Adv. Math. Fin. App., 2023, 8(3), P. 1007-1022

Advances in Mathematical Finance & Applications

www.amfa.iau-arak.ac.ir

Print ISSN: 2538-5569

Online ISSN: 2645-4610

Doi: 10.22034/AMFA.2021.1910900.1493

Applied-Research Paper

Modelling Optimal Predicting Future Cash Flows Using New Data Mining Methods (A Combination of Artificial Intelligence Algorithms)

Bahman Talebi[a], Rasoul Abdi[a,*], Zohreh Hajiha[b], Nader Rezaei[a]

[a] *Department of Accounting, Bonab Branch, Islamic Azad University, Bonab, Iran*

[b] *Department of Accounting, East Tehran Branch, Islamic Azad University, Tehran Iran.*

ARTICLE INFO

Article history:

Received 2020-10-05

Accepted 2021-01-24

Keywords:

Future Cash Flows

Neural Network Model

Genetic Algorithm

Particle swarm Algorithm

ABSTRACT

The purpose of this study was to present an optimal model for predicting future cash flows using a combination of artificial neural network with genetic algorithms (ANN+GA) and particle swarm algorithms (ANN+PSO). Due to the nonlinear relationship among accounting information, the study aimed to utilize these artificial intelligence algorithms to forecast future cash flows. The variables considered for prediction were accruals components and operating cash flows. The analysis was conducted using data from 137 companies listed on the Tehran Stock Exchange during the period of 2009-2017. The findings of this study demonstrated that both neural network models, optimized with genetic algorithms and particle swarm algorithms, performed well in predicting future cash flows when all the variables presented in this study (a total of 15 predictor variables) were included. However, the results also indicated that the neural network optimized with particle swarm algorithm (ANN+PSO) exhibited a lower error coefficient, indicating better efficiency and higher prediction accuracy compared to the neural network optimized with genetic algorithms (ANN+GA).

1 Introduction

Prediction plays a crucial role in the decision-making process, especially in economic activities where it holds significant importance. The prediction of cash flow, being a vital economic resource, is essential for making informed economic decisions. Balancing current cash and cash requirements is a critical factor for the economic well-being and sustainability of any business entity. Future cash flows hold paramount importance in financial decision-making, stock exchange valuation models, and investment plan evaluations [4]. Both cash flow and profit serve as important indicators for measuring the performance of a business entity. While various models for predicting future cash flows have been reported in available resources and studies, they have not effectively captured the dynamic and nonlinear properties of cash flow modeling [12]. Previous studies have explored different models, primarily regression models, such as those employed by Sarraf et al. [16], Sarraf and Sagafi [15], Mahdavi and Saberi [10],

* *Rasoul Abdi. Tel.: +098-9144205650*

E-mail address: abdi_rasool@yahoo.com

to predict future cash flows. Variations in these studies encompassed the models utilized, the variables presented, linear and nonlinear methods, and the methods of analysis. Therefore, to achieve more accurate predictions, it is necessary to incorporate data mining methods. Consequently, the problem of cash flow prediction remains unresolved and requires further investigation based on previous research. Moreover, due to the nonlinear relationships among accounting information, exploring the nonlinear and complex factors influencing future cash flow prediction and developing new approaches and methods with minimal deviation and error is imperative. Thus, this study, titled "Modelling Optimal Predicting Future Cash Flows Using New Data Mining Methods," aims to provide an optimal model for predicting future cash flows by combining artificial intelligence algorithms. The study considers multiple algorithm combinations, including the neural-genetic hybrid model (ANN+GA) and the combined neural network-particle swarm model (ANN+PSO). The ultimate goal is to develop a powerful model with higher efficiency, explanatory power, and overall performance by examining the explanatory power of the variables.Finally, by examining the explanatory power of the variables, this study introduces a strong model with greater efficiency, explanatory power and high performance.

2 Theoretical Bases & Background

Liquidity and profitability are two crucial issues that are of particular interest to financial scholars and managers. Indeed, some consider liquidity to be more important and believe that if a company is not profitable, it is sick, but if it does not have the liquidity it is dying. In other words, it is possible for a company not to be profitable but survive; however, it cannot survive without liquidity. Cash management is one of the main issues of corporate financial management. Liquidity demonstrates the ability of a corporate to engage in short-term commitments. In other words, liquidity is the relationship between the cash that will be available to the company in the short run and the cash that the company will need. The main purpose of liquidity management is to obtain a desirable liquidity so that the company is not exposed to excess liquidity or an abnormal liquidity shortage. [1] In a world where there are tremendous changes in the economy every moment, predicting future events will be a key factor in capturing profit opportunities, and while traditional techniques such as regression have proved to be ineffective in some cases, many people are interested in predicting future events more accurately. [17] Scholars in the last decade of the twentieth century generally believed in the principle that the assumption of rational investment, which is an indispensable principle in modern financial investment and one of the key assumptions in an efficient market, was not a real principle considering the complex factors affecting stocks markets.

They have concluded that a capital market does not have a clear order and using complex mathematics in nonlinear and dynamic systems can create models that abrogate past theories. As a result, one of the considerations that financial data prediction needs to consider is the non-linearity of financial data. Based on the results of some studies, when data are of financial, seasonal and nonlinear nature, nonleaner model prediction such as neural networks will be more efficient. Many studies have proved that neural networks are more efficient compared with other traditional and linear models because neural networks, as opposed to linear models, reflect nonlinear effects and complex interactions among variables. [7] Recent studies have emphasized the non-linear nature of financial information. Thus, the large number of methods for prediction and the unknown factors affecting the return of wealth have caused uncertainty among investors and creditors. That is why they are trying to come up with predictive methods making their estimates closer to reality and more accurate. [12] Consequently, several models have been developed and used to predict cash flows, some of which are based on cash accounting information and some are based on accrual accounting information. Due to the increasing growth of studies related to cash flow prediction in Iran, innovation in designing and modelling for cash flow prediction seems necessary. Most research conducted in the area of linear regression models were based on profit and accruals, while few nonlinear models with new methods have been used. As a result, this study, on the one hand, uses multiple algorithmic combinations of artificial intelligence models to predict future cash flows, and on the other hand, new variables derived from accrual models of research (Farshadfar and Monem[6]; Larson et al., [9] and Peng, [12]) to introduce a strong efficient model with innovative greater explanatory power.

- **Genetic Algorithm**

Genetic model is one of the search algorithms that randomly finds the solutions for the problems. This algorithm falls into the category of trial and error algorithms. This algorithm is used to solve complex optimization problems for which specific rules can be considered. [8] Genetic algorithm-based models require four basic elements for a research:

•Primary population: An initial set of members (chromosomes) that are usually coded as sequences of genes (bits) and provide answers to the problem.

•Evaluation Function: A method for measuring the fitness of each member (answer).

•Selection: The process of selecting the right members for reproduction and re-combination.

•Genetic operators: Used to produce new members and to evolve gradually.

Before a genetic algorithm can be run, it must firstly find the appropriate coding (or representation) for the problem in question. The most common way of displaying chromosomes in genetic algorithm is in form of binary strings. Each decision variable is formed in binary shape and then the chromosomes are produced by getting together these variables. Although this is the most widely used method of coding, other methods such as representation with real numbers are expanding. A fitness function must also be devised to assign value to any coded solutions. During the run, the parents are selected for reproduction and combined together using mutation and combination operators to produce new offspring. This process is repeated several times to produce the next generation of population. Then, the population is checked and the above process will be terminated if convergence criteria are met. The very high power of the genetic algorithm in selecting the optimal points always assures the user that the proposed optimal points will be the optimal points for the problem. What makes the genetic algorithm so exceptional is searching method of this model. The major difference between this method and other available search methods, such as artificial neural networks, is that the genetic algorithm initiates the search process simultaneously from several points in the response space, while other optimization methods are always at risk of losing better answers due to their activity initiation from one point. [11]

- **Particle Swarm Algorithm**

This algorithm is one of nature-inspired optimization algorithms and dynamic computation techniques based on the initial population which simulates the social behaviour of a group of birds in finding food. This method was first applied in 1995 as an optimization method by Eberhart and Kennedy. Since using this algorithm requires only some basic computational operators, it is simple to implement and

economically cost effective. The particle swarm algorithm begins with a population of random solutions of the problem called particles. The primary algorithm is inspired by simulating the social behaviour of a group of birds finding food. The bird strives to find food through social interaction with other birds of the group. Each bird leaves its current location according to its previous location or by contacting a member of the group that has found the best food. In this way, each particle tries to move toward the final solution by adjusting its path and moving to the best personal and collective experience of the group. In this algorithm, each particle has a particular velocity. These particles fly in the search space at a rate which is dynamically tuned to their past behaviour. As a result, particles tend to move to the best and most suitable search area during the search phase.
Mass movement of particles is an efficient technique for solving optimization problems that performs on the basis of probability laws and population. The beginning of the flight pattern of birds is such that a group of particles (solutions) are randomly created and try to find the optimal solution by updating generations. At each step, each particle is updated using its current location and two best values. The first is the best situation the particle has already achieved. This position is known and saved as pbest. The next best value used by the algorithm is the best situation already achieved by the particle population. This position is represented by gbest. All particles know their best location, the best particle location of the group as well as the value of the target function corresponding to each location. At each step of the algorithm, behaviour of particle is selected randomly in a combination of three possibilities of following the path of the particle itself, returning to its previous best location and going to the best previous or current location of the group particles. The particle behaviour is formulated in this algorithm as follows. [2]. Every researcher has employed a particular approach, method and perspective to investigate the cash flow prediction models; however, considering studies related to the topic of this study " Modelling Optimal Predicting Future Cash Flows", there have been no comprehensive and comparative studies using new methods of analysis. Therefore, a number of researches related to the topic of this study abroad and inside the country are briefly mentioned.

Talebi et al. show that both regression and neural network models within proposed variables in the present study have the capability of predicting future cash flows. Also, results of neural network models' processes show that a structure with 16 hidden neurons is the best model to predict future cash flows and this proposal neural network model compared with regression model in predicting future cash flows has a better and accurate function. [19] The empirical results of the study Koskinen [13], show that working capital management has a minor positive impact on the overall explanatory power of the model. However, the differences between the groups are not significant and thus, it cannot be concluded that working capital management would impact the prediction of one-period ahead cash flows. Overall, the results suggest that the main role of accruals in adjusting the recognition of cash flows is not dependent on working capital management. The empirical results of the study Senan [18], show that disaggregation of earnings into accruals has greater predictive ability of Forecasting Future Cash Flows (FFCF), while disaggregated earnings into other accrual components has decreased the predictive power. The results of the study are in contrast to the assertion of Financial Accounting Standards Board (FASB) that earnings have the better ability than CFs in FFCF. Peng [12], conducted a study entitled “Designing Dynamic and Nonlinear Models in Cash Flow Prediction” and focused on the problem of modeling and designing new models that can bridge the gap between simple models and complex models such as cash flow prediction models. The results show that simpler models such as the random step model are closer to the expected future cash flow of the market because it results in a better fit for market share prices using the new discount model. Furthermore, the newly developed valuation models of the study can worth investment and there are dynamic and nonlinear features in cash flow prediction. Saghafi et al. [15], in a study entitled “The Application of Artificial Neural Network in Predicting Future Cash Flows”, showed that two structures with 8 and 11 hidden nodes were the best model for cash flow prediction. Sarraf et al. [17], conducted a study entitled “Linear and Non Linear regression models for cash flow forecasting”. Their findings indicate that the accrual regression model predicts cash flow mor efficient than other examined models and among the firm characteristics the highest correlation was found between firm volatility and firm size with accrual regression models. Saghafi and Sarraf [16], in a study titled “A Model for Cash Flow forecasting in Iranian Companies” has shown that random walk model can predict operational cast flow better than reverse accrual model. However, according to the results achieved from the companies in which the government influences their management showed that accrual model was more suitable to the future cash flow. Rozbaksh et al. [14] , in a study titled “forecasting cash flows operations using artificial neural networks in Tehran Stock Exchange” have shown that artificial neural networks have high capability in predicting the future cash flow because the two hidden layers with 15 and 30 neurons in each layer can predict the cash flows with 99.2 % accuracy. Mahdavi and Saberi [10], in a study entitled “Determining the Optimal Model for the Prediction of Operating Cash Flow of Companies Listed in Tehran Stock Exchange” showed that separating the profit into cash components can increasingly predict future cash flow.

3 Research Hypotheses, Methodology of the Research and Statistical Data

3.1 Research Hypotheses

Predicting cash flows and their fluctuations as an economic event has long been the focus of researchers, investors, managers, financial analysts, and creditors. This is due to the use of cash flows in stock valuation models, payables assessment (dividends, interest and other liabilities), risk assessment, performance assessment of management and management experience, evaluation of management's choice of accounting methods and the use of cash flows for making decisions relevant to decision making models. Therefore, if cash flows can be properly predicted, a substantial portion of information needs related to the cash flow will be met. Consequently, the research hypotheses of this study are as follows:

Hypotheses 1: optimized neural network model with genetic algorithm (ANN+GA) is a suitable model to forecast future cash flow.

Hypotheses 2: optimized neural network model with particle swarm algorithm (ANN+PSO) is a suitable model to forecast future cash flow.

Hypotheses 3: Combined neural network with particle swarm model(ANN+PSO) compared with combined neural network with genetic model(ANN+GA) has lower error coefficient (better efficiency and prediction accuracy) in predicting future cash flows.

3.2 Research Methodology

This study aimed to provide a model for predicting future cash flows of listed companies in Tehran Stock Exchange by combining multiple algorithms of artificial intelligence models in Iranian capital market. Therefore, this study was statistically modeled one and methodologically a descriptive (quasi-experimental) correlation study in which the relationship between variables was analyzed according to the research purpose. In order to collect and calculate the information needed to analyze the relationship between the data, mainly the Rahavard Novin software and databases were used. Artificial intelligence models were implemented using Matlab Comprehensive Programming Environment.

3.3 Data

In this study, all accepted companies in "Tehran Stock Exchange" are statistical population. To choose samples from companies in statistical population; the companies that have the following requirements had been selected.

1- The sample companies should be accepted in Tehran Stock Exchange from the beginning of 2009.

2- The companies' financial statements or other required data should be available from 2009 to 2017.

3-For comparison purposes, those companies whose financial year did not end in March were excluded.

4- The investment companies and other financial intermediaries were excluded due to their different functional characteristics.

Finally, according to the above-mentioned requirements, among all accepted companies in "Tehran Stock Exchange" 137 companies (959 Year – Company) were selected as a sample for this study. In the present study, the combined models: particle swarm optimization Algorithm(ANN+POS) and genetic optimization Algorithm (ANN+GA) was used to improve the accuracy of predictions and overcome the limitations of multilayer perceptron neural network models (limiting the number of input variables).

3.4 Variables and Research Model

In this research, based on theoretical and research background, Farshidfar and Monem, [6]; Larson et al., [9] and Peng, [12] variables and accrual model have been used to propose an optimal model for cash flow in which dependent variable is future cash flows and independent variables are other proposed variables in Table 1. The final model of the research and its components are as follows in equations 1 and 2 bellow.

$$\begin{aligned} \boldsymbol{CF_{it+j}} = \boldsymbol{\gamma_0} &+ \boldsymbol{\gamma_1 CF - Cerr_t} + \boldsymbol{\gamma_2 CF - Cpaid_t} + \boldsymbol{\gamma_3 CF - NCerr_t} + \boldsymbol{\gamma_4 CF - NCpaid_t} \\ &+ \boldsymbol{\gamma_5\,\Delta COA_{it}} + \boldsymbol{\gamma_6 \Delta COL_{it}} + \boldsymbol{\gamma_7 \Delta NCOA_{it}} + \boldsymbol{\gamma_8 \Delta NCOL_{it}} + \boldsymbol{\gamma_9 \Delta FINA_{it}} \\ &+ \boldsymbol{\gamma_{10} \Delta FINL_{it}} + \boldsymbol{\gamma_{11} \Delta INV_{it}} + \boldsymbol{\gamma_{12} \Delta Ap_{it}} + \boldsymbol{\gamma_{13} \Delta AR_{it}} \\ &+ \boldsymbol{\gamma_{14} DEP_{it}}\ \&\ \boldsymbol{AMORT_{it}} + \boldsymbol{\gamma_{15} OTHER_{it}} + \boldsymbol{\varepsilon_{it}} \end{aligned} \tag{1}$$

$$\begin{aligned} &\boldsymbol{TACi,t = \Delta WCi,t + \Delta NCOi,t + \Delta FINi,t} \\ &\boldsymbol{CFO\ it = CF - Cerr_{it} + CF - Cpaid_{it} + CF - NCerr_{it} + CF - NCpaid_{it}} \end{aligned} \tag{2}$$

Table 1: Research Variables, Operational Definitions and Their Measuring Methods

Variable	Symbol	How to Measure
Dependent and Output variable		
future cash flows	CF_{it+j}	(the firm's net cash flow from operations of the next year)
Independent and input variables		
total accruals	TACit	Sum accruals (current operating + non-current operating + financing)
changes in working capital accruals	ΔWCit	Changes in working capital accruals during the year
changes in non-current operating accruals	ΔNCOit	Changes in non- current operational accruals during the year
changes in financing accruals	ΔFINit	changes in financing accruals during the year

Table 1: Research Variables, Operational Definitions and Their Measuring Methods

Variable	Symbol	How to Measure
cash flows from operations	CFO_{it}	the firm's net cash flow from operations, as disclosed in the statement of cash flows
changes in current operating assets accruals	ΔCOA_{it}	Changes in (total current assets – cash - current investments) during the year
changes in current operating liabilities accruals	ΔCOL_{it}	Changes in (total current liabilities - short-term facilities) during the year
changes in non-current operating assets accruals (investment)	$\Delta NCOA_{it}$	Changes in (total non-current assets - long-term investments) during the year
changes in non-current operating liabilities accruals (investment)	$\Delta NCOL_{it}$	changes(total non-current liabilities - long-term facilities) during the year
Changes in financing assets accruals	$\Delta FINA_{it}$	Changes in investments (short-term + long-term) during the year
Changes in financing liabilities accruals	$\Delta FINL_{it}$	Changes in receivable facilities (short-term + long-term) during the year
cash flows received from sales of goods and providing services	CF $Cerr_{it}$	cash flows received from customers (Net sales -increase/+ decrease of Net received accounts+ increase/-decrease Perreceived sales –cost of fuei claims)
cash flows Payments For the purchases of goods and services	CF $Cpaid_{it}$	cash flows Payments For purchases (Net purchases +increase/-decrease of inventories -increase/+ decrease Net payment accounts increase/-decrease of Prepayment of goods
other received cash flows (except for sales of goods and providing services)	CF $NCerr_{it}$	(related Revenue - increase/+decrease of related receivables Revenue +increase /-decrease related Per-received revenues)
Other Payments cash flows (except for purchase of goods and services)	CF $NCpaid_{it}$	cash flows Payments For costs (Total costs items with the exception of non- cash costs , interest and tax - increase /+decrease of payable costs)
Changes in inventories	ΔINV_{it}	Changes in inventory
Changes of payable accounts	ΔAP_{it}	Changes in account payable
Changes of receivable accounts	ΔAR_{it}	Changes in account receivable
tangible and intangible assets depreciation cost	DEP_{it} & $AMORT_{it}$	Depreciation and amortization
Other accruals	$OTHER_{it}$	Other accruals, calculated as earnings before interest, tax, depreciation and amortization. (EBITDA) – (CF + ΔAR + ΔINV – ΔAP – DA).
source: Pang, [12], Farshadfar and Monem [6], Larson et al [9], and standard No 2 of IRAN accounting.		

3.5 Implementation of Neural Network

At this point, using the training data, we implement the aforementioned artificial intelligence algorithms. It should be noted that different parameters of each algorithm must be adjusted in implementing artificial intelligence algorithms. Selecting the values of these parameters is often experimental and trial and error method, so different combinations must be examined to find the right combination of parameters. Although during the setup phase we were looking for the "best combination of parameters", because of the random nature of this process, it cannot be assured that it is "the best" because there are several other combinations with similar accuracy which can be used. Therefore, we are looking for a "proper combination" of parameters with a desirable performance. The neural network architecture of the research is presented in Fig 1.

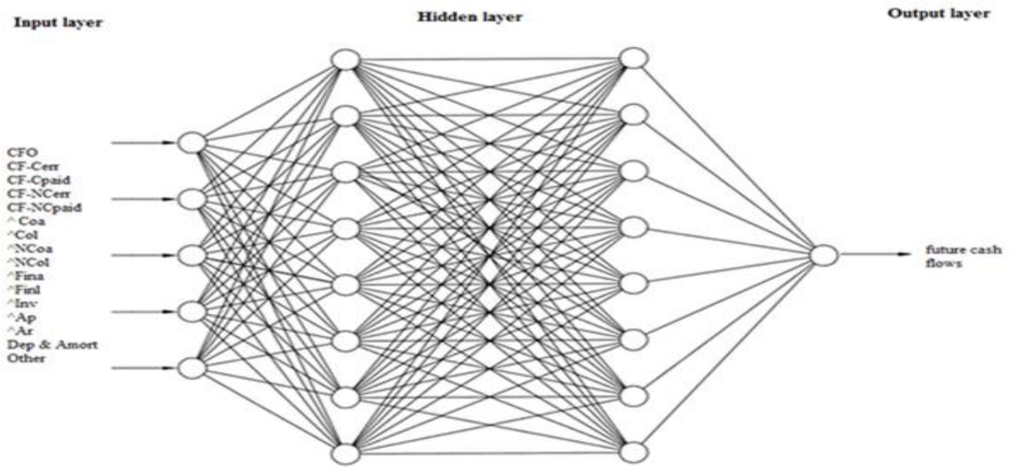

Fig 1: The Neural Network Architecture Research

3.5.1 Implementation of Optimized Neural Network with Genetic Algorithm

In this study, in line with Asadi and Naghdi's [3] study, the steps of combining and developing a neural network fusion model with genetic algorithm for prediction were implemented as follows:

Stage 1: The number of populations in each generation and the number of generations in the first stage was determined and an initial random population was generated.

Stage 2: At this stage the artificial neural network shape was determined using the value of genes in each created population.

Step 3: Designed network was trained using normalized input data. After the network was trained, the network validation and training steps were performed in this step.

Step 4: After the prediction was performed using the designed grid, the mean squared criterion would be calculated. Calculating this criterion determined the objective function of the problem, which was the minimization of the mean squares error in this study.

Step 5: In order to generate the next generation, operators such as genetic and evolutionary ones i.e. combination and gene mutation as well as the roulette cycle were used to select the next generation in the genetic algorithm. At this stage, elitism was also used to convey some of the best current population to the next generation.

Step 6: At this stage the newly created population was replaced with the previous population to create a new generation. At this point, the value of 1 was added to the generation number. The above steps were repeated until the generation number reached its maximum value. [3] A summary of the main parameters and fitting stages of the neural network optimized by the genetic model in this study is presented in Table 2.

Table 2: Summary of the Main Parameters and Fitting Stage Information ANN+GA

parameter	description
the type of neural network	Neural Network model Optimized with Genetic model
Data Normalization	Mapminmax function
Random data separation	Dividerand function

Table 2: Summary of the Main Parameters and Fitting Stage Information ANN+GA

parameter	description
The number of neurons in the input, hidden, and output layers	12, 12-5 and 1 neuron
Activation function and training function	The hyperbolic tangent function and the Levenberg Margo-det function
Number of generations	100
The probability (rate) of a mutation and an in-tersection	0.50 and 0.45
Initial population	Randomly an answerable
Selection process	The use of the roulette wheel suitable to the fit of each chromosome whose function is the objective function of the ideal programming model.
Stopping condition	Maximum generation: 100 generations
Number of chromosomes in each generation	15 chromosomes
Elitism percent	20%
numbers of collection	Training 65% , accreditation 10% and test 25%
evaluation criteria and choosing the best struc-ture to experimental results and test	Mean squared error (MSE) Root-mean-square error(RMSE) Normalized Mean Squared Standard Error (NMSE) Mean Absolute Error(MAE) Mean Absolute Percentage Error(MAPE) correlation coefficient(R2)
Choosing the most efficient variables to fore-casting	Sensitivity Analyze

Table 3: Information of the Main Parameters and Fitting Stages (ANN+PSO).

Parameter	Description
Network type	Optimized neural network with particle swarm pattern
Normalization of data	mapminmax Function
Removing random data	Dividerand function
The number of Neurons of hidden, input and output layers	12 neuron, 4-12 neuron and 1 neuron
Size of population	20
Weight and inertia coefficient	{0- 1} and 0.5
Number of particles	85
Cognitive and collective learning coefficient	1.5 and 2
The number of repetitions(Epochs)	500
numbers of collection	Training 65% , accreditation 10% and test 25%
evaluation criteria and choosing the best structure to experimental results and test	Mean squared error (MSE) Root-mean-square error(RMSE) Normalized Mean Squared Standard Error (NMSE) Mean Absolute Error(MAE) Mean Absolute Percentage Error(MAPE) correlation coefficient(R2)
Choosing the most efficient variables to forecasting	Sensitivity Analyze

3.5.2 Implementation of Neural Network Optimized with Particle Swarm Algorithm

In recent years, due to the limitations of mathematical methods, lots of studies have been conducted on the use of evolutionary algorithms for optimization. One of the most effective methods is the particle swarm algorithm. According to some researchers, this algorithm works faster compared with the genetic algorithm. It is inspired by the simultaneous flight of birds which is composed from a series of simple relationships. In this study, in line with the steps of combining and developing the integrated neural network model with particle aggregation algorithm (bird flight) were implemented for prediction as follows:

1.Establishing the initial population and its evaluation
2. Determining the best personal memories and the best collective memories
3. Updating speed and position and evaluating new responses
4. If stop condition is not met, return to 2
5. End. [3]

A summary of the main parameters and fitting steps of the neural network optimized with the research particle swarm model is displayed in Table 3.

4 Research Findings

4.1 Descriptive Statistics

Table 4: Descriptive Statistics of Research Variables

Variable/Statistic Index	Mean	Median	Minimum	Maximum	Standard Deviation
CF	0,169	0,138	-1,007	1,356	0,206
CFO	0,143	0,123	-0,723	1,147	0,165
COA	0,451	0,562	-0,011	2,213	0,151
COL	0,154	0,069	-0,854	1,054	0,214
NCOA	0,129	0,513	-0,754	3,333	0,201
NCOL	0,201	0,134	-0,087	2,041	0,176
FINA	0,098	0,017	-0,314	0,821	0,037
FINL	0,112	0,035	-0,425	1,678	0,142
CF_Cerr	1,070	0,867	-0,023	11,017	0,858
CF_Cpaid	0,992	0,776	0,002	10,336	0,860
CF_NCerr	0,117	0,057	0,000	1,812	0,170
CF _NCpaid	0,048	0,034	0,001	0,622	0,048
INV	0,065	0,062	-1,512	1,563	0,141
Ap	0,041	0,026	-0,289	0,851	0,094
AR	0,065	0,031	-0,721	2,021	0,174
DEP & AMORT	0,021	0,017	0,013	0,126	0,017
OTHER	-0,048	-0,037	-3,321	1,863	0,302
Number of Observations: 959					
source: Calculations of the research					

Descriptive statistics (regression and central tendency indexes) of the research variables is proposed in the Table 4. The main central tendency index is the “mean” which shows the balance point and distribution mean center. The mean index of all variables was positive and the highest index was related to cash flows received from sales of goods and providing services (CF-Cerr) with 1.070 around which most of the data had been concentrated. Generally, the regression parameters are criterion to determine the regression value from each other or their regression value with respect to the mean. One of the most important regression parameters is standard deviation. The amount of this parameter for paid cash flows variable to purchase goods and services (CF-CAPID) is 0.86 and for the variable of tangible and intangible assets depreciation it is 0.017 that shows these two variables have the lowest and the highest regression value among the research variables.

4.2 Fitting Results of Artificial Intelligence Models

The error convergence process of MSE for future cash output in neural network combination with the genetic model is shown in diagram 1. As can be seen, the neural network optimized with genetic algorithm has rapidly converged and from the 53rd generation on, the value of the target function has remained constant, indicating the power of the algorithm for optimization.

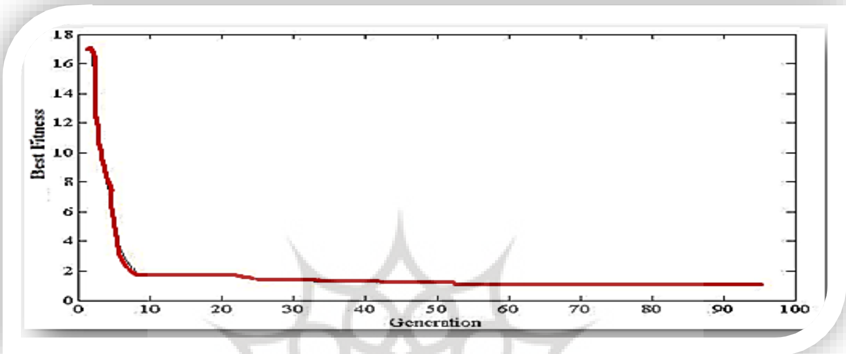

Diagram 1: Neural Network Convergence Optimized with Genetic Algorithm Based on Generation Number.

The convergence diagram of the particle swarm optimization algorithm based on the number of repetitions is presented in diagram 2. The horizontal axis indicates the number of target function evaluations during optimization. As can be seen, the particle optimization algorithm is rapidly converging and from the evaluation number of 150 on, the target function value remains constant, indicating the power of the algorithm for optimization.

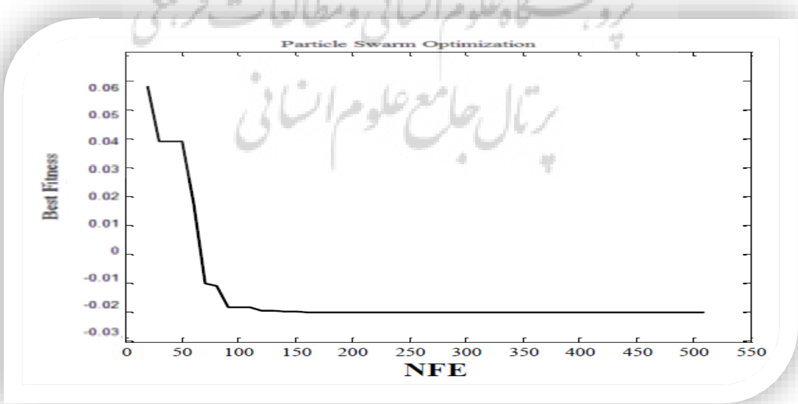

Diagram 2: Convergence diagram of neural network optimized with bird swarm algorithm based on the number of repetitions.

In this study, due to the limitation of the paper space, only the best structural results of both neural network models optimized with genetic algorithm (with 6 hidden neurons) and particle swarm algorithm (with 8 hidden neurons) which are the best predictors of future cash flows are presented in Table 5. Based on the results for the neural network model optimized with the particle swarm algorithm in the structure, the mean square error of standard error and correlation coefficient are (4.68) and (0.95), respectively while for the neural network model optimized with the genetic algorithm in the related structure, the mean square error of the standard error and correlation coefficient are (4.72) and (0.93), respectively. Therefore, the neural network model optimized with the particle swarm algorithm has higher prediction accuracy and better performance compared with neural network models optimized with the genetic algorithm.

Table 5: Comparative Results of the Performance Evaluation Criteria of the Models

Evaluation Criteria and Parameters	ANN+PSO		ANN+GA	
	Amount of Training data	Amount of test data	Amount of Training data	Amount of test data
Mean Squared Error (MSE)	9.04	4.68	10.08	472
Root Mean Squared Error (RMSE)	6.1	4.5	6.1	4.6
Normalized Mean Squared Standard Error (NMSE)	1.05	0.63	1.09	0.67
Mean Absolute Error (MAE)	5.1	3.8	5.04	3.81
Mean Absolute Percentage Error (MAPE)	92	73.9	98	75.4
R Squared(R2)	0.88	0.95	0.84	0.93
source: Calculations of the research				

Table 6: Results of Sensitivity Analysis of Model's Input in Artificial Intelligence Models (ANN+PSO and ANN+GA)

Variable	ANN+GA	ANN+PSO
CF_Cerr	0.145	0.138
CF_Cpaid	0.135	0.142
CF_NCerr	0.118	0.125
CF _NCpaid	0.124	0.138
ΔCOA	0.238	0.247
ΔCOL	0.168	0.179
ΔNCOA	0.129	0.118
ΔNCOL	0.980	0.101
ΔFINA	0.134	0113
ΔFINL	0.116	0.134
ΔINV	0.111	0.129
ΔAp	0.141	0.152
ΔAR	0.127	0.116
DEP and AMORT	0.106	0.124
OTHER	0.109	0.116
source: Calculations of the research		

The results of the sensitivity analysis of the research variables in the best structure of neural network models optimized with particle swarm algorithm and genetic algorithm are presented in Table 6. Sensitivity analysis was used to select the most influential variable in predicting cash flows. Naturally, the higher the sensitivity analysis (weighting factor) of the variable, the greater the impact and weight on the network output and the prediction of future cash flows. The sensitivity analysis process shows the degree of sensitivity of the model to its input variables. In this study, in line with Asadi and Naghdi's [3] study, we obtained the sensitivity coefficient values of the input variables by dividing the total network error in the absence of one variable by the total network error in the presence of all input variables. Accordingly, if the sensitivity coefficient values of a variable are greater than one, the variable plays a large role in explaining the variability of performance evaluation criteria. As can be seen from the results of the Table 6, almost all the weighting coefficients of the variables in both models are larger than one and or almost close to one, indicating the ability of nearly all variables to predict cash flows. In the neural network models optimized with the particle swarm algorithm and the genetic algorithm, the current operating assets accruals (COA) with weight coefficient of (0.247) and (0.238) and non-current operating liabilities (NCOL) with coefficient of (0.098) and (0.101) have the highest and lowest influence on the prediction of future cash flows, respectively. The actual and predicted values of operating cash flows based on neural network models optimized with the genetic model and the particle swarm model are illustrated in Diagram3. As the Figure shows, the results of neural network optimized with particle swarm model compared with neural network optimized with the genetic pattern are closer to the actual results.

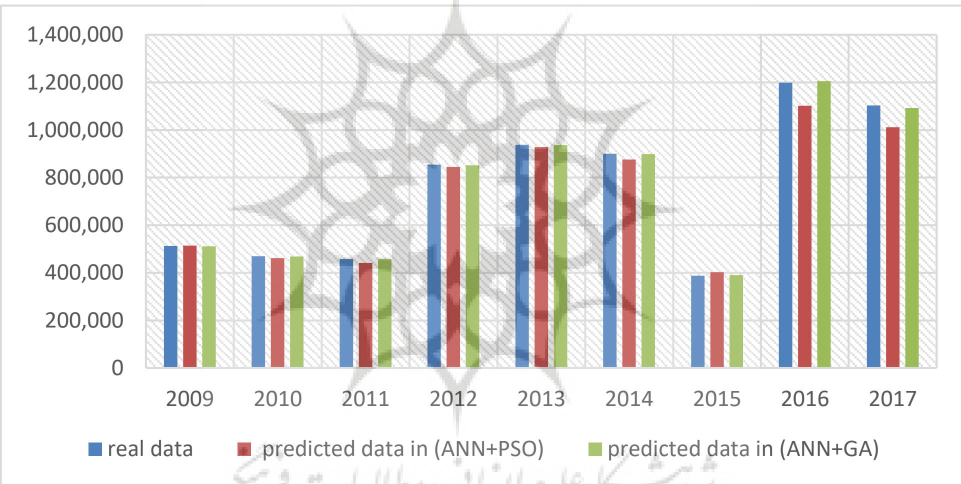

Diagram 3: Comparison of real data with predicted data in (ANN+PSO) and (ANN+GA) patterns

Table 7: Comparative Results of Evaluation Criteria of the Out-of-Sample Performance

Evaluation Criteria and Parameters	**ANN+PSO**		**ANN+GA**	
	Amount of Training data	**Amount of test data**	**Amount of Training data**	**Amount of test data**
Mean Squared Error	1.25	0.89	1.52	1.04
Root Mean Squared Error	0.97	0.71	0.84	0.81
Normalized Mean Squared Standard Error	0.045	0.018	0.051	0.036
Mean Absolute Error	1.1	0.83	1.04	0.94
Mean Absolute Percentage Error	8.7	7.2	9.1	8.4
R SquaredB(R2)	0.81	0.89	0.75	0.86
source: Calculations of the research				

4.3 Out-of-Sample Prediction

So far, the modeling of the research has been such that by using part of the in-sample data, the designed model and the rest of the same data were used for measuring the efficiency and accuracy of the models. In order to test the efficiency and robustness of the designed model, an out-of-sample test was used. Therefore, the modeling was performed using all in-sample data (from 2009 to 2016) and cash flows of 2017 (out of sample) was predicted. That is why after collecting the required data, the comparative results of the predicted data and the actual data of 2017 are shown in Figure 7. As can be seen in Table 7, for both models of artificial intelligence the prediction results of the evaluation criteria are satisfactory. It is worth mentioning that neural network optimized with particle swarm was more successful and efficient than the neural network optimized with genetic model in predicting future cash flows of 1016.

5 Conclusions and Suggestions

The future prediction has been a necessity in everyday life and is a common area of interest in many scientific fields. One of the areas in which prediction has a great importance is economic and financial issues. The effect of stock exchange market in economic development of a country is undeniable. The main task of this market is effective operation of the capitals and optimal allocation of the resources [5]. Accounting is defined today as an information system, and one of the most important information obtained from an accounting system is profit information. Previous studies have examined and presented various models for predicting future cash flows of business entities, and most of these studies have used regression models to predict future cash flows. Therefore, on the one hand, to achieve a more accurate prediction, and on the other hand, to consider the nonlinear and complex factors affecting future cash flow prediction and to find new methods and approaches with the least deviation and probable error with new variables affecting this prediction, it was necessary to conduct a study using new approaches (combining artificial intelligence networks), while separately examining the issue of predicting future cash flows and explanatory power of their variables in each approach, to develop a strong model with more explanatory power and greater efficiency and clarity for both models that can predict future cash flows. The results showed that new data mining approaches (neural network optimized with genetic pattern and bird flight) provided acceptable performance.

Regarding the conformity of the results of this study with other research, it can be stated that there is no similar study with these variables. Therefore, it cannot be fully compared and contrasted with the results of other studies. The results of the research hypotheses showed that both neural network model optimized with particle swarm algorithm (ANN+PSO) and genetic algorithm (ANN+GA) with 15 predictor variables are suitable models for predicting future cash flows. The neural network optimized with particle swarm pattern is closer to the actual results compared with the neural network optimized with the genetic model and has a lower error coefficient (higher prediction accuracy and better performance). Thus, it is recommended that investors, financial analysts and other financial statement users pay more attention to new data mining methods in their prediction to make more rational decisions. With respect to further research, studies under the same title with other new patterns of artificial intelligence (decision tree, backup vector machine, fuzzy artificial neural networks, ant colony, combination of neural network, genetic, bird-flight, etc.) is suggested to be conducted so that the results can be compared with approaches taken. Finally, there may be other variables that can help improve the proposed model so that new variables combined with the current variables of this study can be investigated with the afore-mentioned research patterns.

References

[1] Alborzi, M., Dadras, K., and Gharakhani, D., *ranian Designing a System Dy-namic Model for Corporate Cash Management*, Journal of Development and Evaluation Management, 2016, **23**, P.62-52. (in Persian)

[2] Amirhosseini, Z., and Davarpanah, A., *A Model Designed to Predict the Price of Gold Using particle swarm optimisatin and Genetic Algorithms and Providing Combined Algorithm*, Financial Engineering and Protfolio Management, 2016, **7**(26), P.83-59. (in Persian)

[3] Asadi, Gh., and Naghdi, S., *Designing and Formulating the Forecasting Model of Economic Growth by Ac-counting Approach*, Journal of Accounting Knowledge, 2018, **9**(34), P.39 -63. Doi:10.22103/JAK.2018.11095.2524. (in Persian)

[4] Bahrami Safa, M., *A Study of the Role of Accounting Information Based on Cash and Accrual Basis in Pre-dicting Operating Cash Flows*, Master thesis, Razi University, Faculty of Social Sciences, 2016. (in Persian)

[5] Etemadi, H.,Azar, A., and Baghaee, V., *Application of Neural Networks In Corporate's Profitability Predic-tion*, Journal of Accounting Knowledge, 2012, **3**(10), P.51-70. Doi:10.22103/JAK.2012.444. (in Persian)

[6] Farshadfar, S., Monem, R., *Further evidence of the relationship between accruals and future cash flows*, Accounting and Finance, 2017, **59**(1), P.143–176, Doi:org/10,1111/acfi.12260.

[7] Hamidian, M., Mohhamadzadeh Moghadam, M, B., Naghdi, S and Esmaeili, J., *Dividend Policy Prediction by Multivariable and Univariate Neural Network Models*, journal of knowledge of investement, 2018, **7**(26), P.169-184, (in Persian).

[8] Hemmatfar, M., *Application of Hybrid Genetic Algorithm and Neural Net-works Method to Activity Based Costing*, Journal of Financial Accounting, 2012, **16**, P.136-159. (in Persian)

[9] Larson, C., Sloan, R., and Zha Gied, J., *Defining, measuring, and modeling ac-cruals: a guide for researchers*, Review of Accounting Studies, 2018, **23**(3), P.827-871, Doi: 10.1007/s11142-018-9457-z.

[10] Mahdavi, Gh. and Saberi, M., *Determining the Optimal Model for the Predic-tion of Operating Cash Flow of Companies Listed in Tehran Stock Exchange*, Journal of Ac-counting Advances, 2010, **2**(1), P.199-225. Doi. 10.22099/JAA.2010.3435. (in Persian)

[11] Mohammadian Hajikord. A., Asgharzade Zaferani, M., and Emamdoust, M., *Credit risk assessment of cor-porate customers using support vector machine and genetic algo-rithm hybrid model - a Case Study of Tejarat Bank*, Financial Engineering and Protfolio Management, 2016, **7**(27), P.17-32. (in Persian).

[12] Pang, Y., *The design of dynamic and nonlinear models in cash flow prediction*, PhD thesis, University of Glasgow, Scotland, 2015.

[13] Koskinen, E., *Working capital management's effect on accrual based prediction of future cash flows: Evidence from North America, Atlto university learning center*, OEV Publication only available in digital format locally in the Harald Herlin Learning Centre, 2018, 11, P.56-78.

[14] Roozbacksh, N., Rezaiepajand, P., and Najari, M., *prediction of the current cash flows using artificial neural network in Tehran Stock Exchange*, the first international conference on political epic (with an approach to the middle east revolutions) and economic epic (with an approach to management and accounting), Rudehen, azad university of rudehen branch, 2013. (in Persian)

[15] Sagafi, A., Sarraf, F., Aghaballaiebacktiar, H., *application of artificial neural networks in predicting the future cash flows*, journal of accounting studies, 2015, **3**(9), P.63-80 (in Persian).

[16] Sarraf, F., Sagafi, A., *a model for forecasting the cash flow in Iranian companies*, researches of auditing and accounting, 2014, **31**, P.1-26(in Persian).

[17] Sarraf, F., Sagafi, A., Hassas yeghane, y and Amiri, M., *Linear and nonlinear regression models to estimate cash flow*, Journal of Management Accounting & Auditing Knowledge, 2014, **7**, P.141-155.

[18] Senan, N.A.M., *Ability of Earnings and Cash Flows in Forecasting Future Cash Flows: A Study in the Context of Saudi Arabia*, Academy of Accounting and Financial Studies Journal, 2019, **23**(1), P.40-61. Print ISSN: 1096-3685; Online ISSN: 1528-2635.

[19] Talebi, b., Abdi, r., Hajiha, z., Rezaei, n., *The Evaluation of the Capability of the Regression & Neural Network Models in Predicting Future Cash Flows*, Advances in Mathematical Finance & Applications, 2020, Articles in Press, Accepted Manuscript Available Online from **13** July 2020.
Doi:10.22034/AMFA.2020.1876840.1277.

[20] Zamani, S., Zanjirdar, M., Lalbar, A. *The effect of information disclosure on market reaction with meta-analysis approach*, Advances in Mathematical Finance and Applications, 2022, **7**(3), P. 629-644.
Doi: 10.22034/amfa.2021.1937478.1625

[21] Zanjirdar, M., *Overview of Portfolio Optimization Models.* Advances in Mathematical Finance and Applications, 2020, **5**(4), P.419-435. Doi: 10.22034/amfa.2020.674941.

[22] Zanjirdar, M., Kasbi, P., Madahi, Z., *Investigating the effect of adjusted DuPont ratio and its components on investor & quot; s decisions in short and long term*, Management Science Letters, 2014, **4**(3), P.591-596.
Doi: 10.5267/j.msl.2014.1.003