Adv. Math. Fin. App., 2024, **9**(3), P.1110-1127

Advances in Mathematical Finance & Applications

www.amfa.iau-arak.ac.ir

Print ISSN: 2538-5569

Online ISSN: 2645-4610

Original Research

Presenting a Model for Predicting Various Types of Stock Movement Trends in Water-Intensive Industries Using a Decision Tree

Mohammad Bahmani, Ali Lalbar*

Department of Accounting, Arak Branch, Islamic Azad University, Arak, Iran

ARTICLE INFO

Article history:
Received 2022-12-01
Accepted 2023-12-23

Keywords:
Profit Movement Trend
Industry Movement Trend
Stock Price Trend
Water-Intensive Industries
Decision Tree Approach

ABSTRACT

Stock market activists are trying to find and apply methods so that they can increase the profit of their capital by predict-ing the future stock price. Therefore, it seems necessary that appropriate, correct and scientific-base principles methods are used to determine the future price of stocks for investors. Economists use econometric methods for forecasting in most cases. Therefore, the aim of this research was to present a new approach in predicting the types of stock movement trends in water-intensive industries using the decision tree approach. The local domain of this research was the compa-nies listed on the Tehran Stock Exchange active in water-intensive industries and the time domain was during 2016-2020. In the data section, the research was done by collecting the data of the sample companies by referring to the financial statements, explanatory notes and the stock exchange monthly journal; Based on the systematic elimination method, 72 companies active in water-intensive industries were selected as a statisti-cal sample. In order to describe and summarize the collected data, descriptive and inferential statistics have been used us-ing the decision tree approach; the results showed that by using the decision tree approach, it is possible to predict the profit, industry and stock price trends in water-intensive in-dustries. The results obtained in this research are consistent with the documents mentioned in the theoretical framework of research and financial literature.

1 Introduction

Due to the fact that water-intensive industries are very important in countries like Iran, and also the trend of the stock market movement in these industries is very important, forecasting these trends is very important. Currently, there are many stock price forecasting methods and algorithms in the financial market that predict the future trend of these prices through the historical analysis of stock prices and financial indicators. But in water-intensive industries, factors such as weather conditions, climate changes, political changes, etc. can affect the shares of these industries. Therefore, it is necessary to provide a new approach to predict these trends by considering environmental, political and economic

* Corresponding author. Tel.: *+989183660661*
E-mail address: *a-lalbar@iau-arak.ac.ir*

factors effective in these industries. In this regard, the use of deep learning algorithms and neural networks is able to understand and analyze these factors and can help decision makers in the capital market by providing higher accuracy in forecasting. Accurate prediction of the stock market is valuable for the traders of this market. Forecasting financial time series is one of the most challenging and important issues in forecasting, and researchers try to extract hidden patterns to predict the future of the stock market. The most important challenge in this matter is increasing the prediction efficiency and the accuracy of the results, which researchers are trying to improve by presenting different models and methods; Of course, due to the fluctuating and unstable nature of the stock market, models face limitations and problems in forecasting; In these models, most of the researchers have predicted the stock price and have used criteria that show the closeness of the predicted price to the actual price to evaluate [21]. There is a research gap in this field, among various psychological factors, false self-confidence, optimism and pessimism are the emotions that have been the focus of researches and researches more than others [29, 28, 27, 3] Several studies have measured false confidence based on the type of stock price effect on turnover. According to the findings of these authors, there may be a possible relationship between overconfidence (under-confidence) in the investor and the reaction of the stock market. Other authors have investigated the relationship between turnover and stock price trends, although the results sometimes reach agreement, mostly a positive relationship has been reported in these studies. It seems that the stock price, in particular, has a positive and significant effect on the turnover [13]. In recent years, it has been stated in various sources that stock prices have a non-linear relationship with financial and economic variables, and artificial intelligence techniques are very suitable tools for solving non-linear problems. In some sources, the market index has been used for forecasting. One of the reasons for this is that the market index is a complete symbol of all the companies in the market, and the efficiency of the market is determined based on it, and it gives a more general view to the investors and represents the general trend of the market; choosing the right inputs is one of the main factors influencing the effectiveness of the prediction model; the decision tree algorithm is able to correctly discover hidden information if appropriate inputs are selected [9].

Other recent studies have paid more attention to the psychological state of investors and their impact on the liquidity of the stock market. It seems that optimism and pessimism theoretically have a significant effect on investors' trading strategies. Moreover, it seems that these have particularly asymmetric effects. In fact, in theoretical forecasts, it is assumed that the turnover reacts asymmetrically for analyzing optimism and pessimism and predicting income. According to these forecasts, we assume that both optimistic and pessimistic investors show different reactions based on their expectations. The more optimistic the investors are, the more their turnover increases when predicting the increase in the price of a stock. One hopes that this increase will generate a new leverage factor that will allow for greater profits in the future. When prices fall, he continues trading as usual because he is less sensitive to negative outcomes. At the same time, more pessimistic investors anticipate that any price decline will be accompanied by a new decline and accordingly seek to avoid further losses and therefore reduce their turnover. However, since they are less sensitive to positive outcomes, when they predict that the price of a stock will increase in the future, they still normally continue trading [13].Stock market activists are trying to find and apply methods so that they can increase the profit of their capital by predicting the future stock price. Therefore, it seems necessary that appropriate, correct and scientific-base principles methods are used to determine the future price of stocks for investors. Economists use econometric methods for forecasting in most cases. In the meantime, Arima and linear regression processes are considered to be the most widely used methods in forecasting. In recent years, in parallel with the significant progress in the rapid processing of information by electronic machines, the use of nonlinear models among economists has increased dramatically; one of the most important non-linear models that has been widely used in the financial markets in recent years and has achieved favorable results is the decision tree algorithm. The purpose of this research is to use the decision tree algorithm in predicting stock price movement trend and investigate its accuracy.

2 Theoretical Foundamentals and Research Background

Momentum in physics means that if a closed system is not affected by an external force, it will continue its path without changing the amount of instantaneous movement or the force of motion. The concept of momentum has also been used in financial sciences which is referred to as the momentum of stock returns (price). The momentum of stock returns refers to the increase (decrease) in stock returns following the previous increase (decrease) in stock returns. Past researches have shown that investors achieve additional returns by using stock returns momentums. In other words, the strategy of buying stock that has been winner in the past (stocks with the highest return in the past) and selling stock that has been loser in the past (stocks with the lowest return in the past). momentum is due to sequential covariance and not positive correlation itself, that results in behavioral factors. He argues that firm-specific returns and behavioral models cannot explain much of the momentum. However, the existence of the momentum phenomenon of stock returns means predicting the price and return (using the price and past stocks return) in different time horizons, and this concept violates the assumptions of an efficient market and shows that capital market investors does not analyze published information well [12]. One of the main reasons for the existence of momentum in the market is the behavioral tendencies and biases of investors. The existence of such behavioral disorders in the stock market indicates that investors do not always act rationally and make irrational decisions under the influence of their psychological characteristics. Academic studies show that portfolio managers and professional market participants believe that buying and selling based on strategies such as momentum can lead to additional returns. Today, these strategies are the dominant investment strategies in the world's stock markets and are widely used by individual and institutional investors [2]. The decision tree, like the tree structure, consists of different nodes such as root node, middle node and leaf node. The decision tree is the most widely used technique in data mining for classifying a large amount of data and extracting datasets that have similar patterns. One of the classification tools is the decision tree, which is widely used in this field. In addition to quantitative variables, this tool also classifies qualitative variables. The decision tree structure consists of internal nodes, leaves and roots. Classical forecasting approaches are not able to quickly respond to the needs of organizations in today's changing environment, and they lose their efficiency with the rapid change in the internal and external environment of organizations. All classical approaches prepare strategies based on definite predictions of the future. Predictions and analysis of current conditions and past experiences are used to map the course of events and future process. The more chaotic the environment and the faster the rate of changes, the higher the probability of these predictions not coming true; therefore, it seems necessary to provide a method that can solve this problem. In this article, a new and systematic method called decision tree is used to predict the stock movement trend. Fuzzy decision tree approach is a soft computing approach in artificial intelligence and computer science. By examining its formulation, it can be found that its use in management decisions and solving strategic planning problems is possible, the decision tree approach is used in many fields, such as: pattern recognition, pattern classification, classification, decision-making support system, expert systems, etc., because it is the fastest compared to other methods, including probability maximum. Especially in situations where the sample space is large, in addition to this, it is easy to provide data and it is easier for non-technical people to understand it. Another advantage is that it can classify both numerical and rank data. Decision tree is used successfully in the fields of financial management (in exchanges and transactions), stock market information, administrative review, business law management (project quality analysis, product quality management, feasibility study), banking and insurance (risk review and prediction), environmental sciences (environmental quality analysis, integrated resource analysis, disaster assessment), decision-making in diagnosis and selection of appropriate treatment, etc. The main advantage of the decision tree approach is to show solutions. Inside the country, Bodaghi et al. investigated the effect of financial information on the stock prices movement in winning and losing portfolios using data mining methods: neural networks and decision-making trees and show that the variables related to the profitability of the company has had the greatest impact on the future return of stocks in both winning and losing portfolios[2]. This result is a very important guideline for investors in the stock exchange. Abdul Baghi Attabadi et al. investigated the effectiveness of momentum on industry volatility and show the existence of returns due to short-term momentum behavior in the future periods of one month and two months later in the winning portfolios with high volatility compared to the losing portfolios with low volatility. Also, the average short-term future return of winning companies with high volatility is higher than winning companies with low volatility. Also, in some short-term periods in low volatility conditions, the average future return of the winning companies are more than losing companies with low volatility. Finally, most of the momentum effect is related to the buying position of winning portfolios with high volatility and selling positions of losing portfolios with low volatility, which indicates the intensification of momentum in high volatility conditions[22]. Gholami in investigating the effect of fundamental variables on the power of stock movement among small and large shareholders show the effect of fundamental variables including asset turnover; profitability; the debt ratio and the earnings per share ratio are more visible on the strength of the size of the stock movement among small shareholders than large shareholders[8]. Asadi and Emami investigated the design of trading strategies based on the effect of momentum and return and by using the important floor and ceilings of the past stocks. The research results indicate a significantly higher efficiency of using these strategies compared to the purchase and maintenance method[1]. Susan Abadi, investigated the effect of investors' trust and behavior on price trend and market transactions in companies with the opportunistic behavior nature in financial reporting[5]. Gholamian and Davoudi, in their research entitled predicting the price trend in the stock market using the random forest algorithm, investigated the mentioned issue; they show that the accuracy of the proposed method in estimating the market trend is 64% and it is more accurate than the two compared methods of logistic regression and completely random method[9]. Pakrani in his research titled: stock price movement trend prediction using XCS based on genetic algorithm and reinforcement learning, investigated the mentioned issue; the results indicate that the proposed model has a higher prediction accuracy compared to the random walk model[4]. Taheri investigated the profitability comparison of single-variable and combined momentum strategies. The results show that the single-variable momentum strategies of profit, income and price individually are all profitable. But in the combined strategies, only the combined momentum strategies of price and profit, price and income and also the combined momentum strategies of price, profit and income were profitable for all periods of 3, 6, 9 and 12 months[23]. Abroad as well, Y Shi et al. investigated stock movement prediction with sentiment analysis based on deep learning networks. They found that sentiment features extracted from comments were indeed effective for stocks with higher price-to-book value and lower risk amounts[24]. F Su and X Wang investigated investor attention and co-movement of stock returns: evidence from the Chinese stock market. Empirical findings show that the joint movement of both parties increases the investors' attention and increases the stock returns, in short, an alternative explanation for the mutual movement is provided by observing the causal relationship between the investors' attention and the cooperation of stock returns[14]. Hu changsheng & Wang Yangfeng investigated the effect of attitude on stock prices and arbitrage limit in Chinese stock market. The existing histories show that they explained that the investor's attitude has an effect on the sensitivity of the stock price to the news related to the specific earnings of the company, and in other words, the sensitivity of the stock price to the good news about the profit and earnings in periods with more sentiment than the ones with less sentiment are far more[16]. Ni et al. addressed this issue in their research entitled "investigation of the impact of the characteristics of the board of directors and the company's financial structure on the speed of stock price movements, the results show that the top 10 shareholders (major shareholders) can increase the speed of price movement change; while institutional investors can reduce the speed of price movements[19]. Irwan has addressed this issue in a research titled the effect of external and internal instability (imbalance) on the relationship between returns and turnover. In this research, imbalance is divided into external and internal organizational imbalance. Foreign and domestic imbalances significantly affect the daily return changes in the Indonesian market. The effect of external imbalance is shown in larger stocks, while internal imbalance in stocks is smaller than a certain limit[17]. Ehtashami et al. in a research titled "Forecasting Stock Trend by Data Mining Algorithm" investigated stock trend forecasting with data mining algorithm. This research has two hypotheses. The hypotheses were analyzed based on the data collected from 180 companies admitted to the Tehran Stock Exchange during the years 1388 to 1394. The results showed that the algorithms are able to predict negative stock returns. However, random forest algorithm is more powerful than decision tree algorithm[32]. Ivani et al. investigated the mentioned issue in a research entitled "Developing a Prediction-Based Stock Returns and Portfolio Optimization Model". The empirical evidence is based on the analysis of 112 unique companies admitted to the Tehran Stock Exchange from 2009 to 2019. Regression analysis as well as six decision tree techniques including CHAID, ID3, CRUSE, M5, CART and M5 have been used to determine. The most effective variables for predicting stock returns. The results show that six decision tree methods work better than the regression model in choosing the optimal portfolio. Further analysis shows that the CART model outperforms the other five decision tree models using Akaike and Schwartz Bayesian. The findings show that the selected model correctly predicts the portfolio return[33].

3 Research Methodology

Forecasting methods are divided into two main groups, qualitative methods and quantitative methods, based on the degree of dependence on mathematical and statistical methods. In quantitative methods whose operations are completely mathematical, past data are analyzed with the aim of predicting the future value of the desired variable. In general, quantitative forecasting methods can be divided into regression and non-regression categories. One of the non-regression methods is the decision tree method. In this research, we use the mentioned method to predict various types of stock movement trends. Therefore, the current research method is analytical-inferential according to the different stages of the research. In addition, the current research is practical in terms of its purpose, and based on the method of data collection is non-experimental and examines the relationships between the variables and describes the variables and finally presents the model. In addition, it is placed in the field of post-event studies (use of past information) and is based on the real information of the financial statements of the companies listed on the Tehran Stock Exchange and other real information that will be generalizable to the entire statistical population by inductive method. Research hypotheses

As follows:

Hypothesis 1: Using the decision tree approach, it is possible to predict the profit trend in water-intensive industries.

Hypothesis 2: Using the decision tree approach, it is possible to predict the industry movement trend in water-intensive industries.

Hypothesis 3: Using the decision tree approach, it is possible to predict the stock prices movement trend in water-intensive industries.The statistical population of this research is the companies listed on the Tehran Stock Exchange in water-intensive industries, which were present in the Tehran Stock Exchange during 2016-2020, which based on the systematic elimination method, 72 companies were selected as the statistical sample of the research to test the statistical hypothesis. Dependent variable of the research and Types of stock movement trend including:

Profit movement trend (RV1): In this type of profit trend continuation, it is argued that the stocks that have recently had a profit surprise will act in the same direction in the near future. Earnings surprise (SUE) is defined as the difference between what the market expects the company to earn and what the company actually earns and is calculated as equation (1):

$$\mathrm{SUE_{it}} = \frac{\mathrm{e_{iq}} - \mathrm{e_{iq-4}}}{\sigma_{\mathrm{i,t}}} \tag{1}$$

Which is equal to equation (2):

$$\text{Surprise in profit} = \frac{\text{tree} - \text{months real EPS} - \text{three} - \text{months predicted EPS}}{\text{Unexpected profits standard in past 8 seasons}} \tag{2}$$

Finally, the profit surprise (SUE) measure is used to calculate the profit trend (RV1).

Industry Movement Trend (RV2): It is an industry movement trend that claims that industries that have performed well or badly in the near past will also provide this performance in the future. For example, Markowitz [31] discovered a strong movement trend effect across industries. In this way, when past winning industries were bought and past losing industries were sold, additional returns were created. The winner industry is the industry whose constituent stocks have the best performance based on this criterion, and the loser industry is the industry whose constituent stocks have the worst performance in terms of the standard unexpected profit variable, then calculated the cumulative unexpected profit for the shares of each industry for the next 4 seasons and formula (3) is used to obtain the performance of each industry:

$$\mathrm{ACAR_{IN,Z,X}} = \sum_{t=1}^{T} \frac{\sum_{I}^{N} \mathrm{SUEi,t}}{\mathrm{N}} \tag{3}$$

IN: indicates industries

Z: indicates micro-periods or seasons.

K: indicates the number of seasons in which the performance of industries is tested.

If there is an industry movement trend strategy, industries that have performed well in the past will continue to perform well in future seasons, while industries that have performed poorly in the past seasons will act badly in the near future. For this purpose, the average cumulative unexpected profit (ACAR) has been calculated for industrial portfolios. In other words, if the strategy of the industry movement trend causes unexpected profits, then the winner industries in the past seasons will be winners in the coming seasons and the losers will remain losers. Finally, the industry movement trend (RV2) is equal to the value obtained for ACAR.

Stock price movement trend (RV3): Another type of movement trend is the movement trend of the stock price, in which the stocks that perform better compared to the market returns are selected and are kept for a certain period of time; with this approach, additional returns are obtained. In fact, according to this approach, stocks that have had abnormally higher returns in the past period will have higher returns in the future. Abnormal return is the difference between the actual return and the expected return of the stock.

The actual return of each ordinary share is determined according to the following:

A) Share price volatility during the investment period

B) Cash profit per share

C) Benefits arising from the pre-emptive right to purchase shares

D) Dividend or bonus dividend

Usually, different models have been used for the expected return of the share. Financial and investment researchers have recommended certain models for specific conditions. In most of the researches related to over- or under-expected reaction, two models of market adjusted return and market model have been used.

For example, Albert, Henderson, Keller, Thomas, Kriznoskiv, Zeng have used the market adjusted return model. In the share price movement trend, the adjusted return model of the market is also used. According to this model, it is assumed that the expected return is the same for all securities and the return of each security is similar to the market return, hence we will have:

$$E(R_{i,t}) = Rmt \tag{4}$$

$E(R_{it})$: the expected return of share i in the period t

R_{mt}: the market return in the period t

Abnormal return for each stock is defined by model (5):

$$\mu_{i,t} = R_{i,t} - R_m \tag{5}$$

R_{mt}: return in period t

R_{it}: Real return on share in period t

Then, the cumulative abnormal return adjusted to the market return (CAR) in the observation period, i.e. the period in which its effect on the current return of the share is investigated, is calculated as equation (6):

$$CAR = \sum_{t=1}^{T} \mu_{I,t} \tag{6}$$

Based on this approach, stocks that have had a higher CAR in one period will have higher returns in the future. Finally, the stock price movement trend (RV3) is equal to the value obtained for CAR.

Independent variables As follows:

Imbalance in sales (purchase) orders: The information related to the Rial value of buying and selling for the period of 4 trading days before and after the assembly has been collected from the library information bank of Tehran Stock Exchange Organization. To calculate the imbalance index in orders, the following equation is used; also the basis for identifying buying and selling transactions is using the closing price of the previous day, so that if a transaction is made at a price higher than the closing price of the previous day (positive domain of price volatility), it is considered as buying transaction, and if a transaction is made at a price lower than the closing price of the previous day (negative domain of price volatility), we call it as sale transaction; in model (7) we have:

$$\mathrm{NEI} = \frac{\mathrm{B.V} - \mathrm{S.V}}{\mathrm{S.V} + \mathrm{B.V}} \tag{7}$$

NEI: the imbalance index in sell (buy) orders, which is equivalent to DOI in regression equations.
B.V= Rial value of share purchase
S.V= Rial value of share sales; in model (8), we have:

$$\mathrm{B.V} = \sum_{t=1}^{n} (\mathrm{Bvol}_{i,t} \,.\, \mathrm{N}_{i,t} .\, \mathrm{P}_{i,t}) \tag{8}$$

Bvol_{it} : Purchase order volume
N_{it} : Number of purchase orders
P_{it} : share purchase price; in model (9) we have:

$$\mathrm{S.V} = \sum_{t=1}^{n} (\mathrm{Svol}_{i,t} \,.\, \mathrm{N}_{i,t} .\, \mathrm{P}_{i,t}) \tag{9}$$

Svol_{it} : Purchase order volume
N_{it} : Number of purchase orders
P_{it} : share purchase price

Number of Transactions: It is the number of times a company's shares are used in a certain time interval in Tehran Stock Exchange Organization.
Price to Profit Per Share: It is obtained by dividing the price of the share by the amount of profit per share [11].
Moving Average: It is an index that shows the average price of a share during a period of time with special emphasis on new prices, during five days.
Momentum: In this research, the momentum return is obtained from the difference between the cumulative return of stocks in one previous period (last month) and the cumulative return of stocks in the previous nine periods (last nine months) [11].
Relative Strength Index: it measures the speed of change of price movements. In order to use this indicator, average profits (price increase) and average losses (price decrease) and relative strength in the desired period (usually period 14 is used) are calculated and according to formulas (10) and (11), RSI is calculated. [11]:

$$RSI = 100 - \frac{100}{1+RS} \tag{10}$$

$$\mathrm{RS} = \frac{\textbf{Average incremental change of closing prices of N period Average downward change of closing prices}}{\textbf{he number of periods used in the calculation}} \tag{11}$$

The numerical value of RSI can varies between 0-100, which according to experts, RSI values between 0 and 30 indicate the position and time of buying the mentioned company's shares, and RSI values between 70 and 100 indicate the time of selling its shares.
Investors' trust: this variable includes information that leads investors to overreaction. When they encounter positive returns, they increase the trading volume. Since these investors overestimate their

ability to judge, they consider the risk incurred to be less than the actual level, so they increase their turnover. For the research model, trust-based behavior works in such a way that if the market return is positive on the previous day, then there will be an overreaction for the current day and the value of transactions will increase, and so on. Therefore, the trust virtual variable is defined as follows [5]:

$$if\ \ R_{t-1} \geq o \rightarrow Co = 1$$
$$if\ R_{t-1} < o \rightarrow Co = 0$$

Investors' sentiment: In this research, the investors' tendencies variable (sentiment tendency of investors) is measured based on the model proposed by Shamsaddini [30] in his doctoral thesis.
Shamsaddini [30] examined ten variables of stock turnover ratio, dividend surplus, loss aversion effect, long-term return effect, momentum effect, size effect, share capital ratio and value premium from three perspectives of price per share to profit per share ratio P/E, price per share to book value per share P/B and price per share to net cash flow per share P/CF to calculate the investors' sentiment tendency index. The results of his research showed that out of the ten investigated variables, the effect of four variables: stock turnover ratio, share capital ratio, excess dividend and value premium from the point of view of the price to the book value ratio of each share P/B, cannot be proved in Iran's capital market; in other words, his proposed model for calculating investors' sentiment tendency index was composed of six components as described in formula (12):

$$\mathrm{SEN}_{i,t} = \mathrm{D}_1 \times \mathrm{MOM}_{i,t} + \mathrm{D}_2 \times \mathrm{LTR}_{i,t} + \mathrm{D}_3 \times \mathrm{EP}_{i,t} + \mathrm{D}_4 \times \mathrm{PCF}_{i,t} + \mathrm{D}_5 \times \mathrm{size}_{i,t} + \mathrm{D}_5 \times \mathrm{RA}_{i,t} \quad (12)$$

Shamsaddini [30] then used the abnormal return rate to calculate the long-term return effect and the momentum effect, and the profit growth rate to calculate the loss aversion effect, and the model related to formula (1) was described by formula (13):

$$\mathrm{SEN}_{i,t} = \mathrm{D}_1 \times \mathrm{AAR}_{i,t} + \mathrm{D}_2 \times \mathrm{AAR}_{i,t} + \mathrm{D}_3 \times \mathrm{EP}_{i,t} + \mathrm{D}_4 \times \mathrm{PCF}_{i,t} + \mathrm{D}_5 \times \mathrm{size}_{i,t} + \mathrm{D}_5 \times \mathrm{GP}_{i,t} \quad (13)$$

Table1: Defining the components of the investor's tendencies

D1	is a virtual variable; In such a way, if the abnormal return rate of the previous year of the company was higher than the average abnormal return rate of the market, it is equal to one and otherwise it is equal to zero.
MOM	Momentum effect obtained using D1AAR variable
AAR	The abnormal rate of return is equal to the actual rate of return minus the expected rate of return.
D2	Is a virtual variable; In such a way, if the company's abnormal return rate for the previous three years was lower than the average abnormal return rate of the market, it is equal to one and otherwise it is equal to zero.
LTR	The long-term return effect obtained using the D3AAR variable.
D3	Is a virtual variable; In such a way that if the dividend ratio of the previous year of the company was more than the average dividend of the market, it is equal to one and otherwise it is equal to zero.
PE	The ratio of price per share to earnings per share.
D4	is a virtual variable; so that if the ratio (E/CF) of the previous year of the company was higher than the average ratio (E/CF) of the market, it is equal to one and otherwise it is equal to zero.
PCF	The ratio of price per share to net cash flow per share
D5	Is a virtual variable; so that it is equal to one for small companies and zero otherwise.
SIZE	Firm size is equal to the logarithm of total assets.
D6	Is a virtual variable; In such a way that if the company made a loss in the previous year, it is equal to one and otherwise it is equal to zero.
RA	The loss aversion effect is achieved using the D6 GP variable.
GP	Growth rate of profit or loss

In the above formula, the abnormal rate of return is based on the difference between the current year's actual rate of return and the expected rate of return obtained from the capital assets pricing model; the

real rate of return is also obtained by dividing the difference between the stock price at the beginning and the end of the period plus the dividend on the stock price at the beginning of the period.
Momentum strategy: It means the movement of the market and believes that the recent trends will continue and in this strategy for a time horizon of one month, three months and six months, it can create additional returns. An example of the momentum strategy in the market is that the price trend tends to remain when an external force stops it. This strategy includes investing in the direction of the market. It claims that the positive or negative returns of the past will continue in a certain period of the future as well [10].

4 Findings

In order to analyze the data in M5 tree model, the descriptive statistics of the data under study are calculated, the descriptive statistics Table shows the amount of descriptive factors for each variable separately and for a total of 5 years. In this research, 72 companies have been investigated, and the descriptive statistics of the sample are presented in the Table below:

Table (2): descriptive statistics

symbol	Y_1	Y_2	Y_3	X_1	X_2	X_3	X_4
indices	Profit movement trend	Industry movement trend	Stock price movement trend	Imbalance in orders	Number of transactions	Price to profit per share	Moving average
average	-0.188	-0.344	-0.116	-0.39	4302.2	0.66	7330.2
mean	-0.126	-3.47	-0.156	-0.43	14266.5	0.455	4482.06
Max.	9.63	2.12	1.09	0.83	1537363	0.99	56568.3
Min.	-12.8	-8.06	-1.01	-0.96	3	-0.32	751.4
Standard deviation	2.06	2.52	0.473	0.279	11277.7	0.609	788.2
Skewness	-1.40	0.125	0.289	0.882	8.49	1.14	2.48
Kurtosis	1.18	-0.168	-0.591	1.87	6.11	2.35	8.01
observations	360	360	360	360	360	360	360

Continuation of Table 2:

symbol	X5	X6	X7	X9
indices	momentum	Relative strength index	Investors' trust	Investors' sentiment
average	-0.324	50.1	0.46	49.4
mean	-0.200	49.3	0	9.00
Max.	0.83	84.7	1	5631.8
Min.	-3.16	26.2	0	-0.18.5
Standard deviation	0.569	9.90	0.499	31.4
Skewness	-1.03	0.196	0.123	16.3
Kurtosis	1.91	0.024	-1.99	2.12
Observations	360	360	360	360

According to Table (2), the average profit movement trend is equal to -0.188 and the lowest and highest value of profit movement trend score is -12.8 and 9.63, respectively. The degree of asymmetry of the frequency curve is called skewness. If the coefficient of skewness is zero, the society is completely symmetrical, and if this coefficient is positive, it is skewed to the right, and if the coefficient is negative, it is skewed to the left. The dispersion parameter of the amount of kurtosis or flatness of the frequency curve compared to the standard normal curve is called prominence or kurtosis. If the kurtosis is around zero, it means that the frequency curve is balanced and normal in terms of kurtosis, if this value is positive, it is a prominent curve and if it is negative, it is a wide curve; which in the case of the profit movement trend variable, a slight kurtosis is observed.

4-2 Examining Research Hypotheses

At this stage, in order to investigate the factors affecting the types of stock movement trend in water-intensive industries, data mining models have been used using Weka software, and in this research, the decision tree data mining model has been used.

4-2-1 Predicting the Profit Movement Trend in Water-Intensive Industries Using the M5 Tree Model (Hypothesis 1)

The input variables in the current research include imbalance in sales (purchase) orders; number of transactions; price to profit per share; moving average; momentum; relative power index; investors' trust and investors' sentiments are used as inputs to the model; in this research, normalization was done using the minimum-maximum method and the data were placed in a numerical range between zero and one; if a variable for different years under study is in the range between zero and one, so it does not need to be normalized, (such as the variable of investors' trust). In order to standardize the data, assuming that their distribution function is normal, the following equation is used:

$$x_{new} = \frac{x_{old} - \bar{x}}{\sqrt{\sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n-1}}} \tag{14}$$

Where X_{old} is the initial value of the parameter, X_{new} is the normalized or standardized value of the parameter, and the input and output parameters are in the normal range using the mentioned formula. The performance evaluation of the decision tree model was also carried out with 80% of the training data; that the evaluation criteria of model performance at this stage include RMSE and R, which are given in Table (3):

Table 3: The Results of Forecasting the Profit Movement Trend in Water-Intensive Industries Using Input Variables and M5 Tree Model

Input	Number of laws	Amount of training (%)	correlation coefficient R	RMSE error
Imbalance in sales (purchase) orders; number of transactions; price to profit per share; moving average; momentum; relative power index; investor trust and investor sentiment	1	80%	0.64	0.266

According to the results of Table (3), the result of the tree model and the amount of correlation coefficients (0.64), the coefficient of determination is almost equal to 0.41; which indicates the average relationship between input and output variables based on the tree model; in addition, due to the relatively low level of root mean square error, it can be said that by using the input variables and the decision tree approach, it is possible to predict the changes in the profit trend in water-intensive industries (with moderate power); therefore, the first hypothesis of the research is accepted. In the following, the linear equation (regression tree) is presented:

$$y1 = B_1 * x1 + B_2 * x2 + B_3 * x4 + B_4 * x7 + \in \quad (15)$$

Y_1: profit movement trend; x_1: imbalance in sell (buy) orders; x_2: number of transactions; x_4: moving average; X_7: investors' trust

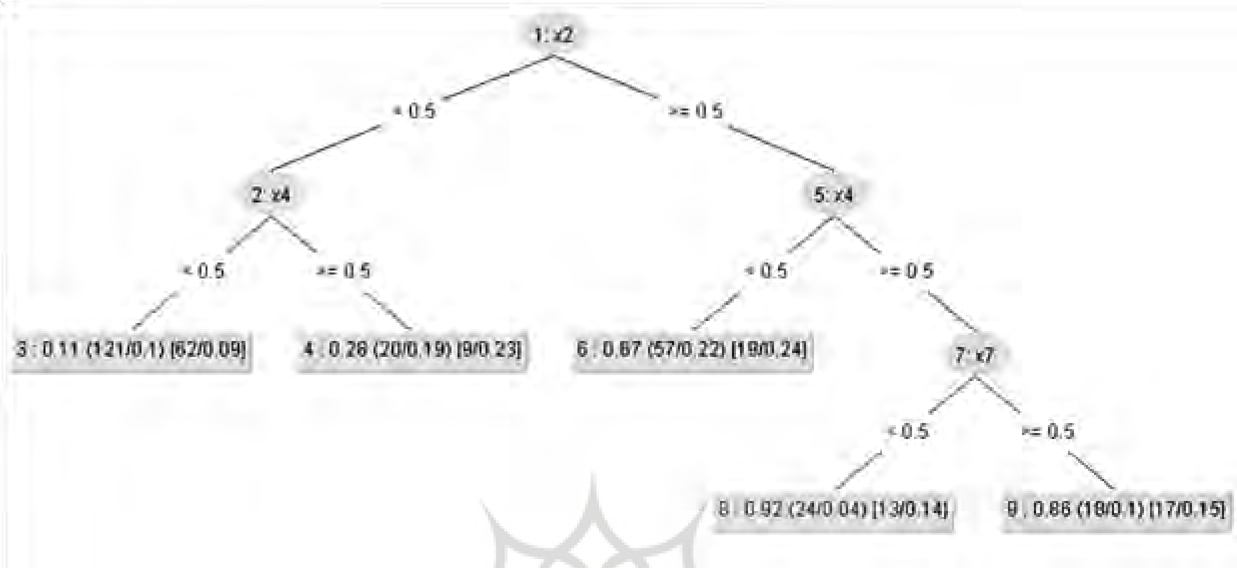

Fig.1: A Tree Model to Investigate the Relationship Between Predictor Variables and the Profit Movement Trend in Water-Intensive Industries

In Figure (1), if the number of transactions is more than 0.5, the model is effective; otherwise, if the number of transactions is less than or equal to 0.5 and the moving average factor is less than or equal to 0.5, it is possible to use the root of the number of transactions to find an optimal value for the dependent variable (here, the profit movement trend in water-intensive industries), found that this amount is optimal or efficient. Hypothesis 2:predicting the industry movement trend in water-intensive industries using the M5 tree model. The performance evaluation of the decision tree model was also carried out with 80% of the training data, that the evaluation criteria of model performance at this stage include RMSE and R, which are given in Table (4):

Table 4: the results of predicting the industry movement trend in water-intensive industries using input variables and M5 tree model

input	Number of laws	Amount of training (%)	Correlation coefficient R	RMSE error
Imbalance in sales (purchase) orders; number of transactions; price to profit per share; moving average; momentum; relative power index; investors' trust and investors' sentiment	1	80%	0.71	0.198

In Table (4), it is possible to predict the industry movement trend in water-intensive industries by using 80% of the data as training data and producing a linear relationship (number of rules); according to the correlation coefficients (0.71), the determination coefficient is almost equal to 0.50; which indicates the average relationship between input and output variables based on the tree model; in addition, due to the relatively low level of root mean square error, it can be said that by using the input variables and the decision tree approach, it is possible to predict the changes in the industry movement trend in water-intensive industries (with medium power); therefore, the second research hypothesis is accepted. Next, the linear equation (regression tree) is presented:

LM num: 1

$$y2 = B_1 * x1 - B_2 * x2 + B_3 * x3 + B_4 * x4 - B_5 * x5 + B_6 * x6 + B_7 * x8 + \in$$

y_2: industry movement trend; x_1: imbalance in sales (purchase) orders; x_2: number of transactions; x_3: price to profit per share; x_4: moving average; x_5: momentum; x_6: relative power index; x_8: investors' sentiment.

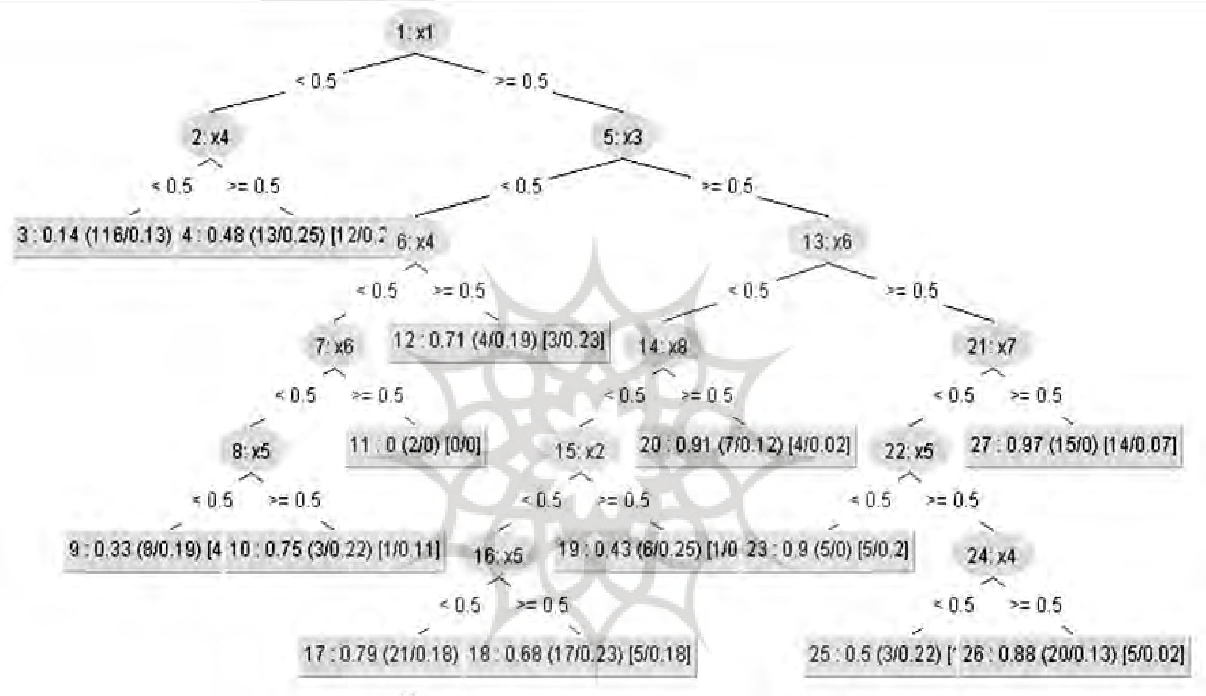

Fig. 2: Tree Model to Investigate the Relationship Between Predictor Variables and Industry Movement Trend in Water-Intensive Industries

In Figure (2), if the imbalance in sales (purchase) orders is greater than 0.5, the model is effective; otherwise, if the imbalance in sales (purchase) orders is less than or equal to 0.5, especially if the factors of the price to profit per share ratio and the relative power index is less than or equal to 0.5, it is possible to find an optimal value for the dependent variable by using the root of the imbalance in the sales (purchase) orders (here, the industry movement trend in water-intensive industries) that this amount is optimal or efficient.

hypothesis 3:Predicting the price movement trend of water-intensive industries using the M5 tree model
The performance evaluation of the decision tree model was also carried out with 80% of the training data; that the evaluation criteria of model performance at this stage include RMSE and R, which are given in Table (5):

Table 5: The Results of Forecasting the Stock Prices Movement Trend in Water-Intensive Industries Using Input Variables and the M5 Tree Model

input	Number of laws	Amount of training (%)	Correlation coefficient R	RMSE error
Imbalance in sales (purchase) orders; number of transactions; price to profit per share; moving average; momentum; relative power index; investors' trust and investors' sentiment	1	80%	0.83	0.105

In Table (5), it can be said that according to the correlation coefficients (0.83), the determination coefficient is almost equal to 0.69; which indicates a relatively strong relationship between input and output variables based on the tree model; in addition, due to the low level of root mean square error, it can be said that by using the input variables and the decision tree approach, it is possible to predict (with relatively high power) the changes in the stock prices movement trend in water-intensive industries; therefore, the third hypothesis of the research is accepted. Next, the linear equation (regression tree) is presented:

LM num: 1

$y3 = B_1 * x1 - B_2 * x2 + B_3 * x3 + B_4 * x4 + B_5 * x7 + + \in$

y_2: industry movement trend; x_1: imbalance in sales (purchase) orders; x_2: number of transactions; x_3: price to profit per share; x_4: moving average; x_7: investors' sentiment

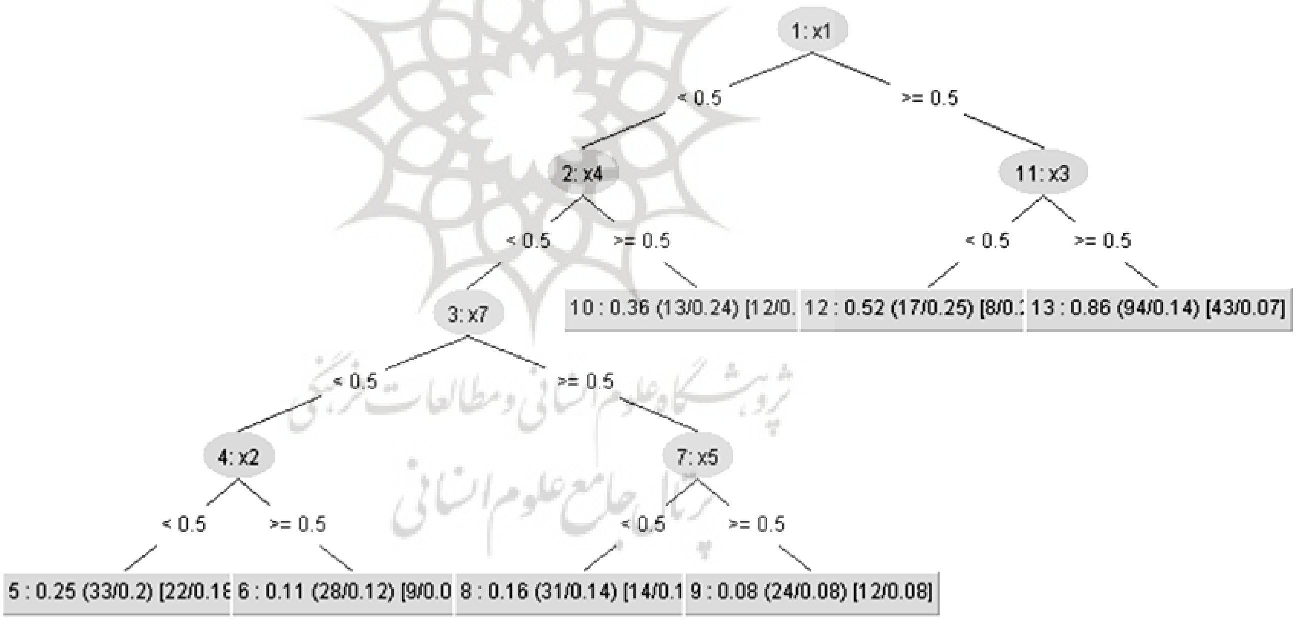

Fig. 3: Tree Model to Investigate the Relationship Between Predictor Variables and Stock Price Movement Trend in Water-Intensive Industries

In Figure (3), if the imbalance in sales (purchase) orders is less than or equal to 0.5, the model is effective; otherwise, if the imbalance in sales (purchase) orders is more than or equal to 0.5, especially if the number of transactions factor is more than 0.5, it is possible to find an optimal value for the dependent variable by using the root of the imbalance in the sales (purchase) orders (here, stock price movement trend in water-intensive industries) that this amount is optimal or efficient.

5 Discussion and Conclusion

The present research seeks to present a new approach in predicting the types of stock movement trend in water-intensive industries by using the decision tree approach and finally we came to the conclusion that the prediction of profit movements using input variables (imbalance in sale (purchase) orders, number of transactions, price to profit per share, moving average, momentum, relative strength index, investors' trust and investors' sentiment), is possible; according to the results of the tree model, it can be said that the determination coefficient is approximately 0.41; which indicates the average relationship between the input and output variables based on the tree model; in addition, due to the relatively low level of root mean square error in the model, it can be said that by using the aforementioned input variables and the decision tree approach, it is possible to predict the profit movement trend in water-intensive industries. There are two basic theories regarding behavioral models and stock price forecasting and profit movement trend. The first theory is that the stock market is unpredictable, on which hypotheses such as the efficient market and the random walk theory are formed based on it, and the second theory is that the stock market and its behavioral models are predictable and this claim is used as a basis for various stock price forecasting methods, such as technical analysis, fundamental analysis, etc.; The results of the present research are in line with the behavioral model theory. In this regard, Bodaghi et al.[2] showed that financial information is able to predict the stock price movement in win and loss portfolios using data mining methods such as neural networks and decision trees; which is in line with the results of the present study. Also, in the second hypothesis, we came to the conclusion that forecasting the industry movement trend using input variables (imbalance in sale (purchase) orders; number of transactions; price to profit per share; moving average; momentum; relative strength index; investors' trust and investors' sentiment), is possible; according to the results of the tree model, it can be said that the determination coefficient is approximately equal to 0.50; which indicates the average relationship between input and output variables based on the tree model; in addition, due to the relatively low level of root mean square error in the model, it can be said that by using the aforementioned input variables and the decision tree approach, it is possible to predict the industry movement trend in water-intensive industries. Among the techniques that have found great importance in predicting the industry movement trend were intelligent systems; because it is very difficult to show the behavior of complex systems such as the capital market in a modern economic system completely in a set of simple and linear equations. The main advantage of non-linear systems such as decision trees is in modeling and predicting irregular and non-linear sets; the results of the current research also showed that by using the decision tree approach, it is possible to predict the industry movement trend in water-intensive industries. In addition, due to the high importance of the two transaction variables of order imbalance in predicting the movement of the industry, it can be said that investors are interested in the past price and turnover. Turnover is an important reference point that investors evaluate related news. Therefore, the turnover and the sale and purchase amount are predictors of the amount and continuity of industry movement. In this regard, kamble et al.[25] showed that short-term and long-term forecasting of stock trends can be achieved by using a decision tree which is in line with the results of the present study. In the third hypothesis, we came to the conclusion that the prediction of stock price movement trend using input variables (imbalance in sell (purchase) orders, number of transactions, price to profit per share, moving average, momentum, relative strength index, investors' trust and investors' sentiments), is possible; according to the results of the tree model, it can be said that the determination coefficient is approximately equal to 0.69; which indicates a relatively strong relationship between input and output variables based on the tree model; in addition, due to the low level of root mean square error in the model, it can be said that by using the aforementioned input variables and the decision tree approach, it is possible to predict the stock price movement trend in water-intensive industries. The efficient market hypothesis is based on the belief that it is not possible to beat the market and obtain a higher return than the market average. Also, changes in stock prices are random and actually follow a random step, so it is not possible to achieve abnormal returns (above the market average) using historical information. Also, this hypothesis claims that there is no trend in the market price and efficiency, and it is not possible to profit from the market trends. Nevertheless, during the last two decades, many researches have been conducted that have seriously challenged the validity of efficient hypothesis. The results of the present research are also in this direction and show that by using the decision tree approach, it is possible to predict the stock price movement trend in water-intensive industries. In this regard, Gholamian and Davoudi [9] showed that by using the random forest algorithm, it is possible to predict the price trend in the stock market, which is in line with the results of the present study.According to the results of the first hypothesis, by using the decision tree approach, it is possible to predict the profit movement trend in water-intensive industries. The results of the evaluation criteria on the real data show that the proposed model based on the input variables can overcome the market fluctuations with high accuracy and can be used as a reliable and practical method to investigate the profit movement trend in the stock markets. Therefore, the use of this model by the Stock Exchange helps to accept the companies in the Stock Exchange so that the companies under investigation are evaluated and measured more accurately. According to the results of the examination of the second hypothesis, it is possible to predict the industry movement trend in water-intensive industries using the decision tree approach. According to the obtained results, the present research has useful information and has a relatively high accuracy, therefore it is recommended that researchers use this method in predicting the industry movement trend and according to the historical information of each share, optimize the factors used and use in forecasting. It is also recommended to researchers and students to carry out their future research by using the research results. According to the results of the examination of the third hypothesis, using the decision tree approach, it is possible to predict the stock price movement trend in water-intensive industries. Based on this, the severe shortage of specialists and financial analysts and the creation of a database to record their analytical reports in the stock market is strongly felt. It means people who can analyze different types of information, in such a case, it can be expected that the available information will be reflected on the prices through the actions of informed and knowledgeable people. Therefore, it is suggested that the market regulatory body take the necessary measures as soon as possible to employ the investment consultants that are referred to in the securities market law. Researchers are suggested to investigate the following topics in their future researches:

- Providing a framework for the factors affecting the stock movement trend with meta-synthesis method.
- Studying suitable models for predicting stock movement trend.
- Identifying and ranking factors affecting the stock movement using AHP - MCDM - MADM - MODM decision models.
- Use the indicators used in the research in other data mining methods such as neural networks or support vector regression and compare the results with the current research.
- The ability of financial and non-financial factors to predict the movement of stocks using genetic algorithm.

References

[1] Asadi, GH., Emami S.A., Designing trading strategies based on the effect of momentum and return using the important floors and ceilings of the past of the stock, *Quarterly Journal of Financial Science and Securities Analysis*, 2019;41(68):57-69.

[2] Bodaghi, H., Hejazi, R., Mehrabanpour, M.R., Investigating the effect of financial information on the stock returns momentum in winner and loser portfolios using data-mining methods: neural networks and decision trees, *Financial Engineering and Securities Management Quarterly*, 2021;18(2):432-460.

[3] Bozorg Asl, M., Sahebgharani ,A.A., Prediction of abnormal return based on the profit-industry-acceleration-based model in Tehran Stock Exchange, *Experimental Studies of Financial Accounting*, 2012;10(38):53-67.

[4] Pakrani, A., Prediction of stock price movement trend using genetic algorithm-based XCS and reinforcement learning, *Financial Knowledge of Securities Analysis*, 2017;1(1):39-54.

[5] Susan Abadi, R., The effect of investors' trust and behavior on price trend and market transactions in companies with the opportunistic behaviors nature in financial reporting [Master's thesis], *Arak Azad University*, 2018.

[6] Salehi, M., Farokhi Pilehroud, L., Prediction of profit management using neural network and decision tree in agricultural and textile industries of companies listed on Tehran Stock Exchange, *Agriculture Economy*, 2018; 10(4):89-105.

[7] Taheri, A., Taheri ,G., Investigating the relationship between the relevance of the accounting information value and the stock turnover of companies listed on the Tehran Stock Exchange, *Management Accounting and Auditing Knowledge*, 2019; 41(68): 203-214.

[8] Gholami, E., Investigating the effect of fundamental variables on the strength of stock movement amount among small and large shareholders [Master's thesis], *Arak Azad University*, 2020.

[9] Gholamian, E., Davoudi, M.R., Forecasting price trend in stock market using random forest algorithm, *Journal of Financial Engineering and Securities Management*, 2017; 11(7): 301-322.

[10] Fadainejad, M.I, Sadeghi, M., Evaluation of the usefulness of momentum and reversal strategies, *Management Message*, 2006; (17-18): 7-31.

[11] Falahpour, S., Gul Arzi, GH., Fatureh Chian N. ,Prediction of stock price movement trend using support vector machine based on genetic algorithm in Tehran Stock Exchange, *Journal of Financial Researches*, 2014; 2(2): 269-288.

[12] Ahmed, ASS., Safdar, I., Dissecting stock price momentum using financial statement analysis, *Accounting and Finance*, 2018; (58)3-43. Doi:10.1111/acfi.12358

[13] Dhaoui, A., Bacha. S., Investor emotional biases and trading volume's asymmetric response: A non-linear ARDL approach tested in S&P500 stock market, *Cogent Economics and Finance*, 2017; 5(1): 1274225.
Doi: 10.1080/23322039.2016.1274225

[14] Su.f., Xinyi.W., Investor co-attention and stock return co-movement: Evidence from China's A-share stock market, *The North American Journal of Economics and Finance*, 2021; 58: 4-18. Doi:10.1016/j.najef.2021.101548

[15] Grinblatt, M., Titman,S., Mutual fund performance: An analysis of quarterly portfolio holdings, *Journal of Business*, 1989; 62: 394-415. Doi:10.1086/296468

[16] Hu, CS., Wang, YF., Investor sentiment and asset valuation, *Systems Engineering Procedia*, 2020;3:166-171. Doi:10.1016/j.sepro.2011.11.023

[17] Irwan Adi, Ekaputra., Impact of foreign and domestic order imbalances on return and volatility-volume relation, *Asian Academy of Management Journal of Accounting and Finance*, 2018; 10(1): 1-19.

[18] Masry, M., The Impact of Technical Analysis on Stock Returns in an Emerging Capital Markets (ECM's) Country: Theoretical and Empirical Study, *International Journal of Economics and Finance*, 2017; 3: 7-14. Doi: 10.5539/ijef.v9n3p91

[19] Ni, Y., Huang, P., Chen, Y., Board structure, considerable capital, and stock price overreaction informativeness in terms of technical indicators, *North American Journal of Economics and Finance*, 2019;Accepted 16 July 2019. Doi: 10.1016/j.najef.2018.07.007

[20]Venkatesh, C., Tyagi, M., Fundamental analysis and Stock returns: An Indian Evidence, *Canadian Journal of Business and Economics*, 2015;1(2):26-34.

[21] Lee, X., Yang, L., Xue, F., Zhou, H., Time series prediction of stock price using deep belief networks with intrinsic plasticity, *Control And Decision Conference (CCDC)*, 2017;29th Chinese.

[22] Abdul BaghiAtaabadi, A.M., MirLohi, SM., Abdulahi ,M., Effectiveness of Momentum on Industry Aggregation: Evidences from the Stock Exchange, *Financial Management Strategy*, 2020;1(1):157-175. Doi: 10.22051/jfm.2019.19575.1624

[23] Taheri, F., Comparison of profitability of single variable and combined momentum strategies,*National conference of accounting and management of new techniques, challenges and solutionsof Iran*, 2016.

[24] Shi, Y., Zheng, K., Guo, X., Ren, X., Stock movement prediction with sentiment analysis based on deep learning networks, *Concurrency and Computation*, 2021;6:1-14. Doi:10.1002/cpe.6076

[25] Kamble, RA., Short and long term stock trend prediction using decision tree, *International Conference on Intelligent Computing and Control Systems (ICICCS)*, 2017; 3-23.

[26] Grundy, BD., Martin JS., Understanding the nature of risks and the sources of rewards to momentum investing, *Review of Financial Studies*, 2014;14:29–78. Doi:10.2139/ssrn.94049

[27] De Bondt, WFM., Thaler ,R., Does the stock market overreact, *Journal of Finance*, 1985;40(3):793-805.: Doi:10.1111/j.1540-6261.1985.tb05004.x

[28] Daniel, K., Hirshleifer, D., Subrahmanyam ,A,. Investor Psychology and Security Market Under- and Overreactions, *Journal of Finance*, 1998;53:1839-1885. Doi:10.1111/0022-1082.00077

[29] Barberis, N., Huang ,M., Stocks as Lotteries: The Implications of Probability Weighting for Security Prices, *American Economic Review*, 2008; 98(5): 2066-2100. Doi: 10.1257/aer.98.5.2066

[30] Shamsaddini, K., Setayesh, M.H., An investigation of the relationship between investor sentiment and price stocks in Tehran Stock Exchange (TSE), *Journal of Accounting Advances (J.A.A)*, 2016; 8(1):70:103-125. Doi:10.22099/jaa.2016.3854

[31] Markowitz, HM., The Early History of Portfolio Theory: 1600-1960, *Financial Analysts Journal*, 1999; 55(4): 5-16. Doi:10.2469/faj.v55.n4.2281

[32] Ehteshami, S., Hamidian, M., Hajihac, Z., Shokrollahi, S. Forecasting Stock Trend by Data Mining Algorithm. *Advances in Mathematical Finance and Applications*, 2018; 3(1): 97-105. Doi: 10.22034/amfa.2018.539138

[33] Eivania, F., Seresht, D. J., Aflatooni, A. Developing a Prediction-Based Stock Returns and Portfolio Optimization Model. *Advances in Mathematical Finance and Applications*, 2022; 7(3): 793-809. Doi: 10.22034/AMFA.2021.1877733.1304