Applying Artificial intelligence techniques for predicting amount of CO_2 emissions from calcined cement raw materials

Yakoub Boukhari
Chemistry Department
Ziane Achour University
Djelfa, Algeria
jacoubchimie@yahoo.fr

***Abstract*—This paper aims to predict the amount of carbon dioxide CO_2 emissions from raw material used in cement clinker production during calcination process. The amount of CO_2 emissions is mainly from the decarbonisation thermal process that is directly related to chemical composition, distribution of particle size and time exposed at high temperature. These influencing factors interact with each other, making the calculation of the amount of CO_2 emissions with conventional techniques more difficult. For this reason, several artificial intelligence techniques are applied to predict the amount of CO_2 emissions. The key advantage of the proposed techniques is its ability to learn and to generalise without any prior knowledge of an explicit relationship between target and its influencing parameters. The intelligence techniques applied are deep neural network (DNN), artificial neural networks (ANN) optimised using ant colony optimization (ACO-ANN) and genetic algorithm (GA-ANN).**

The results obtained are promising and show that all intelligence techniques can provide excellent accuracy with high R^2 and low error. DNN predicts the amount of CO_2 emissions very accurately when comparing to other techniques. Overall, the performance accuracy of ACO-ANN technique is higher than the GA-ANN. According to R_2 values, there are more than 99%, 98.5% and 98% of experimental data in testing phases can be explained by DNN, ACO-ANN and GA-ANN respectively with average relative error less than 1.04%. As conclusion, all intelligence techniques can be employed as an excellent alternative to predict the amount of CO_2 emissions.

Keywords—CO_2 emissions, calcination process, deep neural network, artificial neural networks, ant colony optimization, genetic algorithm

I. INTRODUCTION

The cement industry is one of the most important heavy industry sectors in the world. It is a major source of CO_2 emissions. In terms of greenhouse gas emissions, it contributes approximately 5 to 7% of global CO_2 emissions [1]. Total emissions of carbon dioxide CO_2 from cement manufacturing is mainly due to the combustions of fuels in a rotary kiln and the calcinations of raw materials at high temperature [2] during cement production. During cement production, about 62% of CO_2 is released from calcination process [3] of raw materials. The raw materials used to produce cement are generally composed of the following main oxides CaO, SiO_2, MgO, Al_2O_3 and Fe_2O_3. These oxides play an important role in determining the quality of cement products [4]. The amount of CO_2 emissions is an important indicator of control in cement production. It is affected by many factors which are mainly on chemical composition, powder fineness and burning time.

The emission process of CO_2 is still not fully clarified and explained. It also is complicated phenomena [5] which is difficult to express analytically. Moreover, the mathematic relationships between the influencing parameters and the amount of CO_2 emissions is not precisely known. It is well known that the laboratory experiments are generally difficult, complicated, costly and require number of chemical reagents and equipments. Therefore, using artificial intelligence techniques are required to predict the amount of CO_2 emissions and are important now more than ever. These techniques do not need to specify mathematical relationships or prior knowledge about the relationship between a different influencing factors and the amount of CO_2 emissions.

For decades, artificial intelligence techniques are extensively used in many fields to predict the behavior of complex systems [6,7]. In this study, several artificial intelligence tools are used to predict the amount of CO_2 emissions from cement raw materials with different chemical composition and various fineness in the range of time. The different artificial intelligence techniques proposed in the present paper are deep neural networks (DNN), artificial neural network optimised using ant colony optimization (ACO-ANN) and ANN optimised using genetic algorithm (GA-ANN).

Currently, artificial intelligence techniques are exploited in various fields because they help to transform traditional industry towards the real intelligent industry. There are plenty of successful application of DNN in many fields such as prediction of reservoir production [8] and for predicting drug-target interaction [9]. Hybrid ACO-ANN is utilised as useful tool to model the diesel engine emissions [10] and to predict the capital cost of mining projects [11]. Hybrid GA-ANN is successfully applied to predict turbidity and chlorophyll-a variations [12] and sheet metal forming [13].

The present paper is arranged as follows: brief description of artificial intelligence techniques are presented in section 2; materials and methods are established in section 3; the section 4 discuss the results obtained; section 5 is reserved for our conclusions.

II. Brief description of Artificial intelligence techniques

A. Deep neural networks (DNN)

Deep neural network which imitates the human neural system is powerful predictive technique [14]. It is characterized by a more complicated structure than conventional artificial neural networks. The DNN performance is significantly influenced by the pre-training and training stages. One amongst the several pre-training algorithms and straining are Autoencoder AE [15] and Softmax function [16] respectively. DNN proposed is created by stacking autoencoder with a softmax layer (SL). Autoencoder and softmax are first trained individually and separately, one layer at a time. The AE is trained using the initial inputs where softmax layer is trained with the output of the AE layer. Then AE and SL are joined together to form DNN. Finally, to improve the accuracy, DNN is fine-tuned using back propagation learning algorithm.

Autoencoder AE proposed for pre-training of DNN is a kind of an unsupervised learning ANN with three layers: input layer, hidden layer and output layer. It is used to extract hidden features and reconstruct the input of DNN in order to build suitable representation from [17]. The learning process of AE is self-supervised by raw inputs without a class label or outputs. It consists of two processes of encoding and decoding. The encoder is used to compress input vector and a decoder is used to extend the compressed input for reconstructing and extracting the essential features. The new reconstruction data are built according the activation function chosen by both encoder and decoder. The reconstruction error is calculated through cost functions. The network weights and bias are iteratively updates to converge cost function in the AE towards a minimum. The cost function used for training AE is MSE adjusted by adding a regularization term (L_2 regularization term and sparsity regularization) [18] to prevent over-fitting phenomena during learning process. In order to ensure a perfect accuracy and good reconstruction data, scaled conjugate gradient descent is used to train AE [19].

Softmax layer which applies a softmax function is added to autoencoder as supervised learning to produce good solutions for some nonlinear problems and especially to solve multiple classification problems. It is characterised by simple structure and easy integration. Similarly to AE, the scaled conjugate gradient algorithm is used by SL for training. The goal of supervised learning is minimise the cost function by updating weight and bias parameters iteratively. The loss function used by softmax layer is cross entropy (log loss) [20].

B. Artificial neural network optimised by Ant Colony Optimization (ACO-ANN)

The possibility of falling into local minimum problems without obtain a globally optimal solution and slow convergence speed to optimal points of ANN is inevitable. To overcome shortcomings of ANN, ant colony optimization (ACO) is combined with ANN to optimise its parameters for producing powerful technique called ACO-ANN.

Ant Colony Optimization (ACO) algorithm is an optimization tool that is mainly used to optimise and to solve different optimisation problems. It is inspired by the behavior of ants. In nature, ants is able to search for an optimal and shortest trajectory (edge) between their nest and food source. Each ant moves randomly through all to find the food source. When ant find a food leave behind a chemical substance on the ground, called pheromone. The quantity, the quality and the distance of the food source are related to the quantity of the laid pheromone. The pheromone evaporation of pheromone [21] over the time plays an important role to fall into the local optimum and to avoid the rapid convergence to a local optimal. Without evaporation of pheromone, the first paths selected by the first ant is extremely attractive to the other ants, and consequently, the research space could be limited. The path which contained more quantity of pheromone is becoming the main path. In each iteration with sufficient number of ants, the quantity of pheromone is updated and reinforced which help to attain more precise and optimal solutions.

The main advantage of ACO algorithm is to converge rapidly to optimal solutions without falling into a local optimum. The optimal performance of the ACO mainly depends on suitable parameters settings. The main parameters of ACO are number of ants per iteration, number iterations and evaporation rate of pheromone.

Ant colony optimization algorithm is mainly used to train and optimise the weights and biases of ANN. The risk of getting stuck in local optima is sharply reduced by ACO algorithm [22]. The ACO combined with ANN on the basis of considering that BP neural networks have strong predictive ability and that the ant colony optimization has good ability of searching for optimisation.

C. Artificial Neural Network optimised by Genetic Algorithm (GA-ANN)

Like the ACO-ANN technique, genetic algorithms GA is used as powerful optimizing tool to optimise and select the optimal parameters of ANN, called GA-ANN technique [23] .

Genetic algorithm is one of the most important and popular optimization algorithms. It is much more frequently employed for solving the most optimization problems by finding the optimal values of a function [24]. The main idea of GA is to imitate natural ways of evolution. The three evolutionary principles of genetic operators used by GA algorithms are crossover, mutations, natural selection.

The initial population of n individuals (solutions) is randomly generated. Each individual is characterised their genes. The crossover operator is used for combining genetic information (solutions) of two parents to generate two new offspring. Then, the genetic operators continued with the mutation operator to introduce and maintain diversity from one generation of a population to the next one. The idea of selection is to choose two parents from the previous generation

and let them pass their genes to the next population. The process is repeated to produce better solutions in each new iteration until reachs the termination condition [25]. The main parameters of GA are number of population size per generation, number generations (iterations), crossover probability and mutation probability.

III. MATERIAL AND METHODS

A. *Materials and experiments:*

In the present study, the amount of CO_2 emission is considered as a function of chemical composition, grain size and time exposed. The raw materials are blended and preheated to round 300°C to remove water combined in the hydration products and then up to 850° C to remove impurities, which can affect the cement quality.

The grain size selected are set in four sizes 71 µm, 125µm, 250µm and 350 µm. The chemical composition and mix proportions of four raw materials used are summarised as shown in Table 1. Finally, each mixture of raw materials with gain size are burned in the laboratory furnace at 1000° C for different time 5, 10, 15, 20, 30 min. The amount of CO_2 emissions from each mixture of raw materials is calculated before and after burning at 1000 °C.

TABLE I. CHEMICAL COMPOSITION (% BY WEIGHT) FOR EACH RAW

Raw Materials	**Oxides**				
	SiO_2	CaO	MgO	Fe_2O_3	Al_2O_3
Material 1	12,38	80,28	1,38	1,69	4,27
Material 1	3,96	92,62	0,99	0,65	1,78
Material 1	14,06	78,69	1,35	1,68	4,22
Material 1	14,16	78,04	1,36	2,21	4,23

B. *Data collection*

The data extracted from the experimentation are collected in table of 80 rows and 8 columns. Each row in table presents experiment. From 1 to 7 columns are inputs where the last column is output. The size particle, time exposed, SiO_2(%), CaO (%), MgO (%), Fe_2O_3 (%), Al_2O_3 (%) are inputs while the amount of CO_2 emissions is the output.

The total data are randomly divided into two sets: training and testing. For each intelligent techniques, 75% of data are used for training while the remaining 25% (not unseen data) of data are kept out to evaluate the generalisation ability. The most common performance criterion used to evaluate the accuracy of each intelligent techniques are the coefficient of determination R^2 and the mean absolute percentage error (MAPE). The technique performance is perfect when values of R^2 and MAPE are very close to 1 and 0, respectively.

IV. RESULTS AND DISCUSSION

First of all, finding the appropriate parameter settings values of each models play a key role in achieving highest prediction accuracy [26] and in avoiding over-fitting and under-fitting. Often, it is not easy to select these parameters. Consequently, the appropriate parameters of each models are obtained after several tests during training process.

A. *DNN results*

The main appropriate parameters values of the autoencoder and softmax layers are predefined as listed in Table 2.

TABLE II. DNN PARAMETERS

Parameters	**Values**
L_2 weight regularization	0.01
Sparsity regularization	4
Sparsity proportion	0.05
Number of hidden layer	1
Neuron number	18
Encoder transfer function	logsig
Decoder transfer function	purelin
Iteration number of AE	1050
Loss function of softmax layer	Crossentropy
Iteration number of softmax layer	1250

Figure 1 illustrates comparison between experimental and predicted amount of CO_2 emissions obtained by DNN.

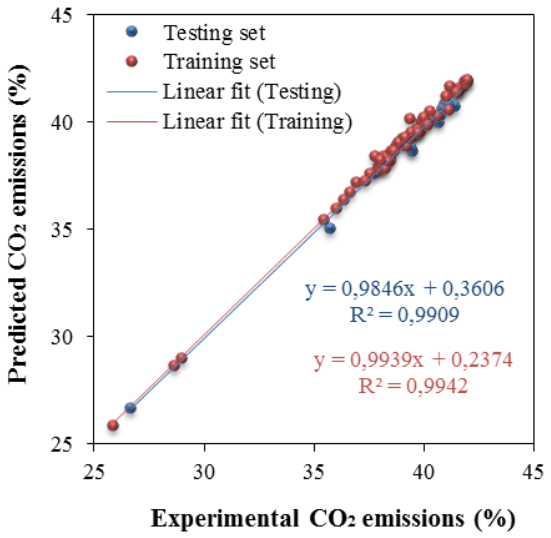

Fig. 1. The comparison between the predicted and experimental CO_2 emissions.

It is observed that almost of points for both phases fall exactly along the linear fit which express perfect equality between experimental and predicted CO_2 emissions. From Figure 1, the values of R^2 calculated for training and testing phases are 0.9942 and 0.9909 respectively. The According to values of R^2, there is a high degree of association between

experimental and predicted values without overfitting or underfitting problems.

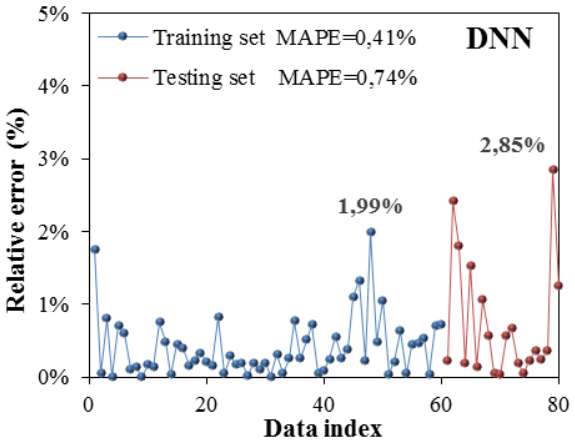

Fig. 2. Distribution of the relative error

Distribution of the relative error of CO_2 emissions for DNN during training and testing phases are given in Figure 2. This figure illustrates clearly that almost of points are closest to zero line. The maximum relative error obtained in training and testing phases are 1.99% and 2.85%, respectively. The average relative error for training and testing phases are 0.41% and 0.74% respectively. These values are indicative of high accuracy in prediction and generalisation.

It is clear that DNN is able to provide excellent performance in both phases, since the values of MAPE and R^2 are near to 0 and 1, respectively. The perfect performance of DNN is due to AE layer which is able to extract the essential information efficiently from data and excellent predicted probability distribution via softmax layer. From the results obtained, it can be concluded that DNN is useful technique.

B. ACO-ANN results

The optimal parameters obtained which leads to good results is summarised in Table 3.

TABLE III. ACO-ANN PARAMETERS

Parameters	Values
Number of hidden layer	1
Neuron number in hidden layer	12
Transfer function of the hidden layer	transig
Transfer function of the output layer	purelin
Numbers of ants	18
Number of iterations	12
Evaporation rate	0.2
Pheromone concentration	0.8

The performance of the ACO-ANN is shown in figure 3 presenting the comparison between the predicted and experimental amount of CO_2 emissions. As a result, the R^2 value obtained in training phase is 0.9890 while the R^2 value in testing phase is 0.9876. The predicted and experimental CO_2 emissions are highly correlated with R^2 closer to unity. It is evident that, ACO-ANN is capable to predict most of the testing and training points closely.

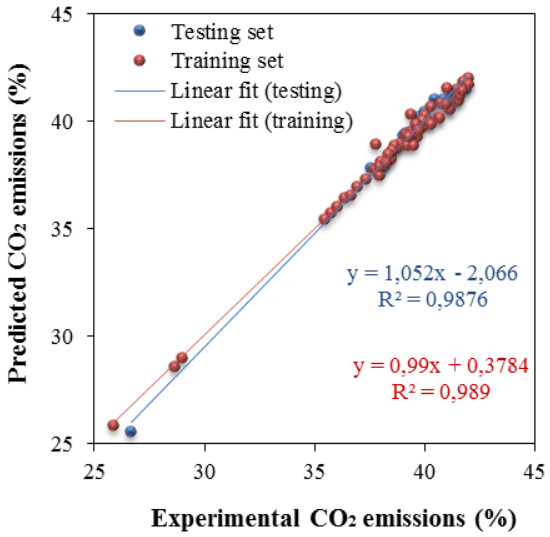

Fig. 3. The comparison between the predicted and experimental CO_2 emissions

Figure 4 show the distribution of the relative error of amount of CO_2 emissions for ACO-ANN during training and testing phases.

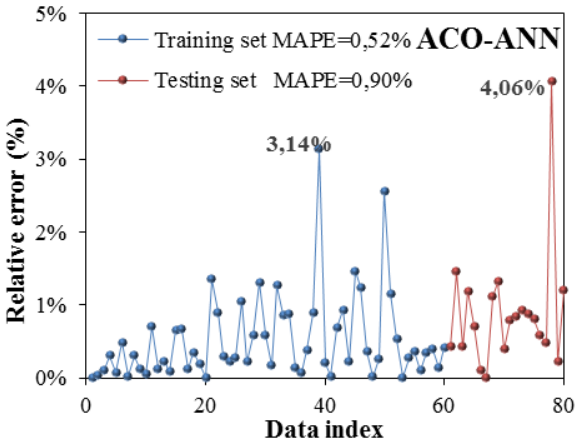

Fig. 4. Distribution of the relative error

As is clearly seen in the figure 4, the average relative error for training and testing phases are 0.52% and 0.9%, respectively indicating the ability of ACO-ANN to find suitable predicted values. Additionally, the relative error does not exceed an error of 3.14% for training phase and 4.06% for testing phase. The performance obtained is reasonable due to the flexibility of ANN in solving complex function and ability of ACO algorithms to optimise initial parameters of ANN.

C. GA-ANN results

It should be noted that optimal architecture of ANN used by ACO-ANN and GA-ANN are the same, which is used mainly for comparisons purposes. The optimal parameters of GA-ANN is listed in Table 4.

TABLE IV. GA-ANN PARAMETERS

Parameters	Values
Number of hidden layer	1
Neuron number in hidden layer	12
Transfer function of the hidden layer	transig
Transfer function of the output layer	purelin
Population size	15
Number of iterations	20
Probability of selecting the best	0.09

Figure 5 presents the comparison between the predicted and experimental amount of CO_2 emissions for two phases. As shown in Figure 5, the GA-ANN can offer good prediction for experimental amount of CO_2 emissions with high values of R^2. However, the R^2 value of the training data is equal to 0.9865. While the R^2 value of the testing data is 0.9834. These values of R^2 indicate a strong linear relationship between the experimental and predicted amount CO^2 emissions during both phases.

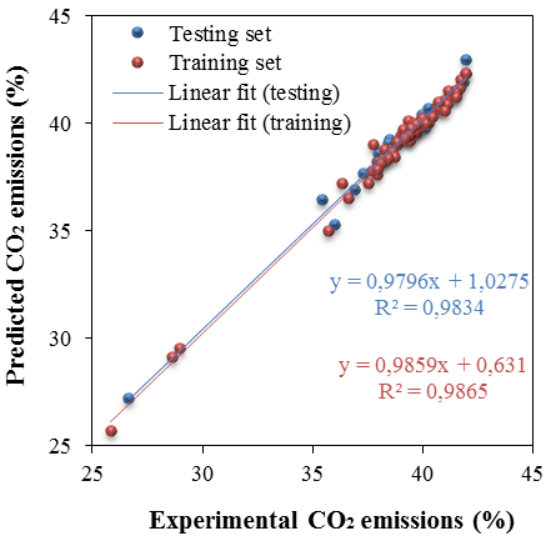

Fig. 5. The comparison between the predicted and experimental CO_2 emissions

Figure 6 a shows the distribution of the relative error between experimental and predicted amount of CO_2 emissions obtained by GA-ANN during training and testing phases. From Figure 6, highest error values of 3.21% and 2.83% are observed in training and testing phases, respectively. The values of MAPE for training and testing phases are 0.71% and 1.04%, respectively. The result obtained from the figure, it still reflects a suitable performance of GA-ANN. The high accuracy of GA-ANN is mostly due to great optimisation of GA and the particularity of ANN in solving non-linear system.

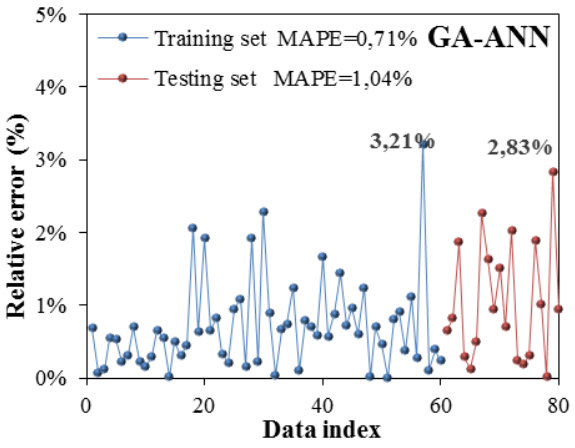

Fig. 6. Distribution of the relative error

D. Comparison between different techniques

The results obtained by DNN, ACO-ANN and GA-ANN during testing phase are compared based on generalisation accuracy. The common performance criterion used to determine the performance are R^2 and MAPE. A comparison of the results obtained from DNN, ACO-ANN and GA-ANN techniques briefly shows in Figure 7.

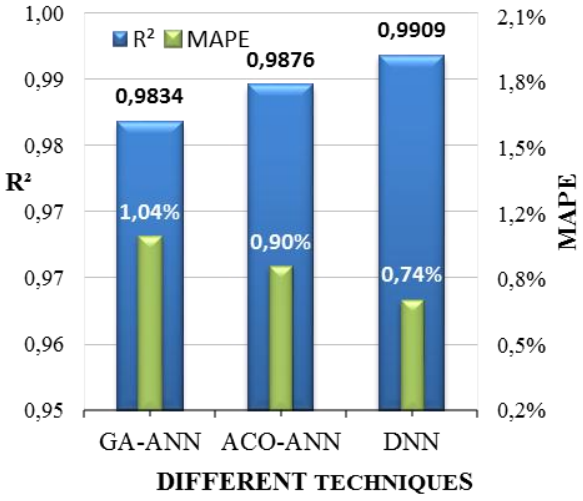

Fig. 7. The comparison between DNN, ACO-ANN and GA-ANN results

The Figure 7 clearly illustrate that all techniques proposed could provide the average relative error less than 1.04% and R^2 more than 0.98. The DNN technique has best generalisation accuracy in a comparison of the other techniques with highest R^2 and lowest MAPE. Whereas, lowest average accuracies at 1.04% is provided by GA-ANN. in addition, the results obtained demonstrate that the DNN is more advanced than ACO-ANN in terms of generalisation accuracy followed by GA-ANN.

TABLE V. OTHER PERFORMANCE INDICES USED

Parameters	GA-ANN	ACO-ANN	DNN
R^2 adjusted	98,25%	98,70%	99,04ù
RMSE	0,4857	0,4080	0,3835

The R_2 adjusted values indicates that 99.04%, 98.7 and 98.25% of data can be explained by DNN, ACO-ANN and GA-ANN respectively. Moreover, values of root mean square error RMSE are very close to 0 mean a higher performance accuracy and the predicted amount of CO_2 emissions are very close to the real data.

V. CONCLUSION

In the current study, the percentage amount of CO_2 emissions from raw materials used in cement clinker production is predicted by the DNN, ACO-ANN, GA-ANN techniques. These techniques are trained and tested using 75% and 25% of experimental data, respectively. The performance of intelligence techniques is judged by R^2 and MAPE as performance criterion.

As a conclusion, the DNN is stamped as superior technique for the prediction of amount of CO_2 emissions with lowest MAPE and highest R^2. The performance accuracy of ACO-ANN technique is higher than the GA-ANN. Each artificial intelligence techniques can provide an useful and interesting techniques to predict the amount of CO_2 emissions produced without needing to actually do the experiment and conventional computational methods. Moreover, it can be used as an alternative technique to calculate the amount of CO_2 emissions. The average relative error obtained by DNN, ACO-ANN AND GA-ANN in testing phase are 0.74%, 0.90% and 1.04%, respectively. According to R^2 adjusted values, there are 99.04%, 98.7 and 98.25% of experimental data can be explained by DNN, ACO-ANN and GA-ANN respectively with very small error.

REFERENCES

[1] E. Possan, W. A. Thomaz, G. A. Aleandri, E. F. Felix, and A. C. P. dos "Santos, CO_2 uptake potential due to concrete carbonation: A case study," Case Stud. Constr. Mater. vol 6, pp. 147-161, June 2017.

[2] M. B. Ali, R. Saidur and M. S. Hossain, "A review on emission analysis in cement industries," Renew. Sust. Energ. Rev. vol 15, pp. 2252-2261, June 2011.

[3] J. Deja, A. U. Bochenczyk, and E. Mokrzyck, "CO_2 emissions from Polish cement industry," Int. J. Greenh. Gas Control, vol 4, pp. 583-588, July 2010.

[4] Z. Cao, L. Shen, J. Zhao, L. Liu, S. Zhong, and Y. Yang, "Modeling the dynamic mechanism between cement CO_2 emissions and clinker quality to realize low-carbon cement," Resour. Conserv. Recycl, vol. 113, pp. 116–126, October 2016.

[5] D. L. Summerbell, C. Y. Barlow, J. M. Cullen, "Potential reduction of carbon emissions by performance improvement: A cement industry case study," J. Clean. Prod, vol. 135, pp. 1327-1339, November 2016.

[6] Y. Boukhari, "Using intelligent models to predict weight loss of raw materials during cement clinker production," Rev. d'Intelligence Artif, vol. 34, pp. 101-110, February 2020.

[7] A. M. Abubakar, E. Behravesh, H. Rezapouraghdam, and S. BahaYildiz, "Applying artificial intelligence technique to predict knowledge hiding behavior," Int. J. Inf. Manag. Sci, vol. 49, pp. 45-57, December 2019.

[8] J. Kim, S. Kim, C. Park, and K. Lee, "Construction of prior models for ES-MDA by a deep neural network with a stacked autoencoder for predicting reservoir production," J. Pet. Sci. Eng, vol. 187, pp. 106800, April 2020.

[9] J. You, R. D. McLeod, and P. Hu, " Predicting drug-target interaction network using deep learning model," Comput. Biol. Chem, vol. 80, pp. 90-101, June 2019.

[10] J. Mohammadhassani, A. Dadvand, Sh.Khalilarya, and M. Solimanpur, "Prediction and reduction of diesel engine emissions using a combined ANN–ACO method,". Appl. Soft Comput, vol. 34, pp. 139-150, September 2015.

[11] H. Zhang, H. Nguyen, X. Bui, T. Nguyen-Thoi, T. Bui, N. Nguyen, D. Vu, V. Mahesh and H. Moayedim, "Developing a novel artificial intelligence model to estimate the capital cost of mining projects using deep neural network-based ant colony optimization algorithm,". Resour. Policy, vol. 66, pp. 101604, June 2020.

[12] I. E. Mulia, H. Tay, K. Roopsekhar, and P. Tkalich, "Hybrid ANN–GA model for predicting turbidity and chlorophyll-a concentrations," J Hydro-Environ. Res, vol. 7, pp. 279-299, December 2013.

[13] W. Liu, Q. Liu, F. Ruan, Z. Liang, and H. Qiu, "Springback prediction for sheet metal forming based on GA-ANN technology," J. Mater. Process. Technol, vol. 187–188, pp. 227-231, June 2007.

[14] R. Katuwal, and P. N. Suganthan, "Stacked autoencoder based deep random vector functional link neural network for classification". Appl. Soft. Comput, vol. 85, pp. 105854, December 2019.

[15] S. Feng, H. Zhou, and H. Dong, " Using deep neural network with small dataset to predict material defects," Mater. Des, vol. 162, pp. 300-310, January 2019.

[16] K. Kannadasa, R. E. Damodar, and K. Venkatanareshbabu, "Type 2 diabetes data classification using stacked autoencoders in deep neural networks," Clin. Epidemiology. Glob. Health, vol. 7, pp. 530-535, December 2019.

[17] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, and P.-A. Manzagol, "Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion," J. Mach. Learn. Res. Vol. 11, pp. 3371–3408, December 2010.

[18] B. A. Olshausen, and D. J. Field, "Sparse Coding with an Overcomplete Basis Set: A Strategy Employed by V1?," Vision. Res, vol.37, pp.3311-3325, December 1997.

[19] M. F. Moller, "A Scaled Conjugate Gradient Algorithm for Fast Supervised Learning," Neural. Netw, vol. 6, pp. 525–533, 1993.

[20] G. D. Angelob, M. Ficcoa, and F. Palmierib, "Malware detection in mobile environments based on Autoencoders and API-images,". J. Parallel. Distrib. Comput, vol. 137, pp. 26-33, March 2020.

[21] R. Jovanovica, M. Tubab, and S. Voß, "An efficient ant colony optimization algorithm for the blocks relocation problem," Eur. J. Oper. Res, vol. 274 pp. 78-90, April 2019.

[22] Y. P. Liu, M. G. Wu, and J. X. Qian, "Predicting coal ash fusion temperature based on its chemical composition using ACO-BP neural network," Thermochim. Acta, vol. 454, pp. 64-68, February 2007.

[23] T. Beltramoa, M. Klockeb, and B. Hitzmanna, "Prediction of the biogas production using GA and ACO input features selection method for ANN model," Inf. Process. Agric, vol.6, pp. 349-356, September 2019.

[24] G. Samira, K. Jarek, and M. Yousef, "A hybrid Genetic Algorithm and Monte Carlo simulation approach to predict hourly energy consumption and generation by a cluster of Net Zero Energy Buildings," Appl. Energy, vol. 179, pp. 626-637. October 2016.

[25] B. Ghorbani, A. Arulrajah, G. Narsilio, S. Horpibulsuk, and M. W. Bo, "Development of genetic-based models for predicting the resilient modulus of cohesive pavement subgrade soils," Soils. Found, vol. 60, pp. 398-412, April 2020.

[26] Y. Boukhari, M. N. Boucherit, M. Zaabat, S. Amzert, and K. Brahimi, "Optimization of learning algorithms in the prediction of pitting corrosion," J. Eng. Sci. Technol, vol. 13, pp. 1153–1164, 2018.