University College of Takestan

Volume 1,Issue 1

Available online at **http://UCTjournals.com**

UCT Journal of Research in Science ,Engineering and Technology

UCT . J. Resea. Scien. Engineer.Techno.(UJRSET)

1(1): *01-08 (2014)*

ISSN:2382-977X

Developing Integrated Model of Data Envelopment Analysis (DEA) and Data Mining to Evaluate Performance of Logistics Service Suppliers

Saeed Tavakolimoqadam [1], Gholamreza Zamanian [2], Faranak Hosseinzadeh Saljooghi [3]

[1] *Department of Industrial Engineering, Zahedan branch, Islamic Azad university, Zahedan, Iran.*
[2] *Assistant Professor in Economics, University of Sistan and Baluchestan, Zahedan, Iran.(corresponding author)*
[3] *Department of Mathematics, University of Sistan and Baluchestan, Zahedan, Iran.*

ABSTRACT

Today, finding material, good, and semi-finished parts suppliers in a competitive atmosphere is significantly easier, therefore there are many options to select from for commercial partners. On the other hand, purchase responsibility as a management decision has challenged company managers due to its complexity and variety of evaluation measures. The present study provides an integrated model using data envelopment analysis (DEA) techniques to develop decision support systems in an auto part supplier, and methods of data mining like neural networks in order to evaluate suppliers. Efficiency scores of each supplier are obtained by solving model of "data envelopment analysis", and then this model, using educational data, trains artificial neural networks to predict and rank suppliers. Results of the selected model provide a complete ranking and an appropriate grouping with an acceptable level of prediction accuracy respectively to evaluation decision making and selection of suppliers. The main objective of this study is to evaluate performance of logistic service suppliers, and also seeks to find an answer about how to use methods of data envelopment analysis and data mining in evaluation of suppliers. Time percentage measure of "timely delivery" of parts has the maximum efficiency and the measure of "ability to reduce price" has the minimum efficiency. In the other hand, once this model is used, there is no need to modeling and resolve data envelopment analysis models with high computational volume. Indeed, application of neural networks has been able to eliminate defect of prediction disability in data envelopment analysis.

Original Article:

Received 08 Oct. 2014
Accepted 17 Nov. 2014
Published 30 Dec. 2014

Keywords:
Supply chain management (SCM), Data envelopment analysis (DEA), Artificial neural networks (ANN), Data envelopment analysis cross efficiency (DEACE)

1.Introduction

One of the most essential issues in supply chain, which has been among main concerns of the management, is decision-making. The necessity of making the right decision and selecting an appropriate option among various possibilities, finding best producers and distributors, best area to attract customers, best commercial partners in integrations and such are of the most important issues of decision-making in supply chain. These decisions include micro as well as macro issues, and sometimes if the wrong decision is made great costs must be paid. Hence, right decision-making in all the issues is important. Generally, in most of decision-making issues there are various goals and factors and the decision-maker try to choose the best option among many available ones. To evaluate suppliers, it is necessary to consider various measures and factors which are in fact input data for evaluation and their output is suppliers ranking according to which selection of suppliers is done. It would be important in many aspects such as using best suppliers, reducing total cost of manufactured products, reducing cost of supplier management, using all the supplier facilities, potential of supplier development to evaluate and select them.

The issue of supplier selection has been considered significantly in recent decays in business and practical experience literature. In fact, selection of an appropriate set of suppliers to work with is essential in success of a company and has been confirmed during a long period of time [3]

Supply sector has a key role in efficiency and effectiveness of organization and also has a direct effect on cost reduction, profitability and flexibility of a company. One of the most important issues of supply chain is purchase. Evaluating vendors in supply chain is one of the most important processes of decision-making that includes not only vendors, but also other decisions like order quantities of each vendor [3]

Method of data envelopment analysis (DEA) is one of the tools to evaluate suppliers and relative comparing of them. However, classic DEA often fails due to data restriction, statistic noises and outlier data and cannot absolutely measure their efficiency degree. Also, it should be noted that methodology analysis and DEA solution interpretation is very critical. Therefore, a general framework should be developed to enable DEA to elicit interesting and important insights systematically and to achieve this perspective of data mining for DEA results structure is developed. Data mining can be suitable for analyzing different DEA models [7]

Data mining approaches such as neural networks and expert systems are used to predict performance of a new manufacturer. The main advantage of data mining algorithms includes using a set of massive data to develop model via effective analysis of procedures and mining incomprehensible classification models and converting them to comprehensible models in evaluation of supplier performance. In this regard, the present study tries to provide an integrated model, mentioned beforehand. Based on the previous research and available information the main concern of SAPCO Company has been evaluation of logistic service suppliers. Since SAPCO Company is connected with many suppliers, based on experiment and elites measure of the company and the field literature, in the present study we are seeking suppliers with the maximum efficiency. What explains the necessity of every scientific research is to express its main goal. In other words, it is based on the goal that it is determined which and how activities are planned and decisions and activities efficiency is measured accordingly. The required goal is the success of each activity and determines the ways of doing it. The main goal of the present study is to evaluate performance of logistic service suppliers in SAPCO Company.

2. Literature Review

Supplier selection is an important part of supply chain because producers spend 60 percent of their time for supplying material and parts. Furthermore, 70 percent of production costs are related to good and service purchase. Therefore, selection and specification the best supplier is an important issue in supply chain which should be examined strategically. In fact, input of supplier evaluation are measures and factors and its output in suppliers ranking based on which best supplier is selected. In table 1, some of the most frequent measure in different industries, which mostly are resulted from 23 measures of Dixon [6] are presented in summery.

Table 1: different measures of supplier selection [6]

Selection Measures	J	I	H	G	F	E	D	C	B	A
Price	√	√	√	√	√	√	√	√	√	√
Quality		√		√	√		√		√	√
Delivery		√		√	√		√	√	√	√
Warranties and claims								√		√
After sales service				√		√		√		√
Technical support					√	√		√		
Educational assistance				√				√		√
Attitude				√		√				√
Performance experience				√						√
Financial position				√				√		√
geographical location				√			√		√	√
Management and organizing				√			√			√
Personnel relations				√						√
Communication system				√						√
Responding to customer request					√			√		
e-commerce ability	√	√	√							
JIT ability			√		√					
Technical ability			√						√	√
Production facilities and capacity				√						√
Packaging ability				√						√
Operational controls				√						√
Ease of use						√		√		
Maintenance ability						√		√		
Past experience				√				√	√	√
Reputation and position in industry				√		√		√	√	√
Mutual agreements				√			√		√	√
Previous cooperation				√		√				√
Eco-friendliness		√								
Product appearance	√									
Technology catalogue	√									

In this study, 8 measures of 23 Dixon measures were selected for SAPCO Company as input variables and efficiency of these 8 measures as output variables. As mentioned before, the reasons of these 8 measure selection are "experience, elites & expert's opinion, literature and SAPCO Company involvement with them and their most roles in this company". These 8 measures are: number of employees, previous cooperation, scores of management systems, ability to reduce price, time percentage of timely delivery of parts, quantity percentage of timely delivery of parts, and sale volume.

In the scope of supply chain management, evaluation of suppliers is mostly based on various measures and indices, as well as single model techniques and some integrated models, and there is no single approach in doing so. Over time with advances of research in the scope, the number and types of evaluation measures and techniques have been developed. Method of data envelopment analysis (DEA) is one of the tools to evaluate suppliers and relative comparing of them. [2 and 9]

Tried to estimate production function parameter in order to achieve maximum efficiency using DEA. In their study, integrated method of DEA and MDE is used [4]
Developed a decision-maker support system using DEA to monitor efficiency based on organization`s key performance in order to evaluate and manage relative performance of the organization [5]
In the present study, our goal is to integrate DEA technique with data mining methods to eliminate restrictions such as disability to predict, impossibility of comparing heterogeneous units, and disability of DEA basic models to specify measurement error; therefore, we use neural networks method.

2.1. Recent research process

1. The rudimentary illustration of suppliers.
20The determination of each supplier performance using data envelopment analysis.
30The ranking of suppliers with use of data mining.
The research model is presented in following steps:

1. **Data collection:** in this step, data collection to evaluate suppliers is done. It is focused on how to produce and collect data. Generally, there are two methods to do so. The first one is when data production process is done under expert or designer supervision. It is known as designed experiment. The second one is when the expert cannot have any effect on data collection process. It is known as observation-based method.
2. **Data preparation and normalization:** in this step an operation is done in order to increase data quality and transform them to an appropriate format in accordance with data mining technique such as making features common size, dealing with remote data, and dealing with missing data.
3. **Evaluation:** primary evaluation of suppliers` efficiency using DEA.
4. **Modeling:** input-output modeling using artificial neural network, which goal is modeling relation between effective inputs on suppliers` efficiency and their achieved efficiency degree by DEA.
5. **Test:** testing neural network and analyzing useful sensitivities associated with variables.

3. Research Method

The artificial neural networks are systems to process information which their mechanism is inspired from biological neural systems. Artificial neural networks, also called briefly as neural networks, are consisted of several interconnected process units, called neurons. Neuron is the smallest unit of information processor in neural networks. Neurons of artificial neural networks are mathematical models that shows neurons performance in organisms` neural networks mathematically. Mathematical neuron model is usually a static neuron which has some inputs and one output. Each of the inputs is multiplied by its related weigh, and then enters neuron. Weighted amounts of inputs are added to each other and then their total weigh is passed through a (usually linear) function.
In this study, multi-layer Perceptron and error back-propagation learning algorithm are used. Error back-propagation algorithm is part of supervised learning methods. This algorithm, based on downward slope optimization, minimizes the mean squares error as a performance index. Error back-propagation algorithm is constituted of two main passes. First one is called forward pass in which input vector is applied to network and its effects are transferred through middle layers to output layer and output vector of the network is formed which is the real response of the network. In this pass, network parameters do not change. Second pass is called backward. In this pass, network parameters are regulated. The regulation is done by error correction. Error signal is formed in output layer. Error vector is equal to the difference between desired response and real response of the network. In the backward pass, error value is distributed from output layer via network layers into the whole network. Since this distribution occurs in the opposite direction of weight communication, the term error back-propagation is selected to explain behavior of this algorithm. Network parameters are regulated such that real response of the network gets closer to the desired response as much as possible [8 and 10]
Figure 1. shows a general algorithm of multi-layer Perceptron type of neural networks with error back-propagation algorithm.

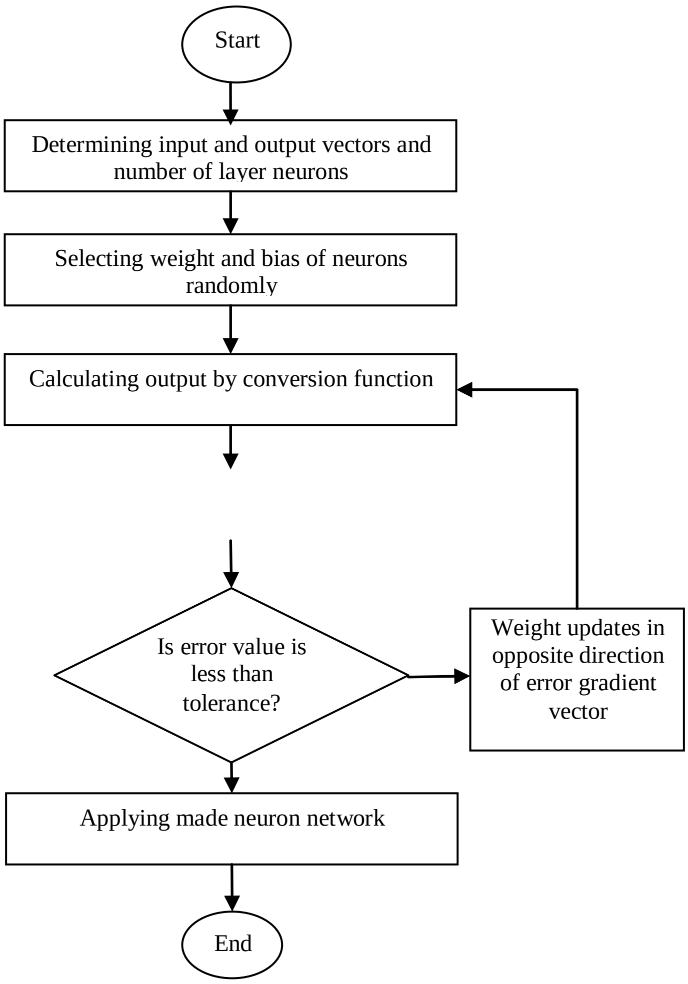

Figure 1: Artificial neuron networks [8]

The research methodology is one of descriptive survey. Statistical population of the study is SAPCO Company and sampling is done randomly.
Following, steps of proposed model implementation are

provided:

First step: defining measures and indices of input/output of auto parts suppliers` evaluation

The evaluation and selection of an appropriate supplier are important commercial objectives of car companies. In this regard, several steps to evaluate potential suppliers are considered. After primary inquiry to get informed about supplier`s preparation to cooperate with the company, resources of information suggested by construction deputy are evaluated, suppliers are examined regarding available measures and indices, and finally after selecting the top company, cost analysis, cooperation contract type, and contract circulation for each part separately are determined.

Table 2: calculation formula of selected input/output indices to evaluate supplier

Index	**Type**	**Monitoring official**	**Data base**	**Calculation formula**
Number of employees (EMP)	Input	Quality deputy	Resource information file	EMP: number of employees (person)
Previous cooperation (COP)	Input	Commerce deputy	Vendors` information file	CO: previous cooperation with SAPCO Company (year)
Scores of management systems (AS)	Input	Quality deputy	Resource evaluation file	AS: based on evaluation questions of suppliers` systems (score from 100)
Ability to reduce price (PRICE)	Input	Financial/economical deputy	Purchase accounting file	$P = \frac{P1 - Q * P2}{Q * P2}$ (*normal* P) PRICE: ability to reduce price P: percentage of price deviant P1: total price of delivered part in calculation interval P2 : part`s base
Time percentage of timely delivery (DELT)	Output	Planning deputy	Orders file	$A_i = \frac{D - S}{L}$, $DELT_i = 100 - ((A_i - 1) * 100)$ A_i: lag of part-manufacture
Quantity percentage of parts timely delivery (DELQ)	Output	Planning deputy	Order and parts receipt file	$DELQ = \frac{Q_1}{Q}$ DELQ: quantity percentage of supplier`s timely delivery Q_1: number of required supplied parts Q : number of ordered parts of supplier for Iran-
Quality	Output	Quality deputy	parts receipt	$PPM = \frac{R1}{D1} * 10^6$ Q: quality index PPM: number of supplier`s returned parts in one million delivered parts R1: total returns of supplier from
Sale	output	Financial/economical deputy	Contracts file	SALE: sum of delivered parts from the contract of part-manufacture multiply by part-manufacturer price

Second step: Data normalizing

The information about 8 input and output indices related to 110 suppliers are the numeric types and are normalized by minimax method below for data mining. Data normalizing increases the power of neural network.

Third step: Determining suppliers` efficiency via DEA crosses efficiency

In this step linear planning method of DEA cross efficiency is used to implement the first step.

Forth step: Model performance modeling using artificial neural networks

In this step, an efficiency modeling of suppliers is done using artificial neural network. Data used to train neural network in this step are as follows, and in fact, modeling neural network do the f function.

Fifth step: Testing neural network and analyzing useful sensitivities related to variables

The evaluating accuracy of results must be done by model test on a new and independent set of data. In doing so, data sets must be divided to training data and test subsets. In this method training data set is used to model learning and test data set is used to evaluate accuracy and results of the model. To train neural network model, information of 100 suppliers are used as training data and information of 10 suppliers are used as test data. Mean absolute percentage error (MAPE) and root mean square deviation (RMSE), as well as best model in certain iteration of training steps in neural network (end of training condition) are determined for each of them.

The mean absolute percentage error (MAPE) and root mean square deviation (RMSE) are calculated according to below formulas for each model:

$$MAPE = 100 \times \frac{1}{N}\sum_{i=1}^{N}\frac{|Y_i - P_i|}{Y_i} \quad (1)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(Y_i - P_i)^2} \quad (2)$$

Where Y_i is the real value and P_i is predicted value and N is the number of data.

After calculating model efficiency using DEA cross efficiency, we can train neural network.

4. Results

The sensitivity analysis is one of the methods to measure input variables effect on target variable in researches. In sensitivity analysis, values of an input variable in different samples change and its effect on modeling target variable is evaluated. The variable with most effect on the modeling is determined in this way. The used method for sensitivity analysis in this study is as follows:

1. First, for each of the input variables a value is added to them in different samples as noise. (in this study, 0%, (not adding value to variable), 2%, 5%, and 10% are considered.
2. Also, for all the samples, minimum, average, and maximum of input variable are considered.
3. Neural network training is done with new data and RMSE of the new model is calculated.
4. The average of change in RMSE for above models is calculated.
5. The higher the rate, more important the variable will be.

After calculating model efficiency using DEA cross efficiency method, we can train neural network. In this step, information of suppliers to use in neural network training is as follows:

Table 3: a sample of suppliers` information to use in neural network training

Real Efficiency(DEA)	SALE	QUALITY	DELQ	DELT	PRICE	AS	COP	EMP	Sup
0.834	0.0637	0.9745	1	0.9843	0.6426	0.2222	0.6429	0.0385	**Supp99**
0.709	0.0884	0.8832	0.916	0.6043	0.352	0.1667	0.6429	0.0881	**Supp98**
0.686	0.6667	0.0099	0.9219	0.8931	0.4698	0.8889	0.7143	0.6788	**Supp97**
0.596	0.2253	0.0098	0.7545	0.6864	0.2209	0.8333	0.7143	0.6586	**Supp96**
0.811	0.6479	0.0098	0.9949	0.989	0.5165	0.2222	0.7143	0.6069	**Supp95**
0.851	0.0334	0.9161	0.9952	0.9269	0.4362	0	0.7143	0.0456	**Supp94**
0.736	0.6196	0.8862	0.9997	0.4034	0.516	0	0.7143	0.1844	**Supp93**
0.599	0.2088	0.0078	0.7392	0.8246	0.9213	0.6111	0.7143	0.1054	**Supp92**
0.622	0.1882	0.0098	0.297	0.9838	0.3046	0.9444	0.7143	0.0476	**Supp91**

Output variable in neural network model	Input variable in neural network model	**Variable type**

In the step of training neural network, information of 100 suppliers and in the test step, information of 10 suppliers is used. After several tests, a suitable structure of neural network with minimum error as a neural network with two hidden layer, that there are 4 neurons in first layer and 3 neurons on second layer, is obtained. Therefore, structure of neural networks is 4-3-1 that is exhibited in below figure. (Input layer has 8 neurons which are considered to the number of input variables). Parameters of neural networks are as follows:

- The rate of neuron network learning is set on 0.8.
- The number of iteration of training error back-propagation algorithm is determined 5000.

Figure2 shows convergence diagram of neural network in different training iterations. (Horizontal axe of training iterations and vertical axe of RMSE show neural network model).

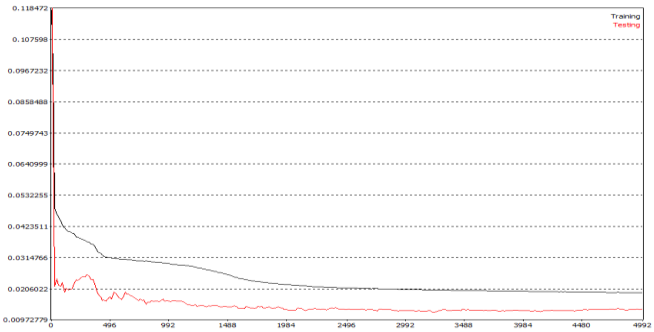

Figure 2: convergence diagram of neural network model

As can be seen in above figure, neural network model can model suppliers efficiency with high accuracy (either in training step or test step). Table 4 shows error measure of MAPE and RMSE for trained model.

Table 4: error measures of MAPE and RMSE

	RMSE	*MAPE*
TRAIN	0.924	2.470156705
TEST	0.015131	1.590018296

More in this section prediction for train and test data is provided. (Horizontal axe represents number of supplier and vertical axe shows its efficiency).

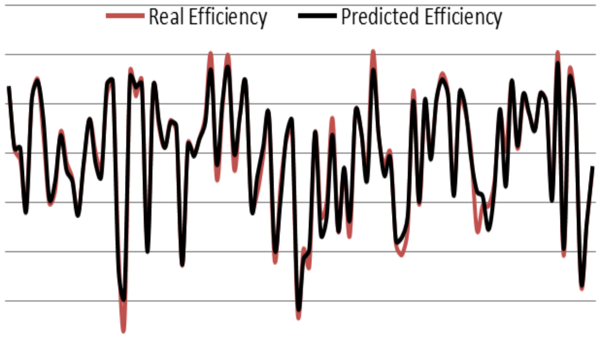

Figure 3: predicting neural network model for train data

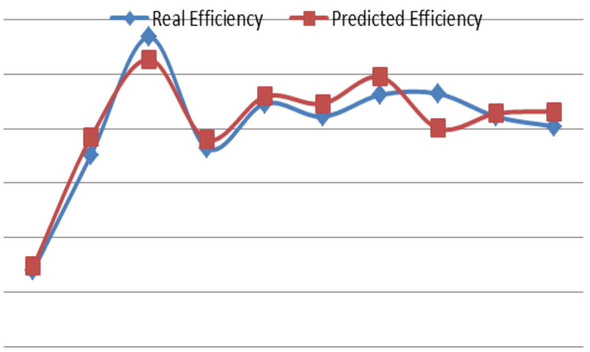

Figure 4: predicting model of neural network for test data

Also, real and predicted values by neural network data for suppliers` efficiency test are presented in table 5:

Table 5: real and predicted values by neural network for suppliers efficiency test

Real Efficiency	Predicted Efficiency
0.67	0.673766
0.776	0.792431
0.884	0.863686
0.782	0.789753
0.823	0.829469
0.811	0.822902
0.831	0.847215
0.832	0.800108
0.811	0.813954
0.802	0.8157

Furthermore, weights of neural networks resulted from train process are obtained as follows.

Table 6: weights of neural networks result from train process

From the input layer				
To be 1th hidden layer	1th neuron	2th neuron	3th neuron	4th neuron
Bias	-3.15242	-2.17868	-4.65374	2.89653
1th neuron	0.934112	1.27835	1.64669	0.6072
2th neuron	-1.00279	2.49111	-0.0254144	0.139794
3th neuron	2.84801	13.2496	-2.18592	0.640613
4th neuron	-0.434773	-1.1879	-0.306722	0.0310427
5th neuron	1.77089	0.43191	-0.253751	-0.831586
6th neuron	1.11495	1.21676	-0.85033	-0.7835
7th neuron	-1.86683	-0.638922	-5.87487	-1.08814
8th neuron	0.505872	4.96869	1.00796	-0.419125

From the 1th hidden layer			
To be 1th hidden layer	1th neuron	2th neuron	3th neuron
bias	-0.2755	0.487408	0.546142
1th neuron	-1.29353	4.97101	0.0479924
2th neuron	-1.26769	-5.24385	0.196997
3th neuron	3.16619	-1.35249	-2.87583
4th neuron	1.22696	-2.20813	-2.10848

To be 1th hidden layer		
1th neuron	bias	0.553515
From the 2th hidden layer	1th neuron	-5.16656
	2th neuron	10.0104
	3th neuron	1.14728

4.1. Sensitivity Analysis of Effective Variables on Suppliers Efficiency

The sensitivity analysis is one of the methods to measure input variables effect on target variable in researches. In sensitivity analysis, values of an input variable in different samples change and its effect on modeling target variable is evaluated. The variable with most effect on the modeling is determined in this way. The used method for sensitivity analysis in this study is as follows:

1. First, for each of the input variables a value is added to them in different samples as noise. (In this study, 0%, (not adding value to variable), 2%, 5%, and 10% are considered.
2. Also, for all the samples, minimum, average, and maximum of input variable are considered.
3. Neural network training is done with new data and RMSE of the new model is calculated.
4. The average of change in RMSE for above models is calculated.
5. The higher the rate, more important the variable will be.

Table 7: rate of changes in RMSE in each step of sensitivity to inputs

RMS After Addition of Noise to The Training Data							
Noise	0%	2%	5%	10%	Set to Min.	Set to Mean	Set to Max.
1th input	0.0192415905	0.0194880705	0.0207337441	0.0237356607	0.0331693448	0.0650786303	0.1330443329
2th input	0.0192415905	0.0197713825	0.0223562047	0.0281222683	0.1041151582	0.0631646690	0.0485646636
3th input	0.0192415905	0.0210835129	0.0283200720	0.0430103024	0.130587185	0.1189184764	0.1630339717
4th input	0.0192415905	0.0193048306	0.0196333213	0.0204821636	0.0416229410	0.0257367668	0.0476791006
5th input	0.0192415905	0.020891431	0.0279547162	0.0408278296	0.2098426597	0.1201877000	0.1043487601
6th input	0.0192415905	0.0198728345	0.0228966216	0.0294355072	0.1832640883	0.1028239138	0.0609109189
7th input	0.0192415905	0.0202389436	0.0248945829	0.0346526273	0.2072796543	0.1139493870	0.0911511115
8th input	0.0192415905	0.0196017638	0.0213787753	0.0254976111	0.0370478518	0.0591725439	0.0893925642

The above table shows average rate of changes in RMSE relative to each of input variables.

Table 8: average rate of changes relative to each of input variables

Parameter		RMSE	Rank
1th input	EMP	0.029967	5
2th input	COP	0.028441	6
3th input	AS	0.064909	2
4th input	PRICE	0.009835	8
5th input	DELT	0.068101	1
6th input	DELQ	0.050626	4
7th input	QUALITY	0.062786	3
8th input	SALE	0.022774	7

The results of sensitivity analysis show that variable of “time percentage of timely delivery DELT” is a variable with greatest effect on suppliers` supply efficiency, and this is reasonable because failure to timely delivery of parts to an automaker company is practically followed by production line suspension and incurring high expenses to management. The second important variable is scores of management system (AS). It is clear that how much more scores a supplier can earn, its supply ability is higher and therefore its efficiency is high too. Quality is another important factor in suppliers` efficiency. The interesting note in this study is that price variable is placed in the last row of effectiveness on suppliers` efficiency. Regarding that data analysis is done for auto industry suppliers in this study, we can conclude that the auto company under study is focused more on delivery time and quality and is ready to pay more to achieve them.

5. Conclusions

The above mentioned model can be used as a support system of decision-making in evaluating suppliers. Receiving input information, it is able to predict efficiency of new suppliers, and thus, an organization can plan its cooperation policies according to suppliers` efficiency. Also, as was observed, this system can be used for simulation variable effects on supply efficiency which provide key information for management decisions to organizations` managers.

References

[1]. Desheng Wu, 2009. Supplier selection: A hybrid model using DEA, decision tree and neural network. Expert Systems with Applications, **36**, 9105–9112.

[2]. 2-Ertek, G., Buyukozkan, G., Akcay, A.E., 2012. analyzing the solutions of DEA through information visualization and data mining techniques: Smart DEA framework, Expert Systems with Applications **39**, 7763–7775.

[3]. 3-Hong Zhang, Chao-lin Gu, Lu-wen Gu, Yan Zhang, 2011. The evaluation of tourism destination competitiveness by TOPSIS & information entropy : A case in the Yangtze River Delta of China. Tourism Management **32**, 443-451.

[4]. 4-Pendharkar, P.C., Troutt, M.D., 2012, DEA based data preprocessing for maximum decisional efficiency linear case valuation models, Expert Systems with Applications 39, 9435–9442.

[5]. 5-Samoilenko, S., Osei-Bryson, K.M., 2013, Using Data Envelopment Analysis (DEA) for monitoring efficiency-based performance of productivity-driven organizations: Design and implementation of a decision support system, Omega, in press.

[6]. 6-Sung Ho Ha, Ramayya Krishnan, A hybrid approach to supplier selection for the maintenance of competitive supply chain, Expert Systems with Applications 34, 2008.1303-1311.

[7]. 7-Seol, H., Lee, S., Kim, C., 2011. Identifying new business areas using patent information: a DEA and text mining approach. Expert Systems with Applications, 38(4), 2933–2941.

[8]. 8-Shahrabi J., Hadavandi, E., Asadi, S., 2013. Developing a hybrid intelligent model for forecasting problems: Case study of tourism demand time series, Knowledge-Based Systems **43**, 112–122.

[9]. M. Shahriari. 2013 .Ranking Network-Structured Decision-Making Units and Its Application in Bank Branches .International Journal of Industrial Mathematics (IJIM) 5 (4), 6

[10]. M Shahriari. 2011. Malmquist Productivity Index for Two-stage Structures and its Applications in Bank Branches. International Journal of Industrial Mathematics (IJIM) 4 (4), 11