Adv. Math. Fin. App., 2023, **8**(1), P. 241-254

Advances in Mathematical Finance & Applications

www.amfa.iau-arak.ac.ir
Print ISSN: 2538-5569
Online ISSN: 2645-4610
Doi: 10.22034/AMFA.2020.1882625.1327

Applied-Research Paper

Improved NARX-ANFIS Network structure with Genetic Algorithm to optimize Cash Flow of ATM Model

Neda Kiani[a], Ghasem Tohidi[b]*, Shabnam Razavyan[c], Nosratallah Shadnoosh[a], Masood Sanei[b]

[a] *Department of Industrial Management, Central Tehran Branch, Islamic Azad University, Tehran, Iran*
[b] *Department of Mathematics, Central Tehran Branch, Islamic Azad University, Tehran, Iran*
[c] *Department of Mathematics, South Tehran Branch, Islamic Azad University, Tehran, Iran*

ARTICLE INFO

Article history:
Received 2019-11-09
Accepted 2020-06-22

Keywords:
prediction model
autoregressive with exogenous input network
Genetic algorithm
ATM device

ABSTRACT

Nowadays, the rapid growth of data in organizations has caused managers to look for a way to analyze them. Extracting useful knowledge from aggregation data can lead to appropriate strategic decision-making for the organization. This paper suggests an application of a hybrid network based on the amount of monthly demand in every ATM device based on a transaction mean of 9 months for 1377 devices to obtain customer behavior patterns, to do so, first designed a basic model based on an auto-regressive with exogenous input network (NARX) then, the optimization of the weight and bias of the designed network is made by the genetic algorithm (GA). As a result, finding the weights of the network represents a nonlinear optimization problem that is solved by the genetic algorithm. Paper results show that the NARX-ANFIS Hybrid network using GA for the learning of rules and to optimize the network weights and weights of the network and the fixed threshold can improve the accuracy of the prediction model. Also, classic models are more efficient and increased benefits and lower financing costs and have more rational inventory cash control. As well, the designed model can lead to increase benefits and decrease costs in the bank so that, exact forecasts and optimal cash upload in ATMs will lead to increase funds on the bank and a rise in customers and popularity of the brand of the bank.

1 Introduction

Cash management plays an important role in the decision-making process. A non-optimal cash upload in the bank's Automatic Teller Machines (ATM) will lead to imposing excessive costs on the bank and a lack of funds in the devices will discontent customers and jeopardize the brand of the bank. By analyzing the information included in this database, banks can provide better services to their customers. Despite many conducted kinds of research concerning the application of data analysis techniques in the banking industry, there are many potential capacities for banks and financial institutions in developing this knowledge [1]. studies and research in the field of cash consumption prediction in ATM devices show can be predicted by modeling the historical behavior of data based on effective parameters and

* Corresponding author. Tel.: *+98-9122951833.*; fax +98-2188385777.
E-mail address: *gh_tohidi@iauctb.ac.ir*

cash withdrawal in these devices with acceptable accuracy. Doing service to Automatic Teller Machine devices, ATM is costly work because of cash management and operating costs [2]. With being high final price and the importance of efficiency, most banks focus on how cash management is done in ATMs, this is, how much cash has been kept in your devices to prevent excess and lack of cash. Therefore, banks as focus on cash management in branches must pay attention to the conditions of time, location, and economics that have done cash management in ATM devices. Also, introduced effective parameters in subject literature are monthly, weekly, and daily consumption and calendar parameters such as weekdays, holidays, and special vacations and financial influential events like the settle of subsidies in Iran and the settlement of salaries and pensions as well as the specific device parameters such as a geographic location near shopping centers or offices and special places, as well as other parameters that are inferred from quantitative studies [3,4].

Generally, to forecast accurately, not only effective parameters should be identified and extracted, appropriate forecasting models and finally, a combination of predictors should be used properly. The dynamic neural network is one of the new tools that are capable to analyze and simulate in nonlinear systems where relationships between system components and parameters are not well-known and descriptive**.** Since ATM transactions are time-dependent and the demand process is not constant and varies based on the timing of the withdrawals, a dynamic neural network is appropriate to predict the cash amount**.** Also, the main purpose of time series modeling is to give specific regularity in time-dependent observations to the prediction of the future; in other words, the most important goal of time series analyzing is to find the model of change and the prediction of the future. this study tries to formulate a possible, dynamic, and proper model that will be enabled to designed the model by the combination of nonlinear autoregressive networks with exogenous inputs, Neuro-fuzzy networks, and Genetic Algorithm to increase forecast accuracy and reduces the standard deviation of forecast error, and consequently, it improves forecasting**.** In the following, we review the literature on modeling, Simutis et al [3] in a paper on the Optimization of Cash Management for ATM Network Information Technology and control, point of view that ATMs are telecommunication devices that provide financial transactions in the public space without the need for humans [3]. Also, Simutis et al., [4] in another paper with the subject, A Flexible Neural Network for ATM Cash Demand Forecasting, ATM daily demand is predicted and optimized cash flow optimization routines for each ATM.

In this regard, Dilijonas et al. [5] Surveyed in the paper Retail Banking Optimization System Based on Multi-Agents Technology that optimized cash management and available service is one of the important factors in the services of ATM networks. Also, according to saying (Simutis et al., [6]) efficient cash management must have based on advanced algorithms that are enabling to predict the exact supply and demand of cash and permit the bank to manage actively using cash in across the network. Darwish [7] examined in his paper that the prediction model of cash demand for ATM networks like demand in the order of cash date for each of the ATM time fluctuations, and often a user's uncertain behavior that to it added, work is challenging. Hence Armenise [8] said that an optimized inventory of ATM cash creates optimal loading strategies and also, it can be caused to minimize the amount of daily cash inventory and ensured the same time cash distribution services. Ekinci [9] have been examined banks in ATM cash replenishment would be used as the lowest resource (such as cash maintenance in ATMs and trucks carrying money) for unstable demand of customers. Venkatesh [10], Cash demand forecasting in ATMs by clustering and neural networks, has improved cash demand forecasting in ATMs so that cash demand forecasting for a group of ATMs have based on the similarity of daily and weekly cash demand patterns. Recently, Arora et al [11] determined amount cash optimization in bank ATMs is Extremely difficult as demand for unstable cash cause changes in customer behavior, prioritization, withdrawal, time, and so on. Also, according to the point of view of Bhandari et al [12], accurate ATM forecasting for the future is one of the most important attributes to forecast because the business sector and daily needs of people are highly largely dependent on this. Biller et al., [13] Using different seasonal indicators has an important role in predicting cash needs for a bank. According to say Vennila et al [14] "the goal of the bank is to determine the optimal amount of money that must be placed in ATMs with the least opportunity cost and simultaneously decide on the level of satisfaction and unexpected customer needs". Improving the forecasting process is very important for both bank investors and researchers in finance as Farshadfar et al. [29] point out in their paper the forecasting in the financial time series is predicting the series behavior one or a few steps ahead with the help of several variables. Also, in the paper by Rashidi et al. [26] to forecast accurately, not only effective parameters should be identified and extracted, appropriate forecasting models and finally, a combination of predictors should be used properly. The result of the recent research shows that neural networks seem effective in different applications, for this reason, in this paper, we propose to improve a specific type of ANN to provide a good prediction model. The nonlinear autoregressive network with exogenous inputs (NARX) and Adaptive Nero-Fuzzy Inference System (ANFIS) hybrid network, is one of the new tools where the relation between components and system parameters is unknown and indescribable in nonlinear systems and uncertain, able to analyze and simulate. Since ATM transactions are influenced by time and are a nonlinear approximation, this network is an efficient model for device transactions. The rest of the paper is organized as follows. Section 2, explains about used models and the learning model process. Section 3, introduced the data set, and Section4, discusses how data is gathered and prepared. And also, is discussed in section 5 about the proposed methodology to accurately forecast customer behavior and maximize opportunities and minimize risks used in ATMs. Finally, section 6 concludes the work.

2 Related Models

2.1 Structure of NARX Network

The Multilayer perceptron, MLP, network is considered to be a static network component. In this type of grid, there are no recurring elements or feedbacks and no delay, and the outputs are obtained directly from the inputs and with the feeder connections. These types of networks are not efficient in complex time series modeling, which is influenced by input-triggering variables, and other neural networks need to be used [4]. The nonlinear autoregressive network with exogenous inputs (NARX) is a recurrent dynamic network, with feedback connections enclosing several layers of the network. The NARX model is based on the linear ARX model, which is commonly used in time-series modeling. In the regression vector of NARX models [15, 16], in dynamic networks, the network output not only is related to the input but also to the network input or output in the present and past. These networks are divided into two groups, 1) networks that have only feed-forward connections, and 2) networks that have recursive components. The nonlinear autoregressive network with exogenous inputs (NARX) is a recurrent dynamic network [15, 16]. These networks are divided into two categories, networks that only have feed-forward connections and networks that have feedback. The NARX neural network is a dynamic feedback network in which external inputs affect the outputs. Mathematically, consists of a finite number of past values of process inputs and outputs that the defining equation for the NARX model is.

$$y(t) = f(y(t-1), y(t-2), \dots, y(t-d_y), u(t-1), u(t-2), \dots, u(t-d_u)) \tag{1}$$

Where *f* is a nonlinear function and u(*t*) is the input of NARX, Y (*t*) is the output and also feedback of NARX. A vital task is to find the required number of lagged observations d_y, d_u to generate the auto-regressive structure for the model identification in time series. Where the next value of the dependent output signal *y* (*t*) is regressed on previous values of the output signal and previous values of an independent (exogenous) input signal. You can implement the NARX model by using a feedforward neural network to approximate the function *f*. A diagram of the resulting network is shown in Fig. 1, where a two-layer feedforward network is used for the approximation. This implementation also allows for a vector ARX model, where the input and output can be multidimensional [15, 16]. There are many applications for the NARX network. It can be used as a predictor, to predict the next value of the input signal. It can also be used for nonlinear filtering, in which the target output is a noise-free version of the input signal. The use of the NARX network is shown in another important application, the modeling of nonlinear dynamic systems [15, 16].

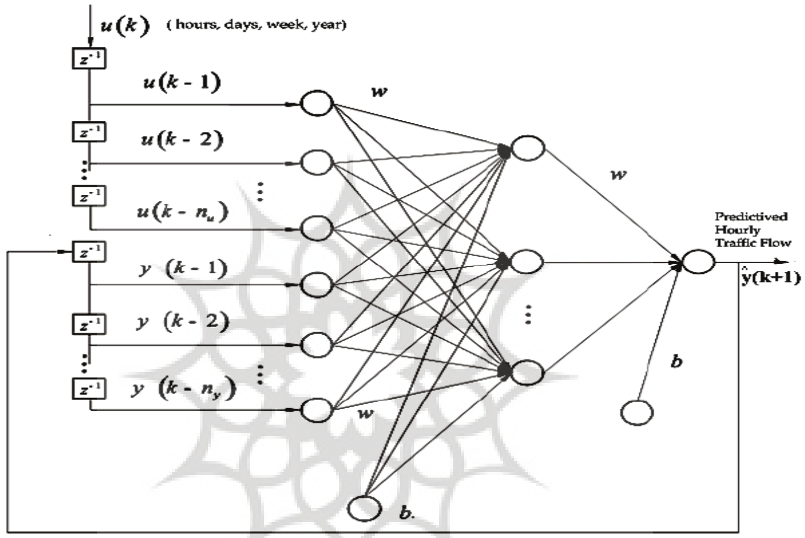

Fig.1: The Architecture of the NARX Network

Before showing the training of the NARX network, an important configuration that is useful in training needs explanation. You can consider the output of the NARX network to be an estimate of the output of any nonlinear dynamic system that you are trying to model. The output is fed back to the input of the feedforward neural network as part of the standard NARX architecture, as shown in Fig. 2.

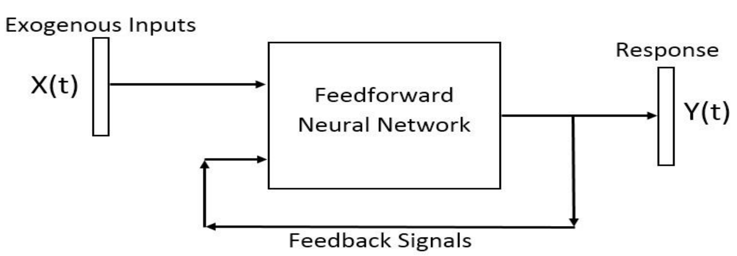

Fig.2: Total NARX network structure

Each time a neural network is trained can result in a different solution due to different initial weight and bias values and different divisions of the data into training, validation, and test sets. As a result, different neural networks trained on the same problem can give different outputs for the same input. To ensure that a neural network of good accuracy has been found, retrain several times [15, 16].

2.2 Nero-Fuzzy System

Neural networks, due to their specific structure, have a powerful capability to learn, adapt and generalize, but in many cases, they consume much time to train. On the other hand, linguistic concepts and inventive rules are utilized in fuzzy systems design [17], and these systems do not need to be trained, but the selection of the best membership functions is the main difficulty of these systems. Furthermore, the lack of a need to train eliminates the adaptation ability. Therefore, the advantages and disadvantages of neural networks and fuzzy systems have made researchers combine them. In this way, while benefiting from both of them, they will be able to overcome the shortcomings of each one through other features. The neural networks, which use fuzzy inputs and outputs, and generalized fuzzy rules to Apply addition or multiplication are known as fuzzy neural networks, and the fuzzy systems that are trained similarly to the neural networks, are called neural networks. Since neural networks (in particular, multilayer perceptron neural networks) are in turn a special case of adaptive networks, so if, in a fuzzy system, membership functions are trained in a comparative network method using past information and adapted to them, the resulting system will be an adaptive fuzzy inference system (adaptive neural fuzzy network). In fact, in an adaptive fuzzy inference system, the experimental-only selection of membership functions is generally eliminated, although the experience may reduce the training time [18].
The adaptive fuzzy inference system uses a feed-forward network to optimize the fuzzy inference system parameters and performs the defined tasks properly. The most common fuzzy inference system with this adaptation located in an adaptive network is the fuzzy Sugino-Takagi system. The Sugino model is considered in this study. An example of the Sugino fuzzy inference system model with two inputs, one output, and two membership functions for each input is shown in Fig. 3(a). We can express two functions of If-Then rules for each model [19]:

$$if\ x = A_1\ and\ y = B_1\ then\ f_{1(x,y)} = p_1 x + q_1 y + K_1 \tag{2}$$
$$if\ x = A_2\ and\ y = B_2\ then\ f_{2(x,y)} = p_2\, x + q_2\, y + K_2 \tag{3}$$

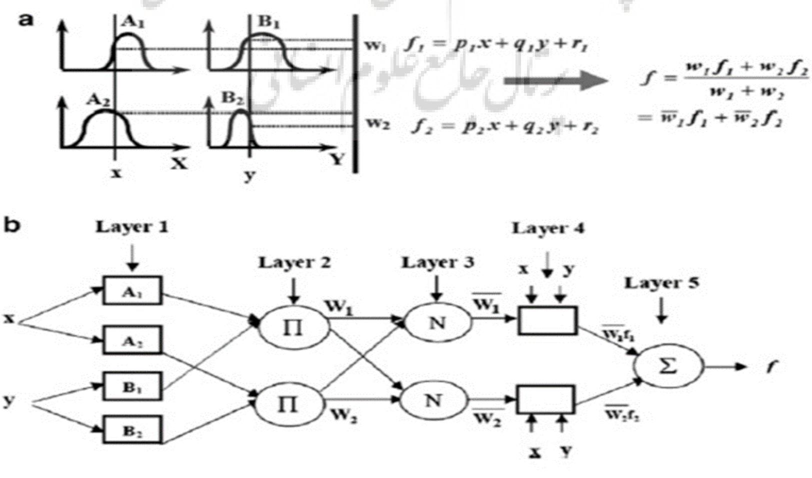

Fig.3: a) Sugeno’s fuzzy inference system model b) an adaptive fuzzy inference system architecture

And, As can be seen in Fig. 3 (b), The ANFIS model calculations are performed in five layers. Details and formulations in ANFIS are given by Jang in 1997. The ANFIS model uses a back-propagation learning algorithm or a hybrid algorithm that integrates backpropagation and least-squares techniques to generate fuzzy rules and membership function types [20].

2.3 Learning Algorithm of the NARX - ANFIS

Genetic algorithms (GA) are computerized search and optimization algorithms based on the mechanics of natural genetics and natural selection. Since the GA algorithm is usually sensitive to its parameters, the solutions presented depend largely on their parameters. Trial and error methods are used to adjust the problem parameters. Accordingly, Wibowo et al., [21] and Rajasekaran et al., [22], and Wulandhari et al., [23], can be described the parameters: 1) Randomly generate an initial population, 2) Compute the fitness of each chromosome in the current population, 3) Create new chromosome by the selection, crossover and applying mutation, 4) Substitute these new chromosomes for some bad chromosomes in the current population and 5) If the end condition is satisfactory, then stop; otherwise repeat step 2.

2.4 Concepts of the NARX-ANFIS Based on the GA Model

Proposes to combine GA with ANN to form an improved prediction model. As shown in Fig. 4.

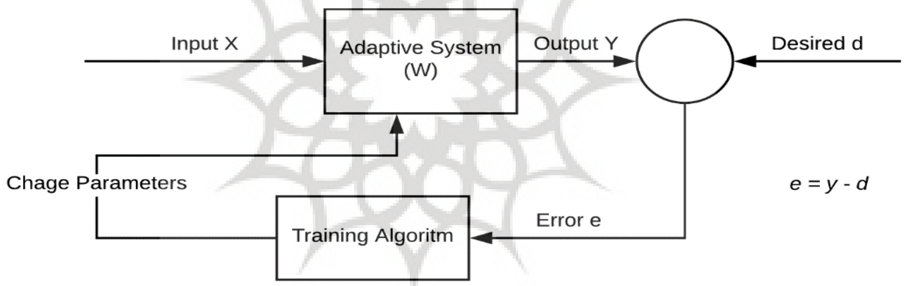

Fig.4: Training an ANN Network to GA algorithm or Error-correction learning

The major algorithm procedures are summarized in 4 steps:
Step1: Dynamic neural network is one of the new tools that are capable to analyze and simulate in nonlinear systems where relationships between system components and parameters are not well known and descriptive. Since ATM transactions are time-dependent and the demand process is not constant and varies based on the timing of the withdrawals, a dynamic neural network is appropriate to predict the cash amount in the tenth month according to studies carried out in the literature and introduction. The amount of cash in the tenth month is appropriate. In this network, the time steps ($t = 1, \ldots, 9$) are the system input. Also, to forecast the cash demand in the tenth month using NARX neural network, continuously we should determine the withdrawal rate of the 1377 units within 9 months and then calculate the system delay time.
Step2: since neural network learning is very important and very effective in training speed, the neural-fuzzy network has been used to train the network. As shown below:

$$y(t) = \overset{\text{ANFIS}}{f}\left(\overbrace{y(t-1), y(t-2), \dots, y(t-d_v), u(t-1), u(t-2), \dots, u(t-d_u)}^{\text{NARX}}\right) \quad (2)$$

Step3: To adjust the network weights and biases and Improve the designed network, a genetic algorithm is used to minimize the errors with a minimum cost function. The goal of meta-heuristic (genetic) algorithms is to search for sets of optimal weights and biases that can minimize the network's error. Doing minimizes the error rate by optimizing weights and biases. Then compute the fitness value through the fitness function and find better individuals by the selection, crossover, and mutation operations. As shown in Fig. 4. Step4: in this step, the improved NARX-ANFIS is used for prediction. First, the best NARX-ANFIS weight and parameter can be optimized locally during training. The results indicate that optimized NARX-ANFIS can achieve accurate predictions and excellent performance. It can be seen from Fig. 5 that the algorithm consists of the improved learning process of a time series neural network using a Genetic algorithm.

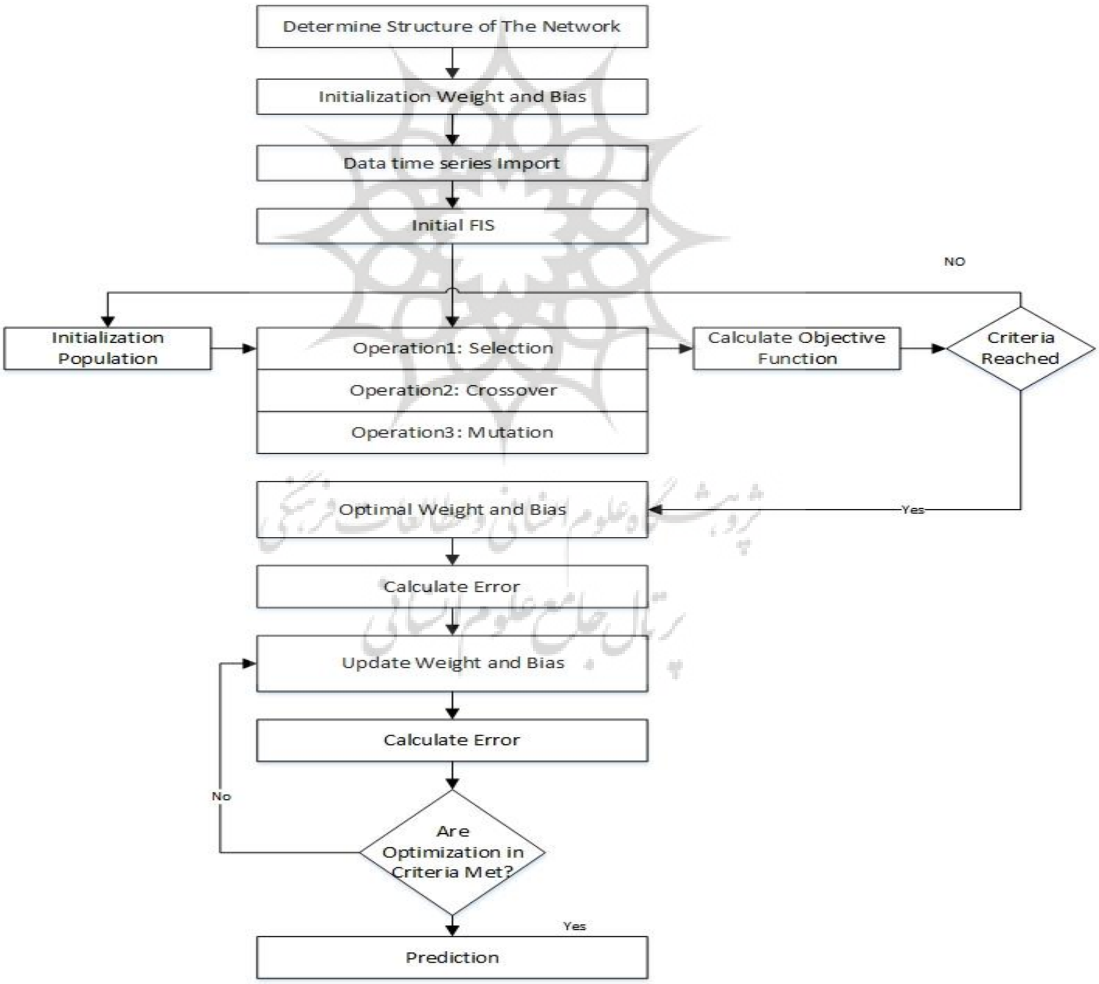

Fig.5: NARX-ANFIS optimized network using of GA in stage learning

2.5 Model Selection Methodology

Some traditional network training algorithms, such as backpropagation, use some form of gradient search and may get trapped in local optima. In contrast, EAs do not use any gradient information and are likely to avoid getting trapped in a local optimum by sampling simultaneously multiple regions of the space. A straightforward combination of evolutionary algorithms and neural networks is to use the EAs to search for weights that make the network perform as desired. In this approach, each individual in the EA is a vector with all the weights of the network. Assessing the fitness of each network involves measuring the accuracy of classification or regression on the training set, so for each fitness evaluation, the training set is passed through the network. This can be inefficient if the training set is large, but the fitness may be estimated using a sample of the training set. Although the fitness would change over different samples, EAs are known to search well using such noisy evaluations [27, 28].
The goal of meta-heuristic (genetic) algorithms is to search for sets of optimal weights and biases that can minimize the function J. The function J minimizes the error rate by optimizing weights and biases based on the GA algorithm searches for a set of network weights, W, which minimizes the objective function, J, where J is defined as the prediction error between the actual response and predicted response for a given data:

$$\mathrm{J} = \sum_{\mathrm{i}=1}^{\mathrm{N}} (\varepsilon_{\mathrm{i}})^2 \quad = \mathrm{J}(\mathrm{W}, \mathrm{B}) \tag{3}$$

That is, $W^*,\ B^* = argmin(\ J)$ Gain the best J based on optimizing weight

Where $\varepsilon = y(k) - d(k)$ (4)
$y(k)$: $actual\ value$
$d(k)$: $forecast\ value$

3 Dataset

The proposed algorithm is applied to the real cash demand bank dataset. In the time series data, each data sample, in this paper 1377 ATM device, shows a time step of different. Features related to values distinguish with time – for example, this paper surveyed withdrawals and amount demand for receiving cash in 9 months in total 12393 transactions done and selected inputs and also is designated the amount of delay on this basis. After collecting data month-to-month from 1377 ATMs in 9 months of 2016 (April to December), since the dimensions of input variables are commonly different before data is entered into the network should be normalized. That is, defined in the interval of -1 and +1 and data would become non-dimensional. Studies have shown that data normalization leads to an increase in network efficiency. The normalization function is defined as [24]:

$$X_n = \frac{2(X - X_{min})}{X_{max} - X_{min}} - 1 \tag{5}$$

In this function X is not normalized value data and x_n is normalized value data, X_{min} and $X_{max}$are respectively the smallest numerical value and the largest numerical value in the data vector.
Thus, normalized Data is done preparation step and identified outliers and missing data and modified these.

4 Empirical Results

4.1 Model Parameters Setting

The first step in creating the neural network and predicting the model is to determine the number of input neurons, in this paper, based on transactions at 1377 ATMs of Saderat Bank, in Iran, in total 12393 transactions done in the total 9 months selected inputs and also is designated the amount of delay On this basis. Since the number of epochs means the number of rounds of the training algorithm implementation, training your network on each item of the set once is an epoch, as well as a training rate, is used to stop overeating and 100% network training when the accuracy of the network for train data would reach this value, would end network training. Based on trial and error on the network, in this paper, the number epoch is 100000. We execute the results of the time series network designed based on GA using 8 input neurons, 1 output neuron, and 5 hidden neurons, during a month. Also, in the final stage, after parameter adjustments identify new weight and bias that The initial value of weight and bias earns of genes of the best individual in GA. Parameter adjustments of the genetic algorithm are still in the state of empirical adjustment, and settings need to be adjusted based on the scale of the problem and the application scenarios. After many experimental attempts for an optimal model, the training parameters of GA are selected as shown in Table 1. The method of solving the genetic algorithm in this paper is as follows, 1) Initialize population P, this Paper uses the real number coding, and the initial population takes to collect data month-to-month from 1377 ATMs in 9 months of 2016 (April to December), this is the amount of demand receiving cash in 9 months in total 12393 transactions. Each test runs the algorithm along 1377 generations on populations made of 1000 individuals. 2) Select and Compute fitness: each evaluation function, and sort them; we can choose the network by the probability value that shows in the Formula [25];

$$p_s = \frac{f_i}{\sum_{i=1}^{N} f_i} \tag{6}$$

f_i is the fitness value of individual i, it can be used to measure the sum $(\varepsilon_i)$of squared errors.

$$f(i) = \frac{1}{\varepsilon_i} \tag{7}$$

Then, it was set up with standard parameters[1]as follows: Tournaments 2, Crossover 0.4, Mutation 0.15, and Elitism 1. Evaluate new pop: Put the individuals into the new population P, and calculate the new evaluation function of the individual. Decide satisfactorily: If you find a satisfactory individual, then the end, or switch to setup parameters.

Table 1: Model parameters of genetic algorithm (GA) used for the training and testing of models.

Parameter Type	**GA Parameter Basic parameter**
Basic parameter	number of generations= 1377 population size = 1000 crossover probability = 0.4 mutation probability = 0.15
Genetic parameter	number of crossover = [50, 1] number of mutation = [20, 1]

Source: Research Results.

As shown in Table 1, the minimum number of crossover generations is 1, and the maximum is 50; the minimum number of mutation generations is 1, and the maximum is 20. According to the initial network

[1] Parameter has been chosen by the simple qualitative analysis, according to common values adopted for them, without any in-depth quantitative analysis for their optimization.

structure, network parameters need to be trained. The genetic process of selection, crossover, and mutation is performed on the whole population according to the fitness value, which is calculated as the mean square error. The generation will not stop until the number of generations reaches 50. The genes of the best individual in the current population are the initial weight and bias of the neural network. After achieving the required performance indicators, you will eventually decode the group's best individual so you can get the optimized network connection weights. The best cost function is 0.028426. It is shown in Fig. 6.

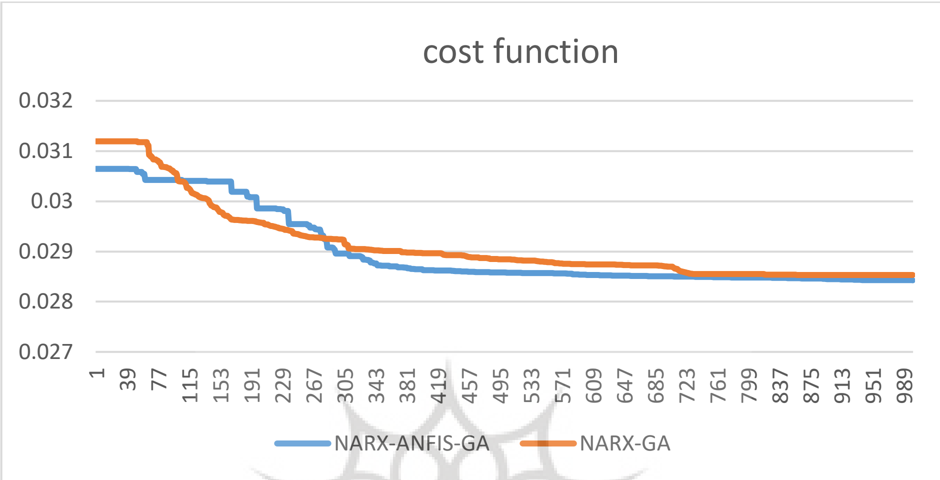

Fig.6: convergence graph of the cost function (Source: Research Results.)

4.2 Prediction Results

To evaluate the performance of the NARX-ANFIS based on GA, we must calculate the prediction error. The error function indicates how the prediction of our network is close to the target values and, therefore, what adjustment should be applied to the weight and bias in the learning algorithm in each iteration. One of the most important and most widely used types of efficiency function is the Mean Squared Error (MSE) and the root mean square error (RMSE) given that we evaluate by [24]:

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}(e_i)^2 \tag{8}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(e_i)^2} \tag{9}$$

e_i = amount of observed errors

Table 2: Performance of Designed Model

Result model of proposed model NARX-ANFIS-GA	S. V	Mean	MSE	RMSE	R
Train Data (80% data)	0.028438	0.00028082	0.00080805	0.028426	0.97408
Test Data (20% data)	0.033305	0.0001012	0.0010874	0.032975	0.9611

Source: Research Results.

The best neural network structure, a network with 5 hidden layers and 8 neurons is based on the definition of regression, and mean squared error is the best structure. In this training, MSE and RMSE, and R-squared showed for train data and test data in Table 2 based on the NARX-ANFIS-GA network.

Thus, in the next step, data was analyzed based on the proposed model, which showed in Table 3, and compared earned results. Finally, the evaluated results of the Training data are shown in Fig. 7.

Table 3: compare of Designed Model

Result model of proposed model	Train data	Error St.D	Error Mean	MSE	RMSE
NARX-ANFIS-GA	Train Data (80% data)	0.028438	0.00028082	0.00080805	0.028426
NARX-GA	Train Data (80% data)	0.086514	0.0014265	0.00748	0.086487
NARX	Train Data (80% data)	0.029251	0.0026201	0.0086161	0.029353
ANFIS-NARX	Train Data (80% data)	0.028729	1.051e-09	0.0047175	0.02872

Source: Research Results.

As a result, shows the designed model is efficient and reliable that MSE and RMSE functions are low and R-Squared is close to one. As a result, the obtained R-squared shows the effectiveness of the designed model. The network output is shown in Table 2. Figure 7 shows performance by designing a model that shows the distribution of data as normal around a regression line and also, a)R-Squared is close to one that shows the effectiveness of the designed model; b)as shown in Figure 7 cash withdrawal of 80 percent of 1377 ATMs before and after optimization; c) and normal and histogram diagrams show that the normal distribution of data is ensured; d) MSE and RMSE functions is low that this shows the efficient and reliable of the model. as shown in Table 3 the designed model plays the best performance between the hybrid models in the training stage.

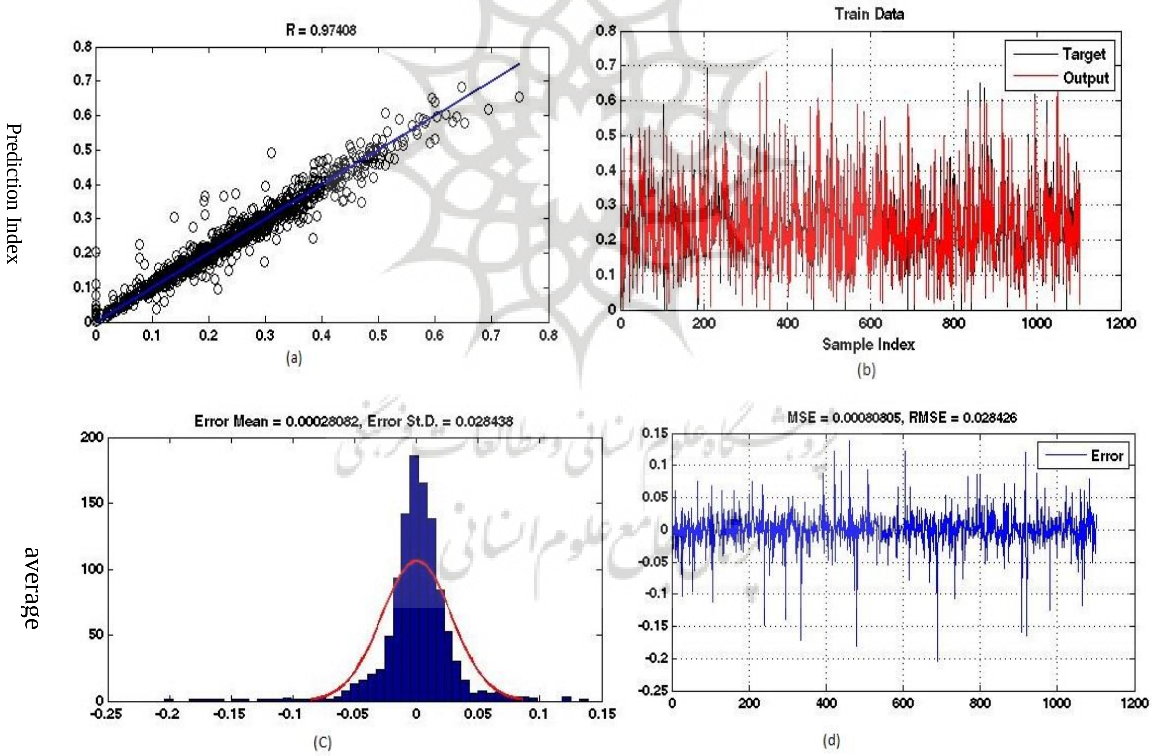

Fig.7: Train Data of Diagrams (Source: Research Results.)؛ a)R-Squared is close to one that shows the effectiveness of the designed model; b)as shown in Figure 7 cash withdrawal of 80 percent of 1377 ATMs before and after optimization; c) and normal and histogram diagrams show that the normal distribution of data is ensured; d) MSE and RMSE functions is low that this shows the efficient and reliable of the model.

As regards, the forecasted model is made for January month to survey the model accuracy, is predicted for December month and finally obtained output is compared to actual data, according to the obtained formula below, and also, the graphics diagram is shown in Fig. 7 Since the difference between forecast data and real data is very close, the accuracy of the designed model is approved.

$$\text{Error Value} = (\text{Desired Output} - \text{Actual Output}) \quad (16)$$

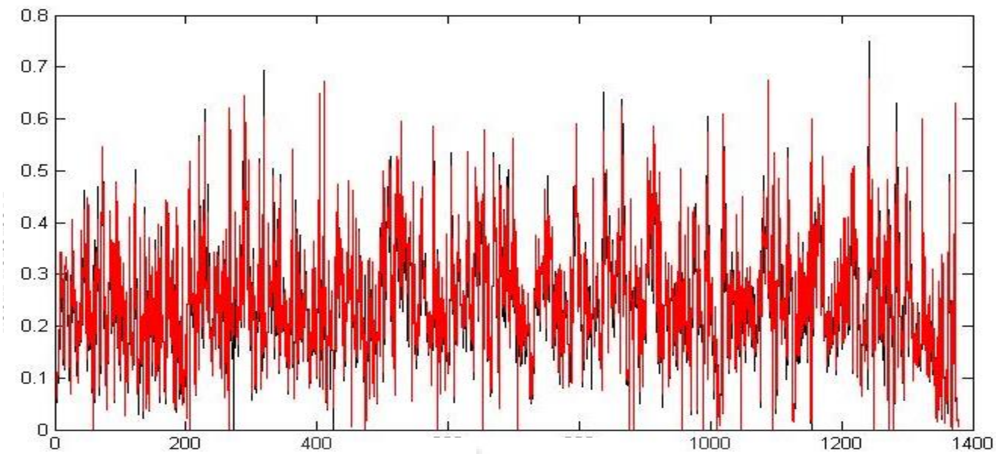

Fig.8: An absolute error between the true and estimated output (Source: Research Results.)

5 Conclusions

The traditional statistical models such as ARIMA and the exponential smoothing model, do have not a high ability for predicting in an environment of chaos. Time-series prediction using NARX-ANFIS-GA networks is a powerful tool for predicting, unlike traditional statistical methods, also it can estimate non-linear structures and there are assumptions slightly about using models for different problems. In this paper, we proposed an optimization algorithm. It is based on the routine of monthly activities, the routine duration of activities, and the number of routines of elapsed activity based on this, considered transactions of 9 months of 2016, April to December, between 1377 ATMs, the results show that changing the process of inventory cash in ATMs depends on time withdrawal, device location and its place (placing in crowded streets, urban areas, places of recreation, supermarkets and tourist locations in the area, etc.), the socioeconomic characteristics of users, holidays, daily, weekly, and monthly withdrawal times.

In this research, the mean monthly withdrawals, first or end of each month are influenced by payroll time, seasonal demand in a particular area, and the unsustainable behavior of users. The sensitivity analysis shows that the model NARX-ANFIS-GA efficiency is 97 percent. Moreover, this model can predict user behavior and According to forecasts made, it is the amount of cash loading in each device. This can lead to increase benefits and decrease costs in the bank. Also, non-optimal cash upload in ATMs will lead to imposing excessive costs on the bank and a lack of funds in the devices will discontent customers and jeopardize the brand of the bank. For this purpose, should be uploaded optimal cash that can fulfill the need of clients. This paper, using real data in the hybrid network, could increase the performance of the network and correctly identify effective parameters and make better decisions in similar places also, the designed model can be partly responded to this need. In future work, the forecasting performance might be enhanced more by using other methods such as time-series regressions and hybrid models of neural networks to improve the quality of solve model.

References

[1] ATMmarketplace.com, *ATM software trends and analysis,* Net World Alliance, Tech. Rep, 2010.

[2] Abor, J., *Technological innovation and banking in Ghana: An evaluation of Customer perception,* American Academy of Financial Management, 2004, P.416–421.

[3] Simutis, R., Dilijonas, D., Bastina, L., Friman, J., and Drobinov, P., *Optimization of cash management for ATM network,* Information Technology and Control, Kaunas, Technologija, 2007, **36**(1) A, P.117 – 121.

[4] Simutis, R., Dilijonas, D., Bastina, L., and Friman, J., *A Flexible Neural Network for ATM Cash Demand Forecasting,* in Proc. 16th Int. Conf. on Computational intelligence, Man-Machine System and Cybernetics, Spain, 2007, Dec. 14-16, P. 162-165.

[5] Dilijonas, D., and Bastina, L., *Retail Banking Optimization System Based on Multi-Agents Technology,* in Proc. 16th Int. Conf. on Computational intelligence, Man-Machine System and Cybernetics, Spain, 2007, Dec. 14-16, P. 203-208.

[6] Simutis, R., Dilijonas, D., and Bastina, L., *Cash demand forecasting for ATM using neural networks and support vector regression algorithms,* in EurOPT 2008 - Proc. of the 20th Euro Mini Conf. on Continuous Optimization and Knowledge-Based Technologies, 2008, May 20-23.

[7] darwish, Saad M., *A Methodology to Improve Cash Demand Forecasting for ATM Network*, International Journal of Computer and Electrical Engineering, 2013, **5**(4), August.

[8] Armenise, R., Birtolo, C., Sangianantoni, E., and Troiano, L., *Optimizing ATM Cash Management by Genetic Algorithms, International Journal of Computer Information Systems and Industrial Management Applications*, 2012, **4**, P. 598-608.

[9] Ekinci, Y., Lu, J-C., Duman, E., *Optimization of ATM Cash Replenishment with Group- Demand Forecasts, Expert Systems with Applications*, 2014, Doi: 10.1016/j.eswa.2014.12.011.

[10] Venkatesh, k., Ravi, V., Prinzie, A., Poel, D., *Cash demand forecasting in ATMs by clustering and neural networks,* European Journal of Operational Research, 2014, **232**, P.383–392. Doi:10.1016/j.ejor.2013.07.027.

[11] Arora N., Kumar, J., Saini, R., *Approximating Methodology: Managing Cash in Automated Teller Machines using Fuzzy ARTMAP Network,* International Journal of Enhanced Research in Science Technology & Engineering, ISSN: 2319-7463, February, 2014, **3**(2), P. 318-326.

[12] Bhandari, R., Gill, J., *An artificial intelligence ATM forecasting system for hybrid neural networks,* International Journal of Computer Applications, 2016, **133**(3), P.13-16. Doi: 10.1.1.734.7389.

[13] BİLİR, C., DÖŞEYEN, A., *Optimization of ATM and Branch Cash Operations Using an Integrated Cash Requirement Forecasting and Cash Optimization Model*, Business & Management Studies: An International Journal,), 2018, **6**(1). Doi:10.15295/bmij.v6i1.219.

[14] Vennila, A., and Rathnaraj, S.N., *Impact on Customer Perception towards ATM Services Provided by the Banks Today: A Conceptual Study*, International Journal of Scientific Research and Management, 2018, **6**(01). Doi: 10.18535/ijsrm/v6i1.em06.

[15] Demuth, H., Beale, M., Hagan, M, *Neural Network Toolbox™ 6-User's Guide, Mathworks Inc*, 2009.

[16] Hagan, M.T., De Jesus, O., and Schultz, R., *Training Recurrent Networks for Filtering and Control, Chapter 12 in Recurrent Neural Networks: Design and Applications, L. Medsker and L.C. Jain, Eds.,* CRC Press, 1999, P. 311–340.

[17] Gross, G., and Galiana, F.D., *short term Load Forecasting, Proc*. Dec. 1987, IEEE, **75**(12), P. 1558-1573.

[18] Arora, N. and Saini, J.R., *Estimation and approximation using neuro-fuzzy systems, International Journal of Intelligent Systems and Applications,* 2016, **8**(6), P.9. Doi: 10.5815/ijisa.2016.06.02.

[19] Sugeno, M., and Kang, G.T, *Structure identification of fuzzy model, Fuzzy Sets Systems*, 1988, **28,** P.15-33.

[20] Jang, J.S.R., Sun, C.T. and Mizutani, E, *Neuro-fuzzy and soft computing, Prentice-Hall, New Jersey*, 1997.

[21] Wibowo, A., Desa, M.I., *Kernel Based Regression and Genetic algorithms for Estimating Cutting Conditions of Surface Roughness in End Milling Machining Process,* Expert System with Applications, Elsevier, 2012, **39**(14), P.11634-11641. Doi:10.1016/j.eswa.2012.04.004.

[22] Wulandhari, L.A., Wibowo, A., and Desa M.I., *Condition Diagnosis of Multiple Bearings Using Adaptive Probabilistic Based Genetic Algorithms and Back Propagation Neural Networks*, Neural Computing and Applications, Springer, 2015. Doi:10.1007/s00521-014-1698-6.

[23] Rajasekaran, S., G.A. Vijayalakshmi, *Neural Networks, Fuzzy Logic, and Genetic Algorithms: Synthesis and Applications,* Prentice Hall of India Private Limited, New Delhi, 2007.

[24] Haykin, S. *Neural Networks a Comprehensive Foundation, Second edition*, Prentice Hall International Inc, 1994.

[25] De j., Kenneth, A., Spears, W. M., Gordon, Diana F., *using genetic algorithms for concept learning.* In: Genetic Algorithms for Machine Learning. Springer US, 1993, P. 5-32.

[26] Rashidi Baghi, M., Taherinia, M., *Prediction the Return Fluctuations with Artificial Neural Networks' Approach,* Advances in Mathematical Finance and Applications, 2019, **4**(2), P.103-114.
Doi: 10.22034/amfa.2019.580643.1149.

[27] Holland, J., *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Application to Biology,* University of Michigan Press: Ann Arbor, MI, USA, 1975.

[28] Hassan, R., Cohanim, B., De Weck, O., Venter, G., *A comparison of particle swarm optimization and the genetic algorithm. In Proceedings of the 46th AIAA/ASME/ASCE/AHS/ASC structures,* structural dynamics and materials conference, Austin, TX, USA, P.18–21 April 2005, P. 1897.

[29] Farshadfar, Z., Prokopczuk, M., *Improving Stock Return Forecasting by Deep Learning Algorithm. Advances in Mathematical Finance and Applications,* 2019, **4**(3), P.1-13. Doi: 10.22034/amfa.2019.584494.1173.