

Forecasting Hydrogen sulfide Level Based Neural Computation

Jabar H. Yousif

Abstract—this paper aims to design and implement an environmental monitoring and forecasting system based on neural computing approach. The output information is used for feeding the alarming systems. The data are collected in real-time through pollution monitoring sensors at Sohar region. Air pollution is a serious problem and coming from different sources that can lead to a catastrophic, which is needed to be monitored and controlled. The proposed work is monitored and managed the alerts on the emissions of Hydrogen sulfide (H₂S) in Oman. It is forecasting the Level of H₂S in the Sohar region based Neural Computation. The SOFM is used to compute and predict the ratio of H₂S and then issue an alarm to take the proper decision which helps to implement the necessary precautions. The experiments are giving evidence that the predicted values are closely to true values that gained from real sensors with accuracy of 78% and less MSE of 0.03865.

Index Terms— Environment pollution, air quality standards, Environmental management system, Neural computing SOFM.

I. INTRODUCTION

THE public health and welfare are significant consequences when air quality standards are not met. However, the objective of an Air Quality Management Plan is to create the foundation for a successful reduction of air pollutant emissions. In the current time of industrialization almost most countries are promoting investments in industrial growth and diversifying the economy based on industrialization. Meanwhile, industries have a positive impact on the economy and income of the country. On the other side, it leads to another kind of challenge which is the amount of pollutants and waste releasing from these industries. Besides, increasing using of mechanical machines and modern transportation like airplanes, cars, will lead to increase the amount of air pollutants. It must be pointed out that the interest in industrial and economic development to develop the society became the priorities of governments without paying attention to the negative effects on the environment. Through industrial development and the growth of civilian life, the pollution is consider to be a real hazard on human health and life. At the begging of the 1960s, the researchers began to concern about the meaning of environmental pollution which is considered as a serious problem because it threatens all objects on the Earth's surface. Air pollution is exposed particles or biological compounds that might damage or harm humans and other organisms. Hydrogen

sulfide H₂S gas is found naturally in the environment or as a result of human industrial activities. It can be founded in most of the oil refineries. The gas imitation is considered as a serious and common problem in the oil industry. Hydrogen sulfide is considered toxic and dangerous gases colorless and highly flammable, it able to poison different devices within the human body, and the nervous system remains the hardest hit. The toxic of hydrogen sulfide is equivalent to the amount of toxic of carbon monoxide. The safe amount of H₂S must be (0.13 PPM) part per million. Consequently, many environmental problems can appear and affect the lives of a lot of people, which causes loss about million people per year. Moreover, many of the serious diseases are emergence that affects the life in general, as well as large economic losses caused by the impact of air pollution on agricultural crops, plants and animals. H₂S can absorb by the human body and works like interior poison which affects eyes, nose, throat, lung and can spoil the nervous system. Also, Hydrogen sulfide has an impact on the all elements of environment such as human, animal, plant, soil, minerals, etc. It reacts with the iron steel, forming sulfur flammable and corrosive to the production, packaging and transport pipes. Moreover, air pollution is spreading rapidly not only stand its impact on the source, but goes beyond that to reach neighboring and remote areas. This leads to not being able controlling air pollution after it released. It is imperative that the pollution is controlled and treated before it is launched into the atmosphere. In recent years, air pollution has become a major worldwide concern. Therefore, the researchers focus on the study of air pollution and its harms on humans and environment. The researcher are proposed several methods to decrease the side effects of air pollution [1]. Neural Computing (NC) is the study of computational methodologies that simulates the behavior of biological, ecological and physical systems. It is a range of techniques and methods, which are designed to model large, complex, and dynamic real life problems. It is used to discover a mathematical model for such complex and dynamical systems. NC generally includes the implementation of Artificial Neural Networks (ANN), Evolutionary Computing, and Genetic Algorithms (GA). The NC has the ability to calculate and decide within a precise and imprecise data range. Besides, it can simulate the human abilities like learn from past experience [2]. The Self-Organizing Feature Map (SOFM) is an unsupervised neural model. The SOFM is implemented in many real-world problems application. Besides, SOFM can be used as a standard

analytical tool in the financial analyses and control theory [3, 4]. The proposed model is used to monitor and manage alerts system for the emissions of H₂S in Oman. It is forecasting the Level of Hydrogen sulfide (H₂S) in the Sohar region based Neural Computation.

This paper discusses the implementation of SOFM neural technique for monitoring the gas levels. The environmental monitoring unit is responsible for observing and managing the entire industrial zone. The implementation of intelligent decision-making agent is enabling analyze of collected data from remote monitoring sensors.

II. THE SOURCES OF AIR POLLUTION

Before starting on building an air pollution system, the developer needs to understand the sources of air pollutants in order to implement the appropriate control measures that fit the type of air pollution in the region. In rustic areas, the understanding of air pollution sources is easy, because the number of pollutants is very low. However, in urban areas it becomes a complex issue as a result of increasing the number of sources of air pollution. These sources affect air quality and include natural processes such as volcanoes, which produce sulfur and chlorine. As well as forest fires and the burning of waste farmers that produce smoke and carbon monoxide and carbon dioxide. Animal waste emits methane as part of the digestion process. Trees also emit volatile organic compounds, such as pine trees. The largest number of pollutants is the result of the residual of industrial plants, and vehicles which produce nitrogen oxides, carbon monoxide, and carbon dioxide, sulfur dioxide and particulate matter. There are major factors and activities, which are responsible for releasing pollutants into the atmosphere. Figure 1 illustrates air pollution sources and their relation and effects on each other. Data quality is consider as important factor for scientists and organizations. Many methods and solutions have been discovered for achieving the standard levels of data quality [5]. Information processing in the environmental systems is very crucial matter that should be done carefully in order to meet the standard level of control systems. The data processing is including different phases like gathering, delivery, destruction and integration of data.

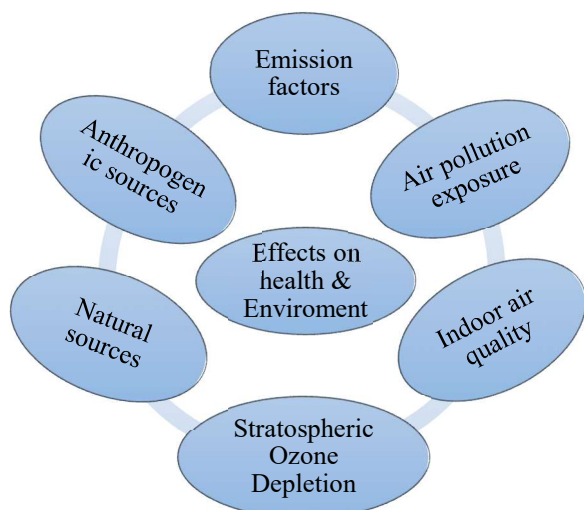


Figure 1: air pollution sources & relations with other factors

III. THE AIR QUALITY MANAGEMENT SYSTEM

Environmental management system (EMS) is a framework designed to achieve the company's goals to preserve the environment from pollution by systematic controlling the methodology of its operations and output. A number of EMS frameworks are existed. The ISO 14001 international standard is the most well-known model as depicted in Figure 2. Besides, the following models:

- European Eco-Management and Audit Scheme (EMAS).
- The Responsible Care model is developed by the American Chemical Council (ACC).
- US Department of Justice (DOJ) "Seven Key Compliance Program Elements.
- EPA National Enforcement Investigation Center (NEIC) "Compliance Focused".

The common main elements of each EMS models as depicted in Figure 2. The first part is the Environmental policy which helps in issuing a message indicating the institution's commitment to preserving the environment. It must unify the vision of environmental concern and compliance with relevant laws and regulation. It must have a policy for monitoring and continual improvement with significant impacts. The second part of the framework is the planning, which is determining the goals and objectives of the institution based on respecting the regulation and policy of the country in preserving the environment.

In addition, defines clearly the environmental aspects and requirements. Besides, illustrates the attributes of products, activities and services based environmental management program. Whereas, large number of EMS implementation is established using a third party certification. However, there is another levels of EMS implementation. Besides, on-going risk management system can be applied if all of the EMS elements are employed. The three levels of EMS implementation based third party include the following:

- Third Party Certification to ISO 14001.
- Complete EMS Implementation Without Certification.
- Partial EMS Implementation Strategies.

The next step is to develop monitoring and measurement procedures. This will help to follow the development of the different EMS programs for accomplishing the stated objectives and targets of the organization. Besides, EMS auditing procedures it must be clearly defined based on the intermittent reviews of the performance of the system. In addition, the implementation of corrective and preventive actions needs to be obviously identified. Besides the Records & EMS Audits must be professionally implemented and managed for ensuring a continues improvement of the system.



Figure 2: Air Quality Management System Phases

IV. RELATED WORK

Sheng-Tun Lia [6] proposed an environmental system in Taiwan for controlling and improving air quality called National Air Quality Monitoring Network (TAQMN). This project implements data mining techniques to expose the unseen knowledge of air pollution dissemination for huge data retrieved from monitoring stations in TAQMN. In order to eliminate noises and recognize the trend of data, multi-scale wavelet transforms is implemented. Besides, two-level SOFM model is adopted to classifying the clusters on multi-dimensional wavelet-transformed space. The distribution of suspended particulate PM10 of the air quality situation in the current areas of pollution is included new zones differ than the present pollution regions. This made the current study very important reference for government agencies in assessing the present situation for developing air pollution policies in the future. Nitin [7], proposed a cloud solution for air pollution handling. They used the cloud computing technologies for helping the companies and governments covering the cost of handling pollution. As well as, carrying out their corporate social responsibility. The used of cloud will save a lot of money, and offer a skill and talent solution. In the coming future, the cloud applications based pollution will play an important role because of scaling, on demand services and including mobile application for free pollution environments. Boger [8] proposed ANN modeling of the nitric acid construction plant predicted NO_x emission levels with slight errors range of 0.6% relative error on NO_x concentration prediction. And 0.006 kg/hour on daily emission in the 20-45 kg NO_x/hour range. The analysis of ANN and AA-ANN models discovered well-known and unknown associations in the plant function. Input data sets of

daily plant feature averages are used for training ANN models, which is added further information in comparison with using 5 minutes data. For the sake of improving the operational knowledge and better efficiency, the control loop set points in the plant database is used. This will provide additional information for future analysis. A rich review and literature survey of using soft computing in air pollution applications is presented in [9].

V. NEURAL NETWORK MODELING

Neural Network technique is a new method of computing that simulate the behavior of human brain. Neural model is a robust tool to model, linear and nonlinear data. The main features of using neural computation are parallel computing, standardization of data, ability to learn, train, and generalization. The artificial neural network is implemented in many applications such as data classification and filtering, data association and conceptualization [10].

The processing unit of the neural architecture is called “neuron”, which feeds data sets as input and constructing a unique output. Usually, neurons classified into three groups input, hidden and output layers. The input layer is used to feed network with a number of data set [11]. Self-Organizing Feature Map (SOFM) is unsupervised learning methodology that its input is a high dimensional data and its output is a two-dimensional feature map. Each node in the SOFM represents some point in the input space. The learning process of SOFM is utilized by choosing a winner neuron and then changed the weights of its closer neighborhood based hebbian rule. An illustration of self-organizing feature map neural network architecture is depicted in Figure 3.

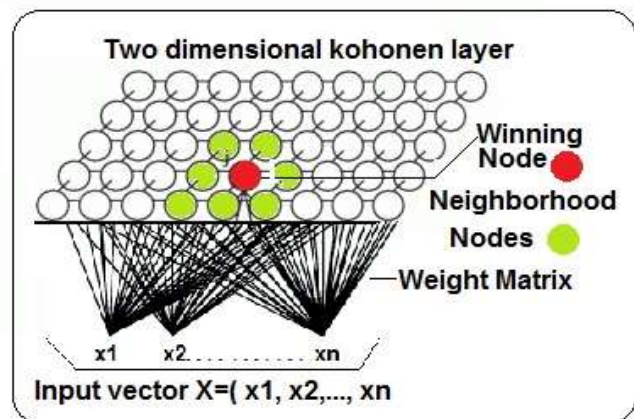


Figure 3: Self-Organizing Feature Map Architecture

The connection configuration of neurons into multi layers is called the architecture of neural network. The transfer function of each neuron must to be defined between each two connected layers in the network. Besides, the weight of all neurons should be also defined initially and then adjust the weight of all neurons in order to train the input data sets. The generalization of unseen data samples is adopt after the training of data is completed. Usually, a suitable encoding method must be implemented in order to transfer the input text into a suitable digital form. However, the data sets are divided into three forms

of data. 40% of data sets are used for training the neural network. And 20% of data sets are used for the cross validation process which helps to determine the error in a test data sets. 40% of data sets are used for testing the functionality of the neural network [12].

VI. EXPERIMENTS AND RESULTS

This paper demonstrates the using of SOMF to monitor and analysis the ratio of Hydrogen sulfide (H₂S), which resides in the air as a result of factory waste and other natural sources. It is Colorless gas has a characteristic odor similar to rotten eggs and is composed as a byproduct of decomposition. Hydrogen sulfide gas is highly toxic and reacts with enzymes in the bloodstream. And this will lead to inability of cell respiration and increase concentration of the closure of lungs. Table 1 shows the effects of various Hydrogen Sulfide (H₂S) Levels.

Table 1. The sound effects Levels of (H₂S) on Human

Level in PPM	"The Conditions and effects on Humans"
0.13	The minimum perceptible odor
4.60	Identified Simply, moderate odor
10	Irritation of eye.
27	A strong and unpleasant odor
100	Cough and eye irritation, defeat smell after 2M to 5M.
200-300	After one hour, The eye inflammation and
500-700	Loss of consciousness, stop breathing and
1000-2000	death in a few minutes after unconsciousness and cessation of breathing

The experiment is implemented momentum as a learning rule with step size of 1 and rate of 0.7. For the sake of optimizing the values of the network parameters, the genetic algorithm is utilized [14]. There are different techniques for determining the networks performance. The mean squared error (MSE) is usually implemented [15]. It is computed as a double of the average cost, which is computed as in equation 1:

$$MSE = \frac{\sum_{j=0}^P \sum_{i=0}^N (d_{ij} - y_{ij})^2}{NP} \quad (1)$$

P is the number of output neurons, N determine the number of neurons in the training process. The desired output is d_{ij} and y_{ij} is the predicted output of the network. Besides, the root mean square error (RMSE), which measures the distance between estimated and actual outcomes. In addition, the mean absolute error (MAE), that estimates the summation absolute errors and they determined as in equations 2 and 3, respectively.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - f_i)^2} \quad (2)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - f_i| \quad (3)$$

The final MSE of predicting the H₂S is 0.03865802120 at epoch 999 at run 3, and the correlation relation between the desired output and the model output is 0.7982193199. The average MSE value of standard deviation boundaries of H₂S which shows that the training graph and cross validation line is clearly smooth trained and closely to each other from the epoch 150. The average of MSE with standard deviation boundaries for three runs of H₂S is summarized in Table 2.

Table 2. the best SOMF network results for H₂S

Best Networks	Training	Cross Validation
Run #	3	1
Epoch #	999	999
Minimum MSE	0.03865802120	0.03569192659
Final MSE	0.03865802120	0.03569192659

In order to evaluate the accuracy of proposed SOMF predicting system and how well it fit the actual results, the determination coefficient R^2 is used. It is calculated as in equation 4:

$$R^2 = 1 - \frac{\sum_i (y_i - f_i)^2}{\sum_i (y_i - \bar{y})^2} \quad (4)$$

where y_i symbolized the values of the actual output, and the f_i determined the predicted values. \bar{y} is the arithmetic mean value of the observed targets. The better model in predicting future outcomes that achieved R^2 closer to 1. The experiment achieved an accuracy value of (79%). Figure 5 illustrates the comparison results between the actual value of H₂S and output of SOMF model. Whereas Figure 6 depicts the final MSE value for three times run of training data set. This to test the stability of the network and then the generalization of unseen data.

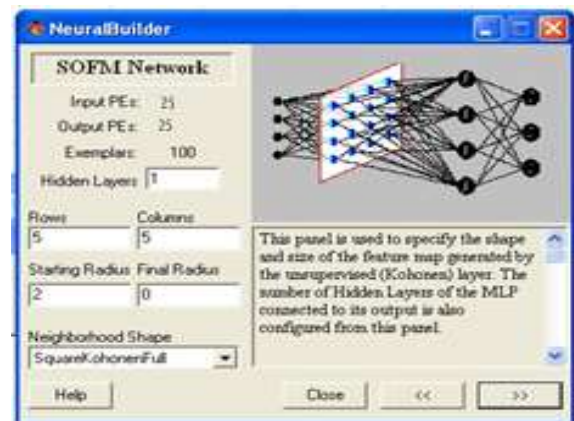


Figure 4: SOFM implementation panel using NeuroSolutions

VII. COMPRESSION STUDY

In order to implement the comparative study, experiments must be conducted under the same conditions and using the same data sets. Therefore, the comparison will consider as a difficult process. The comparison of this work with other

researchers will be done away from the type of variables and the volume of data used.

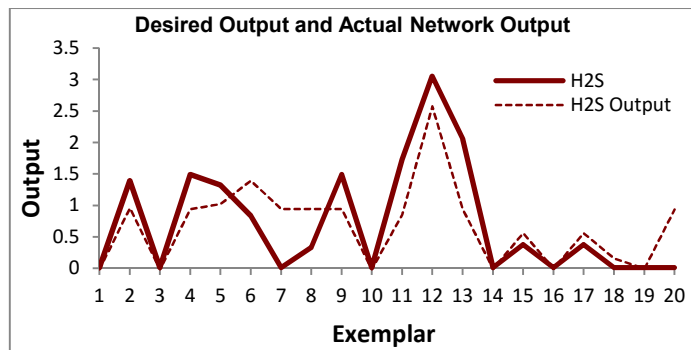


Figure 5: the comparison results of actual value of H2S and SOMF output

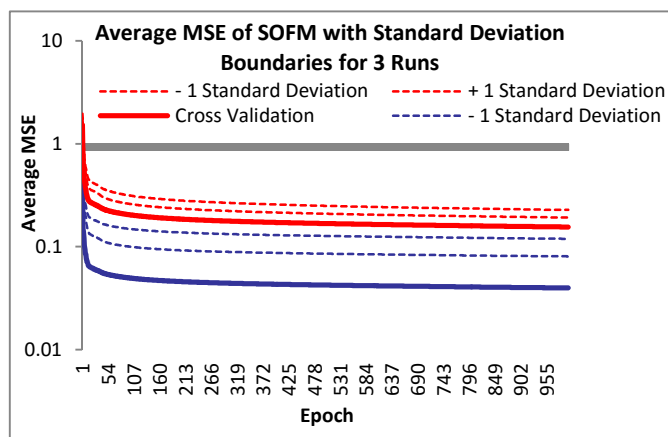


Figure 6: average of MSE with standard deviation boundaries for 3 runs of H2S

Besides, the performance measurements are depended on different factors like the research type, the neural topology, etc. Hence, MSE, RMSE, MAE, determination factor (R^2) are used as a main performance measurement factors. Besides, the best curve fitting can be obtained using the regression models [16]. Santosh K. Nanda, et al. 2011 [17], they developed an environmental predicament for developing countries, which is considered as an intelligent system to predict pollution parameters with less computational cost. They design an environmental engineering application based Legendre neural network for predicting air quality variables. They got a good performance as compared to regression models. The other work by Kevin R. McCullum, et al [18], examined the possibility of using neural model based real-time data for forecasting ambient H2S in the City of Edmonton. This model is used as a quality control tool for identifying actual H2S monitor operational and calibration problems. The neural model approach based multilayer perceptron (MLP) is utilized a historical ambient data for four years. The neural model achieved an R^2 of 0.62, which is considered as a promising technology for ambient H2S prediction, especially through the virtual-monitor approach. The proposed work in this paper is achieved an accuracy of 79% and MSE of 0.03865.

VIII.CONCLUSION

This section is summarized the results of this work as follows:

- 1- The proposed work is implementing unsupervised approach based SOMF model, which is used to analyze and predict the desired data of H2S levels in the Oman region. The proposed model predicate the H2S levels quickly and accurately with less time. As well as, it is used a small number of data for the purpose of learning and generalization of unseen data.
- 2- A comparison between the current work and some of other experiments, as in the [17] and [18], the proposed method got better and faster results, in comparison with the work in [18] which is achieved R^2 of 0.62. Whereas, the proposed model obtained accuracy of 79 %, which means 80 percent of testing data are simulated and predicted correctly. Besides, it is achieved final MSE values for predicting the H2S is 0.03865802120 at epoch 999, third run time, and the correlation relation between the desired output and the actual network output is 0.7982193199. The average of MSE values of H2S shows that the training graph is very close with cross validation line starting at epoch 150.
- 3- Data cleaning is implemented in this paper based intelligent technique in order to analysis the collected data set. The current work is addressing the individual gas concentration with time and date.
- 4- The implementation of intelligent decision-making agent is enabling analyze of collected data from remote monitoring sensors. And effective alarm system is built for catastrophic conditions, which will send a suitable warning alarm to the controller.

REFERENCES

- [1] Abdul-Wahab S.A., 2005. Monitoring of Air Pollution in the Atmosphere around Oman Liquid Natural Gas (OLNG) Plant. Journal of Environmental Science and Health, Part A: Toxic/Hazardous Substances and Environmental Engineering. Vol. 40, No. 3, pp. 559-570.
- [2] Jabar H. Yousif, "INFORMATION TECHNOLOGY DEVELOPMENT", LAP LAMBERT Academic Publishing, Germany ISBN 9783844316704, 2011.
- [3] Teuvo Kohonen, "Self-Organizing Maps" (3rd edition) Springer, ISBN 3540679219, 2001
- [4] Zadeh, L.A. (2001). Applied Soft Computing. Applied Soft Computing 1, 1-2.
- [5] Acken, A.H., Lewis, R.: Final Particulate Matter National Ambient Air Quality Standards. ENRLS update 4(15) (2006).
- [6] Sheng-Tun Lia, and Li-Yen Shue, Data mining to aid policy making in air pollution management, Expert Systems with Applications, 27 (2004) ,pp 331-340. doi:10.1016/j.eswa.2004.05.015
- [7] Nitin B Raut, Jabar H. Yousif, Sanad Al Maskari, and Dinesh Kumar Saini. Cloud for Pollution Control and Global Warming. Journal of Lecture Notes in Engineering and Computer Science, Vol. 2190, Issue: 1 pp: 681-684, 2011.
- [8] Boger, Zvi. "Artificial neural networks modeling to reduce industrial air pollution." In Applications of Soft Computing, pp. 63-71. Springer Berlin Heidelberg, 2009.
- [9] Niharika et al, [2014], A survey on Air Quality forecasting Techniques, International Journal of Computer Science and Information Technologies, Vol. 5 (1), 2014, 103-107
- [10] Dinesh Kumar Saini and Jabar H Yousif. Environmental Scrutinizing System based on Soft Computing Technique. International Journal of Computer Applications 62(13):45-50, January 2013. ISSN 0975 - 8887, Published by Foundation of Computer Science, New York, USA. DOI:10.5120/10143-4952.

- [11] Feng, Xiao, et al. "Artificial neural networks forecasting of PM 2.5 pollution using air mass trajectory based geographic model and wavelet transformation." *Atmospheric Environment* 107 (2015): 118-128.
- [12] Gori, Marco , Maggini, Marco , Rossi, Alessandro , Neural network training as a dissipative process , *Neural Networks journal* , Vol 81, PP 72-80 , 2016 doi:10.1016/j.neunet.2016.05.005
- [13] Andrzej Rusiecki, Robust Learning Algorithm Based on Iterative Least Median of Squares, *Neural Processing Letters*, v.36 n.2, p.145-160, October 2012
- [14] Mirosław Kordos , Andrzej Rusiecki, Reducing noise impact on MLP training, *Soft Computing - A Fusion of Foundations, Methodologies and Applications*, v.20 n.1, p.49-65, January 2016
- [15] Eren Bas , Vedide Rezan Uslu , Erol Egrioglu, Robust learning algorithm for multiplicative neuron model artificial neural networks, *Expert Systems with Applications: An International Journal*, v.56 n.C, p.80-88, September 2016.
- [16] Hassan S. Uraibi, Habshah Midi, Bashar A. Talib and Jabar H. Yousif , Linear Regression Model Selection Based On Robust Bootstrapping Technique. *American Journal of Applied Sciences* Vol. 6 (6): 1191-1198, ISSN 1546-9239, 2009. DOI: 10.3844/ajassp.2009.1191.1198.
- [17] Santosh Kumar Nanda, Debi Prasad Tripathi, S.S. Mahapatra, "Application of Legendre Neural Network for Air- Quality Prediction.", *International Conference on Engineering and Technology*, May 2011.
- [18] Kevin R. McCullum , et al, Prediction of Ambient Hydrogen Sulphide Using Artificial Neural Networks .
http://files.abstractsonline.com/CTRL/70/0/0AA/A8D/26C/40F/787/E16/7F4/081/4A8/41/a49_1.pdf



Dr. Jabar H. Yousif is an Assistant Prof. at Faculty of computing and Information Technology, Sohar University, Oman. Ph.D. Information Science & Technology, M.Sc. & B.Sc. in Computer Science. Postdoctoral fellowship - Future Virtual Reality Laboratory. I have more than 20 years teaching experience. I published more than 25 papers & books in the fields of Artificial Intelligent, Cloud Computing , Soft Computing, Artificial Neural Networks, Natural Language Processing, Arabic text Processing & Virtual Reality. Editorial board & reviewer for a number of scientific journals and conferences.