interdisciplinary sub domain of computer application. The outcome of the data processing (DP) is joining or gathers information from a data set and normalizes format for future uses. Apart from this data review step, it includes DB and DBM stuff, data prior processing also includes considerations, confusion metrics and complexity calculation, after-processing of searched structures, grammatical representation. DM is also known as knowledge invention in databases process [1].

Revised Manuscript Received on March 16, 2020.

\* Correspondence Author

**Jambi Ratna Raja\***, HoD Department of Computer Engineering in G S Moze College of Engineering, Balewadi, Pune

**Dr. Rahul K. Pandey**, Research guide in Maharishi University of Information Technology, Lucknow

**Dr. Biplab K Sarkar**, HoD, PVPIT, Research guide in Savitribai Phule Pune University

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](https://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

The main thing of all these techniques is used to locate all points in lower manifold, and the graph representation required for understand of the manifold. Nearer points joined by weight edges tend to have the same labels and vice versa. In this way, the tags associated with data can be propagated throughout the data mining [2]. . The main target of DM using learning to identify incorrect or missing edges, prediction of possessions of nodes and clustering nodes based on their tie-up patterns. These jobs arise in many networks and biological pathways [5].

## II. RELATED WORKS

Several recent techniques are using the machine learning approaches presented for the air quality prediction and E-nose systems.

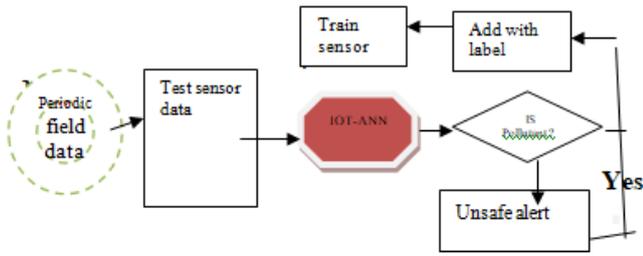
In [12], the authors used 2 procedures of land work on regression estimate absorption of Nitrogen oxides gas and area LA. Apart this to utilizing the meteorological and pollutant detection parameters, base things like populace, land and interspaces from the coastal regions were used. The outcomes proposed that in the of predication NO<sub>x</sub> concentration, the ubiquitous Kriging has better than land-use regression.

In [13], this work the authors executed a 1 year research work on ozone application in the city of Malaga of Spain country. The regression for the prediction of ozone concentration employing the parameters was used. Diffusion techniques used and statistical techniques such as Kriging in template air pollution face several boundaries. Outcome of diffusion technique is allied with i/p records, and it is essential that the data records with more rate are there about the way the pollutants diffuse in the weather. The standard numerical models of Kriging procedure also have been referred for spatial analysis defined as; its average is stable for the temporal variations [14].

In [15], the authors have proposed neural networks for air-pollution prediction. The interrelated things of pollution such as traffic, hours and days of the week last three times of years, the air speed and direction, temperature, rainfall.

## III. METHODOLOGY

Therefore in this research work, we proposed novel bio-inspired metaheuristic optimization hybrid solution IOT-ANN motivated by the living behaviors of microalgae, photosynthetic species, is introduced to optimize ANN training performance. It is starts on the refinement process and the motion of microalgae. The proposed method is specially designed forth E-nose system performance optimization. Figure 3.1 shows the proposed E-Nose system using the IOT-ANN approach.



**Figure 1: The working of the proposed E-nose system using sensor dataset**

As demonstrated in figure 1, the periodic test data is checked by IoT-ANN classifier, which gives the results in form of normal or unsafe conditions. If the pollutant identified, then an alert is issued. Otherwise, the latest test data steps are added to training information with the forecast label. We present design of the IoT-ANN model in this paper. Before submitting the IoT-ANN algorithm, first, we represent the proposed Improved Optimization Technique (IoT) corresponds to each answer in problem area by elaborating the characteristics of algae.

### A. IOT Algorithm

The algorithm consists of 3 necessary steps which are "Evolutionary Process", "Adaptation," and "Helical Movement". Entire population controlled algal colonies. A territory is called collection of algal cells combined together (Eq. (1) and (2)). At a time when one algal is spilt to generate two new one cells, they live adjacently, otherwise two are separated, and the first four cells stay and so forth. Other outer factors such as shear force either or invalid states may scattered, and each scattered part now becomes one life proceeds. Settlement current point pick up point is termed as the colony of optimums, divided of the optimum algal cells.

$$\text{Population of Algal Colony} = \begin{bmatrix} X_1^1 & \dots & X_1^D \\ \vdots & & \vdots \\ X_N^1 & \dots & X_N^D \end{bmatrix} \quad (1)$$

$$i^{\text{th}} \text{ algal colony} = [X_i^1, X_i^2, \dots, X_i^D] \quad (2)$$

Where  $X_i^j$  algal cells in  $j^{\text{th}}$  the dimension of  $i^{\text{th}}$  the colony.

**A. Evolutionary process:** Adequate stage, experience adequate light, enlarges & regenerates to produce two cells in time period, same way as mitotic division. Other side, the algal colony not getting plenty for survives for a while but in the end no longer exist. The rise of kinetic the algal province was calculated with the Monod model given in:

$$\mu = \frac{\mu_{max} S}{K_s + S} \quad (3)$$

Where  $\mu$  is SGR,  $\mu_{max}$  the maximum accurate growth rates are the nutrient concentration, which is the fitness value ( $f^i(X_1)$ ) in period  $t$  and  $K$  is the substrate half-saturation constant.  $\mu_{max}$  We assumed as 1.  $K$  was calculated of the province specified in the following equation:

$$G_i^{t+1} = \mu_i^t G_i^t \quad i = 1, 2, \dots, N. \quad (4)$$

**B.B. Adaptation:** In adaption algal colony cannot grow enough in an atmosphere, try to adapt atmosphere and as outcome, dominant species change the state. It is the task in

that lacking grown algal colony works for defragment itself to the most significant algal territory in the atmosphere. This task continues until the change occur the method. The first starvation mark as 0 for each artificial alga. Increment in the starvation period  $t$  when the algal cell not receives proper light. The artificial seaweed is the top starvation value (Eq. (8)) has so would be (Eq. (9)).

$$\text{Starving}^t = \max A_i^t \quad i = 1, 2, \dots, N; \quad (8)$$

$$\text{Starving}^{t+1} = \text{Starving}^t + (\text{biggest}^t - \text{Starving}^t) * \text{rand}. \quad (9)$$

Where  $A_i^t$ , starvation point where  $i^{\text{th}}$  algal colony in  $t$  time, starving the algal colony at the top point starvation merit in time  $t$ ?

**C. Helical movement:** Due to adequate light for survival algal cells, colonies usually dip and try to reside inside the water surface. All of them swim liquid provides step ahead movement that is not allowed by gravity and a viscous drag. Actions of algal cells differ. As the frequency of helical movements expands up their search friction surface of a viable algal cell gets larger. Therefore, cells are near to the ground area, the more power it has, and it finds more chance to move inside liquid.

Other side, friction area lacking; motion from the liquid stay long time. Due to that, their universal search is more. Anyhow, they can also move less in section to its energy. IoT, the attraction moderate a progress demonstrated as 0, and viscous drag is present as shear force.

(Equations. (10) and (11)).

$$\tau(X_i) = 2\pi r^2 \quad (10)$$

$$\tau(X_i) = 2\pi \left( \sqrt{\frac{3G^i}{4\pi}} \right)^2 \quad (11)$$

Where  $\tau(x_i)$  is the friction surface? Using the helical movement of the algal cell of three dimensions is determined randomly. In that determine a linear progression in Eq. (12) and the other two dimensions provide angular displacement in Eq. (13) and (14). Eq. (12) Is used for 1-D problems and algal colony moves in the one direction. In 2-D problems, algal movement is sinusoidal, because Eq. (12) and (14) are used. In multiple, movement is helical Eq. (12)–(14) is used. Friction or rough surface and distance to the light source establish the step size of the movement:

$$X_{im}^{t+1} = X_{im}^t + (X_{jm}^t - X_{im}^t) (\Delta - \tau^t(X_i)) P \quad (12)$$

$$X_{ik}^{t+1} = X_{ik}^t + (X_{jk}^t - X_{ik}^t) (\Delta - \tau^t(X_i)) \cos \alpha \quad (13)$$

$$X_{il}^{t+1} = X_{il}^t + (X_{jl}^t - X_{il}^t) (\Delta - \tau^t(X_i)) \sin \beta \quad (14)$$

Where  $X_{ik}^t, X_{il}^t$ , and  $X_{im}^t$  are x, y, z coordinates of  $i^{\text{th}}$  the  $t$ ;  $\alpha, \beta, \in [0, 2\pi]$ ;  $P \in [-1, 1]$ ;  $\Delta$  is shear force;  $\tau^t(X_i)$  or area of  $i^{\text{th}}$

### B. IOT-ANN Training Algorithm

Using the algorithm designed for IoT above section, we build IoT-ANN framework for the air quality prediction. Figure 2 demonstrates the IOT-ANN training system. As observed in figure 2, the steps of the IoT technique merged in the ANN training process in which the MSE. MSD finds out the values for measures.

It used to calculate the mean of the square of the error.

MSE also called as risk estimator, as it estimated value of the squared error. The Variation occurs arise due to irregularities of estimator. The MSEesses the function mapping a part of information to a sampled) or an expected (i.e., a function mapping an example of random variable). The definition MSE varies according to the estimator.

**Predictor:**

It  $\hat{Y}$  of  $n$  and  $Y$  is the vector of observed values of the variable being expected, then the within-sample MSE of the predictor is computed as

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \tag{15}$$

**Estimator:** Estimator  $\hat{\theta}$  concerning an unknown parameter  $\theta$  is defined as

$$MSE(\hat{\theta}) = E_{\hat{\theta}} [(\hat{\theta} - \theta)^2] \tag{16}$$

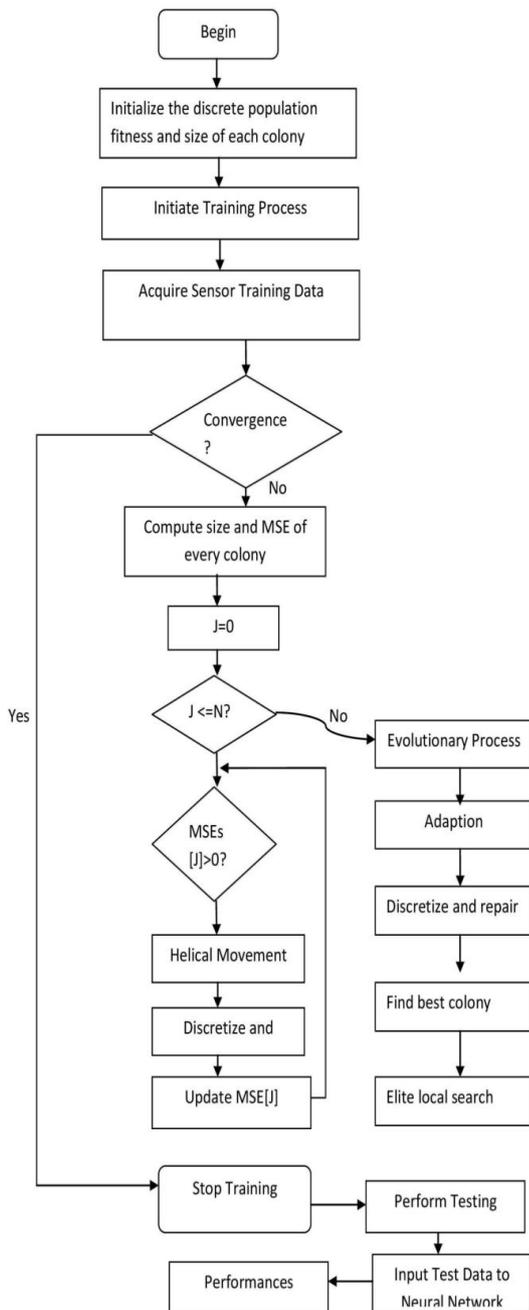


Figure 2: Proposed IOT-ANN Flowchart

In similarity to standard deviation, taking MSE also known as RMSE or RMSD, which has the same sections as; unbiased estimator is the of the changes identified deviation. The MSE acquire quality of an estimator (i.e. the part of data to a feature of the sampled). The precision MSE varies the condition estimator to estimate for prediction. IOT-ANN finds the answer of each problem in by representing properties of it. Same as the real one, artificial one go ahead starting of light with helical of swimming, and adjust to it, can specie can regenerate by division.

**IV. SIMULATION RESULTS**

We implemented and evaluated the proposed approach with essential ANN and ANN-PSO using the standard real-time AQ [31]. dataset calculate the data of 9358 measurements gathered from the five metal oxide chemical sensors. The sensors are kept in notably polluted areas in Italian city. Ground Truth hourly averaged concentrations for CO, Non-Metanic Hydrocarbons, Benzene, Total Nitrogen Oxides (NOx) and Nitrogen Dioxide (NO2) and is supplied by a co-located reference certified analyzer (CA). We divide this dataset into training (70 %) and testing (30%). The results are quantified for ANN, ANN-PSO, and IOT-ANN with varying number of hidden layers. We changed the hidden layers to observe the effects of increasing the hidden layers. First, we computed the error rate in figures 3 to 5 of the total forecasting span of 200 minutes for each the AI technique. The IoT-ANN prediction error performance shows a significant reduction as compared to ANN. The error rate for IoT-ANN at some period is higher compared to the ANN-PSO, but as the span progressed, the performance becomes better for the IoT-ANN. The average error rate of IoT-ANN is less compared to the ANN-PSO. As compared to ANN, the optimization-based ANN variants always deliver the near-perfect solution and hence reduced error rate. Convergences and balancing problem addressed by IoT-ANN thus shows the reduction in error rate compared to IOT-ANN.

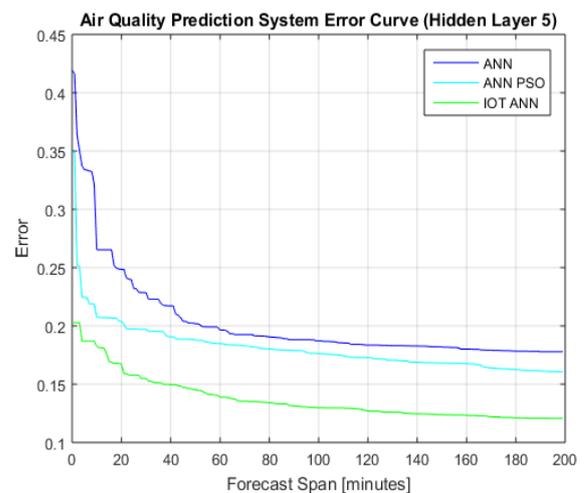
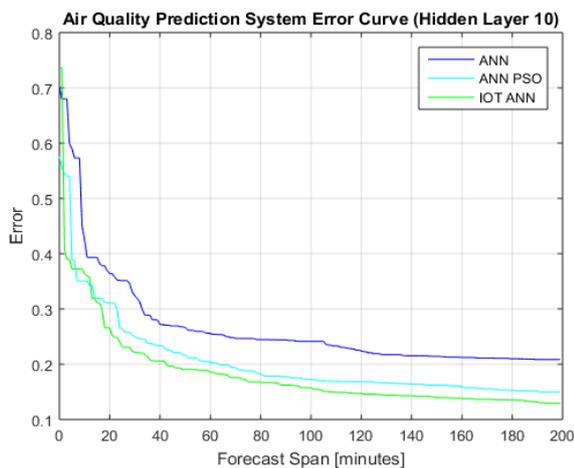
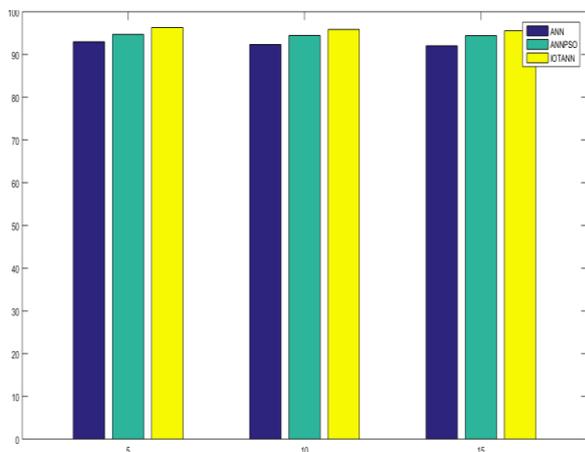


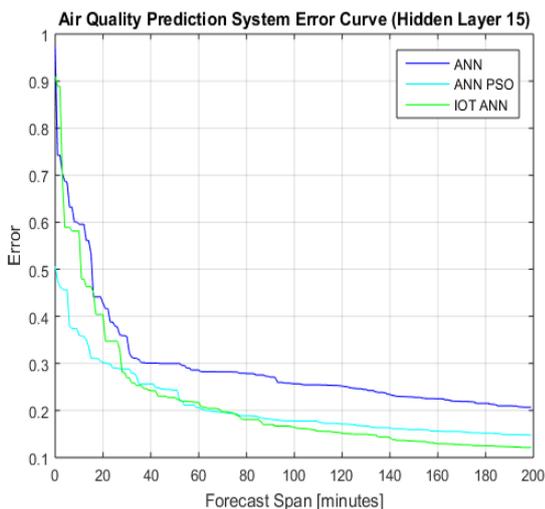
Figure 3: AQ prediction error performance using five hidden layers



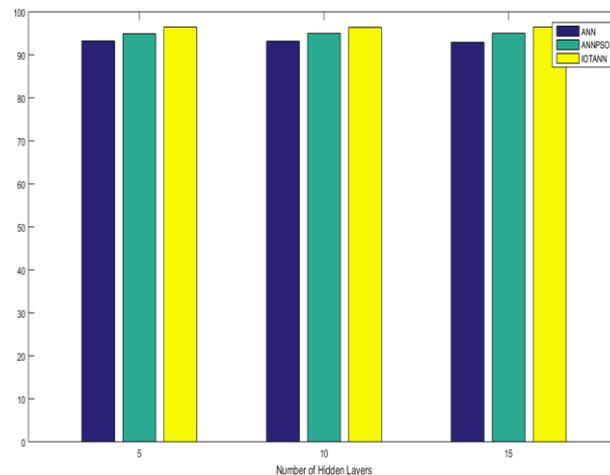
**Figure 4: AQ prediction error performance using ten hidden layers**



**Figure 4.6: AQ training accuracy performance**



**Figure 5: AQ prediction error performance using 15 hidden layers**

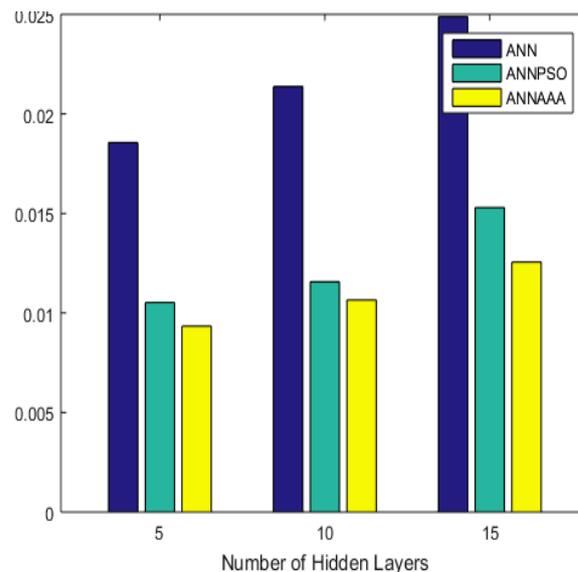


**Figure 4.7: AQ testing accuracy performance**

The average accuracy for training and testing phases measured after the successful training and testing evaluations. Figure 4.6 (table 1) demonstrates the training accuracy results, and figure 4.7 (table 4.2) shows the measurements of the testing accuracy outcome.

As observed the outcomes of dataset AQ dataset results, as the hidden layers increases, the performance of accuracy increases. The ANN method delivered less accuracy compared to other methods. The IoT-ANN shows the improved accuracy result compared to ANNPSO and ANN technique.

Finally, the comparative graphs on average training error and testing error result in figures 4.8 (table 4.3) and 4.9 (table 4) respectively. As the count of hidden layers grow, the performance error rates increase. The ANN error performance is very high compared to ANN-PSO and IOT-ANN due to the obvious reasons discussed above. The above results are summarized in tabular form below. All the average reading performances using the QA dataset for the E-Nose system show that the proposed AI model increases the returns with a minimum error rate.



**Figure 4.8: AQ average training error rate performance**

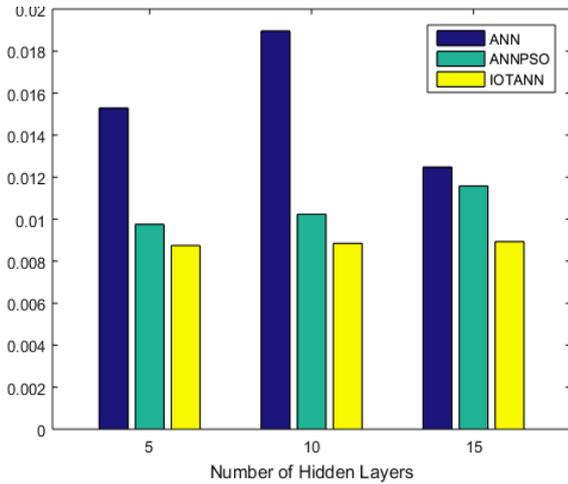


Figure 4.9: AQ average testing error rate performance

Table 1: QA training accuracy tabular analysis

Hidden layers	ANN	ANN-PSO	IOT-ANN
5	92.9702	94.6850	96.3255
10	92.3320	94.4539	96.8521
15	92.0355	94.4189	96.9396

Table 2: AQ testing accuracy tabular analysis

Hidden layers	ANN	ANN-PSO	IOT-ANN
5	93.2275	94.8955	96.4939
10	93.1926	95.0064	96.4102
15	92.9309	95.0216	96.4879

Table 3: AQ training error tabular analysis

Hidden layers	ANN	ANN-PSO	IOT-ANN
5	0.0203	0.0182	0.0137
10	0.0267	0.0205	0.0185
15	0.0296	0.0208	0.0216

Table 4: AQ testing error rate tabular analysis

Hidden layers	ANN	ANN-PSO	IOT-ANN
5	0.0177	0.016	0.0121
10	0.0181	0.0149	0.0129
15	0.0207	0.0148	0.0121

As discussed in the literature, some works reported on air quality prediction using the ANN approach based on different real-time datasets. We compared our results with such methods in terms of the prediction accuracy and error rate in table 5. All such existing plans based on the basic ANN model for the prediction system and our study shows that we improved the performance in IoT-ANN as compared to the essential ANN significantly.

Table 5: Comparative study with similar methods

Method	Year	Accuracy (%)	Error Rate
[28]	2014	93.45	0.02139
[29]	2014	92.12	0.0223
[30]	2019	93.77	0.02034
<b>IOT-ANN</b>		<b>96.7</b>	<b>0.01733</b>

## V. CONCLUSION AND FUTURE WORK

For real-time predictions, the ANN model is suitable as compared to other ML techniques. Thus we initiated our research of designing the E-nose system using essential ANN. The reduced accuracy and error performances motivated to develop, and PSO based ANN approach in this research work called ANN-PSO. The problems of ANN-PSO further addressed by the proposed model of this research IOT-ANN. The IoT-ANN system is based on a bio-inspired algorithm to achieve reliable training during the E-Nose prediction. The optimization-based techniques may take a longer time to converge, hence leads to an unreliable training solution in some cases. To evaluate the effectiveness of IoT-ANN, the two real-time chemical sensor devices based datasets are used in this research work. We implemented IoT-ANN, ANN, and ANN-PSO using both the datasets and measured the performances in terms of accuracy, sensitivity, specificity, error rates, and training error. The results are promising for the IoT-ANN, which shows that it outperforms both ANN and ANN-PSO. The overall accuracy of IOT-ANN is improved by approximate 3-4 % as compared to ANN-PSO, with a significant reduction in error rates and training time. For future work, it will be interesting to check real-time deployment and streaming data application with a proposed E-nose model. The deep learning framework for the E-nose is also the future direction of this research work.

## REFERENCES

- Castelnuovo, I.A.; di Pierro, D.; Coletta, M.; Di Francesco, P. Application of electronic noses for disease diagnosis and food spoilage detection. *Sensors* 2006, 6, 1428–1439.
- Persaud, K.C. Medical applications of odor-sensing devices. *Int. J. Low. Extrem. Wounds* 2005, 4, 50–56.
- Peris, M.; Escuder-Gilabert, L. A 21st century technique for food control: Electronic noses.
  - Anal. Chim. Acta* 2009, 638, 1–15.
  - Ampuero, S.; Bosset, J.O. The electronic nose applied to dairy products: A review. *Sens. Actuators B Chem.* 2003, 94, 1–12.
  - Schaller, E.; Bosset, J.O.; Escher, F. “Electronic noses” and their application to food. *Food Sci. Technol.* 1998, 31, 305–316.
  - Dubreuil, B.; Bonnefille, M.; Neitz, S.; Talou, T. Prospective Experiments of e-Nose for Cosmetics Applications: Recognition of Sweat Odors. In *Artificial Chemical Sensing: Olfaction and the Electronic Nose (ISOEN 2001): Proceedings of the Eighth International Symposium*; The Electrochemical Society: New York, NY, USA, 2001.
  - Hodgins, D. The development of an electronic nose for industrial and environmental applications. *Sens. Actuators B Chem.* 1995, 27, 255–258.
  - Shine Wei Lin, Chih Hong Sun, and Chin-Han Chen. Temporal data mining using the genetic algorithm and neural network a case study of air pollutant forecasts. *Geo-spatial Information Science*, 7(1), 2004.
  - Min Li and Md. Rafiul Hassan. Urban air pollution forecasting using artificial intelligence-based tools. *Air Pollution*, 2009.
  - Soawalak Arampongsanuwat. Prediction of pm10 using support vector regression. In *International Conference on Information and Electronics Engineering*, IPCSIT, volume 6, 2011.

11. Ming Dong, Dong Yang, Yan Kuang, David He, and SerapErdal and Donna Kenski. Pm2.5 concentration prediction using hidden semi-Markov model-based times series data mining. Office of Air Quality Planning and Standards, Research Triangle Park, 2009.
12. Mercer, L.D.; Szpiro, A.A.; Sheppard, L.; Lindström, J.; Adar, S.D.; Allen, R.W.; Avol, E.L.; Oron, A.P.; Larson, L.; Liu, L.-J.S.; et al. Comparing universal kriging and land-use regression for predicting concentrations of gaseous oxides of nitrogen (NOx) for the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air). *Atmos. Environ.* 2011, 45, 4412–4420.
13. Dueñas, C.; Fernández, M.C.; Cañete, S.; Carretero, J.; Liger, E. Assessment of ozone variations and meteorological effects in an urban area in the Mediterranean Coast. *Sci. Total Environ.* 2002, 299, 97–113.
14. Nonnemacher, M.; Jakobs, H.; Viehmann, A.; Vanberg, I.; Kessler, C.; Moebus, S.; Möhlenkamp, S.; Erbel, R.; Kessler, C.; Moebus, S.; et al. Spatio-temporal modelling of residential exposure to particulate matter and gaseous pollutants for the Heinz Nixdorf Recall Cohort. *Atmos. Environ.* 2014, 91, 15–23.
15. Cai, M.; Yin, Y.; Xie, M. Prediction of hourly air pollutant concentrations near urban arterials using artificial neural network approach. *Transp. Res. Part D Transp. Environ.* 2009, 14, 32–41.
16. Pisoni, E.; Farina, M.; Carnevale, C.; Piroddi, L. Forecasting peak air pollution levels using NARX models. *Eng. Appl. Artif. Intell.* 2009, 22, 593–602.
17. Yeganeh, B.; Motlagh, M.S.P.; Rashidi, Y.; Kamalan, H. Prediction of CO concentrations based on a hybrid Partial Least Square and Support Vector Machine model. *Atmos. Environ.* 2012, 55, 357–365.
18. Alexander Nikovb, Ferhat Karacaa, Omar Alaghaa, Atakan Kurtc, and Hüseyin Hakkoymazc, "AirPolTool: A Web-Based Tool For Istanbul Air Pollution Forecasting And Control", Proceedings of the Third International Symposium on Air Quality Management at Urban, Regional and Global Scales. 26-30 September 2005, Istanbul – Turkey.
19. Vikas Singh, Claudio Carnevale, Giovanna Finzi, Enrico Pisoni, Marialuisa Volta, "A cokriging based approach to reconstruct air pollution maps, processing measurement station concentrations and deterministic model simulations", *Environmental Modelling & Software* 26 (2011).
20. Madhavi Anushka, Elangasinghe, Naresh Singhal et al., "Development of an ANN-based air pollution forecasting system with explicit knowledge through sensitivity analysis," *Atmospheric Pollution Research* 5 (2014) 696-708.
21. Navneet led Benjamin, Sarita Sharma, Umesh Pendharker, and JK Shrivastava, "Air quality prediction using the artificial neural network," *International Journal of Chemical Studies* 2014; 2(4): 07-09.
22. Mahmoud Reza Delavar, Amin Gholami, and Gholam Reza Shiran et al., "A Novel Method for Improving Air Pollution Prediction Based on Machine Learning Approaches: A Case Study Applied to the Capital City of Tehran," *International Journal of Geo-Information*, Feb.2019.
23. S. De Vito, E. Massera, M. Piga, L. Martinotto, G. Di Francia, "On-field calibration of an electronic nose for benzene estimation in an urban pollution monitoring scenario, Sensors, and Actuators", *B: Chemical*, Volume 129, Issue 2, 22 February 2008, Pages 750-757, ISSN 0925-4005.

### AUTHORS PROFILE



**Jambhi Ratna Raja, Kumar** working as HoD in Computer Engineering Department in G S Moze College of Engineering, Balewadi, Pune and also Pursuing PhD from Maharishi University of Information Technology, Lucknow.

**Dr. Rahul K. Pandey**, is Research guide in Maharishi University of Information Technology, Lucknow

**Dr. Biplob K Sarkar**, is HoD in PVPIT, Pune and also research guide in Savitribai Phule Pune University and also in research guide in Maharishi University of Information Technology, Lucknow