Advances in Mathematical Finance & Applications, 5(2), (2020), 197-213
DOI: 10.22034/amfa.2019.585043.1184

Published by IA University of Arak, Iran
Homepage: www.amfa.iau-arak.ac.ir

Stock Price Prediction Using the Chaid Rule-Based Algorithm and Particle Swarm Optimization (PSO)

Aliasgar Davoodi Kasbi, Iman Dadashi*, Kaveh Azinfar
Department of Accounting, Babol Branch, Islamic Azad University, Babol, Iran

Article Info
Article history:
Received 11 February 2019
Accepted 06 May 2019

Keywords:
stock price
particle swarm optimization algorithm
Chaid rule-based algoritm

ABSTRACT

Stock prices in each industry are one of the major issues in the stock market. Given the increasing number of shareholders in the stock market and their attention to the price of different stocks in transactions, the prediction of the stock price trend has become significant. Many people use the share price movement process when comparing different stocks while investing, and also want to predict this trend to see if the trend continues to increase or decrease over time. In this research, stock price prediction for 1170 years -company during 2011-2016 (a six-year period) of listed companies in stock exchange has been studied using the machine learning method (Chaid rule-based algorithm and Particle Swarm Optimization Algorithm). The results of the research show that there is a significant relationship between earnings per share, e / p ratio, company size, inventory turnover ratio, and stock returns with stock prices. Also, Chaid rule-based algorithm has a good ability to predict stock prices.

1 Introduction

The role of capital in economic growth across the world has been explicitly accepted. The provision of capital through the stock market contributes significantly to accelerating economic growth. The stock market plays a significant role as a genuine instrument in moving and allocating savings. Most studies have found the existence of developed stock markets to be effective in boosting economic growth [13]. Uncertainty in the capital market means the difference between the expected values and the amounts that actually occur. Designing different analytical and forecasting methods in the capital market has been less significant due to the high level of this amount and the need to know future prices with greater certainty or less uncertainty [22, 46]. The particular importance of the capital market in economic development through the effective management of resources and the optimal allocation of non-negotiable resources is undeniable. Investment in the capital market requires decision making, which requires obtaining information about the future market price of the stock market. Therefore, if the stock market future trend can be achieved with appropriate methods, the investor can maximize the return on investment [16, 26, 34]. Today, due expansion of economic activity, financial markets and the prosperity of investing in equity markets, especially stock exchanges by natural and legal persons, access to accurate and timely information and their accurate and realistic analysis are the most important tool for making the right decisions and obtaining expected benefits and using fi-

* ***Corresponding author Tel.:*** *+989124400915*
E-mail address: *idadashi@gmail.com*

nancial facilities in an optimal and desirable manner. According to [14], investing in the stock offered in the stock market is one of the lucrative options in the capital market. In today's society, information plays an important role in human life, and the more advanced the community, the more and the better it uses the information. One of the reasons for advancement in developed societies is the efficient and effective use of information. An important factor in economic development is the development of financial markets. The main role of financial markets is to attract and direct savings and dispersed liquidity in the economy to its optimal paths, leading to an optimal allocation of financial resources. More developed financial markets are able to finance investment not only through better absorption of domestic resources but also through attracting foreign capital [23, 32]. Financial markets have an important role in the economy of every country and are effective factors in economic growth. The stock market is also one of these markets, which leads funds toward investment opportunities affect a large part of the economy in the world and cause a great deal of concern for the governments [35]. In today's economy, the impact of the stock exchange on investment, the financing of enterprises and the growth and development of countries is quite evident. Regarding the nature of the stock exchange and the risk of investing therein, policy makers have always been looking for ways to predict the stock price index to minimize the risk of decision making [39]. The achievement of long-term and continuous economic growth requires the equipping and optimal allocation of resources at the national economy level, and this is not easily possible without the help of financial markets, especially the extensive and efficient capital market. Investing in stocks listed on the stock exchange is one of the lucrative options in the capital market. However, the evaluation and forecasting of stocks or any other securities requires a historic trend and specialist expertise. Different theories have been proposed regarding the valuation and prediction of stock prices in organized markets. In the early 20th century, a group of experienced experts in valuing securities believed firmly that it was possible to provide an image for predicting future stock prices by studying and analyzing the historical trend of stock price changes. More scientific studies, focusing on the precise identification of stock price behavior, have led to a tendency toward stock price valuation models. At first, the theory of random steps was introduced as a starting point in determining the behavior of stock prices. Then, attention was paid to the characteristics and structure of the capital market, which resulted in these studies and studies leading to an efficient capital market hypothesis. This hypothesis was considered by the scientific gatherings due to its specific composition. In the efficient capital market, stock prices are believed to reflect the current information about that share, and stock price changes do not have a specific and predictable pattern. Stock price prediction has always been a challenging issue for researchers [9]. The arguments raised until the 1980s clearly determined the behavior of stock prices in the market until the New York Stock Exchange developments in 1987, questioned the validity of the market capitalization assumptions and models such as the randomness of prices. In the 1990s and beyond, more attention was paid by experts to a turbulent behavior along with discipline and attempts to design nonlinear models in order to predict stock prices were of increasing importance [29]. Stock price prediction is of particular importance to shareholders in order to earn the most profit and always seek out reasonable and accurate strategies for forecasting. Therefore, the investor in the stock market also needs the necessary, powerful and trusted tools to predict the price. Data mining techniques can perform this prediction well. Data mining techniques, in addition to data collection and management, also include analyzes and predictions that help us to predict the discovery of existing patterns and unknown relationships between data, and most people use it for decision making and future planning. [37]. Given the increasing number of shareholders in the stock market and their attention to the price of different stocks in transactions, the prediction of the stock price trend has become significant. Many people,

when comparing different stocks, use the share price movement when comparing different stocks, and also want to predict this trend in order to know how long the increasing or decreasing trend will continue [1] Thus, the investor in stock exchange needs the necessary, powerful and reliable tools to predict stock prices.

2 Theoretical Foundations

The existence of sufficient market information and its timely and quick reflection on the price of securities is closely related to market efficiency. In the efficient market, information that is being distributed on the market quickly impacts prices. In such a market, the price of securities is close to its intrinsic value. In other words, the important feature of the efficient market is that the market price is a good indicator of the true value of the securities; therefore, the market is an efficient market, in which the price of securities, such as the price of ordinary shares, reflects all the information available on the market. The efficient market should be sensitive to new information. Influence of information on prices is the core of the market. This means that, as new information becomes available, immediate responses will arise, and thus prices will change. In the financial literature, research on how stock prices react to market-based information and company specific information has become so important. In addition to the "direction" and "the rate of stock price change," another dimension in the market efficiency hypothesis has been formed that is the price adjustment rate to the reflection of new information [47]. Different theories have been proposed for stock market forecasting in organized markets. In the early 20th century, a group of experienced experts in valuing securities firmly believed that it was possible to analyze the historical trend of stock price changes, providing an image for predicting future stock prices. More scientific studies, with an emphasis on accurate identification of stock price behavior, led to a tendency toward stock price valuation models. At first, the theory of random steps was introduced as a starting point in determining the behavior of stock prices. Then, attention was paid to the characteristics and structure of capital markets, which resulted in these studies and studies leading to a capital market hypothesis. In the capital market, the stock price is believed to reflect the current information about that share, and stock price changes do not have a specific and predictable pattern [47]. The ideas posed by the 1980s were a good determinant of stock market behavior until New York Stock Exchange developments in 1987 seriously questioned the validity of capital market assumptions. In the 1990s and beyond, more attention was paid by specialists to a poignant behavior accompanied by discipline, and attempts were made to design nonlinear models to predict stock prices. Many recent studies have shown that the stock market is in fact a nonlinear and chaotic system which is dependent on political, economic, and psychological factors [15]. In order to overcome the limitations of traditional analysis techniques in predicting nonlinear patterns, in the past two decades, experts have been using intelligent techniques, especially artificial neural networks and genetic algorithms, to improve Stock price forecast[44]. Using common mathematical methods such as simple mean, balanced mean, regression, etc., the results are efficient when there is a specific mathematical relationship among independent and dependent variables and the data There is a certain mathematical relation [43]. How the investors react to the received information plays a crucial role in deter-mining the return of stock exchange market. Supply and demand based upon incorrect decisions lead to the price deviation of inherent values [24]. Stock market is affected by news andinformation. If the stock market is not effi-cient, the reaction of stock price to news and information will place the stock market in overreaction and under-reaction states[36]. Using statistical tests and reviewing trends and relationships between the variables, planning can be done to invest in it and its performance or

inefficiency can be tested [3]. Using intelligent techniques such as artificial neural networks, genetic algorithm,Differential Evolution algorithm 6, Particle Swarm Optimization Algorithm, Paddy field Algorithm 8, ant colony optimization 9, Bee Colony and Firefly Algorithm [20]. In these systems, the prediction is based on the experience, that is, the time series of that variable is considered, and the factors and relations affecting the formation of the variable are shown in the values of that variable, so the previous values of the variable can be used as the most important source for explaining the changes, and prediction can be done only by studying the process of these changes [13].

3 Literature Review

In recent years, different models have been used to predict stock prices by researchers, since artificial intelligence techniques that include neural networks, genetic algorithms, and fuzzy logic have achieved successful results in solving complex problems, have been used mostly in this regard [28]. The most common starting point for investors when buying a stock is to examine the trend of stock price changes. Internal factors are factors affecting stock prices in relation to operations and company decisions. Jamshidi et al. [22] concluded in his research that the ratio of the cost to income and the rate of return on the stock price is effective on the stock price [22]. One of the internal factors that affect financial performance is the size of a business unit. The size of a company is often measured by the amount of a company's assets. Financial literature suggests that the size of companies is an effective factor in the risk and returns, and other major variables of a company, when the size of a company is large, multiple groups will be shareholders, and therefore the liquidity of the equity transactions of such a company will be more than small companies. The diversity of the products of the large company is generally more than that of small companies, therefore, the business and operational risks of companies are reduced because profitability is accompanied by stability, the ratio of large companies' dividend returns is also stable and increasing [18]. The behavioral consistency hypothesis was used in accounting research in the 1960s. This hypothesis states that investors only pay attention to the financial statements of companies and do not care where this information came from. In other words, the hypothesis suggests that the relationship between stock prices and accounting profit is just a mechanical and habitual relationship, and this can mislead investors by choosing the procedures and methods of accounting. According to the behavioral stability hypothesis, investors are act similarlyin achieving and processing information. This is one of the criticisms of this hypothesis. The opposite of the behavioral stability hypothesis is the market-efficient hypothesis, the hypothesis that the stock prices are determined by investors, thus fully reflecting all available information. The behavioral stability hypothesis asserts that investors are incapable of identifying accounting practices, so different accounting methods can affect stock prices, even if cash flows do not change, unknowing decision makers are associated with changes in reported earnings figures that are not associated with changes in cash flows. [12]. In Iran, the development of the Tehran Stock Exchange has been set up with the aim of mobilizing and equipping private savings toward generation and attraction of investors' public participation. One of the important issues in the stock market is stock prices as a signal in directing the volume of liquidity and effective allocation of capital that, if matched to the inherent value of the stock, becomes a powerful tool in the efficient allocation of resources, the changes in the price of stockis related to the systematic changes in the company's fundamental values and the investor's irrational behavior has no effect on stock returns [42] . In the method of fundamental valuation of stock prices affected by size and quality, important factors such as income, profit, sales, price-to-profit ratio, etc. are determined. The fundamental analysis is based on the assumption that each securities(and the

market in general) has an intrinsic value or true value based on investor estimates. This value is a function of the company's fundamental variables that are combined to create expected returns and risks associated with it. By evaluating these fundamental variables which determine the value of securities, one can estimate the intrinsic value. Then, the estimated inherent value can be compared to the current price of the securities market [38]. Lakshmi et al [27] in their article titled, "A new statistic for measuring the level of dependence toward the fluctuation of the stock price," in India, examined four indicators, including NIFTY, S & P500, FTSE100 and DAX, during the sample period from January 1996 to March 2015, and they showed that stock fluctuations are predictive of daily data. Ashraf et al. [5] in an examination of the comparative evidence of the relationship between stock prices and the value of accounting information based on IFRS in Germany and the United Kingdom, stated that there was a long-term relationship between accounting variables and stock price changes in countries with common laws such as in the United Kingdom. Sujeewa et al. [40] in his research, studied the impact of specific factors on stock prices among companies listed on the Columbia Stock Exchange. The results indicate a positive and significant relationship between the specific factors of the company, the profit of each share, the interest of each share, the value of the asset of each share with the value of the stock. Bartov et al. [11] investigated the behavior of investors in dealing with the reflection of profitability of unexpected items in the form of profit and loss statement. They showed that investors showed a reaction to disclosure of profit and loss of unexpected items. Therefore, the disclosure of unexpected items in the context of profit and loss statement is directly related to the price of the stock. Their findings are consistent with the behavioral stability hypothesis. Islami et al. [22] studied the effect of internal factors on changes in stock prices of investment companies listed in Tehran Stock Exchange. The statistical population of the study includes all investment companies in the financial intermediation industry and related activities during the period of 2008 to 2014. The three variables of earnings per share, cash dividends per share, and the ratio of price to earnings, the liquidity ratio affect supply and demand, and ultimately stock prices more than other factors. AhmadKhan et al. [9] predict stock prices with Artificial Neural Network and colonial competition algorithm based on chaos theory. Therefore, using the daily pricing information of Irankhodro's shares between 2010 and 2016, we taught the neural network with different optimization algorithms. In order to evaluate the performance of the approaches, three perspectives were used: the accuracy of prediction (statistics for error measurement R2, RMSE), the amount of memory used and the time of implementation. The results suggest that the proposed approach has a better performance than other approaches. Roshandel et al. [37] tested and compared the methods of data mining in stock price prediction. In this paper, different data mining algorithms that have been used to predict share prices in the stock market up to now, have been compared and reviewed. One of these algorithms is artificial neural network, fuzzy neural network and genetic algorithms. Bigazadeh et al. [14] in their review of stock price prediction using the genetic algorithm state that investing in stocks listed on the stock exchange is one of the lucrative options in the capital market.The stock market is a nonlinear and chaos system that is influenced by political, economic, and psychological conditions, and nonlinear systems such as genetic algorithms can be used to predict stock prices. In this paper, the design and presentation of a stock price prediction model using genetic algorithm and reduction of stock price prediction error by using exponential function have been investigated. The sample of this study is Saipa Company. The comparison of the mean square error in the two models of linear function and exponential function represents this factor that exponential function model has a lower error rating. Considering the noticeable and evident accuracy of exponential function model in Saipa share price prediction, this

model, as a more precise simulation model, can be used to predict the price of this share. Nasirzadeh et al. [34] evaluated the ability of the data mining model to predict stock prices. This study examines the accuracy of the sVR support vector estimator models. The minimum LARS of the ANFIS neural network model is used to predict stock prices at three daily, weekly and monthly levels. The results show that all three models have the ability to predict stock prices, but SVR support and neural-fuzzy network support vector estimators are more capable of forecasting stock prices by daily and weekly data. Mehrara et al. [30] have studied the data mining techniques in predicting stock prices. They have argued that gaining economic growth and incentives to investment, get accelerated in a country when the country has active and reliable capital markets. The existence of active stock markets has always pushed many investors to expedite the flow of capital and funds to productive sectors. It is possible to use more precise methods that can be used in any kind of prediction. Data mining techniques such as the neural network and decision tree and logical separated analytical regression are the general additive model of the Boosting method. The stock price forecast for the listed companies in Tehran Stock Exchangeusing data mining techniques and according to financial criteria is among important subjects that there are not enough researches related it. Aghajan et al. [3] in the study of the relationship between financial ratios and stock prices in Tehran Stock Exchange showed that there is a significant relationship between financial ratios and stock prices. In order to capitalize on capital markets, investors always seek to find a share ideal for investment and reasonable prices for buying and selling. Therefore, all of the proposed forecasting models have always sought to answer the three basic questions: what stock, within which time period and with what price should be bought or sold. Before examining the answers given to these questions, we have to answer a more serious question. Is it possible to predict financial markets for stock prices? Further, it should be noted that if the capital market is to be predicted, we should examine the different dimensions of the capital market and the methods presented in each field for forecasting. The purpose of this study is to examine the possibility of predicting prices. If there is a possibility to predict the price, more return can be achieved while some others face some losses. To perform this test, a non-parametric method called a turning test is used.

4 Research Methodology

The results of the research show that stock prices in Tehran Stock Exchange are predictable. According to the mentioned theoretical principles and the aim of the research, the following hypotheses are developed.

1) Are financial ratios capable of predicting stock prices according to the chaid rule-based algorithm?

2) Are financial ratios capable of predicting stock prices according to the particle swarm algorithm?

3) Which algorithm particle swarm and the chaid rule-based has more ability to predict stock prices?

In machine learning problems, decision trees such as CHAID algorithm provide a non-linear hierarchical rule-based model that can be nonlinear and can handle nonlinear problems. In this Paper, the model of the PSO algorithm is a linear model, and the linear model parameters are obtained by PSO algorithm. As a result, this algorithm can handle linear problems. In this paper, we use R2016b MATLAB software to implement algorithms. This software is one of the most important applications of machine learning domain. This software has a variety of tools for implementing the machine learning algorithms. In today's world, there are many issues that can be addressed by algorithms and ap-

propriate solutions for them.One of the applications of algorithms in industry and commerce is that it is necessary to allocate rare resources with the most advantageous method, so that with limited resources we can achieve maximum and minimum profits. This research is applied based on the purpose and in terms of the type, is a field-library study using historical information as an after-event (using past information). The statistical population of this research includes all companies accepted in Tehran stock exchange which have the following conditions.

1-During the course of study, there should be no change in the financial period.

2- They should not be one of investment companies, financial intermediaries, banks, insurance and leasing.

3. Their considered data be available.

Finally, due to the limitations mentioned. 1170 year-company has been selected as a statistical population between 2011 and 2016, which, considering the availability of information, all companies have been considered as the statistical sample. Financial ratios are considered to be useful tools in accounting for financial statements of companies, which, by presenting the ratio of some important accounting items, obtain an understanding of important assets in the period between the results of operations and the financial position of a company. A variable is a concept that is assigned more than two or more values or numbers. In other words, the variable refers to features that can be viewed or measured. And replaced two or more values or numbers.

Table 1: The Variables Used in This Research

Operational definition	Quantitative variable
Net profit margin [36].	Net profit / net sales
Return on Assets [6].	Net Profit / Total Assets
Return on Equity[40].	Net Profit / Equity
Earnings Per Share [10].	Net Profit / Number of Shares
Current ratio[24].	Current asset / Current Debt
Immediate ratio [25].	(current asset-inventory-prepayment) / current debt
Ratio of working capital [8].	(current asset – current debt) / assets
Financial leverage [41].	Debt / Asset
Debt to equity ratio[8].	Debt to equity ratio
Inventory turnover ratio [8].	Sales / inventory of goods
Fixed asset turnover ratio[8].	Sale / fixed assets
Asset Turnover Ratio [8].	Sale / Asset
Turnover of Receivable Accounts [8].	Sales / Receivable Accounts
Ratio of market value to book value of assets[36].	Company market value / Net book worth of assets
Systematic Risk [45].	The systematic risk of the degree of return variance is a specific investment relative to the return changes of the total investment of the market and is measured by the β index.
Return on Equity [19].	Use of published annual stock returns information
Unconditional conservatism [30].	{Value of equity market / Equity book value} * (-1)

Table 1: Continue

Quantitative variable	Operational definition
Conditional conservatism [30].	{Book value of assets / net profit difference and operating cash flow plus depreciation expense} * (-1)
Cash / asset ratio [8].	Cash / Assets
Operating cash ratio [33].	Operational cash / assets
Ratio of Current Assets [8].	Current Assets / Assets
Ratio of Current Debt to Equity[8].	Current debts/ equity
Long-term debt ratio [32].	Long-term debt / Total assets
Size of the company [31].	The natural logarithm of the total assets
Dividend Profit Ratio [32].	Dividend Profit per Share / Total Assets
Tax avoidance [7].	From the division of the tax on profit before tax.
The price / profit ratio [31].	From dividing the stock price by the profit of each share.
Conditional conservatism [30].	{Book value of assets / net profit difference and operating cash flow plus depreciation expense} * (-1)
The dependent variable	
Stock price	Stock price information of the companies at the end of the fiscal year
Research methodology	Using the particle motion algorithm and Chaud rule-oriented algorithm
Research period	2011- 2016 (6-year period)
Research design	Two-stage approach, 1- Variable selection test 2- Stock price forecast

5 Variable Selection

5.1 Feature Selection Method Based on Lars's Algorithm

Suppose we have sets of variables n+1 that are set up $x_1,x_2,...,x_n,y$. In the approximation of function 18 (regression 19), we seek a function (or model) that approximates the value of the dependent variable y as a linear function of x_j, $j = 1,..., n$ the independent variable. Vector $x \in R^n = (x_1, x_2,..., x_n)^t$ is a vector of independent variables. The goal is to find f as to have $y = f(x)$. Suppose we know that $f(x)$ belongs to linear models, that is, $f(x) = \sum_{j=1}^{n} x_j \beta_j + p =< x, \beta > + p < .,. >$ in which $< .,. >$ is the internal multiplication of two vectors and P is a constant sentence. It can also be written that $f(x) = x^t\beta + p$ that $\beta \in R^n$ and p describe the function $f(x)$. Without loss of generality, let's consider $p = 0$. To construct the model, we look for the value of β. For theis purpose, suppose that we have the pair of company-year samples are many as follows:

$$(x_1, x_2, \dots, x_n)_1 = X_1 \leftrightarrow y_1$$
$$X_2 \leftrightarrow y_2$$
$$\vdots \qquad \vdots$$
$$X_m \leftrightarrow y_m$$

We want to find β so that $\forall i \quad X_i^t \beta = y_i X\beta = y$. If we want to represent matrices, β is the solution of $X\beta = y$. In this method, as well as the method of selecting an expanding feature, first we set all the coefficients Bj to zero and select the independent variable that has the most correlation with the dependent variable (x_{j1}). Then, we take the maximum length of a step that can be taken in the direction of this variable, as long as there is another variable such as x_{j2} with the same correlation as with the current remainder. Instead of continuing in the direction of x_{j1} , Lars continues, in a direction that has an angle with both variables, as long as the third variable x_{j3} enters the "maximum correlation set". Then, we continue in the direction of the same angle of the three variables x_{j1} ,…, x_{j2} and x_{j3} which is called Least Angle Direction. This problem is shown in Figures 2 and 3.

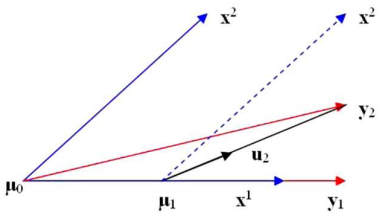

Fig. 1: Equal Angle Between Two Variables

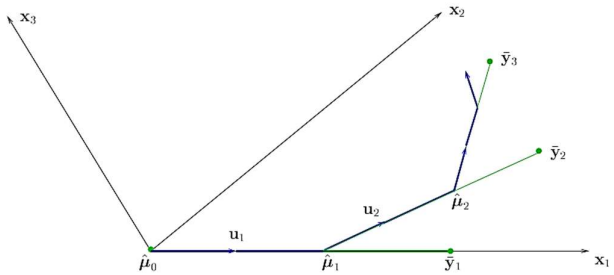

Fig. 2: The Geometric Routine of the Lars Algorithm

Efron et al. [17] studied Least Angle Regression(LARS) The advantages of this algorithm are 1) we need only M steps that M is the number of independent variable,2) Selecting a new variable dependent on the previous selected set of variables. The second advantage is the key advantage of this algorithm that leads to select independent variables dependent on each other. Let's now summarize the LARS algorithm with respect to the correlation r1 between x and Y, and Rx correlation matrices of the variables is summarized as follows:

1. $A = \phi, \qquad s_A = \phi$
2. $m = \arg \max |r_j|, \qquad s_m = sign\{r_m\}, \qquad r = s_m r_m$
3. $A \leftarrow A \cup \{m\}, \qquad s_A \leftarrow s_A \cup \{s_m\}$
4. $Calculate \quad a == \left[1'_A (D_A R_A D_A)^{-1} 1_A\right]^{\frac{1}{2}} \quad where \quad D_A = diag(s_A), \quad R_A \subset R_X$
   $Calculate \quad w_A = a(D_A R_A D_A)^{-1} 1_A$
   $and \quad for \quad j \in A^c \quad : \quad a_j = (D_A r_{jA})' w_A$

Which r_{jA} is a vector of correlation between X_j and active variables. (Note: When there is only one active variable:) $a = 1, \quad w = 1, \quad a_j = r_{jm}$

5. $for \quad j \in A^c \quad Calculate \quad : \quad \gamma_j^+ = \frac{r - r_j}{a - a_j}, \gamma_j^- = \frac{r + r_j}{a + a_j}$

$\gamma_j = \min(\gamma_j^+, \gamma_j^-), \qquad , \qquad \gamma = \min(\gamma_j, j \in A^c)$

If m is the index of arg min, we have Y=Ym.

$if \quad \gamma_m = \gamma_m^+ \quad then \quad s_m = +1 \quad esle \quad s_m = -1$

$for \quad j \in A^c \quad Modify \quad r \leftarrow r - \gamma a \quad , \quad r_j \leftarrow r_j - \gamma a_j$

Repeat 3,4,5 steps.

Table 2: Selected Independent Variables with Weight (Significance).

	Selected independent variables	Weight
X1	Earnings per share	5/277
X2	P to e ratio	3/953
X3	size of the company	130/347
X4	Inventory turnover ratio	8/331
X5	Stock returns	2/485

After selecting the independent variables of the problem, these independent variables were given to construct the model to the particle and chaud motion algorithm. Subsequently, these algorithms were investigated.

5.2 Mass Particle Swarm Optimization Algorithm

The PSO is of optimization algorithms that operate on the basis of random primary population generation. This algorithm is based on the simulation of the behavior of massive (group) bird flying or collective movement (grouping) of fish. Each member in this group is defined by the vector of the velocity and position vector in the search space. In each time repetition, the new particle position is defined according to the vector of velocity and position vector in the search space. At each time interval, the new particle position is updated according to the current velocity vector, the best position found by that particle, and the best position found by the best particle in the group. This algorithm was initially defined for continuous parameters, but given the fact that in some applications we deal with discrete parameters, this algorithm is also expanded in a discrete state. The particle swarm optimization algorithm is introduced in an enhanced binary state with (BPSO).

Suppose we have a d-dimensional search space. ith particle in this d-dimensional space is described with the position vector Xi as follows:

$$X_i = (x_{i_1}, x_{i_2}, x_{i_3}, \ldots, x_{i_d}) \qquad (1)$$

The vector of velocity of ith particle is also determined by the vector Vi as follows:

$$V_i = (v_{i_1}, v_{i_2}, v_{i_3}, \ldots, v_{i_d}) \quad (2)$$

We define the best position ith particle have found with P(i.best):

$$P_{i.best} = (p_{i_1}, p_{i_2}, p_{i_3}, \ldots, p_{i_d}) \quad (3)$$

The best position that the best particle has found in the whole particle is defined by P g.best as follows:

$$P_{g.best} = (p_{g_1}, p_{g_2}, p_{g_3}, \ldots, p_{g_d}) \quad (4)$$

We use the following for updating the location of each particle:

$$V_i(t) = w * V_i(t-1) + c_1 * rand_1\left(P_{i.best} - X_i(t-1)\right) + c_2 * rand_2\left(P_{g.best} - X_i(t-1)\right) \quad (5)$$

$$X_i = X_i(t-1) + V_i(t)$$

Where *w* is the inertial mass index (motion in the same path), which indicates the effect of the velocity vector in the previous repetition on the velocity vector in the current repetition (Vi (t+1)). C1 is the constant coefficient of training (moving in the direction of the best value of the particle being examined), c2 is the constant coefficient of training (motion on the path of the best found particle in the total population), rand_1 and rand_2 are two random numbers with uniform distribution in the interval 0 to 1, Vi(t-1) is the velocity vector in the repetition (t-1)th, Xi(t-1) is the position vector in the repetition (t-1). To avoid excessive increase in the velocity of a particle moving from one location to another (the velocity divergence), we limit the velocity changes to the range V_min to V_max; $V_{min} \leq V \leq V_{max}$ the upper and lower limit is determined by the type of problem.

5.3 PSO Algorithm Implementation Steps

Random generation of the initial population is simply a random determination of the initial site of the particles with uniform distribution in the solution space (search space). An initial population random generation stage exists in almost all probabilistic optimization algorithms, but in addition to the initial random particle location, the algorithm also allocates a value for the initial particle velocity. The initial range for particle velocity can be extracted from the following equation.

$$\frac{X_{min} - X_{max}}{2} \leq V \leq \frac{X_{max} - X_{min}}{2} \quad (6)$$

Selecting the number of primary particles: Increasing the number of primary particles reduces the number of repetitions needed to converge the algorithm. This reduction in the number of repetitions does not mean that program execution time is reduced to convergence. Although the increase in the number of primary particles leads to a reduction in the number of repetitions, but an increase in the number of particles causes the algorithm to spend more time in the particle evaluation stage, which in turn increases the time the algorithm is implemented to reach convergence does not decrease despite the decrease in the number of repetitions. So increasing the number of particles cannot be used to reduce the runtime of the algorithm. There is another misconception that reducing the implementation time of the algorithm can reduce the number of particles, but in order for the algorithm to reach the optimal solution, the number of repetitions should increase. If we consider the convergence condition to be unchanged at the expense of the best member in several consecutive repetitions, it ultimately does not reduce the runtime of the program to achieve the optimal response. Reducing the number of particles may result in clinging to local minima, and the algorithm fails to reach the main minimum. If

we consider the convergence condition to be the number of repetitions, although the algorithm will be reduced by decreasing the number of initial particles, the answer will not be the optimal solution for the problem, since the algorithm has been implemented imperfectly. In summary, the number of primary population is determined by the problem. In general, the number of primary particles is a compromise between the parameters involved in the problem. Experimentally, selecting a primitive particle population of 20 to 30 particles is a good choice, which works well for almost all testing issues. You can consider the number of particles a little more than necessary to have a bit of safety margins when faced with local minima.

5.4 Chaid Rule-Based Algorithm

This algorithm, called Chi-square automated interaction detection, uses the criterion of the similarity between characteristics as a separator criterion. This algorithm uses the criterion of chi2 (X2) to select the best feature in each decision node. The criterion X2 examines the relationship between variables and is defined as follows.

$$X^2 = \sum_j \frac{(x_{ij} - E_{ij})^2}{E_{ij}}, \qquad i = 1, \dots, n \tag{7}$$

In which x_{ij} and E_{ij} in order of observation and the expected value are the jth class and n is the number of attributes. The feature that has the highest value X2 is selected as the separator feature.

Table 3: Mean of Smape Error Rate for Evaluating Performance with Test Data Every Two Years

error	Current year		Next year	
SMAPE	PSO	CHAID	PSO	CHAID
Fold				
1	0.086	0.22	0.094	0.25
2	0.085	0.23	0.11	0.29
3	0.086	0.24	0.097	0.25
4	0.085	0.25	0.101	0.22
5	0.086	0.25	0.093	0.24
6	0.087	0.25	0.099	0.23
7	0.086	0.24	0.099	0.25
8	0.086	0.24	0.101	0.27
9	0.087	0.23	0.098	0.27
10	0.086	0.25	0.098	0.24
Mean	0.086	0.24	0.099	0.25

But what we need to worry about is the occurrence of a phenomenon called over-fit. Therefore, in order to review the generality of the models presented, the error rate is obtained to predict the dependent variable of the stock price in each three years for the testing company-years (the company-years that were discarded by the 10-fold cross-validation method and the algorithm has not seen them so far). For each error criterion, 10 errors, each reported by the 10-Fold Cross-Validation method, are obtained, which are shown in Table 11 along with their mean for every three years. Similarly, it is concluded that the obtained models are generic, that is, for the companies that have not seen them,

they are doing well, and also the problem of overfitting has not happened, since the difference between the criteria of the error of the training and evaluation data is negligible. In addition, the PINSVR linear model has a relatively large error and cannot properly predict the dependent variable of the stock price in comparison to the other two algorithms [46]. To evaluate the results of the research, the mean absolute magnitude of the symmetric error (SMAPE) has been used.

$$SMAPE = \frac{1}{n} \frac{\sum_{i=1}^{n} |d_i - y_i|}{\sum_{i=1}^{n} (d_i + y_i)}$$

As seen above, the PSO algorithm has a higher ability to predict stock prices with a percentage error of SMAPE of 0.86 and 099, than the chaid algorithm.

6 Conclusions and Discussion

There are various ways to predict stock prices. These methods are effective when they have full and accurate information. Due to the complexity of the stock exchange and the impact of changes in economic, social, and social conditions on it, it is difficult to predict stock prices. Investigating the results of the research carried out in this regard, except for the research of Toloi et al, shows that using traditional methods and instruments of forecasting has a high error, and it often fails when compared with newer versions and non-linear models, which can be due to the complexity of the stock exchange , the increase in the variables affecting the stock price, as well as the lack of specific mathematical relationships among these variables. Artificial intelligence algorithms may be the best way to predict stock markets, because it learns by experience and does not come to a deadened. These algorithms are ideal tools that, in addition to exploiting statistics, also consider mental aspects; therefore, it is possible to predict the price of stock with this algorithm. The results of this research are consistent with most studies conducted in this field ie stock price prediction with intelligent algorithms. Therefore, the main hypothesis of this research is confirmed. AI algorithms are an efficient method for predicting stock prices. It is suggested to capital market analysts to use the power of Chaud rule-oriented algorithm to analyze stock prices, but in relation to other variables such as the size of the company. They also showed that there is a significant relationship and the stock prices of larger firms have less effect than small companies in terms of profit information, which indicates more information disclosure in the large companies in Iran. Also, the results of the research are in line with the work of Banimahd et al. [39] is a significant relationship between the earnings per share and the share price. Mehrara et al. [30] have studied the data mining techniques in predicting stock prices. They have argued that gaining economic growth and incentives to investment, get accelerated in a country when the country has active and reliable capital markets. The existence of active stock markets has always pushed many investors to expedite the flow of capital and funds to productive sectors. It is possible to use more precise methods that can be used in any kind of prediction. Data mining techniques such as the neural network and decision tree and logical separated analytical regression are the general additive model of the Boosting method. The stock price forecast for the listed companies in Tehran Stock Exchangeusing data mining techniques and according to financial criteria is among important subjects that there are not enough researches related it. Jamshidi et al. [23] concluded that E / P ratio was effective on stock price index and volume of transactions. Huddart et al. [21] investigated the effect of the psychological factors caused by the fluctuation of stock prices in the past on the decision to purchase

shares and the volume of trades and confirmed this effect. Aghajan et al. [2] have shown that there is a significant relationship between financial ratios and stock prices in the study of the relationship between financial ratios and stock prices in Tehran Stock Exchange. Based on the results obtained from the research, it was found that the financial ratios have the ability to predict stock prices according to the chaid algorithm and the pso algorithm, as well as the pso algorithm with a lower error rate than the chaid algorithm of higher ability To predict stock price.

It is suggested that according to the preliminary results of the research, the variables of the earnings per share, the ratio of the price to profit of each share, the size of the company, the ratio of inventory turnover and stock returns are the most important in predicting the stock price, therefore, to the managers of capital market companies Iran is recommended to consider the variables mentioned above to decide on the continuation of the company's activity and to encourage investors to invest in the company. According to the secondary results of the research, which shows that the pso algorithm has a high power in predicting stock prices in relation to the Chaid algorithm, it is recommended that capital owners and company decision makers use their ability to predict artificial intelligence algorithms, especially the pso method, in their decisions on stock investing.

References

[1] Abdeh Tabrizi, H., Joohari, A., *The Investigation of the Efficiency of Stock Price Index of TSE*, Financial Research, 1996, **11**(12), P.47-61.

[2] Aghajan Nashtaee, R., Daghigh Rohi,S. *Investigating the relationship between financial ratios and stock prices in Tehran Stock Exchange*. 6th International Conference on Economics, Management and Engineering, Belgium,International Center for Academic Communication of Iran, 2015. Doi:ICOAC01_163 (in Persian).

[3] Alijani, M., Banimahd, B., Madanchi, M. Study and Research on the Six-Year Process of Bitcoin Price and Return. Advances in Mathematical Finance and Applications, 2019, **4**(1), P.45-54. Doi:10.22034/amfa.2019. 577434.1126

[4] Andrea ,B., Matteo, M. *The impacts of oil price shocks on stock market volatility: Evidence from the G7 countries*.Energy Policy, 2016, **98**(c), P.160-169. Doi: 10.1016/j.enpol.2016.08.020

[5] Ashraf E. Elbakrya, J. Nwachukwub, H. Abdouc,d. *Comparative evidence on the value relevance of IFRS-basedaccounting information in Germany and the UK*,Journal of International Accounting. Auditing and Taxation 2016, **28**, P.10–30. Doi:10.1016/j.intaccaudtax.2016.12.002

[6] Akbari, S., Dadras ,M. *Investigating Financial Ratios with Stock Prices in Industries Affiliated to the* Agricultural Sector in Iran's Stock Exchange, Agricultural Economics Research, 2017,**9**(1), P.176-165, (in Persian).

[7] Barzegari Khanagha, J., Jamali, Z. *Predicting Stock Returns with Financial Ratios; An Exploration in Recent Researches*. Journal of Accounting Research, 2016, **6**(2), P.71-92. Doi: 10.22051/ijar.2016.2432.

[8] Alimohamadi, A., Abbasimehr, M., Javaheri, A. Prediction of Stock Return Using Financial Ratios: A Decision Tree Approach. *Financial Management Strategy*, 2015, **3**(4), P.125-146. Doi: 10.22051/jfm.2016.2349.

[9] AhmadKhanBeygi, S., Abdolvand, N. *Stock Price Prediction Modeling Using Artificial Neural Network Approach and Imperialist Competitive Algorithm Based On Chaos Theory*. Financial Management Strategy, 2017, **5**(3), P.27-73. Doi: 10.22051/jfm.2017.14635.1319.

[10] Bahram Far, N., Mehran, K.*The Relationship Between Earnings Per Share, Dividend Profit and Investment in Companies Listed in Tehran Stock Exchange*, Accounting and Auditing Reviews, 2004, **11**(36), P.27-46.

[11] Bartov, E., Partha, M. *Does Income Statement Placement Matter to Investors? The Case of Gains/Losses from Early Debt Extinguishment Leonard N*. Stern School of Business, New York University, New York, 2012. Doi:10.2308/1558-7967-89.6.i.

[12] Benjamid, B.,Moradzadefard, M., Moorwarid Araqi, F. *Positive Change of Earnings Per Share and Share Price: Functional Fixation Hypothesis*, Management Accounting, 2014, **7**(1),P.67-77, (in Persian).

[13] Bayat, A., Bagheri, Z *Stock price prediction using firewall algorithm*. Journal of Financial Knowledge, Securities Analysis, Tenth, Thirty-Fifth Edition, 2017, P.135-146 ,(in Persian).

[14] Bigazadeh Abbasi, A., Bigzade Abbasi, F.,and Porkakani, M. *Stock price forecast using Evidence from the Tehran Stock Exchange using genetic algorithm*. International Conference on New Research in Management, Economics and Accounting, Istanbul, Institute of Managers of the Ideas Capital, Iliya, 2015, (inPersian).

[15] Chia,L.,Yung,L. *A double-threshold GARCH model of stock market and currency shocks on stock returns* ، Article in Mathematics and Computers in Simulation,2008, **79**(3) ,P.458-474. Doi:10.1016/j.matcom.2008.01.048.

[16] Dibachi, H., Behzadi, M.H., Izadikhah, M., *Stochastic multiplicative DEA model for measuring the efficiency and ranking of DMUs under VRS technology*, Indian Journal of Science and Technology,2014, **7** (11), P. 1765–1773.

[17] Efron, B., et al. *Least angle regression* , Stanford University,2004,32. **32**, P.407–499. Doi:10.1214/009053604000000067

[18] Eghdami, I., Seyyed Javadifimeh, S., Maleki Khavari, M., Hosseinzadeh, A. *A survey of small and large companies in terms of financial performance*. Second National Conference on Accounting, Management and Economics, 2014, P.1-13, (in Persian).

[19] Foroughi, D., Faraghie, M . *The Effect of Price Coincidence and Stock Returns on Stock Liabilities in Companies Accepted in Tehran Stock Exchange*, Journal of Management and Finance, 2015, **4**(1), P.85-98,(inPersian).

[20] Golestan, A ., Molaei Zahedi, Z., *Investigating group-based optimization algorithms*. Eighth Symposium on Advancements in Science and Technology, Mashhad, Iran،2001, (in Persian).

[21] Huddart ,S. *Psychological factors, stock price paths, and trading volume, Management science*. University of North Carolina at Chapel Hilland,2006.

[22] Islami, M., Abadi, H., Saeedi, P., Khalilzadeh, S. *Investigating the Effect of Internal Factors on Stock Price Change in Investment Companies Accepted in Tehran Stock Exchange*, Accounting Research, 2017, **6**(4), P. 103-138, (in Persian). Doi: 10.22051/IJAR.2017.2080.

[23] Jandagian, A. *Stock price projections using data mining and neural networks*, Master's Theses, Khajehnesir Al-Din Tusi University of Technology, 2014, (in Persian).

[24] Jamshidi Ashkelini, E. *Study of the relationship between earnings and stock prices in Tehran Stock Ex-*

change. Master's Thesis, Tehran: Tarbiat Modares University, 1998, (in Persian).

[25] Khaleghi, P., Aghaei, M., Rezaei, F. Salience Theory and Pricing Stock of Corporates in Tehran Stock Exchange. *Advances in Mathematical Finance and Applications*, **3**(4), 1-16. 2018. Doi:0.22034/amfa.2019.577140.1120

[26] Khoshroo, A., Izadikhah, M., Emrouznejad, A., *Improving energy efficiency considering reduction of CO2 emission of turnip production: A novel data envelopment analysis model with undesirable output approach,* Journal of cleaner production,2018, **187** (20), P. 605-615. Doi: 10.1016/j.jclepro.2018.03.232

[27] Lakshmi, P.,Maheswaran, S. *A New Statistic to Capture the Level Dependence in Stock Price Volatility*.The Quarterly Review of Economics and Finance, 2017, **65**, P.355-362. Doi: 10.1016/j.qref.2016.12.001

[28] Moradzadeh Fard, M., Darabi, R.,and Shahalizadeh, R.*Integrating artificial intelligence techniques to provide a stock price prediction model*. Financial Accounting and Auditing Research, 2014, **6**(24), P. 89-102, (in Persian).

[29] Monadjemi, S., Abzari, M., Rayati Shavazi, A.. *Modeling of Stock Price Forecasting in Stock Exchange Market*, using Fuzzy Neural Networks and Genetic Algorithms, 2009, **6**(22), P.1-26. Doi:10.22055/jqe.2009.10697.

[30] Mehrara, A., atf, Z., and Asgari, Z. *Data mining techniques in predicting stock prices of stock exchanges*. National Conference on Accounting, Financeand Investment,2012,(in Persian).

[31]Meshki, M., Elahi Rodposhti, S. *Acounting Conservatism and the Value of Cash Holdings*. Empirical Research in Accounting, 2014, **4**(1), P.23-43. Doi: 10.22051/jera.2015.1882

[32]Mehrabian, A. "The sensitivity of the stock market to financial and monetary fluctuations." Economic Research Journal, 2004, **1**, P. 169-186, (in Persian).

[33]Nakhaii, H., Lari Dasht Beyaz, M., Fazelangi Sangani, Z., *Investigating the relationship between returns, operating cash flow ratios, accumulated profit ratios, and dividend ratios with stock price fluctuations in listed companies in Tehran Stock Exchange Quarterly*, Journal of New Research in Management and Accounting, No. Seventh, 2016, P. 107-134, (in Persian).

[34]Nasirzadeh, F., and Nikravesh, Z. *Evaluating the ability of the data mining model to predict stock prices*. 11th National Accounting Conference of Iran, 2013 ,(in Persian).

[35]Nguyen Ngoc Thanh, V. *Determinants of Stock Price Synchronicityevidence from HoChiMinh city stock exchange*.Doctoraldissertation,Vietnam,InternationalUniversityHCMC,2013. Doi:10.22161/ijaers.4.4.14

[36]Parsa, B., Sarraf, F. (2018). *Financial Statement Comparability and the Expected Crash Risk of Stock Prices*. Advances in Mathematical Finance and Applications, **3**(3), 77-93. Doi: 10.22034/amfa.2018.544951

[37]Roshandel, R., Amiri Menash, M. *Review and compare data mining techniques in stock price prediction*. Second National Conference on Computer Science and Information Technology, Babol, Scientific Research Institute of Komeh, Science and Technology, 2016. (in Persian).

[38]Roshan, S., Arian, A., Hosseini, S., Navabi Zand, K., Dari Kandeh, A. *Evaluating Factors Affecting the Earnings Rate (E / P) of the Companies Accepted in Tehran Stock Exchange*. Financial Knowledge Analysis of Securities, 2012, **14** , P. 41-57, (in Persian).

[39]Shaabani Sadrpastadeh, M. *Stock price index prediction in Tehran Stock Exchange*. Islamic Azad University, Central Tehran Branch, Faculty of Economics, Master's thesis, 2011, (in Persian).

[40]Sujeewa, K. *Impact of Firm Specific Factors on the Stock Prices: A Case Study on Listed Manufacturing Companies in Colombo Stock Exchange*. International Journal for Research in Business, Management and Accounting, 2016, **2**(3) P. 9-16.

[41]Saleh Nejad, S., Ghayour, H. *The Effect of Rate of Return on Equity and Return on Ownership and Cash Flow on Shares of Companies Accepted in Tehran Securities Exchange*, Management Quarterly, Seventh Year, 2010, **18**, P. 17-28, (in Persian).

[42]Setayesh, M., Kazemnejad, M. *Effective Factors on Disclosure Quality of the Firms Listed in Tehran Stock Exchange*, Journal of Accounting Advances, 2012 ,**4**(1), P.49-79. Doi: 10.22099/jaa.2012.514.

[43]Toloi Ashlaghi, A, and Haghdoust, S. *Modeling stock price prediction using neural network and comparing it with mathematical prediction methods*. Journal of Economic Research, Summer, 2007, **25**, P. 237-252, (in Persian).

[44]Yu, Lean,W., Lai, K. *Neural network-based mean–variance–skewness model for portfolio selection*. Computers & Operations Research, 2008, **35**(1), P 34-46. Doi:10.1016/j.cor.2006.02.012

[45]Yahyizadehfar, M., Hijabi, R. *The Relationship between Price to Profit with Systematic Risk and Stock Returns in Tehran Stock Exchange*. Business Strategies, 2010, **8** (41), P. 325-338, (in Persian).

[46]Yang, S., and Fu, L. *Critical chain and evidence reasoning applied to multi-project resource schedule in automobile R&D process*, International Journal of Project Management, 2014, **32**(1), P. 166-177. Doi:10.1016/j.ij proman.2013.01.010.

[47]Zamani, A., Anvary Rostamy, A., Badavar Nahandi, Y.,and Saeedi, A. *The Effect of Agency Costs on the Speed of Stock Price Adjustment: Evidence of Tehran Stock Exchange*. Financial Management Strategy, 2018, **5**(4), 25-44. Doi: 10.22051/jfm.2018.15774.1403.