Research Journal of Poultry Sciences 5 (3): 39-49, 2012
ISSN: 1993-5285

Improve Poultry Farm Efficiency in Iran: Using Combination Neural Networks, Decision Trees and Data Envelopment Analysis (DEA)

[1]Iman Rahimi and [2]Reza Behmanesh
[1]Department of Mathematics, University of Sistan and Baluchestan, Zahedan, Iran
[2]Department of Accounting, Khorasgan (Isfahan) Branch, Islamic Azad University, Isfahan, Iran

Abstract: Since, poultry meat farming sub-sector has high potential for enhancing the agriculture industry in comparison to other sub-sectors. Therefore, evaluation of Decision Making Units (DMUs) of poultry in provinces and hence improving them is important task to the whole agriculture. Besides, there exist several proposed approaches to resolve this problem. However, a different methodology is proposed due to its powerful discriminatory performance, in this research. For this purpose, combination of Data Envelop Analysis (DEA) and requisite data mining techniques same as Artificial Neural Network (ANN) and Decision Tree (DT) are employed in order to enhance the power of predicting the DMUs evaluation performance because of their well-known efficiency and thereby to present precise decision rules for improving their efficiency. To illustrate the proposed model, all poultry companies in Iran were taken into account. However, in this case there is a small dataset and because the large dataset is necessary to collect data as well as to apply data mining methodology, so, researchers employed k-fold cross validation method to validate the model. Consequently, applied model is supposed to predict efficiency of DMUs and thereby to present decision rules in order to improve the efficiency precisely and accurately according to used optimizing techniques.

Key words: Data envelopment analysis, decision tree, artificial neural network, poultry meat farming, efficiency

INTRODUCTION

Recently, agriculture besides petroleum industries has had considerable effect on economic growth and its stability in Iran. Among various sectors in agriculture, poultry meat farming sub-sector has high potential for enhancing the industry because of having by-product such as egg and manure. Furthermore, consumption of poultry meat is higher in comparison to other meat such as birds, lamb, beef, etc. in Iran. Regarding this sub-sector, newest strategy is to increase productivity of related companies not only in order to reduce costs but also to enhance the product. For this purpose, the key processes and hence their decision making units must be evaluated. Data Envelopment Analysis (DEA) is a broadly used linear programming technique that was developed in operations research and economic literature as a method to determine the relative efficiencies of the Decision Making Units (DMU) (banks, restaurants, public houses, hospitals, schools and corporate performance), this method originally proposed by Charnes *et al.* (1985). Example of these applications were Samoilenko and Osei-Bryson (2008), Samoilenko *et al.* (2010), Cooper *et al.* (2007), Celebi and Bayraktar (2008), Emrouznejad and Anouze (2010) and Bojnec and Latruffe (2007). However, in many situations, it is important to assessing relative efficiency before for this purpose we must classify DMUs in order to evaluate and identify decision rules and then to select the best classifier. Therefore, it is important to develop the process of classifier selection and identification rules. One of the applications of Data Mining (DM) is to use historical data for training a model and to make prediction of new classifier performance with the trained model. Examples of application cases of DM approaches include neural networks and expert systems (Wu, 2009; Yaghoobi *et al.*, 2010). This study aims to improve efficiency of DMUs in chicken industries. The other main aims of this research are included three important objects. In the first place, results yield a favorable classification and selection the best decision rules accuracy rate. Secondly, the results of this study provide insight for selection appropriate classification method for any dataset with many patterns in hybrid structure. Finally, improving decision making performance based on aforementioned results.

Literature

Cluster Analysis (CA): Clustering is a popular data mining technique which involves the partitioning of a set

Corresponding Author: Iman Rahimi, Department of Mathematics, University of Sistan and Baluchestan, Zahedan, Iran

of objects into a useful set of mutually exclusive clusters such that the similarity between the observations within each cluster (i.e., subset) is high whereas the similarity between the observations from the different clusters is low (Samoilenko *et al*., 2010; Samoilenko and Osei-Bryson, 2008). Unlike decision trees which assign a class to an instance (supervised method), clustering procedures are applied when instances are divided into natural groups or clusters (unsupervised method). There are different ways to produce these clusters. The groups may be exclusive, i.e., any instance belongs to only one group probabilistic or fuzzy, i.e., an instance belongs to each group to a certain probability or degree (membership value) hierarchical, i.e., there is a crude division of instances into groups at the top level and each of these groups are refined further up to individual instances (Thomassey and Fiordaliso, 2006). In other literature, overview of two general approaches to clustering was provided: hierarchical clustering, partitional clustering (k-means and k-median) (Samoilenko and Osei-Bryson, 2008). Examples of application of clustering is given by Banfield and Raftery (1992), Ben-Dor and Yakhini (1999), Dhillon (2001), Fisher (1997), Hirschberg and Lye (2001), Johnson (1967), Lai *et al.* (2009), Okazaki (2006) and Wallace *et al.* (2004).

Artificial Neural Networks (ANNs): Another data mining technique is neural networks that are mathematical representations inspired by the functioning of the human brain. Many types of neural networks have been suggested in the literature for both supervised and unsupervised learning (Baesens *et al*., 2003). Neural Network (NN) modeling aims to develop a black box model (an Artificial Neural Network) of the unknown complex relationships in the data. This data mining method is particularly appropriate when there is no known mathematical formula that relates the input and output variables and prediction is more important than explanation. It has been applied extensively by researchers (Samoilenko *et al*., 2010). Neural networks have enormous capabilities for compliance, flexibility and also they are universal nonlinear functional estimators. One of the most effective and flourishing applications of neural networks in data analysis is the MLP Model. MLP models are non-linear neural network models that can be utilized to approximate roughly any function with a high degree of accuracy. The MLP models are a feed forward neural network employed to capture the conjunction between the inputs and outputs. MLP includes a hidden layer of neurons that uses non-linear activation functions such as logistic function (Vahdani *et al*., 2012). On the other word, an MLP is typically composed of an input layer, one or more hidden layers and an output layer, each consisting of several neurons. Because the focus is on classification, researchers will discuss the Multilayer Perceptron (MLP) neural network in more detail. One of the key characteristics of MLPs is that all neurons and layers are arranged in a feed forward manner and no feedback connections are allowed (Baesens *et al*., 2003). Many researchers have used Neural Networks (Celebi and Bayraktar, 2008; Choi and Yoo, 2001; Deng and Pei, 2009; D'heygere *et al*., 2006; Emrouznejad and Anouze, 2010; Golmohammadi *et al*., 2009; Keskin *et al*., 2010; Deng *et al*., 2008; Pao and Sobajic, 1991; Roshdy and Carla, 2004; Wu, 2009; Wu *et al*., 2010; Yaghoobi *et al*., 2010; Khashei *et al*., 2008).

Decision Tree (DT): Another technique of Data mining is decision tree that its application is for classification patterns or piecewise constant functions (Emrouznejad and Anouze, 2010). Application of decision tree is for decision problems that is a tree structure representation such that each non-leaf node is associated with one of the decision variables, each branch from a non-leaf node is associated with a subset of values of the corresponding decision variable and each leaf node is associated with a value of the target (or dependent) variable. There are two main types of DTs: classification trees and regression trees. For a classification tree, the target variable takes its values from a discrete domain (e.g., efficient or inefficient) and for each leaf node the DT associates a probability (and in some cases a value) for each class (i.e., value of the target variable). The class that is assigned to a given leaf node of the classification tree results from a form of majority voting in which the winning class is the one that provides the largest class probability even if that probability is <50%. For a regression tree, the target variable takes its values from an interval domain (e.g., Relative Efficiency Score is between 0 and 1). In this study, researchers will focus on the classification tree which is the most commonly used type of DT (Samoilenko and Osei-Bryson, 2008). Examples of these applications can be given by D'heygere *et al.* (2006), Olaru and Wehenkel (2003) and Sindhu *et al.* (2012).

Data Envelopment Analysis (DEA): Data Envelopment Analysis (DEA) is a widely non-parametric and powerful data analytic tool which is commonly applied in the research and practitioner communities to determine the relative efficiencies of the Decision-Making Units (DMU) (Samoilenko *et al*., 2010; Samoilenko and Osei-Bryson, 2008). Any entity that receives a set of inputs and produces a set of outputs could be designated as a DMU thus any group of such entities could be subjected to

DEA. Consequently, this method has been applied to evaluate productivity and performance of DMU (Emrouznejad and Anouze, 2010; Samoilenko *et al.*, 2010). One of the fundamental assumptions of DEA is that all DMUs in the sample are functionally similar in the sense that all DMUs receive the same number and the same type of inputs and outputs (Samoilenko *et al.*, 2010). In the next step, some corresponding literature to issue is reviewed and contribution of this research is discussed in body of knowledge. Also in another literature, DEA is called a non-parametric analytic tool for measuring the efficiency of Decision Making Units (DMUs). One of the fundamental assumption of DEA is that all DMUs in the sample are functionally similar in the sense that all DMUs receive the same number and the same type of inputs and outputs. The proposed method by charnes is as follow.

Let there are n DMUs and assume that DMUj consumes m inputs xj = (x1j, ..., xmj) to produce s outputs yj = (y1j,..., ypj) for any j: xj>0, xj>0, j = 1,..., n. DEA formulation used in this study for measuring the efficiency of the DMUo often referred to as the DEA CCR Model.

Model 1
Input Oriented-CRS Model:

$$\max_{\mu_k, v_i} \sum_{k=1}^{p} \mu_k y_{ko}$$

$$st \sum_{i=1}^{m} v_i x_{io} = 1$$

$$\sum_{k=1}^{p} \mu_k y_{kj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0, \; j = 1,\ldots,n,$$

$$\mu_k \geq \varepsilon, v_i \geq \varepsilon, \; k = 1,\ldots,p; \; i = 1,\ldots,m$$

DEA efficiency score assigned to given DMU is the optimal value of the objective function. For the given efficiency score equal 1. The given DMU fulfills the necessary condition to be efficient otherwise, it is inefficient. The examples of the application of this method is given by Beasley (1990), Charnes *et al.* (1985), Demirbag *et al.* (2010), Grosskopf and Moutray (2001), Kirigia *et al.* (2001), Santos and Themido (2001), Mukherjee *et al.* (2001), Schaffnit *et al.* (1997), Khouja (1995) and Pare and Sicotte (2001).

In the first place, researches related to main idea were studied and hence it was reviewed how the researchers developed their purposes, methodologies and results were considered for resolving this problem. Then, researchers reviewed some surveys which applied methodology like proposed techniques in this research even though these have been employed to resolve another scopes except of chicken industries DMUs improvement problem.

Tsai *et al.* (2009) and Tsai and Chiou (2009) conducted the methodology using combination neural network and decision tree in order to decrease the financial crisis risks derived from earnings management and help the investors avoid suffering a great loss in the stock market. For this purpose, the researcher developed a neural network model to predict the level of earnings management. In addition, the cases which were correctly predicted by the neural network model were applied to construct a decision tree model to generate useful decision rules. As a result it was demonstrated that integrating the neural network and decision tree models provide both higher rate of prediction accuracy and important decision rules. Also, two important rules are identified to allow investors and creditors for effective earnings management forecasting. Thomassey and Fiordaliso (2006) proposed a forecasting system, based on clustering and decision tree which performs mid-term forecasting. Performances of proposed models were measured using real data from an important French textile distributor. In this survey, it was pointed out that existing forecasting models are generally unusable due to specific constraints of the textile sales; thereby these techniques were demonstrated very effective on prediction. Emrouznejad and Anouze (2010) developed a framework to combine DEA with C and R decision tree in order to evaluate the efficiency of DMUs. In addition to results of forecasting, a set of rules which can be used by policy makers to discover reasons behind efficient and inefficient DMUs were presented. For illustration of integrated model the banking sector in the Gulf Cooperation Council countries were taken. In one survey by Samoilenko *et al.* (2010) a five-step methodology was proposed to augment DEA with Cluster Analysis (CA) and Neural Networks (NN) and hence associated model allows an investigator to figure out whether the difference in the scores of scale heterogeneous DMUs is because of the heterogeneity of the levels of inputs and outputs or whether it is due to their efficiency of conversion of inputs into outputs. One research by Samoilenko and Osei-Bryson (2008) was developed so that a three-step methodology was presented to allow researcher for increasing the discriminatory power of DEA in the presence of the heterogeneity of the sample. In the first place, the Cluster Analysis (CA) was applied in order to test for the presence of the naturally occurring subsets in the sample. Then, DEA was employed to calculate both the relative efficiencies of the DMUs and averaged relative efficiencies of each subset identified in the previous step. Finally, Decision Tree (DT) induction was taken to inquire into the subset-specific nature of the relative efficiencies of the DMUs in the sample.

To conclude, based on experience neural networks have better performance in classification model and worse performance than decision trees in the prediction model. Decision makers should be careful in identifying their objectives before choosing the appropriate models; therefore researchers need a global model that performs better in the classification and prediction model than NNs and DTs. The purpose of combining of several models is improving classification and forecasting accuracy. Based on empirical findings combining several methods can be an effective and efficient way to improve forecasting.

MATERIALS AND METHODS

As it can be seen, many methods have been proposed and used for classification and determining the best decision rules for efficient DMUs (Keskin *et al*., 2010). As it was indicated, both selection and evaluation of DMUs carried out in major of researches because the selection is based on evaluation results thus it is necessary to employ the methodologies which have ability to assess the performance of DMUs in the first place and then efficient DMUs will be forecasted. This problem requires consideration of a variety of attributes. Various studies have been performed for effective evaluation and selection of DMUs by using several approaches such as linear weighting methods, mathematical programming models, statistical methods (Celebi and Bayraktar, 2008). All models consider the most factors of DMUs such as cost, human force, number of chicken, type of company, volume of products, value of input/service, delivery time, size of contract farmers, price, location however, quality can be seen as an essential factor for DMUs evaluation among various criteria (Begum *et al*., 2012) to analyze the weights of them in order to evaluate the DMU. However, MCDM Methods are able to rank the DMUs, nevertheless these can not be applied as forecasting tools. Exploring among historical data can aid us to discover hidden relationship and hence to perform new approach for predicting future results. So, forecasting tools are efficient and flexible for working out the DMUs selection. Therefore among these tools, data mining methods such as ANN, DT are suitable because of their well-known accuracy rate.

In order to conduct the research, four-step methodology is proposed:

- Employing DEA technique to figure out the efficiency or inefficiency of DMUs
- Clustering the DMUs according to important criteria (output-input data for DEA)
- Labeling clusters in order to apply ANNs along with training and testing process and using k-fold validation along ANN's process
- Applying DT analysis on data classified by ANNs in order to enhance discrimination of data and hence best prediction and best decision rules to determine efficient DMUs

Before applying data mining we must evaluate the DMUs and so among the techniques, DEA is a non-parametric method for measuring the efficiency and productivity of DMUs according to scores of efficiency and inefficiency of them. Functional similarity of the DMUs in a sample is one of the fundamental assumptions of DEA. In other word, these DMUs must be similar in terms of utilization of the inputs and production of the outputs. Thus, DEA requires us to make sure that we compare apples and apples and not apples and oranges (Samoilenko and Osei-Bryson, 2008; Samoilenko *et al*., 2010). On the other hand data mining techniques allow DMUs to discover meaningful and previously hidden information from large databases (Emrouznejad and Anouze, 2010). Besides, in order to discover new relations and rules, researchers have to consider DMUs productivity as outputs and effective factors on their efficiency as inputs. Therefore, it is fundamentally to classify the DMUs inside clusters with more similarity due to DEA characteristics.

For the purpose of using homogeneity in DEA, researchers employed cluster analysis approach in accordance with aforementioned effective factors on DMU efficiency as well as DEA scores of them. In this study, researchers focus on basic method to cluster (k-means algorithm) data. This tool is effective not only to find and describe patterns in data in order to make prediction but also to build an explicit representation of the knowledge. The k-means method is simplistic but reasonably effective to carry out the training for the decision trees. It allows dividing instances into disjoint clusters from numeric attributes (Thomassey and Fiordaliso, 2006). Thereafter, clusters must be labeled based of both their efficiency (0: inefficient DMU; 1: efficient DMU) and the number of cluster figured out by cluster analysis techniques. The most reasons for applying clustering are to find a set of natural groups (i.e., segmentation) and the corresponding description of each group, firstly. Then, second cause is to improve the performance of other predictive modeling and DM techniques when there are many competing patterns in the data. The model is shown in Fig. 1.

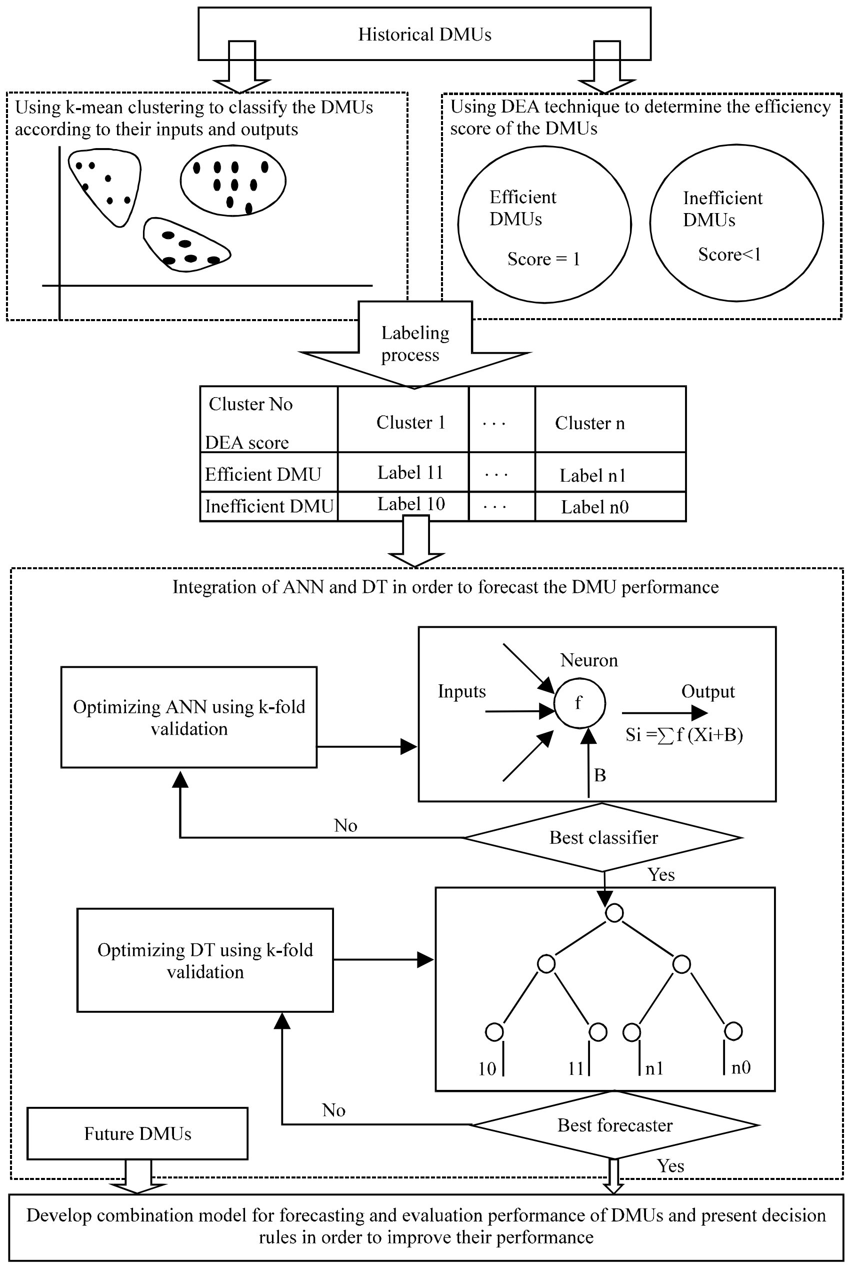

Fig. 1: Theoretical (conceptual) framework

In the second part, the tools such as ANNs were useful to predict the patterns made up previous step; the ANN is assumed as classification tool in order to forecast the patterns and thereby efficiency situation of DMU. The several researches demonstrated the ANNs as classification tool more stable and suitable than regression tool (Emrouznejad and Anouze, 2010). According to structure of mentioned MLP, in the hidden layer as there is no method to decide the optimal number of hidden nodes directly, researchers choose 2 different numbers of hidden nodes and then the experimental design based on 2 settings is applied in order to optimize the accuracy of prediction in ANNs. As it was indicated, to resolve problems such as small data, k-fold validation is taken into account as an optimizing method for more accurate prediction by ANNs.

In the third step, the purpose is to present decision rules for being efficient DMUs according to historical data trained and tested such as effective factors and efficiency in ANN tools. It is easy to understand and map well to a set of rules for decision trees. In addition, decision trees have the ability to build models with datasets including numerical and categorical data. The process of building decision trees starts with the best classifier selection previous phase (ANN phase) going through the decision trees training phase and ends with the phase of testing the decision trees. In the beginning of the best classifier selection phase, the testing data in ANN are used to test the MLP classifiers generated from ANN. After testing the classifiers, the classifier which provides the highest accuracy rate as the best classifier is chosen.

RESULTS AND DISCUSSION

Data: The numerical example was considered and it is shown in Table 1 taken from poultry industries of Iran. The data was derived from information and statistical center of Iran. The data presents five input variables which represent capabilities of each DMU to generate three output factors that represent the performance outcomes of DMUs in the measurement process. In this study, province was considered as DMU and technical efficiency has been used for DEA Method. The inputs and outputs are determined as follows:

Inputs:

- Chicken type 1: it infers the genetic characteristics which are Lyne chickens (an especial type of chicken in Iran)
- Chicken type 2: it infers other chickens except of Lyne chickens
- Employee number in each DMU
- Feed: Chicken food consumed for poultry industries in each DMU
- Full auto-system: number of full automate system poultry house in each DMU

Outputs:

- Produced egg
- Produced chicken
- Produced manure

Implementation: It must be pointed out that researchers applied software GAMS, MATLAB and MINITAB in order to analysis of the example. As it was said first, DEA was used to determine technical efficient and technical inefficient DMUs based on inputs and outputs by multiplier CCR Model. Then researchers clustered each DMU by k-means clustering according to outputs and inputs in order to increase discriminatory power of DEA. For this purpose, researchers combined results of DEA and clustering to make the classification groups so as to achieve accurate prediction for efficiency of DMUs. As it is shown, Table 2 presents results of DEA, Table 3 presents results of clustering and Table 4 shows results of integrated DEA and clustering. Consequently, 4 patterns are made according to combination so that we

Table 1: Data for 26 DMUs

Province (DMU)	Chicken type 1	Chicken type 2	Employee number	Feed	Full auto system	Egg	Produced chicken	Manure
1	250000	3099305	797	169633	9	78677	1071	61653
2	8000	275600	71	7389	4	6320	3	2277
3	0	79600	22	3778	5	2069	3	882
4	1697562	1811690	974	158136	33	75666	1818	46225
5	1146200	1610210	428	55764	14	30194	1423	24421
6	1846727	5110764	1488	244600	24	118041	3216	98881
7	24000	48800	13	2800	2	1656	0	950
8	17500	173305	88	10644	1	82954	99	2949
9	1087763	3456632	894	186468	13	4404	1796	62220
10	163842	25300	152	3792	1	2061	196	1251
11	108500	300000	130	23893	4	11320	196	3988
12	0	420580	89	17645	1	8193	54	4314
13	858217	890362	783	60290	7	32257	161	18472
14	1963212	1308230	950	134203	14	63101	1634	41487
15	95900	3281050	712	163324	22	79187	436	49767
16	0	101000	28	4153	0	2027	2	966
17	150000	148200	49	10059	0	4211	27	2445
18	29544	285221	75	10087	7	5838	433	4713
19	0	0	7	360	1	154	0	180
20	471615	622630	317	40488	4	20038	550	12114
21	19400	0	28	2629	1	963	1	1040
22	0	306406	104	10172	2	5171	226	6822
23	414360	910680	423	60677	13	29013	42	18699
24	167300	415761	250	26011	5	13046	429	7237
25	103900	517200	74	15299	2	8194	60	2640
26	0	241500	62	3812	2	1776	61	1705

Table 2: Results of DEA

Inefficient DMUs	Efficient DMUs
2 (52%)-4 (96%)-7 (97%)-10 (59%)-12 (82%)-19 (98%)-22 (79%)-23 (82%)-24 (50%)-26 (90%)-	6-8-9-11-13-14-15-16-17-18-20-21-1-3-5--25

Table 3: Results of clustering

Cluster 1 (high production)	Cluster 2 (mid production)
1-4-5-6-10-15-16	2-3-7-8-9-11-12-13-14-17-18-19-20-21-22-23-24-25-26

Table 4: Results of combined DEA and clustering

	Clusters	
Efficient and inefficient	Cluster 1	Cluster 2
Efficient DMUs	3-7-8-9-12-14-17-18-19	1-5-6-10-15-16
Inefficient DMUs	2-11-13-20-23-24-25-26	4

Table 5: Training set and ANN structure

Net (FF-Bpp)	Activation functions	Parameter
Number of hidden layers	-	2
Number of output layer	-	1
Neuron Layer 1	Log-sig	20
Neuron Layer 2	Tan-sig	20
Train function	Trainlm	-
Performance function	-	MSE

give lable 1 to efficient DMU, lable 0 to inefficient DMU, lable 1 to cluster1 and lable 2 to cluster 2, so pattern 1 is efficient DMU in cluster 1 (lable 11), pattern 2 is inefficient DMU in cluster 1 (lable 10), pattern 3 is efficient DMU in cluster 2 (lable 21), pattern 4 is inefficient DMU in cluster 2 (lable 20).

The DM techniques such as ANN and DT were applied to all data including inputs as well as outputs. The outcome to DM techniques is groups (patterns values) in classification models calculated by DEA and clustering. Table 5 shows the structure of ANN and all used functions.

DM techniques require large dataset for effective application however, the small dataset was provided in this study. Therefore data partitioning method such as cross validation tool was employed in order to validate the model. Researchers took k-fold cross-validation technique to validate the utilization of small dataset in the case.

To do the k-fold cross-validation, researchers divide the data set into k subsets and repeat run DM models k times where k is an integer and divisible by the number of records in the dataset. Each time, one of the k subsets is used as the test set and the other k-1 subsets are merged as a training set. The advantage of this method is that the limited number of data point can be used in both training set and the validation set as many times as possible.

Table 6: Results of performance evaluation on ANN training

Parameters	V1	V2	V3	V4	V5
Epoch	26	29	40	21	42
Grad	0.217	0.321	0.016	0.0306	0.107
MSE	0.0663	0.0801	0.0522	0.0881	0.0558

In the example, researchers divide the whole data set of 26 cases into five segments, each containing five DMUs in four subsets and the fifth one containing six DMUs. Then, DM is trained with the first four segments of 20 cases and used to predict the fifth segment of another six DMUs. This process repeats five times to accomplish five-fold cross-validation. Also, this process was repeated on decision trees training and testing. Table 6 shows results of performance evaluation in training phase in each fold. Table 7 shows results of training and testing error report in ANN prediction. It must be demonstrated that a prediction error of 15% was regarded as acceptable for a service process or production process and from 20-30% it was regarded as reject (Cook and Chiu, 1997; Malinov *et al*., 2001). As shown in Table 8, due to this reason that total prediction errors calculated by ANN (training and testing subsets) is <15%, it must be stated that this model can be regarded as valid.

In the decision trees training phase, researchers used the samples with the correct-classified output labels based on the best MLP classifier and the samples with the incorrect-classified output labels as the observations in building decision trees. This aim is to apply the MLPs outputs to select more discriminative data for training decision trees which could provide better prediction performance. The observations in building decision trees are derived from the neural network models' training phase and are collected to be the training sample subset to build decision trees (Tsai *et al*., 2009). Also, in the decision trees testing phase, the testing sample subset is based on the best MLP classifier as well but originated from the neural network models' testing phase. Further, the best rules in decision trees are identified as well for DMU efficiency prediction.

Figure 2 shows one of the classification trees in k-fold training that was best classification tree. And also, best decision rules are presented. As it can be seen, decision trees identify crucial variables from original dataset and hence generate meaningful rules. In this study, best decision tree figured out full auto-system, manure and chicken food as key attributes in order to make best prediction and thereby decision rules.

Table 9 shows the results of error prediction in testing decision tree. As it is presented, predicted outcome of ANN were applied as outcome of DT so that classes were divided to Correct Prediction (CP) and Incorrect Prediction (IP) in order to make accurate prediction in DT results. As it cab be seen,

Table 7: Error report of training and testing classification on ANN

Training results					Testing results				
Fold	Class	DMU number	Error#	Error (%)	Fold	Class	DMU number	Error#	Error (%)
V1	1	7	2	28.6	V1	1	2	0	0
	2	8	0	0.0		2	2	0	0
	3	5	0	0.0		3	1	0	0
	4	1	1	100.0		4	Nan	Nan	Nan
	Overall	21	3	14.3		Overall	5	0	0
V2	1	7	2	28.6	V2	1	2	2	100
	2	8	0	0.0		2	2	0	0
	3	5	0	0.0		3	1	0	0
	4	1	0	0.0		4	Nan	Nan	Nan
	Overall	21	2	9.5		Overall	5	2	40
V3	1	7	2	28.6	V3	1	2	0	0
	2	8	1	12.5		2	2	0	0
	3	5	0	0.0		3	1	0	0
	4	1	0	0.0		4	Nan	Nan	Nan
	Overall	21	3	14.3		Overall	5	0	0
V4	1	7	3	42.9	V4	1	2	1	50
	2	8	0	0.0		2	2	1	50
	3	5	0	0.0		3	1	0	0
	4	1	0	0.0		4	Nan	Nan	Nan
	Overall	21	3	14.3		Overall	5	2	40
V5	1	7	1	14.3	V5	1	2	1	50
	2	8	0	0.0		2	2	1	50
	3	5	0	0.0		3	1	0	0
	4	1	0	0.0		4	Nan	Nan	Nan
	Overall	21	1	4.8		Overall	5	2	40

Table 8: Total error prediction of training and testing classification on ANN

Percentage	V1	V2	V3	V4	V5	Total
Error	11.5	15.4	11.5	19.2	11.5	13.82
Accuracy	88.5	84.6	88.5	80.8	88.5	86.18

best classification tree is achieved from first fold because the error rate is 0% and it can be best predictor for future data. Therefore the decision rules are according the best tree which are extracted in order to have efficient DMUs.

According to Fig. 2, there are two useful rules to predict efficiency and inefficiency of DMUs of poultry industries:

- If produced manures (C) are ≥1 21560 kg and full auto-systems (A) are <28 units in DMU as a result, provinces (DMU) are efficient that have high production (categorized in Cluster 1)
- If produced manures (C) are <21560 kg and full auto-systems (A) are <2 units or >5 units in DMU as a result, provinces (DMU) are efficient that have mid production (categorized in Cluster 2), otherwise (with full auto-systems between 2 and 5) are inefficient. On the other hand, inefficient provinces with mid production will be converted to efficient if they consume chicken food (B) >3306 kg

Analysis of variance for prediction error: ANOVA experiments were applied in order to compare the ANN model and Combination Model (ANN and DT), Table 10 shows result of ANOVA, according p-value (0.012) <0.05, it is demonstrated that there is meaningful difference

Table 9: Error report of testing classification on DT

Folds	Class	DMU number	Error#	Error (%)
V1	11CP	2	0	0
	11IP	Nan	Nan	Nan
	10CP	2	0	0
	10IP	Nan	Nan	Nan
	21	1	0	0
	20CP	Nan	Nan	Nan
	20IP	Nan	Nan	Nan
	Overall	5	0	0
V2	11CP	2	0	0
	11IP	Nan	Nan	Nan
	10CP	Nan	Nan	Nan
	10IP	2	1	50
	21	1	0	0
	20CP	Nan	Nan	Nan
	20IP	Nan	Nan	Nan
	Overall	5	1	20
V3	11CP	2	0	0
	11IP	Nan	Nan	Nan
	10CP	2	1	50
	10IP	Nan	Nan	Nan
	21	1	0	0
	20CP	Nan	Nan	Nan
	20IP	Nan	Nan	Nan
	Overall	5	1	20
V4	11CP	2	1	50
	11IP	Nan	Nan	Nan
	10CP	2	0	0
	10IP	Nan	Nan	Nan
	21	1	0	0
	20CP	Nan	Nan	Nan
	20IP	Nan	Nan	Nan
	Overall	5	1	20
V5	11CP	2	1	50
	11IP	Nan	Nan	Nan
	10CP	2	0	0
	10IP	Nan	Nan	Nan
	21	1	0	0
	20CP	Nan	Nan	Nan
	20IP	Nan	Nan	Nan
	Overall	5	1	20

Table 10: ANOVA (Analysis of Variance)

Source	DF	SS	MS	F-value	p-value (sig.)
Model	1	3036	3036	6.73	0.012
Error	68	30665	451	-	-
Total	69	33701	-	-	-

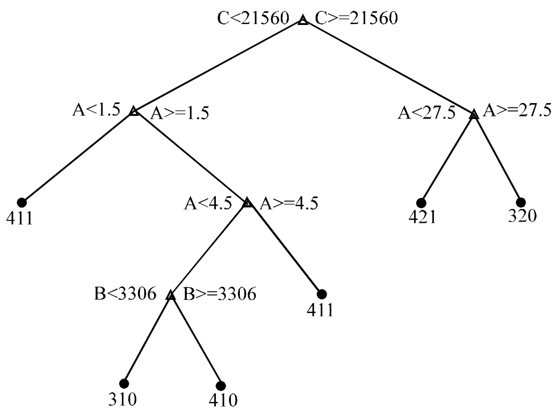

Fig. 2: Best classification tree k-fold training based on k-fold training

between ANN Model and ANNDT Model. So, results show accuracy in Combination Model has been increased. In other word, error percent in ANNDT is reduced from 24-16% in testing process.

Therefore, the Combination Model of ANN and DT presents better performance in forecasting the efficiency of DMUs in comparison to ANN in this real world case.

CONCLUSION

In this study, researchers studied on poultry meat farming industries as a sub-sector of agriculture because of its effects on economic of Iran. Several researches were conducted on this issue in order to measure the efficiency of poultry units and thereby improve their efficiency and productivity. In this study, the methodology was employed in order to evaluate efficiency as well as to enhance this according to forecasting by data mining techniques such as Decision Trees (DT) and Artificial Neural Networks (ANN). The proposed model was constituted of combination ANN and DT because the results shows performance of this model is better than used ANN according to results of ANOVA experiments. In the first place, technical efficiency of poultry units (DMUs) was calculated based on data envelop analysis (DEA) also, each DMU was clustered by k-means clustering analysis according to their attributes (outputs and inputs) in order to increase discriminatory power of DEA. Then, 4 patterns were made by combination results of DEA and clustering for forecasting process. Consequently, researchers designed the intelligent system (ANNDT) to forecast the efficiency of DMUs and then to achieve the best rules in order to improve their efficiency. According to this proposed model, accuracy of forecasting and presenting the rules in order to improve the efficiency was 84%. For implementing this model, poultry meat farming in provinces of Iran were considered as DMUs. As we know, however, one of the limitation about the research is small data set for next researches, fuzzy structure data mining has been propose such as fuzzy neural networks and fuzzy decision tree, also using combination other data mining techniques such as tabu search and genetic algorithm with applied model in this study in order to enhance the accuracy of forecasting would be proposed.

APPENDIX 1

Decision rules from best decision tree according to 5-fold cross validation.

Decision tree for classification:
1 if C<21560 then node 2 else if C>=21560 then node 3 else 411
2 if A<1.5 then node 4 else if A>=1.5 then node 5 else 411
3 if A<27.5 then node 6 else if A>=27.5 then node 7 else 421
4 class = 411
5 if A<4.5 then node 8 else if A>=4.5 then node 9 else 410
6 class = 421
7 class = 320
8 if B<3306 then node 10 else if B>=3306 then node 11 else 410
9 class = 411
10 class = 310
11 class = 410

REFERENCES

Baesens, B., R. Setiono, C. Mues and J. Vanthienen, 2003. Using neural network rule extraction and decision tables for credit-risk evaluation. Management Sci., 49: 312-329.

Banfield, J. and A. Raftery, 1992. Identifying ice floes in satellite image. Naval Res. Rev., 43: 2-18.

Beasley, J., 1990. Comparing university departments. Omega, 8: 171-183.

Begum, I.A., M.J. Alam, J. Buysse, A. Farija and G.V. Huylenbroeck, 2012. Contract farmers and poultry farm efficiency in Bangladesh: A data envelopment analysis. Appl. Eco., 44: 3737-3747.

Ben-Dor, A. and Z. Yakhini, 1999. Clustering gene expression patterns. Proceedings of the 3rd Annual International Conference on Computational Molecular Biology, 1999, Lyon, France, pp: 11-14.

Bojnec, S. and L. Latruffe, 2007. Measures of farm business efficiency. Ind. Management Data Syst. Voldataz, 108: 258-270.

Celebi, D. and D. Bayraktar, 2008. An integrated neural network and data envelopment analysis for supplier evaluation under incomplete information. Expert Syst. Appl., 35: 1698-1710.

Charnes, A., C.T. Clark, W.W. Cooper and B. Golany, 1985. A developmental study of data envelopment analysis in measuring the efficiency of maintence units in the US Air forces. Ann. Oper. Res., 2: 95-112.

Choi, Y.S. and S.I. Yoo, 2001. Text database discovery on the web: Neural net based approach. J. intelligent inf. Syst., 16: 5-20.

Cook, D.F. and C.C. Chiu, 1997. Predicting the internal bond strength of particleboard: Utilizing a radial basis function neural network. Eng. Appl. Artif. Intell., 10: 171-177.

Cooper., W.W., L.M. Seiford and K. Tone., 2007. Data Envelopment Analysis, A Comprehensive text With Models. Springer, New York, USA.

D'heygere, T., P.L.M. Goethals and N.D. Pauw, 2006. Genetic algorithms for optimization of predictive ecosystems models based on decision trees and neural networks. Ecol. Modell., 195: 20-29.

Demirbag, M., E. Tatoglu, K.W. Glaister and S. Zaim, 2010. Measuring strategic decision making efficiency in different country contexts: A comparison of british and Turkish firms. Omega, 38: 95-104.

Deng, W.J. and W. Pei, 2009. Fuzzy neural based importance-performance analysis for determining critical serice attributes. Expert Syst. Application, 36: 3774-3784.

Deng, W.J., W.C. Chen and W. Pei, 2008. Back-propagation neural network based importance-performance analysis for determining critical service attributes. Expert Syst. Appl., 34: 1115-1125.

Dhillon, I.S., 2001. Co-clustering documents and words using bipartite spectral graph partitioning. Proceedings of the 7th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, August 29-29, 2001, ACM, New York, 269-274.

Emrouznejad, A. and A.L. Anouze, 2010. Data envelopment analysis with classification and regression tree-a case of banking efficiency. Expert Syst. J. Knowledge Eng., 27: 231-246.

Fisher, M.R., 1997. Segmentation of angler population by catch preference, participation and experience: A management-oriented application of recreation specialization. North Am. J. Fisheries Management, 17: 1-10.

Golmohammadi, D., R.C. Creese and H. Valian, 2009. Neural network application for supplier selection. Int. J. Product Dev., 8: 252-275.

Grosskopf, S. and C. Moutray, 2001. Evaluating performance in Chicago public high schools in the wake of decentralization. Eco. Edu. Rev., 20: 1-14.

Hirschberg, J.G. and J.N. Lye, 2001. Clustering in a data envelopment analysis using bootstrapped efficiency scores. Department of Economics, Working Papers Series 800.

Johnson, S.C., 1967. Hierarchical clustering schemes. Psychometrika, 32: 241-254.

Keskin, G.A., S. Ilhan and C. Ozkan, 2010. The Fuzzy ART algorithm: A categorization method for supplier evaluation and selection. Expert Syst. Appli., 37: 1235-1240.

Khashei, M., S.R. Hejazi and M. Bijari, 2008. A new hybrid Artificial neural networks and fuzzy regression model for time series forecasting. Fuzzy Sets Syst., 159: 760-786.

Khouja, M., 1995. The use of data envelopment analysis for technology selection. Computers Ind. Eng., 28: 123-132.

Kirigia, J., L. Sambo and H. Scheel, 2001. Technical efficiency of public clinics in kwazulu-natal province of south Africa. East Afr. Med. J., 78: 1-13.

Lai, R.K., C.Y. Fan, W.H. huang and P.C. Chang, 2009. Evolving and clustering fuzzy decision tree for financial time series data forecasting. Expert Syst. Appli., 36: 3761-3773.

Malinov, S., W. Sha and J.J. Mckeown, 2001. Modelling the correlation between processing parameters and properties in titanium alloys using artificial neural network. Comput. Mater. Sci., 21: 375-394.

Mukherjee, K., S. Ray and S. Miller, 2001. Productivity growth in large US commerical banks: The initial post-deregulation experience. J. Banking Finance, 25: 913-939.

Okazaki, S., 2006. What do we know about mobile internet adopters? A cluster analysis. Inf. Management, 43: 127-141.

Olaru, C. and L. Wehenkel, 2003. A complete fuzzy decision tree technique. Fuzzy Sets Syst., 138: 221-254.

Pao, Y.H. and D.J. Sobajic, 1991. Neural networks and knowledge engineering. IEEE Trans. Knowledge Data Eng., 3: 185-192.

Pare, G. and C. Sicotte, 2001. Information technology sophistication in health care: An instrument validation study among Canadian hospitals. Int. J. Med. Inform., 63: 205-223.

Roshdy, S.Y. and N.P. Carla, 2004. Combining genetic algorithm and neural networks to build a signal pattern classifier. Neurocomputing, 61: 39-56.

Samoilenko, S. and K.M. Osei-Bryson, 2008. Increasing the discriminatory power of DEA in the presence of the sample heterogeneity with cluster analysis and decision trees. Expert Syst. Appli., 34: 1568-1581.

Samoilenko, S., K. Muata and O. Bryson, 2010. Determining sources of relative inefficiency in hetergenous samples: Methodology using cluster analysis, DEA and neural networks. Eur. J. Oper. Res., 206: 479-487.

Santos, J. and I. Themido, 2001. An application of recent developments of data envelopment analysis to the evaluation of secondary schools in porugal. Int. J. Services Technol. Manage., 2: 142-160.

Schaffnit, C., D. Rosen and J.C. Paradi, 1997. Best practice analysis of Bank branches: An application of DEA in a large Canadian bank. Eur. J. Oper. Res., 98: 269-289.

Sindhu, S.S.S., S. Geetha and A. Kannan, 2012. Decision tree based light weight intrusion detecteion using a wrapper approach. Expert Syst. Appli., 39: 129-141.

Thomassey, S. and A. Fiordaliso, 2006. A hybrid sales forecasting system based on clustering and decision trees. Decision Support Syst., 42: 408-421.

Tsai, C.F. and Y.J. Chiou, 2009. Earnings management prediction: A pilot study of combining neural networks and decision trees. Expert Syst. Appli., 36: 7183-7191.

Tsai, M.C., S.P. Lin, C.C. Cheng and Y.P. Lin, 2009. The consumer loan default predicting model-An application of DEA-DA and neural network. Expert Syst. Apli., 36: 11682-11690.

Vahdani, B., S.H. Iranmanesh, S.M. Mousavi and M. Abdollahzade, 2012. A locally linear neuro-fuzzy model for supplier selection in cosmetics industry. Appli. Math. Modell. (In Press). 10.1016/j.apm. 2011.12.006.

Wallace, L., M. Keil and A. Rai, 2004. Understanding software project risk: A cluster analysis. Info. Manage., 42: 115-125.

Wu, D., 2009. Supplier selection: A hybrid model using DEA, decision tree and neural network. Expert Syst. Appli., 36: 9105-9112.

Wu, J.J., J.L. Zheng, C.F. Zheng and C.Y. Xi, 2010. The study on neural network-based supplier selection and evaluation. Proceeding of the 3rd International Conference on Information and Computing, June 4-6, 2010, Washington, DC, USA., pp: 31-34.

Yaghoobi, R., M.B. Aryanezhad and F.H. Lotfi, 2010. Application of multi-layer recurrent neural network and fuzzy time series in input/output prediction of DEA models: Real case study of a commerical bank. Proceedings of the 40th IEEE International Conference on Computers and Industrial Engineering(CIE), July 25-28, 2010, Awaji, pp: 1-6.