Neural network based prediction of cone side resistance for cohesive soils

Tammineni Gnananandarao[1[0000-0002-3332-8083]], Rakesh Kumar Dutta[2[0000-0002-4611-9950]] and Vishwas Nandkishor Khatri[3[0000-0002-8624-465X]]

[1] NIT Hamirpur, Hamirpur, Himachal Pradesh, India. E-mail: anandrcwing@gmail.com

[2] NIT Hamirpur, Hamirpur, Himachal Pradesh, India. E-mail: rakeshkdutta@yahoo.com

[3] IIT Dhanbad, Jharkand, India. E-mail: vishwas@iitism.ac.in

Abstract. The assessment of soil properties for the design of structure requires a wide range of tests. Sampling difficulty, time and cost constraints forces the practitioners to adopt correlations existing among the in-situ tests and the physical or mechanical properties of soils. This paper presents the application of neural network to predict the cone side resistance (q_s) obtained in the cone penetration test (*CPT*) for the cohesive soil based on plasticity index (*PI*), consistency index (*CI*) and the under drained shear strength (S_u). Feed-forward back propagation algorithm was used for this purpose for the development of neural network model which was developed using 50 in situ dataset collected from the literature. Finally, the cone side resistance obtained from the developed neural network model was compared with the measured cone side resistance obtained from the *CPT* tests reported in literature. Further, the sensitivity analysis was performed to study the impact of plasticity index, consistency index and the under drained shear strength on the cone side resistance. The results of this study reveal that the developed neural network model was able to predict cone side resistance accurately.

Keywords: Cone penetration test; Plasticity index; Consistency index; Drained shear strength; Cone side resistance; Neural network.

1 Introduction

Cone penetration is considered to be one of the best in-situ tests for the ground investigation especially for the footings resting on the soft clay, soft silt and fine to medium sand deposit as well as in classifying the soils. Conducting the cone penetration test in the field is cumbersome, time consuming and costly. Therefore, there is need to develop alternate ways to determine its value based on some simple test such as plasticity index, consistency index and un-drained shear strength. In this context, application of neural network to develop such correlation based on the actual field data collected from the literature can be an option. This paper presents a neural network based model to predict the cone side resistance (force required to push the friction jacket in the cone penetration test equipment) of the soft clay and fine to medium sand deposit based on the plasticity index, consistency index and undrained shear strength (shear stress which a soil can resist without dissipation of pore water pressure). Cone side resistance was the output in this case. Sensitivity analysis was also conducted on the input parameters affecting the cone side resistance. Comparison of the developed neural network model was also made with the model obtained from the multiple regression analysis. Based on the trained weights and biases, finally, an equation is proposed.

2 Background

Application of soft computing techniques in geotechnical engineering is gaining momentum in the recent decade. Researchers were applying these soft computing techniques in various areas such as prediction of residual strength of clay [1], recompression index [2], coefficient of consolidation [3], hydraulic conductivity of bentonite-soil mixes [4], resilient modulus for unbound granular material [5], modulus of elasticity [6], bearing ratio of the clay [7], Deviator stress of sand reinforced with waste plastic strips [8], free swell index for the expansive soil [9], bearing capacity [10-11], settlement [12] of the footings, diameter of jet grouting columns [13], strength and elasticity modulus of granite [14], uniaxial compressive strength of sandstone [15] and abrasiveness index of some Indian rocks [16] using artificial neural networks. These studies have revealed the prediction efficacy of the soft computing techniques. Till date, no study has been reported in literature to predict the cone side resistance for the cohesive soils. This paper attempted to fill this gap. In this paper, a neural network based model was developed to predict the cone side resistance in the cohesive soils from the data collected from the literature. The plasticity index, consistency index, un-drained shear strength was considered as an input while cone side resistance was the output. It is pertinent to mention here that the consistency index (CI) and plasticity index (PI) are two independent input parameters. CI includes the effect of natural water content of the soil whereas the same is absent in PI.

3 Neural Networks and Data Set

Neural networks architecture is good at mimicking the nervous system and the human brain. This technique has the ability to co-relate the input variables with the output variable in order to solve the linear or the non-linear problems. The structure of neural network comprise of number of elements (processing) and nodes or neurons which are generally arranged in layers (input layer or output layer) with one or more hidden number of layers in between. 50 records collected from literature [17] were used as dataset in this study and the range of the dataset are shown in Table 1.

3.1 Neural Network Model and Activation Function Selection

The first step in neural network modelling is to decide the optimal number of layers as well as the neurons in the hidden layer. There is no well-defined procedure to get an optimal neural network system and the parameters setting.

Table 1. Range of data collected from literature

Input & output parameters	data set			
	Minimum	Maximum	Mean	Standard deviation
PI (%)	8	24	15.59	4.06
CI	0.19	1.46	0.87	0.29
S_u (kPa)	17.85	104.38	54.69	23.57
q_s (kPa)	1	121	40.56	26.86

Also the prediction from the neural network is very much dependent on the initial weights and given input parameters. Be that as it may, a tedious trial and error strategy still stays appropriate. Based on the guidelines reported by [18], the authors fixed the optimal hidden layer neurons which are 2/3 of the size of the input parameters. After deciding the hidden layer neurons, the major concern was when to stop the training as excessive training results in noise and inadequate training leads to poor predictions. The authors decided to adopt a trial and error procedure to decide the optimum number of epochs for the training dataset. Additionally mean square error was also calculated between the actual and the predicted values for different epochs and the optimum number of epochs (2000) was chosen based on Figure 1. Therefore, the neural network model for our experiment had a structure of 3-2-1 as shown in the Figure 2 for modeling.

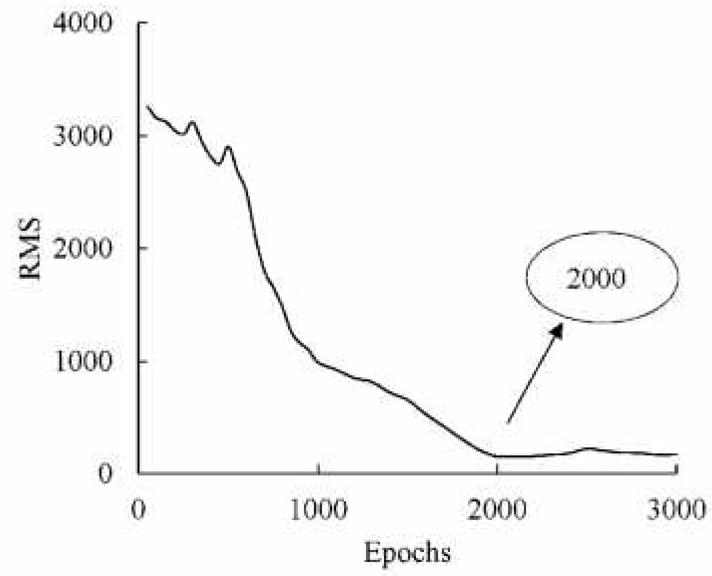

Fig. 1 Deviation of mean square error with number of epochs.

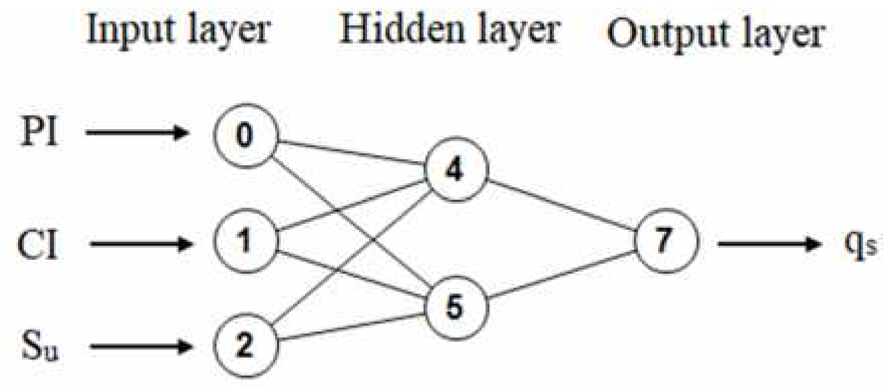

Fig. 2. Artificial neural networks architecture for cone side resistance

Every neuron in the neural network is equipped with an activation function which specifies the output of the neuron corresponding to the given input. This activation functions scale the output of the neural network into proper ranges and also helps to introduce nonlinearity into it. This ability of the activation function makes the neural network powerful. Numerous activation functions are available. Out of which transfer functions are the most common choice for neural network application. The objective of this study is to analyze the performance of the neural network using different activation functions for the neurons in the hidden as well as in the output layers. For the dataset used in this study, linear, sigmoid, sigmoid symmetric, sigmoid symmetric stepwise, gaussian, gaussian symmetric, elliot, elliot symmetric, linear piece, linear piece symmetric, sin, sin symmetric, cos symmetric activation functions which are available in the open source AgielNN software were used.

3.2 Performance Measures Used

After identifying the model, the next step was to assess its performance in predicting the cone side resistance using test data set. There is no general consensus among researchers as far as choosing the best performance measure is concerned. Therefore, accuracy of the cone side resistance prediction is considered as one of the criteria in order to assess the quality of prediction. The various error models used along with their range and interpretation are given in Table 2. The precision of the anticipated cone side resistance was decided using error models such as the coefficient of determination (R^2), correlation coefficient (r), *MSE*, *RMSE*, *MAE* and *MAPE* for the training as well as testing data. The activation function which gave the best measured statistical results was used to select the best activation functions among all. In this study, sigmoid function was chosen based on the above mentioned criteria.

Table 2. Error models with mathematical expressions

Statistical coefficient	Mathematical expression	Range	Interpretation
Correlation coefficient (r)	$r = \frac{\sum q_{st_i} \times q_{sp_i} - \bar{q}_{st} \times \bar{q}_{sp}}{(n-1) S_{q_{st}} S_{q_{sp}}}$	-1 ≤ r ≤ +1	Closer to 1 accurate prediction Closer to zero implies a weak correlation
Coefficient of determination (R^2)	$R^2 = 1 - \frac{\sum_i (q_{sp_i} - q_{st_i})^2}{\sum_i (q_{sp_i} - \bar{q}_{sp_i})^2}$	0 to 1	Closer to 1 accurate prediction
Mean square error (*MSE*)	$MSE = \frac{1}{n} \sum_{i=1}^{n} (q_{st_i} - q_{sp_i})^2$	-	Smaller values represents a better model
Root mean square error (*RMSE*)	$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (q_{st_i} - q_{sp_i})^2}$	-	
Mean absolute error (*MAE*)	$MAE = \frac{1}{n} \sum_{i=1}^{n} \left\| q_{st_i} - q_{sp_i} \right\|$	0 to + ∞	Provides average size of anticipating error when negative signs are overlooked
Mean absolute percentage error (*MAPE*)	$MAPE = \left[\frac{1}{n} \sum_{i=1}^{n} \left\| \frac{q_{st_i} - q_{sp_i}}{q_{st_i}} \right\| \right] \times 100$	<10%	Excellent accurate prediction
		Between 10 to 20%	Good prediction
		Between 20 to 50 %	Acceptable prediction
		>50%	In accurate prediction

Note: q_{st}, q_{sp} target and predicted cone side resistance, $\bar{q}_{st}, \bar{q}_{sp}$: mean of the target and predicted cone side resistance respectively, $S_{q_{st}}$, $S_{q_{sp}}$: standard deviation of the target and predicted cone side resistance respectively, n : number of observations

4 Results and Discussions

Statistical results for the sigmoid activation function are presented in Table 3.

Table 3. Statistical values for the training and testing data

Statistical values for the training data								
	Activation function	*r*	R^2	*MSE*	*RMSE*	*MAE*	*MAPE* (%)
Training	Sigmoid	0.99	0.94	50.39	7.09	5.33	19.47
Testing	Sigmoid	0.99	0.96	26.97	5.19	4.33	18.79

Study of Table 3 indicates that, the statistical values obtained in the present study are within the range shown in Table 2. It means the predicted cone side resistance is reasonable when compared with the actual value (Figures 2 and 3). After each successive completion of neural network process using respective activation function, weights and biases were obtained which were presented in the Table 4 for the sigmoid activation function.

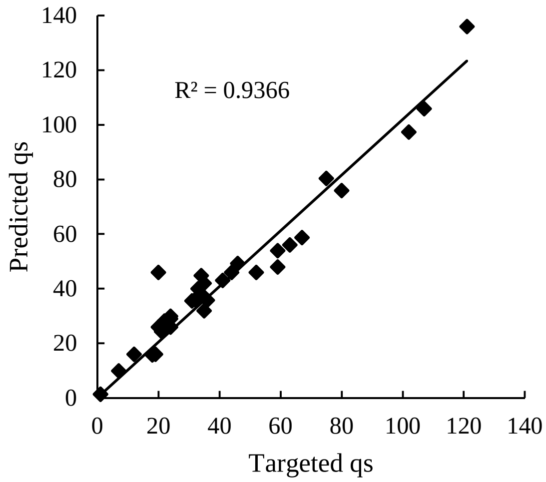

Fig. 3. Measured verses predicted cone side resistance for the training dataset

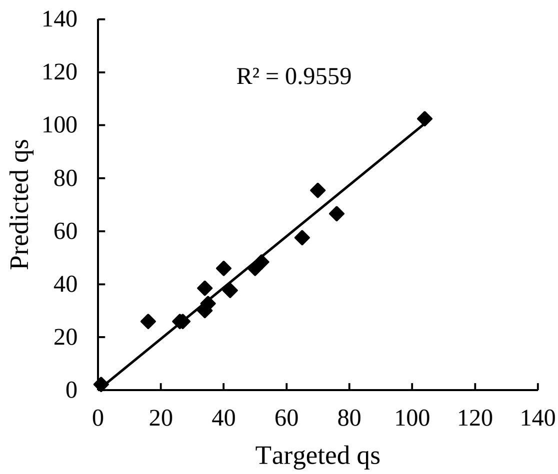

Fig. 4. Measured verses predicted cone side resistance for the testing dataset

Table 4. Trained weights and biases for the sigmoid activation function

	Weights(w_{ji})				Biases	
hidden neurons	*PI* (%)	*CI*	*Su* (kPa)	q_s (kPa)	b_{hk}	b_0
Hidden 1	-0.24	-1.34	2.29	-5.68	-5.82	33.15
Hidden 2	-3.45	6.11	7.46	9.76	-10.30	--

Further the plot between the measured cone side resistance and predicted cone side resistance generates the correlation coefficient (R^2) of 0.93 for training and 0.96 for testing as shown in Figures 3 and 4. The statistical values of mean absolute percentage error was 19.47 for training and 18.79 for testing data set respectively which implies that the predicted cone side resistance is within the permissible degree of accuracy.

4.1 Sensitivity Analysis

In order to study the individual contribution of plasticity index, consistency index, un-drained shear strength on the cone side resistance, a sensitivity analysis was carried using a method reported by [19]. This method was based on weight formation (Table 4). In this analysis the relative importance of individual plasticity index, consistency index, un-drained shear strength was measured. The results of this analysis reveal that the under drained shear strength is the most important input variable for the prediction of cone side resistance. It's impact on the cone side resistance was about 57 % and the impact of the other input parameters was shown in Figure 5.

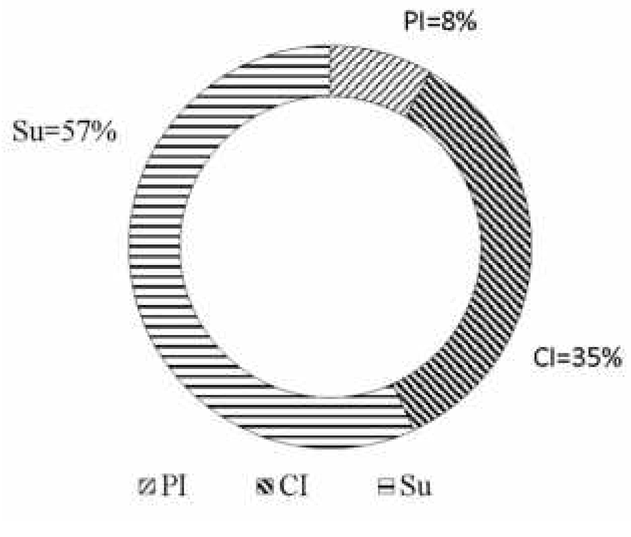

Fig. 5. Impact of input parameters on the output parameter in percentage

4.2 Model Equation

The equation for the cone side resistance prediction can be formulated based on the weights and biases obtained in the trained network which are given in Table 3. The model equation for the cone side resistance is as follows.

$$A = -5.82 - 0.24PI - 1.34CI + 2.29Su$$

$$B = -10.30 - 3.44PI + 6.11CI - 7.45Su$$

$$E = 33.15 - \frac{5.68}{\left(1+e^{A}\right)} + \frac{9.76}{\left(1+e^{B}\right)}$$

$$q_s = \frac{1}{\left(1+e^{E}\right)} \qquad - \ (1)$$

The q_s (kPa) will be in the range of [0 to 1] since the activation function used was sigmoid. Hence, the de-normalization of the output is carried out in order to obtain the actual value. The de-normalized equation is given as

$$q_s\left(kPa\right) = 0.5\left(q_s+1\right)\left(q_{s\max}\left(kPa\right) - q_{s\min}\right) + q_{s\min} \qquad - \ (2)$$

Where q_{smax} is the maximum predicted cone side resistance, q_{smin} is the minimum predicted cone side resistance respectively.

4.3 Richardson's Regression Model Equation

Using the regression analysis, a model equation was proposed based on the Richardson's algorithm. The model equation (3) for the cone side resistance was presented below. For solving this equation (3) the required parameters were similar to the one used for the *ANN* modelling.

$$q_s = e^{\left(3.41PI - 0.69CI + 5.38S_u - 0.29\right)} \qquad - \ (3)$$

The predicted cone side resistance obtained from equation (3) and the measured cone side resistance were plotted in Figure 6.

4.4 Comparison of ANN Model with Richardson's Regression Model

Study of Figures 3 and 6 reveal that the prediction from the *ANN* model and the Richardson's regression model were comparable based on the correlation coefficient (R^2). Further, examining these figures reveals that the prediction accuracy of the *ANN* model was superior to the one using Richardson's regression model.

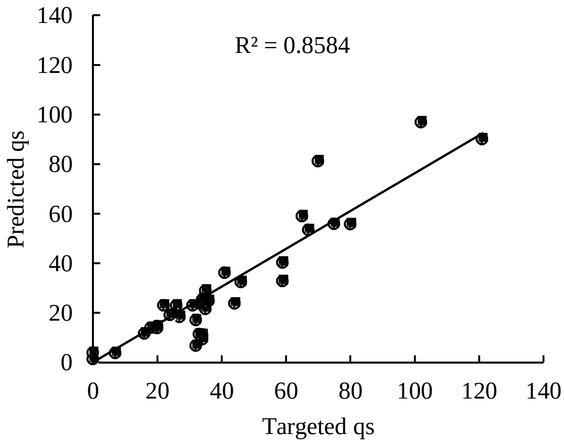

Fig. 6. Measured verses predicted bearing capacity plot for Richardson's regression model

The statistical results (R^2=0.94, *r*= 0.99 *RMSE*=7.09, *MAE*=5.33, *MSE*= 50.39, and *MAPE*=19.47 %) for the training and (R^2=0.96, *r*= 0.99, *RMSE*=5.19, *MAE*=4.33, *MSE*= 26.97, and *MAPE*=18.79 %) for the testing dataset indicated that the neural networks are able to predict the cone side resistance accurately. Further, *ANN* model was found to perform better than the Richardson's regression model. The sensitivity analysis result indicated that the under drained shear strength is the most important parameter affecting the cone side resistance. Finally, an equation based on the trained weights was proposed for use by the geotechnical engineering professionals.

Notations

ANN	= Artificial neural networks
qs	= Cone side resistance
CPT	= Cone penetration test
PI	= Plasticity index
CI	= Consistency index
S_u	= Undrained shear strength
r	= Correlation coefficient
R^2	= Coefficient of determination
MSE	= Mean square error
RMSE	= Root mean square error

MAE	= Mean absolute error
MAPE	= Mean absolute percentage error
q_{st}	= Target cone side resistance
q_{sp}	= Predicted cone side resistance
$\bar{q}_{st}$	= Mean of the targeted cone side resistance
$\bar{q}_{sp}$	= Mean of the predicted cone side resistance
$S_{q_{st}}$	= Standard deviation of the targeted cone side resistance
$S_{q_{sp}}$	= Standard deviation of the target predicted cone side resistance
n	= Number of observations
f	= Optimum activation function
b_o	= Bias at the output layer
h	= Number of neurons in the hidden layer
w_k	= Connection weight between kth neuron of hidden layer and the single output neuron
b_{hk}	= Bias at the kth neuron of the hidden layer
m	= Number of neurons in the input layer
W_{jk}	= Connection weight between jth input variable and kth neuron of hidden layer
X_j	= Normalized input variable j in the range [-1, 1]

References

1. Khana, S.Z., Suman, S., Pavani, M., Das, S.K.: Prediction of the residual strength of clay using functional networks. Geoscience Frontiers, 7(1), 67- 74 (2016).
2. Kordnaeij, A., Kalantary, F., Kordtabar, B., Mola-Abasi H.: Prediction of recompression index using GMDH-type neural network based on geotechnical soil properties. Soils and Foundations, 55(6), 1335- 1345 (2015).
3. Pham, B.T., Nguyen, P.M.D., Bui, K.T.T., Prakash, I., Chapi, K., Bui, D.T.: A novel artificial intelligence approach based on Multi-layer Perceptron Neural Network and Biogeography-based Optimization for predicting coefficient of consolidation of soil. Catena, 173, 302- 311 (2019).
4. Mishra, A.K., Kumar, B., Dutta, J.: Prediction of Hydraulic Conductivity of Soil Bentonite Mixture Using Hybrid-ANN Approach. Journal of Environmental Informatics 27(2), 98-105 (2016).
5. Saha, S., Gu, F., Luo, X., Lytton, R.L.: Use of an Artificial Neural Network Approach for the Prediction of Resilient Modulus for Unbound Granular Material. 2672(52), 23-33 (2018).

6. Sharma, L.K., Singh, R., Umrao, R.K., Sharma, K.M., Singh, T.N.: Evaluating the modulus of elasticity of soil using soft computing system. Engineering with Computers, 33(3), 497- 507 (2017).
7. Taskiran, T.: Prediction of California bearing ratio (CBR) of fine grained soils by AI methods" Advances in Engineering Software. 41(6), 886-892 (2010).
8. Dutta, R.K., Dutta, K., Jeevanandham, S.: Prediction of deviator stress of sand reinforced with waste plastic strips using neural network. International Journal of Geosynthetics and Ground Engineering 1(11), 1-12, (2015).
9. Dutta R.K., Singh, A., Gnananandarao, T.: Prediction of Free Swell Index for the Expansive Soil Using Artificial Neural Networks. Soft Computing in Civil Engineering, 3(4), 47-65 (2019).
10. Gnananandarao, T., Dutta R.K., Khatri, V.N.: Artificial neural networks based bearing capacity prediction for square footing resting on confined sand. In: Indian Geotechnical Conference, 14-16 December, IIT Guwahati, Assam, India, (2017).
11. R.K. Dutta, R. Rani, T. Gnananandarao, "Prediction of ultimate bearing capacity of skirted footing resting on sand using artificial neural networks." Journal of Soft Computing in Civil Engineering, 2(4), 34-46, 2018.
12. T. Gnananandarao, R.K. Dutta, V.N. Khatri, "Application of artificial neural network to predict the settlement of shallow foundations on cohesionless soils." Indian Geotechnical Conference, 15-17 December, IIT Madras, Chennai, India, (2016).
13. Ochmańskia, M., Modonib, G., Bzówka, J.: Prediction of the diameter of jet grouting columns with artificial neural networks. Soils and Foundations, 55(2), 425-436 (2015).
14. Armaghani, D.J., Mohamad, E.T., Momeni, E., Monjezi, M., Narayanasamy, M.S.: Prediction of the strength and elasticity modulus of granite through an expert artificial neural network. Arabian Journal of Geosciences, 9(48), 1-16 (2016).
15. Armaghani, D.J., Amin, M.F.M., Yagiz, S., Faradonbeh, R.S., Abdullah, R.A.: Prediction of the uniaxial compressive strength of sandstone using various modeling techniques" International Journal of Rock Mechanics and Mining Sciences. International Journal of Rock Mechanics and Mining Sciences, 85, 174-186, (2016).
16. Tripathy, A., Singh, T.N., Kundu, J.: Prediction of abrasiveness index of some Indian rocks using soft computing methods. Measurement, 68, 302-309, (2015).
17. Cheshomi, A.: Empirical relationships of CPTu results and undrained shear strength. Journal of GeoEngineering, 13(2), 49-57 (2018).
18. Boger, Z., Guterman, H.: Knowledge extraction from artificial neural network models. IEEE Int Conf Comput Cybern Simul 4, 3030–3035, (1997).
19. Garson G.D: Interpreting neural-network connection 704 weights. AI Expert 6(4), 46–5 (1991).