Physica A 489 (2018) 78–93

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Integrating dynamic fuzzy C-means, data envelopment analysis and artificial neural network to online prediction performance of companies in stock exchange

Mustafa Jahangoshai Rezaee *, Mehrdad Jozmaleki, Mahsa Valipour

Faculty of Industrial Engineering, Urmia University of Technology, Urmia, Iran

HIGHLIGHTS

- Proposing an approach for online prediction performance of units.
- Development of fuzzy c-means to be dynamic behavior.
- Integrating DFCM, DEA and ANN.
- Applying proposed approach in a real case study on stock exchange market.
- Evaluating and predicting companies using financial ratios in six years.

ARTICLE INFO

Article history:
Received 22 February 2017
Received in revised form 1 July 2017
Available online 6 August 2017

Keywords:
Prediction
Dynamic fuzzy C-means
Self-organizing map
Data envelopment analysis
Multi-layer perceptron
Stock exchange market performance

ABSTRACT

One of the main features to invest in stock exchange companies is their financial performance. On the other hand, conventional evaluation methods such as data envelopment analysis are not only a retrospective process, but are also a process, which are incomplete and ineffective approaches to evaluate the companies in the future. To remove this problem, it is required to plan an expert system for evaluating organizations when the online data are received from stock exchange market. This paper deals with an approach for predicting the online financial performance of companies when data are received in different time's intervals. The proposed approach is based on integrating fuzzy C-means (FCM), data envelopment analysis (DEA) and artificial neural network (ANN). The classical FCM method is unable to update the number of clusters and their members when the data are changed or the new data are received. Hence, this method is developed in order to make dynamic features for the number of clusters and clusters members in classical FCM. Then, DEA is used to evaluate DMUs by using financial ratios to provide targets in neural network. Finally, the designed network is trained and prepared for predicting companies' future performance. The data on Tehran Stock Market companies for six consecutive years (2007–2012) are used to show the abilities of the proposed approach.

1. Introduction

Investment is an important factor for development in the current century. In this regard, the most important investment way that can lead to the various industries and economic activities, is Stock Exchange Market. Furthermore, the management

* Corresponding author.
E-mail address: m.jahangoshai@uut.ac.ir (M. Jahangoshai Rezaee).

http://dx.doi.org/10.1016/j.physa.2017.07.017

of transferring the risk and distributing it, information transparency, price discovery, and creation of competitive market are the other basic functions of Stock Exchange. The importance of stock exchange markets in developed countries is very much so that nowadays it is one of the most important indicators of economic growth of these countries. Because of the increasing of variables affecting decision making, nowadays, managers, investors, stakeholders prefer to have a mechanism that can help them in decision-making. For this reason, the new methods are used by which the prediction of the estimations is closer to reality and has the low-level errors. Therefore, if the future trend of the stock market may be predicted more accurately by appropriate methods, investors can maximize returns from the investments. Since the mid-70s, many endeavors for increasing capabilities of stock pieces, returns, and efficiencies using mathematical methods, econometrics, and intelligent systems have been begun and many researches have been done on stock price and index in several developed countries such as Britain, the United States, Canada, Germany, and Japan. To be shown existence or lacks of stock price information in certain structure and violated the assumption of stochastic steps.

Forecasting and predicting methods used so far have been divided into two general categories, including classical and modern methods. Classical methods include econometrics-based approaches, statistical inferences, and traditional mathematical programming. In recent years, many advances in predicting stock prices in different countries using the soft computing algorithms and artificial intelligence emerged. Various algorithms and methods such as neural networks, meta-heuristic algorithms, and fuzzy inference are in this domain that they have achieved successful results in solving complex problems. Schwartz and Whitcomb [1] have done one of the first studies on stock market prediction using neural network. Using IBM daily prices, they showed that the neural networks are able to identify non-linear patterns in time series and the unknown rules on asset price and stock price changes.

Afterwards, using neural networks have been extended in the financial sphere. Chiang et al. [2] used a back propagation network for predicting net asset price of investment companies at the end of the year. Their network results with the results obtained from traditional econometric methods were compared. They found neural networks are significantly more effective than regression-based methods when the number of data is low. Leung et al. [3] used General Regression Neural Network (GRNN) for forecasting UK exchange rate. They compared GRNN with other forecasting methods including multi-layer feed-forward network (MLFN), multivariate transfer function, and random walk models. The results showed that GRNN has a higher degree of forecasting accuracy and performs statistically better results than other mentioned methods.

Also, Chen and Leung [4] proposed a forecasting approach that combines the strengths of neural network and multivariate time series models. In the proposed approach, first, forecasting exchange rate of UK, USA, and Japan was done by time series and then GRNN was used to correct the forecasting errors. Santos et al. [5] examined forecasting accuracy of exchange rate in Brazil using different approaches. For this purpose, they applied the intelligent systems such as multilayer perceptron and radial basis function neural networks and the Takagi–Sugeno fuzzy system versus the traditional methods such as auto regressive moving average (ARMA) and ARMA-generalized auto regressive conditional heteroskedasticity (ARMA-GARCH) linear models. They found the intelligent-based methods are able to provide more accurate results than the traditional methods. Carriero et al. [6] forecast 33 exchange rates with a large Bayesian Vector Auto Regression using multivariate time series models. Other works were done by Chortareas et al. [7] for forecasting exchange rate volatility using high-frequency data, Sermpinis et al. [8] applying radial-basis functions neural network and particle swarm optimization to forecast foreign exchange rates, Sermpinis et al. [8] for forecasting and trading the EUR/USD exchange rate using the Psi Sigma Neural Network (PSI) and the Gene Expression algorithm, Yuan [9] using polynomial smooth support vector machine to forecast the movement direction of exchange rate. Afterwards, Korol [10] developed a forecasting exchange rate model with the use of fuzzy logic. The paper were evaluated the exchange rate at the time of welfare (2005–2007) and during the financial crisis (2009–2011). Sermpinis et al. [11] combined genetic algorithm and support vector regression (GA-SVR) to predict exchange rates. The authors have done their studies on three countries: USA, UK, and Japan since 1999–2012. The results indicate that the proposed method has high statistical accuracy and high performance. Pinčák and Bartoš [12] proposed an approach for analyzing and forecasting financial market and time series data. They showed their ideas were used about the portfolio selection problem of multi-string structure and the stability of the algorithm was demonstrated on transaction costs in long-term period. Kanjamapornkul et al. [13] investigated Thailand's stock market during the 2008 financial crisis. They redefined the behavior matrix for time series data using Pauli matrix and modified Wilson loop. This matrix was used to detect the 2008 financial market crash by using a degree of cohomology group of sphere over tensor field. Finally the analysis of financial tensor network was provided.

Oztekin et al. [14] applied three methods namely artificial neural network, adaptive neuro-fuzzy inference system, and support vector machine in order to predict daily stock return in a new business. They evaluated their method in stock market from 2004 to 2012. Rubio et al. [15] used weighted fuzzy time series (FTS) model to select uncertain parameters in the portfolio, and improved accuracy by applying some changes in the model. This method has also the ability to estimate the investment risk associated to possibilistic moments of trapezoidal fuzzy numbers. They implemented this method on Spain stock market, the results of which indicated that the proposed method had a higher prediction power than the classic one. Mishra et al. [16] proposed a new prediction based mean–variance model (PBMV) to select constrained portfolio assets. They predicted stock return using artificial neural network, and made use of particle swarm optimization (PSO) for the relevant weights of network. They also utilized Sign test and Wilcoxon rank test in order to compare the proposed method performance with Markowitz model. The findings demonstrated that this method had a better ability to produce Pareto solutions.

In recent years, in addition to increasing use of artificial networks, the use of Fuzzy C-means (FCM) approach has been increased. Using Fuzzy C-means and K-means algorithms in order to analyze data communication performance based on computational speed, Velmurugan [17] showed that FCM algorithm was more accurate and understandable than K-means, but K-means algorithm was faster and yet more vulnerable to noise. Akman [18], first, considered green as the providers' performance and evaluation measures and then, carried out factor analysis to evaluate the selected factors verification. Secondly, the author determined all of the providers of an automobile manufacturer considering some criteria including delivery, quality, expense, clustering services, and the best performance of the providers, and then, evaluated them based on environmental measures (green measures). The best performance of the providers was divided into three groups of green measures including good, average, and weak. Finally, the providers within the weak group were sequenced by using VIKOR in order to develop the plans of green provider. Bai et al. [19] introduces a combination of Rough set theory and FCM approach. The combination approach was based on the measurement of providing methods performance, investment objectives agreement, and identification of all instructions, which can be helpful in making decision about investment in more effective and sensible providing ways. In this paper, a combination of FCM algorithm and Rough set theory has been used to develop the simple rules and the more complete and accurate approach for the analyzers, and decrease the use of rules for investment decisions. Also, Fathabadi [20] implemented dynamic fuzzy c-means algorithm and a three-layer artificial neural network to minimize the power missed in two electric power distribution networks of IEEE69-bus and IEEE33-bus. This method was compared with the classic method of switching. The results indicated that reconfiguration of electric power distribution network using this method was a short process and had a higher power than other methods. Zhang et al. [21] focused on incomplete data clustering and proposed the FCM model with maximum possibility index using non-parametric hypothesis test to describe the missed value of a specific Gaussian distribution based on the nearest neighbor. Then, they optimized this model using alternative three-level optimization and Lagrange multiplier method. Convergence and time complexity of the model was discussed and according to the sample error index, it was proved that the proposed model had an effectively stable convergence, high precision, and considerable strength for incomplete data clustering.

Esfahanipour and Aghamiri [22] presented a method consisting of Takagi–Sugeno–kang (TSK) fuzzy rule-based system to evaluate the stock market, and used fuzzy c-mean clustering and Gaussian function to determine some rules and estimate the degree of membership. In addition, TSK parameters were adapted by Adaptive Nero-fuzzy Inference System (ANFIS). In order to examine their own proposed method, they used some of the Tehran stock exchange indices. Based on the results, this proposed method has a better performance than other methods including multiple regression analysis and back propagation neural networks (BPNN). Pai et al. [23] introduced a new compound system for precise prediction of tourism demand. This system is a FCM model with a logarithm least-squares support vector regression (LLS-SVR) method that makes use of genetic algorithm (GA) to select LLS-SVR parameters. The results of evaluating this method using Taiwan and Hong-Kong's data from 1969 to 2010 demonstrated that this method is better than other prediction methods in terms of precision. Considering the observations, which are uncertain, Yolcu et al. [24] presented a model of time series prediction. This model contained three basic stages: (1) fuzzification of the crisp observations, (2) identification of fuzzy relations, and (3) defuzzification. They introduced ignoring the degree of membership in prediction process as one of the negative features of fuzzy time series, and solved this problem using artificial neural networks. They also used a new proposed high-order method from FCM in the fuzzification step and a multi-user MIMO (FFANN) to identify fuzzy relations. In order to solve the problem associated with the input numbers of ANN in the high-order model, they utilized an operation crossroads for the membership degree value of each observation in fuzzy time series delay time.

On the other hand, evaluating decision-making units (DMUs) is one of the main management problems. Data envelopment analysis (DEA) is a non-parametric method for evaluating DMUs that introduced by Charnes et al. [25]. DEA has been applied in various area such as bank branches, refineries units, industrial units, educational and financial institutes. Eilat et al. [26] integrated DEA and balanced scorecard to evaluate portfolio of R&D projects. Edirisinghe and Zhang [27] used DEA for analyzing 230 firms from various US technology-industries over time to determine a relative financial strength indicator and to predictive stock price returns for each firm. They proposed two-step algorithm including random sampling and local search optimization. Pätäri et al. [28] studied the applicability of DEA for constructing equity investment strategy integrating the value investing and the momentum investing. They used three different models of DEA including the constant return to scale (CRS) model, the cross-efficiency model, and the super-efficiency model. Also, Misiunas et al. [29] used an integrating approach base DEA and neural network to predict the organ recipient functional status.

Since the precision of neural network results greatly depend on the size and complex of training data, the neural network is not sufficient. Therefore, a stronger approach is necessary for financial predictions. In addition, previous researches have shown that combining models perform better when the data structure changes. Thus, combining models are good enough to improve prediction.

Abovementioned methods have often used all data and simply static ones to predict the stock market. Meanwhile, the prediction strength and precision decreases due to inefficient and heterogeneous data. Hence, a method is required, which has the ability to distinguish between data. In this article, an approach is introduced in which dynamic fuzzy c-means (DFCM), DEA, and ANN methods have been integrated for the first time in order to predict stock exchange online data. The classical FCM method is unable to be dynamic behavior such as changing the number of clusters or clusters members. Hence, it is developed to make dynamic features to remove these problems in classic FCM. Then, DEA is used to evaluate DMUs by using financial ratios as outputs to provide efficiency scores (targets) in neural network. Finally, designed network is trained and prepared for predicting companies' future performance. This approach has solved the problem of training datasets for the neural network and has taken advantage of clustering as well.

The paper structure proceeds in this way in Section 2, the research methodologies including the introduction of FCM, DEA and ANN are investigated. Then, in Section 3, the developed version of FCM called DFCM is presented. Furthermore, introducing case study is given in this section. In Section 4, the results of the analysis of implementing the proposed approach for stock exchange companies are stated. Finally, in Section 5, summary of work and conclusions will be presented.

2. Methodology

In this paper, three approaches including FCM, Artificial neural network (ANN) and data envelopment analyses (DEA) will be applied and if needed they will be extended. The details of the used approaches are described as follows.

2.1. Fuzzy C-means

Classical approaches apply statistical methods for clustering similar and homogeneous data. According to available data, the criteria are different. The criteria should be chosen such that the data can be classified in different categories. But, in real applications, a sample may be belonged to different clusters with different membership degree. Fuzzy clustering is a method to assign data to different clusters using fuzzy logic, which provides effective means for divorcing overlapping clusters. The most common fuzzy clustering algorithm is Fuzzy C-means (FCM). This method was developed by Dunn in 1973 and improved by Bezdek in 1981. The association to a cluster is verified by computing the reverse distance to the cluster center. The cluster centers verified by FCM clearly depend on the geometric locations of the data points on the plane or space. In the FCM algorithm, an objective function, which should be minimized, is reflected as:

$$F(U,V,m;X)=\sum_{i=1}^{k}\sum_{j=1}^{n}\left(u_{ij}\right)^{m}\left\|x_j-v_i\right\|^2 \tag{1}$$

where m is the fuzzy factor, k is the number of clusters, $V=(v_1,v_2,\ldots,v_k)^T$ is cluster centers vector containing the centers of the k clusters, n is the number of the data points, $X=(x_1,x_2,\ldots,x_n)^T$ is the data points vector, $U=\left[u_{ij}\right]_{k\times n}$ is the membership matrix involving of the membership u_{ij} which shows the membership of $x_j$in the ith cluster, and $\|.\|$ shows the Euclidean distance norm ($\|Z\|=\sqrt{Z^T\cdot Z}$). m is used to normalize and fuzzify the memberships the sum of which should be equal to 1. Minimization of $F(U,V,m;X)$ is carried out through an iterative techniques such as alternating optimization (AO). When $m>1$, an optimal solution that minimizes $F(U,V,m;X)$ is found as:

$$u_{ij}=\left[\sum_{p=1}^{k}\left(\frac{\left\|x_j-v_i\right\|}{\left\|x_j-v_p\right\|}\right)^{2/(m-1)}\right]^{-1} \tag{2}$$

where $1\le i\le k$, $1\le j\le n$, and the center of the ith cluster is achieved as:

$$V_i=\frac{\sum_{j=1}^{n}\left(u_{ij}\right)^{m}x_j}{\sum_{j=1}^{n}\left(u_{ij}\right)}. \tag{3}$$

After clustering the data, a validity index is used to expression how well the data have been clustered. There are different validity indexes such as Xie–Beni index and modified partition coefficient (PCC) index.

2.2. Multi layers perceptron

Multi Layers Perceptron (MLP) network is one of the most popular artificial neural networks used in various scientific fields; specially in the forecast and prediction. MLP network is formed by an input layer, one or more hidden layers and an output layer. Input layer accepts the data vector or pattern; hidden layers include one or more layers and accept the output from the previous layer, weigh them, and pass through an activation function. Output layer takes the output from the final hidden layer and weighs them, and possibly passes through an output activation function to produce the target values (see Fig. 1).

MLP modeling process is including the following steps:

1. Data collection, definition number of hidden layers and relative weights.
2. Preprocessing data (If necessary).
3. Defining transformation functions.
4. Training network with train data set.
5. Performing network for test data set.
6. Determining, validating, and verifying network efficiency.
7. Reviewing the above steps (if necessary).

There are varieties of activation functions that may be used. Consider the general form:

$$y_i(x)=f\left(b_i+\sum_{i}w_{io}^{T}x_i\right) \tag{4}$$

where $f(X)$ is activation function, x_i activation of ith hidden layer node, w_i is the interconnection between ith hidden layer node and oth output layer node, b_i is bias of ith hidden layer. There are many ways for network error calculation such as MAE

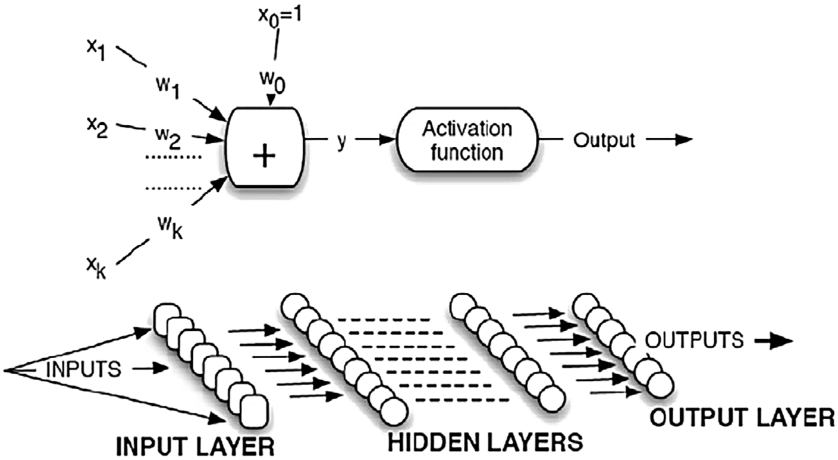

Fig. 1. Scheme of a neuron and an artificial neural network (multilayer perceptron).

(mean average error), MSE (mean square error), and RMSE (root mean square error). RMSE is a frequently used measure of error that general formulation is as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (x_i - x)^2}{n}} \tag{5}$$

where x_i is the output of ith unit.

2.3. Data envelopment analysis

Data envelopment analysis (DEA) allows for the comparative evaluation of efficiencies within a group of decision-making units (DMUs). For this, the ratio of weighted outputs to weighted inputs is computed from historical data. DEA models differ from the consideration of constant or variable returns to scale [25,30] and have been known CCR and BCC models respectively. CCR model can be adjusted based on the output or input and the choice of these models depends on the stuffs of decision-making unit (DMUs) in the production frontier. The input-adjusted model minimizes the using of inputs for a given level of outputs and the output-adjusted model maximizes the producing of outputs for a given level of the inputs. Suppose we have n DMUs and each DMU$_j$ $(j = 1, \ldots, n)$ produces s outputs y_{rj} $(r = 1, \ldots, s)$ by using m inputs x_{ij} $(i = 1, \ldots, m)$. DEA uses the following mathematical model for evaluation of DMU$_o$'s efficiency:

$$\begin{aligned} &\max \frac{\sum_{r=1}^{s} u_r y_{ro}}{\sum_{i=1}^{m} v_i x_{io}} \\ &\text{s.t. } \frac{\sum_{r=1}^{s} u_r y_{ro}}{\sum_{i=1}^{m} v_i x_{io}} \le 1, \quad j = 1, \ldots, n \\ &u_r, v_i \ge \varepsilon, \quad i = 1, \ldots, m, \ r = 1, \ldots, s. \end{aligned} \tag{6}$$

3. Proposed approach

Prediction is a powerful tool in the process of planning that can provide the decision maker with a prediction about the future events according to using experiences and applying statistical, mathematical, or computational methods. Due to the application of prediction methods in different fields such as supply chain, transportation planning, economy, telecommunication, production, weather forecast, earthquake, world crude oil price, stock price prediction, and other similar cases, has promoted research on prediction models and techniques in the last few decades. In many real-world applications, online analysis of data and creating an adaptive clustering technique is required. The aim of this article is to present an online prediction system based on FCM, DEA, and ANN methods. In the first stage of data analysis, homogeneous data should be distributed in relevant clusters.

As was described in the previous section, FCM method is one of the most well-known methods of data clustering, but because the clusters need to be updated for online prediction in different time series, old clustering methods such as FCM will lose their effectiveness. Furthermore, when the new data are heterogeneous or different from the existing data, it is likely that all cluster sets including the data of each cluster and even the number of clusters may be changed. In this case,

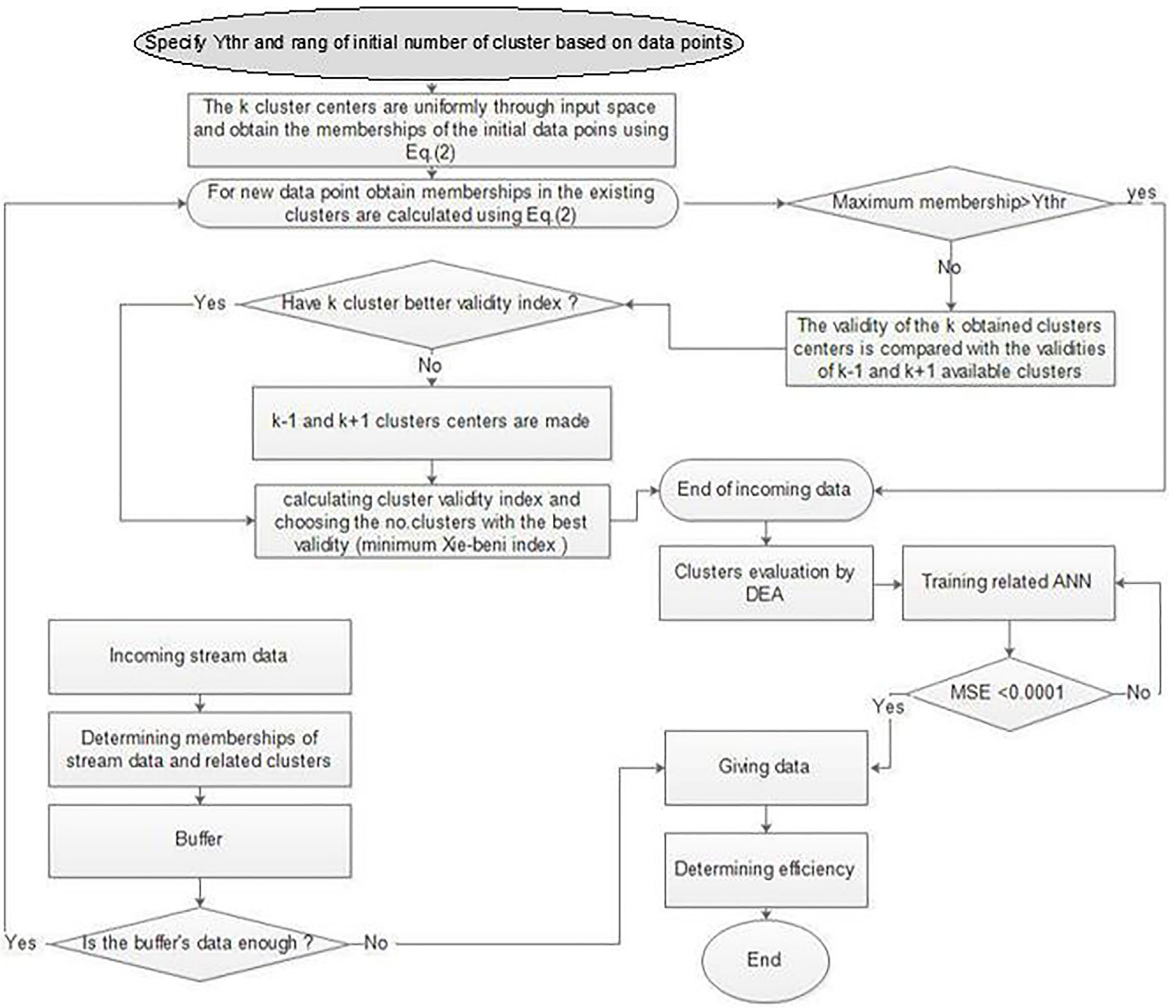

Fig. 2. Flowchart of proposed approach based on DFCM-DEA-ANN.

FCM loses its capability because it cannot update the clusters. Therefore, we have to carry out dynamic clustering. To this reason, we can analyze dynamic data in different fields by making a change in FCM algorithm. This model, named Dynamic Fuzzy C-means (DFCM), can increase or decrease the number of clusters on demand, and solve the heterogeneous of data distribution problem [31].

In order to using DFCM model, new parameters should be defined. The first parameter is membership threshold (*Ythr*). This parameter is the maximum acceptable level for the degree of memberships. FCM error (*EFCM*), the next parameter, is the maximum acceptable difference between two centers of the clusters obtained in two successive stages using FCM algorithm. K is the number of initial clusters, which is used in Eq. (2). In this paper, Xie-Beni index is used as a validity index. It is used to determine how to better show the data in obtained clusters.

Xie-Beni index is one of the indices that is widely used because not only is it related to data membership degree, but also it is calculated based on geometric distance. This index is defined as:

$$I \cdot V_{XB}(U, V; X) = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n}(u_{ij})^2 \left\|x_j - v_i\right\|^2}{k \cdot \left(Min_{i\neq j}\left\{\left\|v_i - v_j\right\|\right\}\right)} = \frac{F(U, V, 2; X)}{k \cdot \left(Min_{i\neq j}\left\{\left\|v_i - v_j\right\|\right\}\right)}. \tag{7}$$

DFCM clustering algorithm is described as follows:

Firstly, membership threshold, FCM error, and the number of initial clusters are estimated and initialized by the decision maker. Then, initial input data are entered in FCM algorithm and clustered. Afterwards, new input data are entered in the algorithm and the membership degree of new input data are calculated based on Eq. (2). If the maximum degree of membership is bigger than or equal to membership threshold, it means that the data belong to one of the clusters. However, if this degree is smaller than *Ythr*, it must be examined to see whether or not there is a better choice available. k , the number

Table 1
Summary of data characteristics used in case study.

Year		Criteria						
		TDTA	CLTA	CFTA	WCTA	NITA	CATA	EBIE
2007–2008	Mean	0.630216	0.546978	0.026544	0.065178	0.109212	0.612156	0.165022
	Variance	0.159574	0.162235	0.000651	0.156385	0.018725	0.044384	0.019449
2008–2009	Mean	0.623031	0.539064	0.03406	0.079374	0.099116	0.618438	0.15411
	Variance	0.067889	0.0665	0.001306	0.071647	0.017888	0.041814	0.019804
2009–2010	Mean	0.639334	0.555828	0.031706	0.067941	0.09828	0.623768	0.152187
	Variance	0.091062	0.083893	0.001027	0.085927	0.016002	0.041881	0.013997
2010–2011	Mean	0.644352	0.55662	0.037232	0.071782	0.10194	0.628402	0.158112
	Variance	0.105275	0.091729	0.00185	0.09204	0.033454	0.042997	0.033577
2011–2012	Mean	0.638093	0.542529	0.041185	0.102419	0.693138	0.644948	0.773091
	Variance	0.114897	0.095113	0.002128	0.089097	30.73679	0.039859	31.19343

of obtained clusters centers, is compared with $k-1$ and $k+1$ centers of the existing clusters and then, clusters centers are selected by an iterative process and based on Xie-Beni index. If $k-1$ or $k+1$ present a better validity index, $k-1$ or $k+1$ of clusters centers are created using new data points. DFCM process continues until no new data are received.

In the proposed method, online data are first entered into DFCM. This trend continues until online data reach the final cluster numbers. Afterwards, data in each cluster are evaluated using the DEA method. The most important point in this part of our proposed approach is not implementing the DEA method using the whole data since after implementing the DFCM method, DMUs with similar characteristics are placed in one cluster and finally, the outliers are easily recognizable. After the outliers are removed and DMUs are homogenized by the mentioned method, the DEA approach will be usable precisely. After the DMUs of each cluster are evaluated using this method and assigned efficiency values by using financial ratios, the next stage, i.e. neural network training will begin. Data are used to the network, as a training sample should be refined and homogenized as much as possible. Hence, training datasets and the network error rate are two important parameters in prediction ability of the network. In order to solve this problem in the proposed approach, the clusters are separately given to the neural network as training data (input). The efficiency of DMUs in the previous stage is considered as the neural network target. The network will be trained for clusters obtained by the DFCM method.

After network training is done, the next stage, i.e. making prediction for the new data are started. In order to identify the new data to which of the created clusters in this stage belong, they are firstly entered into the first stage. After the new homogeneous data are distributed to the relevant clusters, the data enter the neural network in a way that the network trained by ith cluster in the previous stages can make prediction for new data of ith cluster.

After different periods when the number of new data is enough, the collected data are again added to the first stage so that the data, number of clusters, and other stages are updated.

In this approach, it should be considered that prediction for the new data are made after clustering and by the network trained by that cluster rather than the one trained by the whole data because due to the heterogeneity of the data, network precision will decrease and it will contain significant errors. Fig. 2 summarizes the implementation stages of the proposed approach.

On the other hand, the used financial data that have been applied in this paper are expressed. These data are belonged to Tehran stock exchange market.

Total debt divided by total assets (TDTA): This ratio shows the percentage of funds that have been funded by debt, creditors prefer low debt ratio, because if it is high, financial risk and financial difficulty probability goes up. It is an input variable in DEA model.

Current liabilities divided by total assets (CLTA): This ratio shows the company's reliance degree on short-term financial resources for the acquisition of the company assets. Value of this ratio is placed under the influence current liabilities and permanent sources of funds and creditors preferring a lower debt ratio. It is an input variable in DEA model.

Cash flow divided by total assets (CFTA): This ratio indicates that how profitable the total assets of the company is used in generating profits in the form of cash flow and a higher ratio indicates high profitability and ability to generate high cash flow. It is an output variable in DEA model.

Working capital divided by total assets (WCTA): Whatever this ratio be higher, the ability of liquidity is better. It is an output variable in DEA model.

Net income divided by total assets (NITA): This ratio is the net profit divided by the sum of assets and shows profit that has been made for every unit of company assets. In other words, it measures the company's success in obtaining a net efficiency toward sales revenue or toward investments. It is an output variable in DEA model.

Current assets divided by total assets (CATA): It can determine the limits, or amount that company has invested their funds in items of current assets. In other words, this ratio indicates the company's ability to use all resources at its disposal. It is an output variable in DEA model.

Earnings before interest and taxes divided by total assets (EBTA): It is clear that investors prefer more profit. It is an output variable in DEA model.

Table 1 summarizes the data characteristics according to mentioned financial ratios.

Table 2
Results of using DFCM method for companies' financial data.

Initial number of clusters	Membership threshold	Proposed number of clusters by DFCM	Sum of data-distance to center of clusters
4	0.5	5	1117.031713
	0.7	5	1725.498948
	0.9	5	1725.498954
5	0.5	6	1117.001712
	0.7	6	1298.297829
	0.9	5	1552.214739
6	0.5	7	1725.498955
	0.7	6	1410.346636
	0.9	6	1410.396644

Table 3
Results of efficiency scores using DEA.

Cluster number	Number of DMUs in cluster	Maximum efficiency	Minimum efficiency	Average	Variance
1	223	1	0.2	0.730135	0.032042
2	195	1	0.06	0.146	0.018388
3	157	1	0.17	0.741847	0.043487
4	304	1	0.47	0.743651	0.017532
5	258	1	0.65	0.836047	0.008355
6	191	1	0.12	0.356283	0.036899

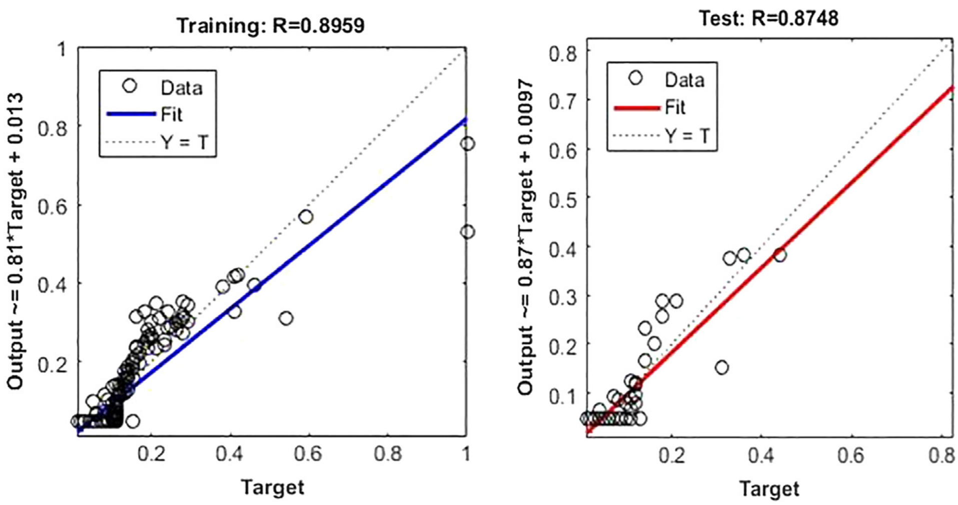

Fig. 3. Accuracy results of train and test data for clustered data (without using proposed approach).

4. Results and analysis

Regarding the proposed model, which was described in the previous sections, for DFCM clustering, the number of initial clusters and different membership thresholds are examined, which are illustrated in brief in Table 2. As was mentioned, in DFCM method, first, the initial number of clusters and threshold are given to algorithm and then, DFCM algorithm find the optimum number of clusters according to given threshold. This is repeated with different initial number of clusters and different thresholds. According to Table 2, the combination of six clusters (with initial five clusters) with membership threshold of 0.5 was selected as the optimum combination because the total distances of the points from the centers are less than other combinations. When this algorithm was run, at most six clusters were created, and the number of DMUs in clusters is 191, 258, 304, 157, 195, and 223, respectively. The advantage of using DFCM instead of FCM is the use of the latest available data (online data) as well as self-compatibility. According to online data, the DFCM algorithm decides whether or not a new cluster is required. Obviously, using this approach causes the clustering is appropriate and the placement of dissimilar data in each cluster effectively decreases.

In the next stage where DMUs were evaluated by CCR model (after clustering), the efficiency scores of DMUs in each cluster were entered the neural network separately. Maximum and minimum efficiency as well as the mean and variance of each cluster are briefly shown in Table 3.

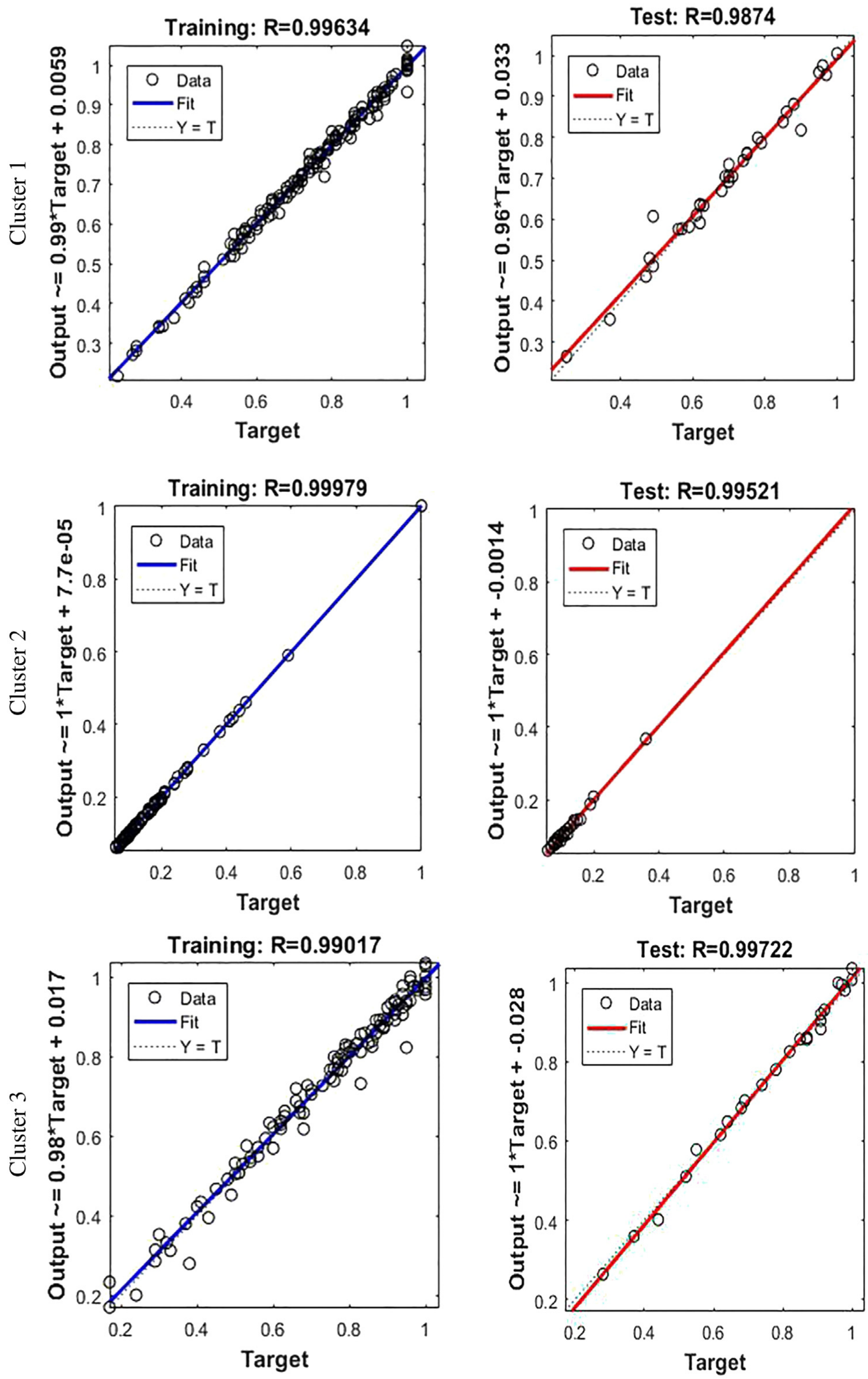

Fig. 4. Accuracy results of train and test data for clustered data (using proposed approach).

After data were evaluated with DEA, all TDTA, CLTA, CATA, NITA, WCTA, CFTA, and EBTA indices in the neural network were considered as input and DMUs efficiency scores were considered as the target.

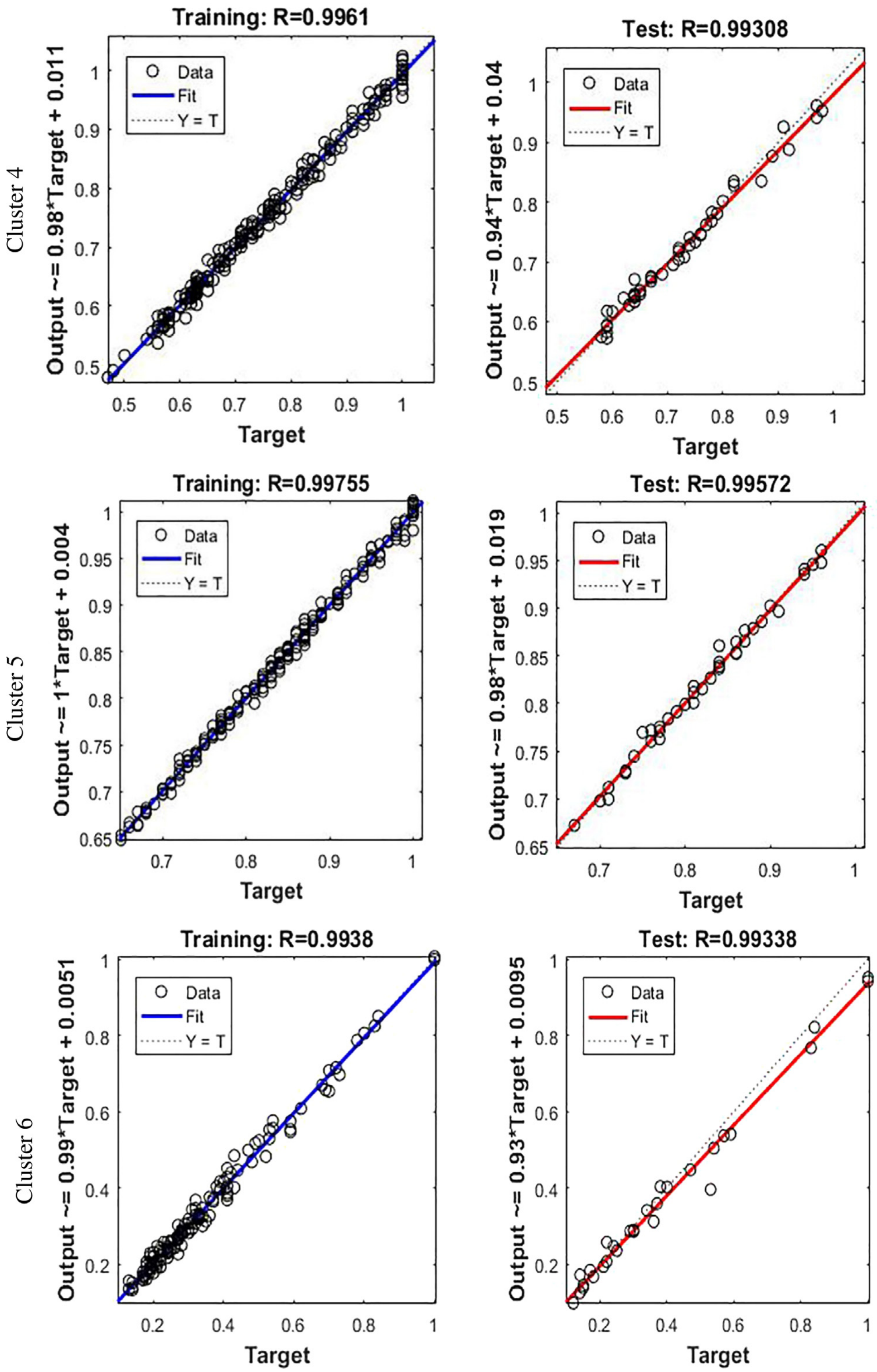

Fig. 4. (*continued*)

In this article, MLP neural network has been employed to train datasets in a way that tansig, purelin and satlin functions (functions in Matlab artificial neural network toolbox: See Fig. 5) have been selected as transfer functions, and some

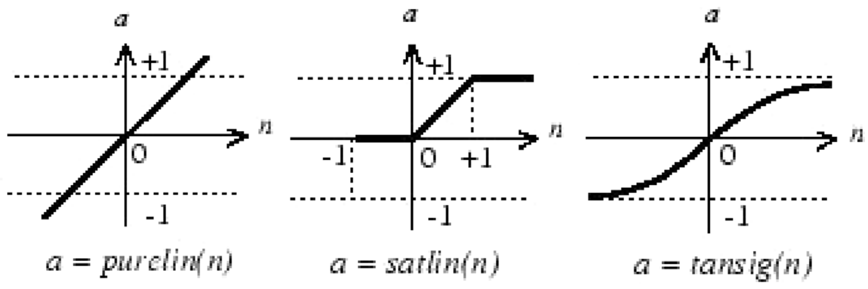

Fig. 5. ANN transfer functions.

Table 4
Sensitivity analysis of ANN parameters for sample cluster (Cluster 5).

Number	Parameter					
	Number of layer	Number of neuron	Transformation function	RMSE	*R* train	*R* valid
Cluster 5	1	3	tansig	0.0013	0.99234	0.98617
			purelin	7.5736e−04	0.93976	0.9579
			satlin	0.001	0.98707	0.98322
		4	tansig	7.6801e−04	0.99158	0.99025
			purelin	8.4602e−04	0.95684	0.96898
			satlin	0.0015	0.98685	0.97708
		5	tansig	5.4366e−04	0.99643	0.99166
			purelin	0.0011	0.9526	0.94391
			satlin	9.2499e−04	0.99069	0.98838
	2	3-1	tansig–tansig	5.6072E−04	0.98242	0.9845
			tansig–purelin	6.3737e−04	0.9903	0.98989
			tansig–satlin	6.8473e−04	0.89164	0.86153
		3-2	tansig–tansig	6.4048e−04	0.96137	0.96663
			tansig–purelin	1.1428e−04	0.96464	0.9127
			tansig–satlin	4.4381e−04	0.85573	0.83365
		4-1	tansig–tansig	0.001	0.98982	0.9896
			tansig–purelin	3.3662e−04	0.98868	0.98863
			tansig–satlin	0.0014	0.89234	0.87459
		4-2	tansig–tansig	5.3135e−04	0.97137	0.95778
			tansig–purelin	6.9433e−04	0.98848	0.98743
			tansig–satlin	5.8467e−04	0.87526	0.89414
		4-3	tansig–tansig	9.5455e−04	0.95783	0.94997
			tansig–purelin	5.8352e−04	0.95971	0.97457
			tansig–satlin	8.1355e−04	0.8465	0.86199
		5-1	tansig–tansig	9.0678e−04	0.99414	0.98579
			tansig–purelin	6.1888e−04	0.99645	0.98462
			tansig–satlin	6066.1e−04	0.90164	0.8923
		5-2	tansig–tansig	0.0011	0.99225	0.98893
			tansig–purelin	7.6885e−04	0.99162	0.99338
			tansig–satlin	3.4970e−04	0.89565	0.83218
		5-3	tansig–tansig	0.001	0.9851	0.9868
			tansig–purelin	5.8825e−04	0.99095	0.98838
			tansig–satlin	3.7722e−04	0.86821	0.84915
		5-4	tansig–tansig	8.1611e−04	0.95928	0.93679
			tansig–purelin	3.6551e−04	0.962	0.92706
			tansig–satlin	6.6295e−04	0.85952	0.82442

combinations of them, based on the number of middle layers in each cluster, have been evaluated separately. Furthermore, the percentages of train, valid and test data were considered 0.15, 0.15, and 0.75, respectively. Our data are divided into two sections. 70% of them are used for training and 30% of them are used for test. In the training process, RMSE error measurement index has been used. Note that the output layer is not a part of the middle layer. Figs. 3 and 4 depict the ability of the proposed approach to train and finally predict the performance scores of companies. Fig. 3 shows when all data are given to neural network, the algorithm cannot be well-trained. While MLP is run corresponding to the six clustered data sets separately, it provides a more accurate fitting of both train and test data.

Table 5
Results of clusters behaviors using ANN.

Number	Parameter				
	Number of layer	Number of neuron	RMSE	R train	R valid
Cluster 1	1	4	0.0012	0.98842	0.99508
		5	0.0024	0.99424	0.98524
	2	4-2	0.0031	0.9839	0.97997
		4-3	8.5875e−04	0.93478	0.9588
		5-1	8.8411e−04	0.9936	0.99251
		5-2	0.0027	0.98467	0.98564
Cluster 2	1	4	1.0926e−04	0.9992	0.97638
		5	1.6110e−04	0.97526	0.89893
	2	4-2	1.6040e−04	0.99432	0.9896
		4-3	9.9712e−04	0.87943	0.9693
		5-1	2.2731e−04	0.99884	0.98491
		5-2	6.4182e−04	0.9339	0.98628
Cluster 3	1	4	6.0094e−04	0.99643	0.99624
		5	0.0011	0.99893	0.99691
	2	4-2	0.0027	0.97345	0.93388
		4-3	4.9467e−04	0.98032	0.89262
		5-1	8.8862e−04	0.99602	0.99634
		5-2	0.0016	0.94893	0.98345
Cluster 4	1	4	6.4609e−04	0.9938	0.98329
		5	0.0011	0.99538	0.9903
	2	4-2	0.0015	0.95652	0.92825
		4-3	7.6681e−04	0.94722	0.95915
		5-1	0.0017	0.99188	0.99002
		5-2	0.002	0.98687	0.98864
Cluster 5	1	4	7.6801e−04	0.99158	0.99025
		5	5.4366e−04	0.99643	0.99166
	2	4-2	0.98743	0.98743	0.98743
		4-3	0.98848	0.98848	0.98848
		5-1	6.1888e−04	0.99645	0.98462
		5-2	7.6885e−04	0.99162	0.99338
Cluster 6	1	4	5.7546e−04	0.99271	0.98944
		5	0.0025	0.99268	0.99132
	2	4-2	0.0029	0.97798	0.9719
		4-3	0.0028	0.93813	0.9643
		5-1	9.6515e−04	0.99237	0.98668
		5-2	5.9099e−04	0.97608	0.97951

The dashed line in each plot denotes the difference between the targets and outputs. Also, the best-fit linear regression between estimated outputs and targets has been showed by solid line. Furthermore, R value is an indication of the relationship between the estimated outputs and targets. $R = 1$ shows that there is a linear relationship between targets and estimated outputs. There is a weak linear relationship between targets and estimated outputs when R obtains small value so that $R = 0$ presents no linear relationship between them. For further explanation, in the training section of cluster 2 in Fig. 4, $R = 0.99979$ which indicates that training process is appropriate. The test result also indicates that the correlation coefficient is high ($R = 0.99521$). So it can be said that the network is well-trained and has a very good performance and shows a strong linear relationship between outputs and targets.

As was said, each cluster was given to ANN separately. Without loss of generality, the fifth cluster is considered for more analysis and is presented the ability proposed approach for online performance prediction. In the fifth cluster, one or two middle layers were selected, and various combinations of the number of neurons and different activation functions were evaluated, which are briefly illustrated in Table 4. According to the results, it can be seen that by considering two middle layers with one and five neurons in each layer, and *tansing-purelin* transfer functions, RMES and the regression of train and test data have the most appropriate values. Therefore, cluster 5 was trained using these parameters. This process occurred separately for all clusters, which can be seen in brief in Table 5. According to the results, the combination of two middle layers, 1–5 neurons, and *tansing-purelin* transfer functions is the best combination in all clusters. Eventually, regarding the fact that data of different companies are usually published in the stock market in different time intervals such as three-month or six-month time intervals. We can make use of network training to make prediction for these types of online data. For verification of this approach, the data of 2012 were considered as new online data that published by stock exchange market in different periods of companies financial year (a year as reckoned for taxing or accounting purposes). Each company data was given to algorithm and was determined appropriate cluster. Then, the trained network of corresponding cluster

Table 6
Verification of proposed approach with considering data of 2012 as online data.

Company	Identified cluster by DFCM	Predicted efficiency	Real efficiency	Company	Identified cluster by DFCM	Predicted efficiency	Real efficiency	Company	Identified cluster by DFCM	Predicted efficiency	Real efficiency	Company	Identified cluster by DFCM	Predicted efficiency	Real efficiency
C001	6	0.175	0.15	C041	5	0.751	0.73	C081	6	0.529	0.47	C121	1	1	1
C002	4	0.868	0.84	C042	3	0.771	0.76	C082	2	0.261	0.22	C122	5	0.672	0.67
C003	2	0.28	0.28	C043	2	0.276	0.23	C083	2	0.259	0.22	C123	1	0.905	0.94
C004	6	0.598	0.55	C044	3	0.938	1	C084	2	0.095	0.08	C124	5	0.755	0.75
C005	6	0.774	1	C045	5	0.718	0.68	C085	6	0.099	0.11	C125	6	0.18	0.16
C006	4	0.282	0.26	C046	4	0.569	0.58	C086	2	0.771	0.57	C126	2	0.153	0.13
C007	2	0.487	0.35	C047	4	0.776	0.78	C087	2	0.637	0.45	C127	2	0.097	0.08
C008	1	0.95	0.93	C048	3	0.76	0.78	C088	6	0.42	0.34	C128	4	0.892	0.86
C009	1	0.353	0.38	C049	4	0.818	0.79	C089	2	0.163	0.15	C129	2	0.071	0.06
C010	6	0.664	0.59	C050	5	0.87	0.85	C090	6	0.193	0.17	C130	4	0.926	0.9
C011	1	0.636	0.64	C051	6	0.265	0.27	C091	2	0.224	0.21	C131	6	0.231	0.23
C012	2	0.181	0.17	C052	2	0.226	0.22	C092	6	0.396	0.39	C132	2	0.192	0.18
C013	2	0.092	0.08	C053	6	0.178	0.18	C093	2	0.073	0.07	C133	4	0.909	0.88
C014	4	0.793	0.79	C054	3	0.883	0.86	C094	6	0.467	0.5	C134	1	0.94	0.98
C015	2	0.238	0.2	C055	2	0.082	0.08	C095	6	0.305	0.34	C135	5	0.749	0.72
C016	1	0.822	0.84	C056	5	0.933	0.92	C096	2	0.132	0.12	C136	3	0.996	1
C017	2	0.15	0.13	C057	2	0.082	0.07	C097	6	0.25	0.23	C137	1	0.818	0.81
C018	2	0.495	0.41	C058	6	0.238	0.21	C098	2	0.061	0.05	C138	6	0.227	0.25
C019	2	0.173	0.17	C059	4	1	1	C099	6	0.209	0.18	C139	2	0.1	0.09
C020	3	0.798	0.79	C060	3	0.929	0.96	C100	6	0.314	0.26	C140	6	0.233	0.23
C021	4	0.609	0.59	C061	4	0.551	0.56	C101	1	0.73	0.74	C141	1	0.72	0.73
C022	1	0.911	0.91	C062	2	0.076	0.06	C102	2	0.106	0.09	C142	6	0.193	0.23
C023	1	0.833	0.86	C063	2	0.179	0.18	C103	6	0.304	0.31	C143	5	0.868	0.85
C024	3	0.62	0.63	C064	4	0.994	0.96	C104	6	0.207	0.18	C144	4	0.638	0.63
C025	4	0.853	0.85	C065	4	0.756	0.75	C105	1	0.636	0.68	C145	5	0.817	0.81
C026	5	0.716	0.71	C066	3	0.16	0.22	C106	6	0.237	0.21	C146	5	0.808	0.81
C027	5	0.89	0.85	C067	4	0.621	0.6	C107	2	0.062	0.05	C147	2	0.062	0.06
C028	5	1	1	C068	3	0.325	0.38	C108	6	0.201	0.14	C148	2	0.098	0.09
C029	3	0.929	0.91	C069	5	1	1	C109	6	0.313	0.27	C149	3	0.936	0.94
C030	5	0.692	0.66	C070	5	1	1	C110	6	0.488	0.38	C150	3	0.74	0.75
C031	3	0.823	0.79	C071	6	0.238	0.18	C111	6	0.311	0.29	C151	4	0.804	0.8
C032	4	0.595	0.61	C072	2	0.312	0.23	C112	6	0.202	0.18	C152	6	0.31	0.3
C033	4	0.939	0.68	C073	3	0.619	0.64	C113	6	0.203	0.17	C153	5	0.813	0.79
C034	4	0.614	0.63	C074	2	0.082	0.07	C114	6	0.461	0.4	C154	4	0.609	0.59
C035	5	0.99	0.99	C075	2	0.09	0.09	C115	6	0.449	0.38	C155	4	0.982	0.97
C036	4	0.599	0.6	C076	3	1	1	C116	2	0.244	0.2	C156	5	0.706	0.69
C037	5	0.869	0.84	C077	6	0.951	0.99	C117	2	0.427	0.33	C157	3	0.896	0.88
C038	4	0.678	0.61	C078	2	0.115	0.1	C118	2	0.336	0.28	C158	1	0.914	0.92
C039	5	0.791	0.8	C079	5	0.746	0.75	C119	1	1	1	C159	5	1	1
C040	3	0.587	0.58	C080	2	0.621	0.61	C120	4	0.989	1	C160	3	0.742	0.45
C161	1	0.724	0.71	C197	2	0.084	0.08	C233	3	0.965	0.99				
C162	5	0.669	0.66	C198	2	0.11	0.09	C234	6	0.235	0.24				
C163	3	0.843	0.86	C199	3	0.64	0.54	C235	4	0.754	0.75				
C164	3	0.528	0.48	C200	5	0.798	0.78	C236	6	0.888	0.48				
C165	4	0.752	0.73	C201	1	0.484	0.48	C237	2	0.142	0.12				
C166	2	0.178	0.13	C202	5	0.75	0.74	C238	2	0.16	0.12				
C167	2	0.081	0.07	C203	2	0.067	0.06	C239	4	0.733	0.78				

(*continued on next page*)

Table 6 (*continued*)

C168	1	0.892	0.86	C204	2	0.208	0.19	C240	4	0.788	0.78
C169	5	0.824	0.83	C205	5	0.698	0.67	C241	2	0.09	0.07
C170	2	0.102	0.08	C206	5	0.884	0.88	C242	4	0.998	1
C171	2	0.087	0.07	C207	1	0.801	0.8	C243	4	1	1
C172	5	0.796	0.8	C208	2	0.118	0.09	C244	5	0.781	0.78
C173	4	0.798	0.8	C209	4	0.627	0.63	C245	2	0.076	0.07
C174	3	0.685	0.67	C210	3	0.894	0.89	C246	2	0.225	0.18
C175	3	0.935	0.92	C211	2	0.081	0.06	C247	5	0.993	0.98
C176	5	0.824	0.83	C212	5	0.847	0.84	C248	4	0.919	0.89
C177	4	0.679	0.68	C213	2	0.173	0.17	C249	5	0.864	0.88
C178	4	0.834	0.81	C214	6	0.525	0.38	C250	4	0.85	0.82
C179	5	0.908	0.92	C215	4	0.774	0.73	C251	2	0.077	0.06
C180	6	0.606	0.5	C216	3	0.78	0.77	C252	5	0.838	0.84
C181	4	0.841	0.83	C217	5	0.766	0.77	C253	2	0.073	0.06
C182	4	0.641	0.61	C218	6	0.558	0.51	C254	5	1	1
C183	4	0.862	0.82	C219	5	0.751	0.76	C255	4	0.9	0.9
C184	5	0.982	1	C220	5	0.826	0.81	C256	4	0.795	0.79
C185	5	0.721	0.72	C221	3	0.175	0.22	C257	1	0.89	0.99
C186	4	0.909	0.93	C222	6	0.228	0.21	C258	2	0.137	0.1
C187	4	0.798	0.72	C223	4	0.674	0.66	C259	4	0.722	0.72
C188	2	0.16	0.14	C224	1	0.438	0.41	C260	2	0.088	0.07
C189	5	0.738	0.73	C225	5	0.755	0.75	C261	5	0.923	0.91
C190	1	0.869	0.87	C226	5	0.931	0.92	C262	2	0.069	0.06
C191	4	0.874	0.86	C227	2	0.082	0.07	C263	2	0.97	1
C192	5	0.924	0.97	C228	6	0.243	0.21	C264	5	0.745	0.73
C193	5	0.795	0.79	C229	2	0.152	0.12	C265	4	0.962	0.95
C194	2	0.126	0.11	C230	3	0.985	0.96	C266	2	0.149	0.14
C195	2	0.067	0.05	C231	5	0.788	0.78				
C196	1	0.686	0.66	C232	4	0.987	1				

Table 7
Comparison of FCM and DFCM clustering method in proposed approach.

DMU	FCM	DFCM	Prediction FCM	Prediction DFCM	Real efficiency	DMU	FCM	DFCM	Prediction FCM	Prediction DFCM	Real efficiency
C011	✓	×	0.694	–	–	C157	✓	×	0.716	–	–
C020	✓	×	0.629	–	–	C159	×	✓	–	1	–
C024	✓	×	0.601	–	–	C160	✓	×	0.746	–	–
C026	✓	✓	0.818	0.716	0.71	C162	×	✓	–	0.669	–
C027	✓	✓	0.743	0.89	0.85	C164	✓	×	0.337	–	–
C028	×	✓	–	1	–	C169	×	✓	–	0.824	–
C029	✓	×	0.711	–	–	C172	×	✓	–	0.796	–
C030	✓	✓	0.789	0.692	0.66	C174	✓	×	0.893	–	–
C031	✓	×	0.798	–	–	C175	✓	×	0.577	–	–
C035	×	✓	–	0.99	–	C176	×	✓	–	0.824	–
C037	×	✓	–	0.869	–	C179	×	✓	–	0.908	–
C039	✓	✓	0.901	0.791	0.8	C184	✓	✓	1	0.982	1
C040	✓	×	0.405	–	–	C185	×	✓	–	0.721	–
C041	✓	✓	0.934	0.751	0.73	C187	✓	×	0.925	–	–
C042	✓	×	0.855	–	–	C188	✓	×	0.955	–	–
C044	✓	×	0.865	–	–	C189	✓	✓	0.861	0.738	0.73
C045	✓	✓	0.723	0.718	0.68	C192	×	✓	–	0.924	–
C048	✓	×	0.752	–	–	C193	×	✓	–	0.795	–
C050	×	✓	–	0.87	–	C199	✓	×	0.527	–	–
C054	✓	×	0.718	–	–	C200	✓	✓	0.968	0.798	0.78
C056	×	✓	–	0.933	–	C202	×	✓	–	0.75	–
C060	✓	×	0.592	–	–	C205	✓	✓	0.889	0.698	0.67
C066	✓	×	0.330	–	–	C206	×	✓	–	0.884	–
C068	✓	×	0.581	–	–	C207	✓	×	0.839	–	–
C069	×	✓	–	1	–	C210	✓	×	0.737	–	–
C070	×	✓	–	1	–	C212	×	✓	–	0.847	–
C071	✓	×	0.753	–	–	C216	✓	×	0.622	–	–
C076	✓	×	0.969	–	–	C217	×	✓	–	0.766	–
C079	✓	✓	0.924	0.746	0.75	C219	✓	✓	0.707	0.751	0.76
C101	✓	×	0.664	–	–	C220	×	✓	–	0.826	–
C105	✓	×	0.550	–	–	C221	✓	×	0.300	–	–
C106	✓	×	0.557	–	–	C225	×	✓	–	0.755	–
C122	✓	✓	0.691	0.672	0.67	C226	×	✓	–	0.931	–
C124	✓	✓	0.897	0.755	0.75	C230	✓	✓	0.745	–	–
C134	✓	×	0.371	–	–	C231	✓	✓	0.945	0.788	0.78
C135	✓	✓	0.899	0.749	0.72	C233	✓	×	0.737	–	–
C136	✓	×	0.751	–	–	C236	✓	×	0.941	–	–
C141	✓	×	0.801	–	–	C244	✓	✓	0.785	0.781	0.78
C143	×	✓	–	0.868	–	C247	×	✓	–	0.993	–
C145	×	✓	–	0.817	–	C249	×	✓	–	0.864	–
C146	×	✓	–	0.808	–	C252	✓	✓	0.995	0.838	0.84
C149	✓	×	0.801	–	–	C254	×	✓	–	1	–
C153	×	✓	–	0.813	–	C257	✓	×	0.868	–	–
C156	×	✓	0.646	0.706	0.69	C261	✓	✓	0.942	0.923	0.91
C157	✓	×	0.716	–	–	C264	×	✓	–	0.745	–

was used to evaluate company efficiency. The results showed that this approach is able to apply for this purpose. Table 6 presents the summary of the verification results. Finally, the comparison of prediction method based on two clustering methods including FCM and DFCM is presented in Table 7. In this table, the real and predicted efficiency scores are for the test data. The real efficiency scores are evaluated when the data are available previously but predicted efficiency scores are obtained by proposed approach.

5. Summary and conclusion

Performance prediction of stock exchange firms, as a computational process, helps investors to make the right investment decisions. The previous evaluation methods use historical data. However, they may not be able to show company changes over time. Therefore, a dynamic approach should be provided to mitigate the effects of this matter. For this purpose, this paper proposed an integrating dynamic fuzzy c-means, data envelopment analysis, and artificial neural network. First, historical data of companies have been gathered and clustered by the modified fuzzy c-means that is able to generate dynamic clusters. After that, each cluster members were evaluated using financial ratios by DEA and then were trained by neural network. When the new data of company are received, they are given to algorithm. Algorithm select best cluster and if needed, it updates the clusters and evaluates the company by trained network. When the online financial data of companies is received in different periods and we may not evaluate company simultaneously, this approach can be fruitful. It can help investors to invest in companies before data of all companies can be collected in full. As a result, decision makers will have a faster and more effective decision. Finally, the real case study for Tehran stock exchange market has been used to demonstrate the abilities of the proposed online prediction approach.

Conflict of interest

The authors declare that they have no conflict of interest.

References

[1] R.A. Schwartz, D.K. Whitcomb, Evidence on the presence and causes of serial correlation in market model residuals, J. Finan. Quant. Anal. 12 (02) (1977) 291–313.
[2] W.C. Chiang, T.L. Urban, G.W. Baldridge, A neural network approach to mutual fund net asset value forecasting, Omega 24 (2) (1996) 205–215.
[3] M.T. Leung, A. Chen, H. Daouk, Forecasting exchange rates using general regression neural networks, Comput. Oper. Res. 27 (2000) 1093–1110.
[4] A.S. Chen, M.T. Leung, Regression neural network for error correction in foreign exchange forecasting and trading, Comput. Oper. Res. 31 (7) (2004) 1049–1068.
[5] A.A.P. Santos, N.C.A. da Costa, L. dos Santos Coelho, Computational intelligence approaches and linear models in case studies of forecasting exchange rates, Expert Syst. Appl. 33 (4) (2007) 816–823.
[6] A. Carriero, G. Kapetanios, M. Marcellino, Forecasting exchange rates with a large Bayesian VAR, Int. J. Forecast. 25 (2) (2009) 400–417.
[7] G. Chortareas, Y. Jiang, J.C. Nankervis, Forecasting exchange rate volatility using high-frequency data: Is the euro different? Int. J. Forecast. 27 (4) (2011) 1089–1107.
[8] G. Sermpinis, K. Theofilatos, A. Karathanasopoulos, E.F. Georgopoulos, C. Dunis, Forecasting foreign exchange rates with adaptive neural networks using radial-basis functions and particle swarm optimization, European J. Oper. Res. 225 (3) (2013) 528–540.
[9] Y. Yuan, Forecasting the movement direction of exchange rate with polynomial smooth support vector machine, Math. Comput. Modelling 57 (3) (2013) 932–944.
[10] T. Korol, A fuzzy logic model for forecasting exchange rates, Knowl.-Based Syst. 67 (2014) 49–60.
[11] G. Sermpinis, C. Stasinakis, K. Theofilatos, A. Karathanasopoulos, Modeling, forecasting and trading the EUR exchange rates with hybrid rolling genetic algorithms—Support vector regression forecast combinations, European J. Oper. Res. 247 (3) (2015) 831–846.
[12] R. Pinčák, E. Bartoš, With string model to time series forecasting, Physica A 436 (2015) 135–146.
[13] K. Kanjamapornkul, R. Pinčák, E. Bartoš, The study of Thai stock market across the 2008 financial crisis, Physica A 462 (2016) 117–133.
[14] A. Oztekin, R. Kizilaslan, S. Freund, A. Iseri, A data analytic approach to forecasting daily stock returns in an emerging market, European J. Oper. Res. 253 (3) (2016) 697–710.
[15] A. Rubio, J.D. Bermúdez, E. Vercher, Forecasting portfolio returns using weighted fuzzy time series methods, Internat. J. Approx. Reason. 75 (2016) 1–12.
[16] S.K. Mishra, G. Panda, B. Majhi, Prediction based mean–variance model for constrained portfolio assets selection using multiobjective evolutionary algorithms, Swarm Evol. Comput. 28 (2016) 117–130.
[17] T. Velmurugan, Performance based analysis between k-Means and Fuzzy C-Means clustering algorithms for connection oriented telecommunication data, Appl. Soft Comput. 19 (2014) 134–146.
[18] G. Akman, Evaluating suppliers to include green supplier development programs via fuzzy c-means and VIKOR methods, Comput. Ind. Eng. 86 (2015) 69–82.
[19] C. Bai, D. Dhavale, J. Sarkis, Complex investment decisions using rough set and fuzzy c-means: An example of investment in green supply chains, European J. Oper. Res. 248 (2) (2016) 507–521.
[20] H. Fathabadi, Power distribution network reconfiguration for power loss minimization using novel dynamic fuzzy c-means (dFCM) clustering based ANN approach, Int. J. Electr. Power Energy Syst. 78 (2016) 96–107.
[21] L. Zhang, W. Lu, X. Liu, W. Pedrycz, C. Zhong, Fuzzy C-Means clustering of incomplete data based on probabilistic information granules of missing values, Knowl.-Based Syst. 99 (2016) 51–70.
[22] A. Esfahanipour, W. Aghamiri, Adapted Neuro-Fuzzy Inference System on indirect approach TSK fuzzy rule base for stock market analysis, Expert Syst. Appl. 37 (7) (2010) 4742–4748.
[23] P.F. Pai, K.C. Hung, K.P. Lin, Tourism demand forecasting using novel hybrid system, Expert Syst. Appl. 41 (8) (2014) 3691–3702.
[24] O.C. Yolcu, U. Yolcu, E. Egrioglu, H. Aladag, High order fuzzy time series forecasting method based on an intersection operation, Appl. Math. Model. 40 (19) (2016) 8750–8765.
[25] A. Charnes, W.W. Cooper, E. Rhodes, Measuring the efficiency of decision making units, European J. Oper. Res. 2 (6) (1978) 429–444.
[26] H. Eilat, B. Golany, A. Shtub, Constructing and evaluating balanced portfolios of R&D projects with interactions: A DEA based methodology, European J. Oper. Res. 172 (3) (2006) 1018–1039.
[27] N.C.P. Edirisinghe, X. Zhang, Generalized DEA model of fundamental analysis and its application to portfolio optimization, J. Banking Finance 31 (11) (2007) 3311–3335.
[28] E. Pätäri, T. Leivo, S. Honkapuro, Enhancement of equity portfolio performance using data envelopment analysis, European J. Oper. Res. 220 (3) (2012) 786–797.
[29] N. Misiunas, A. Oztekin, Y. Chen, K. Chandra, DEANN: A healthcare analytic methodology of data envelopment analysis and artificial neural networks for the prediction of oran recipient functional status, Omega 58 (2016) 46–54.
[30] R.D. Banker, A. Charnes, W.W. Cooper, Some models for estimating technical and scale inefficiencies in data envelopment analysis, Manage. Sci. 30 (9) (1984) 1078–1092.
[31] R.P. Sandhir, S. Muhuri, T.K. Nayak, Dynamic fuzzy c-means (dFCM) clustering and its application to calorimetric data reconstruction in high-energy physics, Nucl. Instrum. Methods Phys. Res., Sect. A 681 (2012) 34–43.