Adv. Math. Fin. App., 2024, 9(3), P.983-1005

Advances in Mathematical Finance & Applications
www.amfa.iau-arak.ac.ir
Print ISSN: 2538-5569
Online ISSN: 2645-4610
Doi:10.71716/amfa.2024.23021862

Original Research

The Modeling the Fixed Asset Investing with a Machine Learning Approach By Emphasizing the Role of Financial Criteria

Farzaneh Shamsdoost[a], Omid Mahmoudi Khoshro[a]*,Ataollah Mohammadi Malgharni[a] ,Amir Sheikhahmadi[b]

[a] *Department of Accounting, Sanandaj Branch, Islamic Azad University, Sanandaj, Iran*
[b] *Department of Computer, Sanandaj Branch, Islamic Azad University, Sanandaj, Iran*

ARTICLE INFO
Article history:
Received 2023-02-19
Accepted 2024-01-04

Keywords:
Fixed Asset Growth
Financial Benchmark
Artificial Intelligence
Linear and non-linear Models

ABSTRACT

The purpose of this research is to provide a growth model of fixed assets based on the financial criteria of companies admitted to the Tehran Stock Exchange. The current research is applied in terms of objective classification and descriptive-correlation in terms of method. The research method is deductive-inductive. The statistical population of the current research is all the companies admitted to the Tehran Stock Exchange in the period from 2012-2021 and the financial information of 101 companies are use. Research hypotheses were tested using artificial intelligence algorithm. In this research, investment in fixed assets has been consider as a dependent variable, and financial criteria has been considered as primary independent variables. The results of research hypotheses testing using the methods of linear and non-linear algorithms of artificial intelligence PINSVR and KPLSR in predicting fixed asset investors of companies and by calculating the three errors criteria MAE, MSE and SMAPE in annual fixed assets. The asset forecasting in the next year of companies showed that the error difference between linear models and non-linear models is not so great that it can be claim that linear models are ineffective in predicting asset growth so that artificial intelligence algorithms are capable of predicting investment in company assets.

1 Introduction

One of the measures of firm growth is asset growth. Asset growth indicates that the company has funds to pay debts to related parties or investors. If there is growth in assets, investors will invest their funds in companies that have higher assets. If the company has low assets, investors can see it from another side of the company's financial statements, for instance sales growth. However, high or low assets are not a guarantee of achieving company profits. Hence, in order for convincing investors to trust and continue to invest their funds in a company, management must keep the firm's profit stable or increase from year to year by implementing an earnings management strategy [1]. One of the important factors to solve the economic problems of the countries is the expansion and development of investment, but this alone is not enough and due to the limitation of financial resources, in addition to the issue of

*Corresponding author Tel: *+989181719766*
E-mail address: *omid.khoshroo@gmail.com*

investment development, increasing the efficiency of investment is also one of the important issues. Conceptually, investment efficiency is achieved when the firm invests only in all projects with a positive net present value. In fact, some capital market defects such as information asymmetry and agency costs can lead to the process of over-investment or under-investment. In the sense that neither projects with a positive net present value (investment below the limit) nor projects with a negative net present value (investment below the limit) are reject [2]. A portfolio can be integrated across multiple assets and investments Criteria and Methodologies However, fundamental analysis is about financial review Ratios and perform an efficient evaluation for the investor to choose the asset. In between Different methods for asset selection by valuation method, preference based Competitive advantages can be a powerful tool for investors [3]. Making decisions about favorable and profitable investment opportunities is a sensitive and important issue. These opportunities are, referred to as invisible variables that do not happen by themselves, but must be identify or created. Considering that investment opportunities cause the company's financial resources to be allocated for the purpose of income or cost reduction, therefore regular and principled financial policies may be implemented by the company for investment decisions, economic theories state that long-term growth Duration depends on investment decisions [4]. If the risk management is well implemented in the business unit, it can create a competitive advantage [5]. Dealing with the issue of risk shows that reaction to risk is an important aspect that has received a lot of attention in recent years regarding increasing investment efficiency and reducing investment inefficiency. A manager is considered more capable if he shows his ability well in terms of risk response. Facing multiple risks can affect management ability. From this point of view, how to react to risk is a test field to evaluate the ability of management to influence the efficiency of investment. There is uncertainty due to the presence of risk in the operations of any company, but it is believed that how to deal with risks can (threat) turn the resulting risk into an opportunity for the growth and efficiency of the company's investments, which, of course, depends on the company's risk management [6]. Managers argue that accounting comparability reduces the cost of information acquisition and increases the overall quantity and quality of information available to decision-makers. This should help firm managers to make better investment decisions and be more efficient in research and development investments [7]. One of the improvements in new studies methods Artificial intelligence is employed, using the feature selection as a pre-stage for it is the main classifier model. However, according to research, support vector machine has an acceptable performance in investment forecasting but the accuracy of its performance is significantly affected by the number of features of the input variables. Therefore, the number of features that should be used in machine learning has a significant effect on increasing the accuracy of the results and reducing the cost. Therefore, one of the improvements in new studies has been applied, using feature selection as a pre-step, for this is the main classifier model [8]. Partial least squares (PLS) regression is commonly used for multivariate calibration of instruments [9]. The PLS method has been a popular regression technique in its domain of origin Chemometrics . The method is similar to PCR where principal components determined solely from explanatory variables creates orthogonal, uncorrelated, input variables in a regression model. In contrast, PLS creates orthogonal components by using the existing correlations between explanatory variables and corresponding outputs while also keeping most of the variance of explanatory variables. PLS has proven to be useful in situations when the number of observed variables (N) is significantly greater than the number of observations (n) and high multicollinearity among the variables exists. This situation when (N) is common in chemometrics and gave rise to the modification of classical principal component analysis (PCA) and linear PLS methods to their kernel variants [10]. By examining the aspects of research innovation, it has been observed that, no research has been done in the field of investing in fixed assets with artificial intelligence. Also, the variable selection method to reduce the error rate is consider one of the research innovations and the calculation of linear and non-linear methods in reducing prediction errors is another research innovation, and finally, no model has been presented regarding investment in fixed assets. Based on this, the researcher in this research seek to analyze and present a model for predicting the investment of companies admitted to the Tehran Stock Exchange in fixed assets based on financial quality.

2 Theoretical Bases and Background

One of the most important roles of accounting is the efficient allocation of capital, so it is not surprising that the current literature in this field focuses on the role of accounting in capital allocation decisions. Therefore, one of the common challenges of these studies is how to identify and measure investment efficiency. From the point of view of theory, the efficient allocation of capital means the circulation of capital for its optimal and valuable use. Despite this, it is difficult to observe the flow of capital and distinguish between high and low investment values [11]. Therefore, researchers from finance and accounting fields have developed different methods to identify and measure investment efficiency in advanced economies. In investment discussions, the type of capital decision Therefore, researchers from finance and accounting fields have developed different methods to identify and measure investment efficiency in advanced economies. In investment discussions, the type of decision-making by investors and the factors influencing their decision-making are very important. Financial theories and theories have had two different approaches in the last few decades. The first approach is the neoclassical approach in financial sciences. The basic premise of financial theories and theories according to this approach is the efficiency of the market and the rational behavior of investors in the market. This approach with the pricing model of capital assets and the theory of efficient markets in the limited time in the 1960s, the pricing model with intermediate capitals and Miller and Madelaine's arbitrage pricing theory began in the 1970s [12]. Every aspect of learning could be broadly defined. So that a machine can be simulated to solve all kinds of problems with features such as recognition, language use, abstract formation and general concepts. Types of machine learning systems are dividing into three categories: 1) the system trained with human supervision or not? Includes: supervised, unsupervised, semi -supervised and feedback learning. 2) Does the system see during training activities or not? Includes online learning or offline. 3) The system only compares unknown data with information or instead recognizes patterns during the training process and builds a model that can predict which includes an example to be modeled against. The advantages of artificial intelligence are reducing human error, helping in repetitive tasks, digital assistant, quick decision making, solving complex problems in new inventions and quickly learning tasks [13]. One of the characteristics of successful firms is their competitive power. Competitive power is defined as the firm's economic ability to maintain its share in international markets or increase its share in the market. The competitive environment has an essential informational role that can improve investment efficiency [14]. On the other hand, if the firm's accounting system and financial statements are more comparable with other firms, market participants such as analysts, investors, and legislators can more accurately evaluate the firm's economic performance [15]. The research results of Jahan shad and Khalili , show that there is a negative relationship in the six-month period and a positive relationship in the three-month and annual period between the growth of fixed assets and stock returns [16]. The research results of Darabi and Karimi, show that there is a significant negative relationship between the increase in the growth rate of fixed assets and short-term and long-term stock returns [17]. Haji an et al, in the revaluation of fixed assets and its effect on company performance and stock returns show that there is no relationship between asset revaluation and the company's future performance and the existence of a direct relationship between the revaluation surplus and the company's stock price. Also, revaluation through changing the debt leverage can change the stock returns [18]. Fir mania et al, in their research, investigated the effect of financial leverage, company size, liquidity and operating cash flow on the renewal of fixed assets. The results showed that financial leverage has an effect on the renewal of fixed assets. Companies with high financial leverage try to restore leverage ratios by revaluation of fixed assets. Also, operating cash flow is related to the decision to reevaluate fixed assets [19]. Lopez and Walker, in investigating the relationship between the company's future performance, stock price and return on fixed assets revaluation and the effect of the company's management mechanisms on fixed asset revaluation, showed that the company's future performance, stock price, and stock return decrease following the fixed assets revaluation. Lopez and companies that have higher debt or lower liquidity are more likely to reevaluate fixed assets [20]. Mir Mohammadi and Soleimani Amiri, in examining the relationship between debt surplus and current performance (employment rate, investment in fixed assets and financial stress) of small and medium companies in the capital market of Iran showed that the amount of debt surplus has no effect on the amount of investment in fixed assets [21]. In another research, which aims to evaluate and compare the fractal feature selection method 2 with others. Methods, including relief, 3 Sia face 4 and several other methods, using vector machine classification. The support was done by Leong, the superiority of the fractal method over other filtering methods of feature selection. It was proved [22]. According to the results of many domestic and foreign studies, the superiority of support vector machine methods over methods such as artificial neural networks and also, statistical methods have been proven. Therefore, in the following, we will mention the most researches that the relationship with the topic under investigation is discussed: Gholamreza Mansour far [23] in study examined the Investigated the moderating role of internal and external dimensions of corporate governance on the relationship between information asymmetry and investment efficiency. To achieve the objectives of the research, 106 companies selected for the period of 2009 to 2018. According to the theoretical foundations and research findings, the information asymmetry variable has a negative and significant relationship with investment efficiency. Vahid Taghizadeh et al. [24] in study examined Validate investment efficiency models based on agency theory, information asymmetry, managerial fronting, and company value maximization. For this purpose, 180 companies had use for the period of 1386-1396. The findings showed that free cash flows and financial restrictions have a positive effect on over- and under-investment, respectively. The findings indicated that investment efficiency has a positive effect on economic add value and company value, but this effect was not confirmed through all investment efficiency models. The results showed that the test of all hypotheses based on the native model of investment efficiency was confirmed. Brian Silverstein et al [25] in study examined how do managerial characteristics affect the investment efficiency of companies? Previous research determines the efficiency of a firm's investment as a function of the firm's information environment and internal governance. We examine how managerial opportunism is an agency conflict that distorts firms' investment policy. The results show that managerial opportunism reduces the company's investment efficiency and has negative effects on the company's accounting and financial performance. He, Y., Chen et al [26] studied and evaluated investment performance with a predictive machine learning approach. This paper proposes a nonlinear automatic neural network method to evaluate the investment performance of public enterprises. This method is different from the traditional method based on machine learning, such as linear regression, structural equation, Clustering and principal component analysis. In this article, the regression forecasting method is using to analyze investment efficiency. First, they analyze the relationship between ownership diversity, corporate debt, and investment efficiency of state-owned companies. In the second step, a set of investment efficiency evaluation index systems for state owned companies were made and from a Non-linear an autoregressive neural method was used for verification. Data on the shares of state-owned enterprises in Shanghai and Shenzhen from 2009 to 2018 are taking as samples. The experimental results show that the output value of the nonlinear self-regression neural network is very appropriate to the real data. Based on the regression analysis of the neural network model, this article performs a descriptive statistical analysis of the main variables and the control variables of the evaluation indicators. This confirms the direct effect of ownership diversity on the investment efficiency of SOEs and the indirect effect on the investment efficiency of SOEs through corporate debt leverage. Spyros K et al [27] in a study investigated management tools using machine learning. In this work an innovative approach to exploit the true potentials of machine learning (ML) by the financial industry is presented; using ML technology not as a source of investment ideas but as a consultant for trading decisions. In particular, the artificial intelligent risk management system (AIRMS) is present that introduces one of the first efforts in the literature to utilize supervised ML as a risk management tool. Two AIRMSs systems are develop based on two well-known ML algorithms, i.e., artificial neural networks and decision trees. These two systems are applied into the five major currency pairs (FOREX) using signals obtained from an existing technical breakout trading strategy introduced in previous study by the authors, covering a seven-year period (2010–2016). Technical indicators and times series of the past entry points feed AIRMS in order to classify produced signals from the trading strategy into two classes: profitable and not. Constructing new portfolios using signals classified only as profitable resulted in an increased profit of more than 50% compared to the original ones. In this work, technical improvements are also propos on the application of ML algorithms to financial data related to evaluation metrics and smoothing inputs. The obtained results revealed that the two AIRMSs can achieve impressive improvements to the performance of already profitable portfolios and proved that using ML to build risk management tools is very promising. Marcelina et al [28] in research to obtain empirical evidence regarding the effect of growth, financial leverage, fixed asset turnover, profitability, firm size, firm age, audit quality, board independence and managerial ownership as independent variables on focused on earnings management as a dependent variable .This research uses non-financial companies listed in Indonesia Stock Exchange (IDX) during 2016 to 2018 as population. Samples for this research were obtained through purposive-sampling method, in which 516 non-financial companies with 354 data were taken as samples. Multiple linear regression and hypothesis tests were used as data analysis method in this research. The results show that growth positively affects earnings management. This indicates that companies experiencing higher growth tends to improve earnings management practice. Meanwhile, the other variables do not have effect on earnings management. Wickremasinghe et al [29] in a study examined investigate the effect of risk management practices, how they have applied to accomplish sustainable financial performance in Sri Lanka and offer guidance to all the businesses as to how they can mitigate the risk faced by them. The research has expanded to extensive coverage of business sectors by 65 listed companies and secondary data are obtain from the annual reports publication in CSE. The statistical analysis of the multiple regression technique performed to calculate the results using e view-9 software. Risk management practices in operating cash flow have negative correlation between sustainable financial performances. The study was revealed that investment cash flow risk management practice have no correlation between sustainable financial performances. This will give detailed idea of cash flow risk management practices involvements and the importance to the listed companies in Sri Lanka.

Yang et al [30] in a study examined parametric-insensitive nonparallel support vector regression (PINSVR) algorithm for data regression. PINSVR indirectly finds a pair of nonparallel proximal functions with a pair of different parametric-insensitive nonparallel proximal functions by solving two smaller sized quadratic programming problems (QPPs). By using new parametric-insensitive loss functions, the proposed PINSVR automatically adjusts a flexible parametric-insensitive zone of arbitrary shape and minimal size to include the given data to capture data structure and boundary information more accurately. The experiment results compared with the ε-SVR, ε-TSVR, and TPISVR indicate that PINSVR not only obtains comparable regression performance, but also obtains better estimations. Jorge Daniel et al [31] in a study, they investigated a procedure that optimizes the generalization capacity of KPLS multivariate models using genetic algorithms (GA). It selects the values of the kernel function parameter and the number of components for which the value of the cross-validation coefficient Q2cum is maximum, adds preliminary tests to configure the GA and defines a convergence criterion in terms of dispersion in the estimates. GA has demonstrated a good performance in the task of optimizing KPLS with convergent solutions towards a global optimum. Huerta Ramon et al [32] discussed the profitability of a trading strategy based on training a model to identify stocks with higher or lower expected returns. In this method, a classifier is train on historical sequence sets and tested on future data. The classifier is chose to be a nonlinear support vector machine (SVM) due to its simplicity and effectiveness. The data range from 1981 to 2010. The SVM is train once per month, in order to adjust to changing market conditions. Portfolios are form by ranking stocks using the classifier output. The highest ranked stocks are using for long positions and the lowest ranked ones for short time sales. The Global Industry Classification. Standard is use to build a model for each sector such that a total of eight long-shorts portfolios for Energy, Materials, Industrials, Consumer Discretionary, Consumer Staples, Health Care, Financials, and Information Technology are formed. Without measuring trading costs, but using 91 day holding periods to minimize these, the strategy leads to annual excess returns (Jensen alpha) of 15% with volatilities under 8% using the top 25% of the stocks of the distribution for training long positions and the bottom 25% for the short ones. The results showed that linear SVMs are less efficient than non-linear SVMs. Huang and his colleagues [33] attempt to predict the stock price using the selection method. A cover feature with a system composed of several viewer layers, including a vector machine. The support and the network were nervous. Their results showed that the accuracy of predictions using the method the coverage feature selection is higher than different filtering methods. Lee Ming [34] with the aim of predicting the trend of price changes of the Nasdaq index, step by step a research with the combination of vector machine model and hybrid feature selection resulting from the combination of feature selection. Fisher and forward sequential search supported. As input features of his model used the 30 stocks available in this case, the results of Ming's research showed that the use of the support vector machine with the selection of hybrid features is more accurate in predicting the trend compared to Backpropagation neural network as well as the use of other selection methods have a feature. In the research conducted by Zebrowski [35] the support vector machine was used with actual daily trading volume, along with Fisher's feature selection, in order to predict usage trends became. As an input, he used 7 technical analysis indicators to forecast stock price trends and used by showing the superior performance of his proposed model, he showed the superiority of this model over the model He proved simple support vector machine and support vector machine without feature selection. In a research conducted by Monajmi, Abarzaei and Ra'iti [36] stock prices in the market was predicted by fuzzy neural network and genetic algorithm. The results of their research showed that the proposed hybrid model performs better both in terms of accuracy and speed.

3 Research Methodology and Hypotheses

This research has two hypotheses as follows:

Hypotheses 1: The financial accounting criteria are effective with the amount of investment in fixed assets.

Hypotheses 2: The artificial intelligence algorithm has the ability to predict investment in the company's fixed assets.

The purpose of this research is to provide a growth model of fixed assets based on the financial criteria of companies admitted to the Tehran Stock Exchange. The current research is practical in terms of purpose and in terms of gathering post-event information because past information of the sample companies is used, and in terms of data collection and inference, it is descriptive-correlation and comparative-inductive research approach. Eviews and MATLAB software are used for analysis. In this study, in terms of easier access to information of companies listed on the Tehran Stock Exchange and the high reliability of information, the statistical population includes all companies listed on the Tehran Stock Exchange in 2013 until the end of 2022. To determine the statistical sample, a systematic elimination method will be used and for this purpose, those companies of the statistical community that have the following conditions will be selected as a statistical sample and the rest will be removed:

❖The financial year of the company is the end of March of each year.

❖Has not changed activity or changed fiscal year during the research period.

❖To be present in the stock exchange continuously during the research period.

❖Their financial information is available for the entire period under review.

After applying the above restrictions on the statistical population, 101 companies were selected as the statistical sample.

4 Research method

The purpose of this research is to provide a growth model of fixed assets based on the financial criteria of companies admitted to the Tehran Stock Exchange. Measurements and description of how to check them in a conceptual model as explained in Figure (1), in this research, variables are select by regression method using primary patterns. After selecting the variables, fixed asset investment predicts using intelligence artificial algorithm. Then the average amount of errors is extract and the algorithms are compared according to the errors and finally the best method is introduced for investors to use.

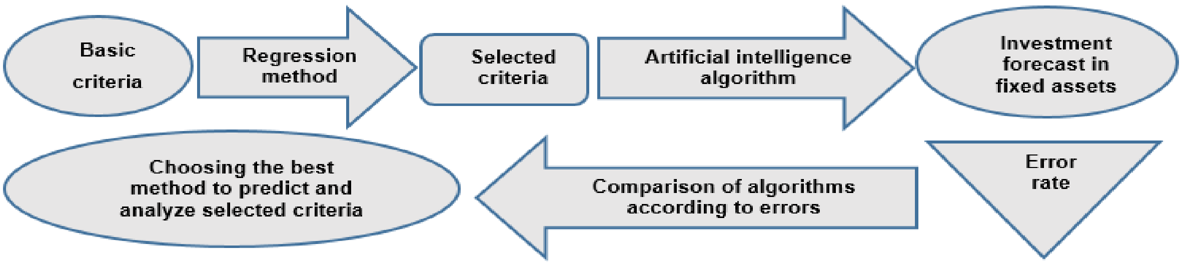

Fig. 1: Research Steps (Made by The Researcher)

The research process will be done according to the conceptual model of Figure (2). The variables of the companies are selected based on the variable selection method. The data is divided into 10 folds and then the data is tested and trained. Then the training data is calculated with four artificial intelligence algorithms: linear and non-linear PINSVR and linear and non-linear pls. In the training phase of the linear and non-linear models after learning, the same training-validation data without the dependent variable are again will be given to them to determine the value of the fixed asset growth variable. To check the learning power of the models using three errors MAE, MSE and SMAPE evaluation criteria in the training phase for each criterion, errors are reported for the current year and the next year.

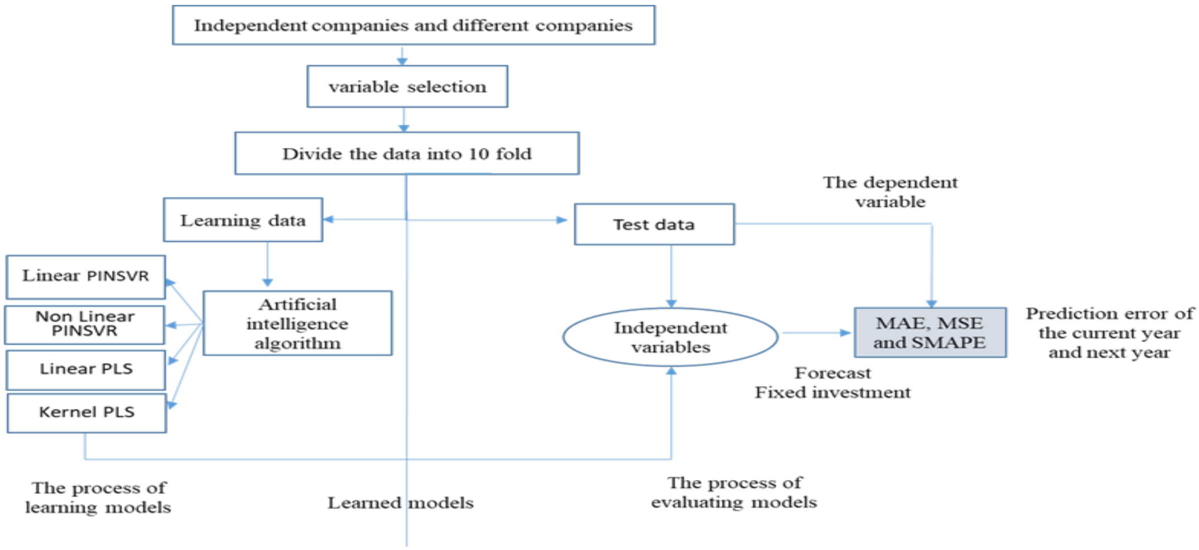

Fig. 2: The Process of Conducting Research (made by the researcher)

4.1 Method of Selecting and Finding the Importance of Independent Variables Based on Relief-F (Variable Selection)

In order to interpret and analyse the data, first, the information related to the descriptive statistics of related and descriptive variables was examined, which includes the calculation of the average, median, and dispersion, etc., which shows the characteristics of each member. EViews and MATLAB software are used for analysis. The target population in descriptive statistics, the information obtained from a group describes the same group then, by introducing the variables from the perspective of "accounting performance" in Table 1 in order to predict the dependent variable of fixed assets growth, the external data in the collected data set is checked. After removing external data from the set of samples, the effective features in predicting the growth of fixed assets are examine from the perspective of accounting performance and finally, two prediction algorithms, PINSVR and KPLSR, are introduced.

The available independent variables are specified in Table (1).

Relief-f method uses a statistical solution for feature selection [37]. This method is an algorithm based on weighting independent variables, which idea is inspired by sample-based algorithms. Figure (3) shows the algorithm of this method, which will be explained further. This algorithm selects a subset of companies from the training sample set D (or in other words, the company-year set along with the independent variable set S, which is the total number of variables N). The user specifies the number of companies (No Sample) in this subset as a predefined value. Suppose there is a set of N observations (company-years) along with the input vector (independent variables) x. All these company-years are represented by a data matrix X, so that the nth row is denoted by $\boldsymbol{x}_n^T$and let it represent the independent variables of the n company and let n=1.2...N and let y represent the dependent variable, i.e., the growth of fixed assets. PINSVR algorithm seeks to find non-parallel proximal linear functions $f_1(x)$ and $f_2(x)$ simultaneously, as well as two different non-parallel proximal linear functions $g_1(x)$ and $g_2(x)$ .These functions are shown below.

Table 1: Independent Variables of the Research

	The name of the primary research variable	Operational calculation method	Authors
1	return assets	The result of dividing net profit by total assets	Chen, et al [26]
2	firm size	The logarithm of the company's total assets	
3	financial leverage	The result of dividing the debt by the asset of financial leverage	
4	return on equity	The result of dividing net profit by equity	
5	sales growth	The difference between the sales of the current year and the previous year divided by the previous year's sales	
6	return on sales	The result of dividing net profit by sales	
7	Sales to assets ratio	The result of dividing sales by total assets	
8	Company value (Tobin's Q)	The sum of the company's market value and book value of debt divided by total assets	
9	Price-to-earnings ratio per share	The result of dividing the stock price at the end of the year by the profit per share	
10	return on stock	The market value of the company at the end of the year, minus the market value of the company at the beginning of the year, plus the company's shares, minus the increase in capital from cash inflows and receivables, divided by the market value of the company at the beginning of the year.	
11	dividend to total assets	The result of dividing the dividend by the total assets	
12	Operating profit margin	The result of dividing operating profit by sales	
13	Systematic risk	Changes in yield relative to market yield Beta $= (\text{cov}\ (Efficiency\ stocks.\text{Efficiency market}))/$ $(\text{var}\ (\text{Market efficiency})\ )$	
14	Acid Ratio	The result of dividing current assets minus inventories by current liabilities	

The dependent variable of this research is investment in fixed assets, which is shown in Table (2).

Table 2: The dependent variable of the research

	The name of the primary research variable	Operational calculation method	Authors
1	Investment in fixed assets	The difference between the first and last fixed assets of the period divided by total assets	Chen, et al [26]

Relief (D, S, NoSample, Threshold)
(1) *T= Ø*
(2) *Initialize all weights, w_i, to zero.*
(3) *For i =1 to NoSample/* Arbitrarily chosen */*
Randomly choose an instance x in D
Finds its nearHit and nearMiss
For j=1to N
$w_j = w_j - diff(x_j, nearHit_j)^2 + diff(x_j, nearMiss_j)^2$
(4) *For* j *=1 to N*
If $w_j \geq$ Threshold
Append feature f_j to T
(5) Return T

Fig. 3: Relief Algorithm1 **[37]**

$$f_1(x) = w_1^T x + b_1 \text{ و } f_2(x) = w_2^T x + b_2 \tag{1}$$

$$g_1(x) = w_3^T x + b_3 \text{ و } g_{2(x)} = w_4^T x + b_4 \tag{2}$$

$$\begin{gathered} L^{g_1}(x.y.f_1) = \sum_{i=1}^{m} max\{0, -(y_i - f_1(x_i) + g_1(x_i))\} \\ L^{g_2}(x.y.f_2) = \sum_{i=1}^{m} max\{0, -(f_2(x_i) - y_i + g_2(x_i))\} \end{gathered} \tag{3}$$

$$R_{emp}^{g_1}[f_1] = \sum_{i=1}^{m} max\left\{0, \left(y_i - f_1(x_i)\right)^2\right\} + c_1 \sum_{i=1}^{m} max\{0, -(y_i - f_1(x_i) + g_1(x_i))\} \tag{4}$$

$$R_{emp}^{g_2}[f_2] = \sum_{i=1}^{m} max\{0, (f_2(x_i) - y_i)^2\} + c_2 \sum_{i=1}^{m} max\{0, -(f_2(x_i) - y_i + g_2(x_i))\} \tag{5}$$

In which the function is greater than other.

$$\begin{gathered} \min_{w_1, w_3, b_1, b_3, \xi} \frac{1}{2} c_3 (w_1^T w_1 + b_1^2 + w_3^T w_3 + b_3^2) + \frac{1}{2} \xi^{*T} \xi^* + c_1 e^T \xi \\ s.t. \begin{cases} Y - (Aw_1 + eb_1) = \xi^* \\ Aw_3 + eb_3 \geq 0 \\ Y - (Aw_1 + eb_1) \geq -(Aw_3 + eb_3) - \xi, \quad \xi \geq 0 \end{cases} \end{gathered} \tag{6}$$

$$\begin{gathered} \min_{w_2, w_4, b_2, b_4, \eta} \frac{1}{2} c_4 (w_2^T w_2 + b_2^2 + w_4^T w_4 + b_4^2) + \frac{1}{2} \eta^{*T} \eta^* + c_2 e^T \eta \\ s.t. \begin{cases} (Aw_2 + eb_2) - Y = \eta^* \\ Aw_4 + eb_4 \geq 0 \\ (Aw_2 + eb_2) - Y \geq -(Aw_4 + eb_4) - \eta, \quad \eta \geq 0 \end{cases} \end{gathered} \tag{7}$$

Where c_1 ،c_2 ،c_3 و c_4 are the input parameters of the problem. Now, we state the geometric explanation of the optimization problem, a simple two-dimensional example of the implementation of the PINSVR algorithm along with the loss function is show in Figure (3). The structural risk in the relationship is minimize due to the existence of the $\frac{1}{2} c_3 \left(w_1^T w_1 + b_1^2 + w_3^T w_3 + b_3^2\right)$ adjustment term. The second term

in the objective function according to the expression $\frac{1}{2}c_3(w_1^T w_1 + b_1^2)$ represents the squared loss function of the error between the values of the decision function
$f_1(x) = w_1^T x + b_1$. The dependent variable for samples x label shown by Y is the third set of the objective function that minimizes the sum of variables ξ. That is, the algorithm tries to ensure that the samples do not cross the line $f_1(x)$ و $g_1(x)$ as much as possible and all are place on the same side of this line. There is a similar explanation for the optimization problem.

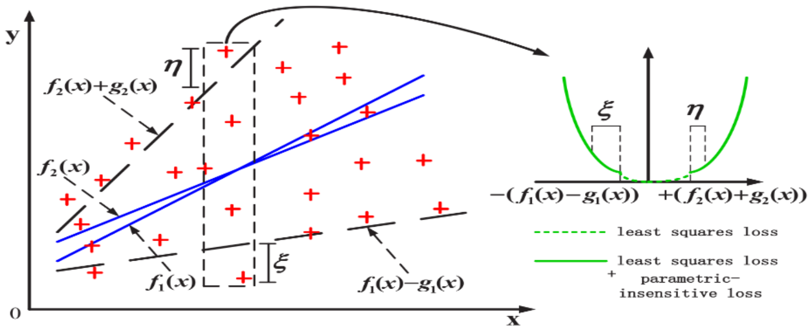

Fig. 4: Geometric 2 dimensions' interpretation and Lagrange function **[37]**

$$\min_{\alpha,\beta} \frac{1}{2c_3}\alpha^T G G^T \alpha + \frac{1}{2}\beta^T \left(G(G^T G + c_3 I_1)^{-1} G^T + \frac{1}{c_3} G^T G\right)\beta + \frac{1}{c_3}\alpha^T G G^T \beta - Y^T (G(G^T G + c_3 I_1)^{-1} G^T - I_2)\beta \quad s.t. \begin{cases} \alpha \geq 0 \\ * \\ 0 \leq \beta \leq c_1 e \end{cases} \tag{8}$$

$$\min_{\alpha^*,\beta^*} \frac{1}{2c_4}\alpha^{*T} G G^T \alpha^* + \frac{1}{2}\beta^{*T} \left(G(G^T G + c_4 I_1)^{-1} G^T + \frac{1}{c_4} G^T G\right)\beta^* + \frac{1}{c_4}\alpha^{*T} G G^T \beta^* - Y^T (G(G^T G + c_4 I_1)^{-1} G^T - I_2)\beta^* \quad s.t. \begin{cases} \alpha^* \geq 0 \\ * \\ 0 \leq \beta^* \leq c_2 e \end{cases} \tag{9}$$

$$u_1 = (G^T G + c_3 I_1)^{-1} G^T (Y - \beta) \tag{10}$$

$$u_3 = \frac{1}{c_3} G^T (\alpha + \beta) \tag{11}$$

PINSVR parametric insensitive non-parallel support vector regression Similar to the linear case, non-parallel proximal nonlinear functions $f_1(x)$ and $f_2(x)$ and two different non-parallel proximal non-linear functions $f_1(x)$and $g_1(x)$ are considered as follows.

$$u_2 = \left(G^T G + c_4 I_1\right)^{-1} G^T (Y - \beta^*) \tag{12}$$

$$u_4 = \frac{1}{c_4} G^T (\alpha^* + \beta^*) \tag{13}$$

$$f(x) = \frac{1}{2}\left(f_1(x) + f_2(x)\right) = \frac{1}{2}(w_1 + w_2)^T x + \frac{1}{2}(b_1 + b_2) \tag{14}$$

$$f_1(x) - g_1(x) = (w_1 - w_3)^T x + b_1 - b_3 \tag{15}$$

$$f_2(x) + g_2(x) = (w_2 + w_4)^T x + b_2 + b_4 \tag{16}$$

The linear PINSVR is extend for the nonlinear case using a kernel trick. The input data is map to the high-dimensional feature space using non-linear kernel functions. In the feature space, a linear regression function corresponds to a non-linear regression function in the input space, which is show in figure (4) of this issue.

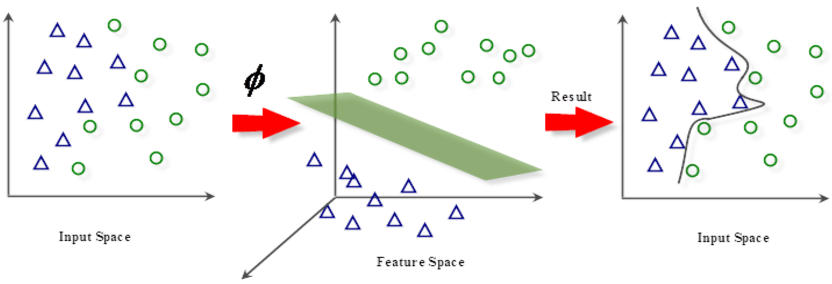

Fig. 5: Entrance Space Concept **[37]**

PINSVR parametric insensitive non-parallel support vector regression Similar to the linear case, non-parallel proximal nonlinear functions $f_1(x)$ and $f_2(x)$ and two different non-parallel proximal non-linear functions $g_1(x)$ and $g_2(x)$ are considered as follows.

$$f_1(x) = k\left(x^T{\text{و}}A^T\right)w_1 + b_1 \text{ و } f_2(x) = k\left(x^T.A^T\right)w_2 + b_2 \tag{17}$$

$$g_1(x) = k\left(x^T.A^T\right)w_3 + b_3 \text{ و } g_{2(x)} = k\left(x^T.A^T\right)w_4 + b_4 \tag{18}$$

Where k is the kernel function,$g_1(x) \geq 0$, $g_2(x) \geq 0$. The basic problem of nonlinear PINSVR is restat as functions presented in the section on Nonparallel Insensitive-Parametric Support Vector Regression (PINSVR).

4.2 partial Least Squares (PLS)

PLS algorithm is a method for modeling the linear relationship between a set of output variables (responses) $(n \times p)$ T ،U and $(N \times p)$ W and a set of input variables $\{t_i\}_{i=1}^p$ ،$\{u_i\}_{i=1}^p$ ،$\{w_i\}_{i=1}^p$ ،$\{c_i\}_{i=1}^p$ is N (regressors). In the first step, PLS creates uncorrelated latent variables that are a linear combination of the main regressors [38]. The main point of this method is that the weights are used to determine the linear combinations of the main repressor that are proportional to the covariance between the input and output variables. The least squares algorithm is performed on the subset of extracted hidden variables. This results in biased but smaller variance estimates of the regression coefficients compared to the original least squares (OLS) regression. PLS regression is an iterative process. For example, after extracting a component, the algorithm starts again using the X and Y matrices. We can get a sequence of models until we reach rank X. However, in practice, the cross-validation method is usually use to avoid under fitting or over fitting those results from using models with too small or too large dimensions. After extracting p components, we can get $(n \times p)$ T ‹U and $(N \times p)$ W matrix and (L × p) C matrix whose columns are the vectors $\{t_i\}_{i=1}^{p}$ ‹$\{u_i\}_{i=1}^{p}$ ‹$\{w_i\}_{i=1}^{p}$ ‹$\{c_i\}_{i=1}^{p}$are create. The PLS regression model can be write as a matrix as follows:

$$Y = XB + F \tag{19}$$

Where B is a matrix (N × L) of regression coefficients and F is a residual matrix (n × L). This linear relationship is the same as used in other regression models.

4.3 Kernel Partial Least Squares (KPLS)

Vector property mapping $\Phi(x_i)^T\Phi(x_j) = K(x_i, x_j)$. Our goal is to build a linear PLS regression model in F space. By modifying the linear algorithm, the NIPALS-PLS nonlinear algorithm changes as follows [38] Consider a nonlinear mapping of input variables $\Phi: x_i \in R^N \rightarrow \Phi(x_i) \in F$ to a feature space F.

(20)

1. randomly initialize u
2. $t = \Phi\Phi^T u$, $t \leftarrow t/\|t\|$
3. $c = Y^T t$
4. $u = Yc$, $u \leftarrow u/\|u\|$
5. repeat^T steps 2 – 5 until convergence
6. Deflate $\Phi\Phi^T$, Y matrices: $\Phi\Phi^T \leftarrow \Phi - tt^T\Phi)(\Phi - tt^T\Phi)^T$, $Y \leftarrow Y - tt^T Y$.

By applying the kernel trick, namely $\Phi(x_i)^T\Phi(x_j) = K(x_i, x_j)$, it can be see that $\boldsymbol{\Phi^T\Phi}$ is the representation of the K kernel gram matrix. Instead of using the data mapping to the high-dimensional feature space, the calculations can be performed in the input space. By define$\boldsymbol{\Phi^T\Phi} = \mathbf{K}$, we have:

$$K \leftarrow (I - tt^T)K(I - tt^T) = K - tt^TK - Ktt^T + tt^TKtt^T \tag{21}$$

$$\hat{Y} = \Phi B = KU(T^TKU)^{-1}T^TY = TT^TY \tag{22}$$

$$R = \Phi^T U(T^TKU)^{-1} \text{ And } T = \Phi B$$

Where (1) is an n-dimensional identity matrix. Similar to the linear case, the matrix of regression coefficients B can be obtained. It is only necessary to mention that due to the anonymity and the high dimension of the feature mapping, this matrix cannot be practical.

4.4 Data Segmentation Method and Model Evaluation Indicators

To train the linear and non-linear models of PINSVR and KPLSR, first data the set of company-year samples divide into training and testing [26]. The training data is use to learn the parameters and super-parameters of the model and the test data is used to evaluate the predictions of the models. 10-Fold Cross-Validation use in this research in order to better and more efficiently evaluate the prediction of the models. According to Figure (4), these samples they divided into 10 repetitions by replacing the evaluation samples. This division is such that one tenth of the total sample was selected as evaluation data in the first iteration and the rest of the data was selected as training-validation data in 9 parts. For the second iteration, according to the figures of the second part they selected as evaluation data and other parts as training-validation data. These repetitions are repeat in the same order up to 10 times. With the help of the 10-fold cross-validation method and using the training-validation dataset for each of the linear and non-linear PINSVR and KPLSR models, one model is learning in each iteration. In simpler terms, in the first iteration, each model independently learns its parameters and super parameters based on algorithms with the help of (9) training-validation parts and builds its own model. Four models made so far. To check how well these models learned from the training-validation data, the exact same nine parameters are feeding into the trained models to predict fixed asset growth .Now, using the following evaluation criteria, such as MSE, the amount of forecast error and the actual value of the fixed asset growth variable are measured. This number is stored as the learning error of the first iteration . This is repeat for other parts and 10 MSE errors are obtain. The average of these errors is stored as the learning phase error of each model. Usually, the error rate on the training data is lower than the error rate on the evaluation data that see in the learning process .Therefore, the learning error cannot be used to compare two algorithms, because the overfitting phenomenon may occur Overfitting phenomenon means that the models predict the training-validation data well, but they predict the real-world data that they have not seen yet very badly and with a large difference. Therefore, in addition to the training data set, a set of data is need for testing, for which the test data shown in each iteration in Figure (6) are for this purpose. These data have not seen by the models so far and the data is different from other iterations . In each iteration, the corresponding model predicts the evaluation data that it has not seen so far, and the estimation criteria of the predicted and actual value of fixed asset growth are calculated, and finally, the average of these errors is stored as the prediction error of each model. Now, if the average error of the learning phase and the error of the evaluation phase are close to each other, it means that the overshooting phenomenon has not occurred.

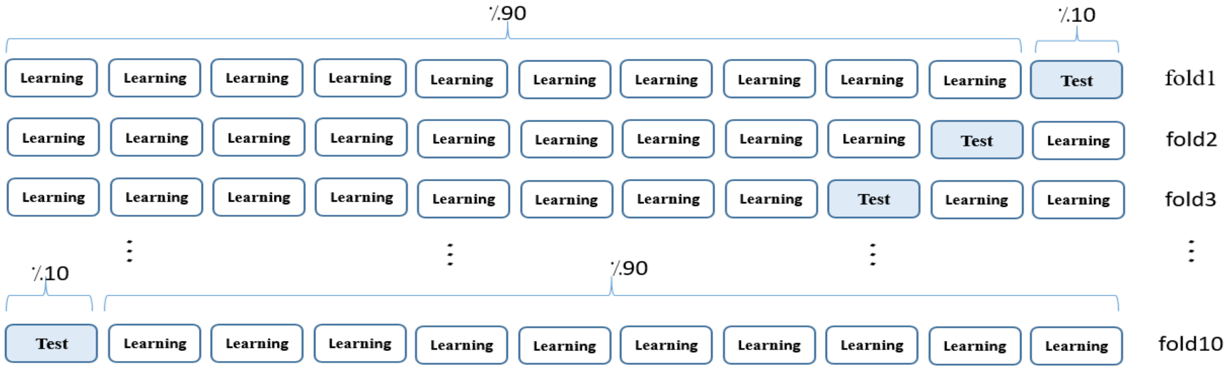

Fig. 6: The steps of selecting two training and test data sets with 10-point cross-validation

After dividing the company-years into two groups of training-validation and test data using 10-point cross-validation to evaluate linear and non-linear models from three evaluation criteria named symmetric mean absolute value of error (SMAPE), mean absolute value of error (MAE), The mean squared error (MSE) is used, which is calculated using the following relations.

$$SMAPE = \frac{1}{n}\frac{\sum_{i=1}^{n}|d_i - y_i|}{\sum_{i=1}^{n}(d_i + y_i)} \tag{23}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - d_i| \tag{24}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - d_i)^2 \tag{25}$$

5 Data Analysis
5.1 Selection of Independent Variables Based on Relief-f

The independent variables are select using the weighting-based with Algorithm Relief-F method in Figure (1), which is bases on the inspired sample .in the other hand, the growth of fixed assets of companies is a dependent variable [39].

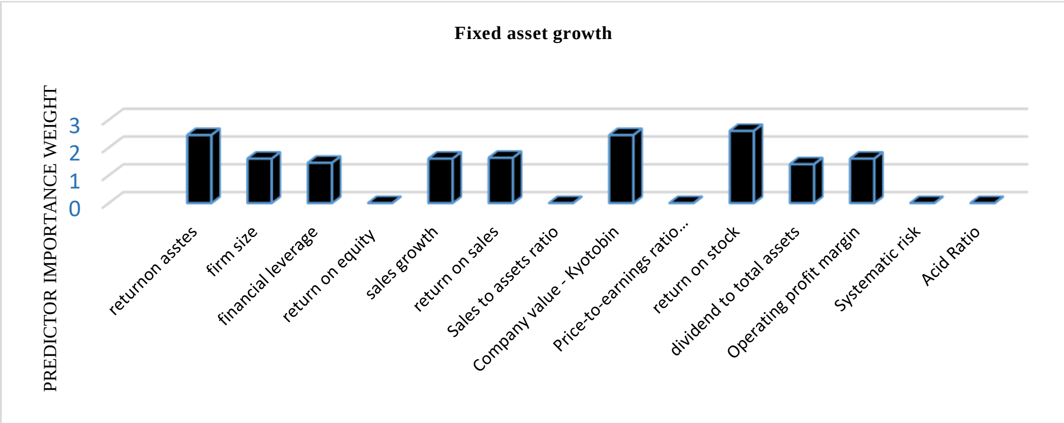

Fig. 7: Accounting Performance Variables Selected Using The Relief-F Method **[39]**

The details of the errors obtained in each iteration for each algorithm with different qualities are show. In table (1) of learning and table (2) of test error and deviation from the quality of each line is written at the end of it.

5-2 Evaluating the Power of Learning and Predicting Models

In this step, using the 10-fold cross-validation method, the training-validation data set give to four algorithms: linear PINSVR, non-linear PINSVR, linear PLS and non-linear PLS, which are call KPLS. These algorithms learn their parameters and hyper-parameters using this data. The set of samples, with accounting performance criteria (selected by Relief-F algorithm) as independent variables (features) along with the dependent variable of fixed asset growth give to four algorithms. In the training phase of the linear and non-linear models after learning, the same training-validation data without the dependent variable are again will be given to them to determine the value of the fixed asset growth variable. The power of linear models is extremely low and all three show this error well. The error difference between linear models and non-linear models is not so great that it claims that linear models are inefficient in the problem of predicting asset growth. As a result, it is not possible to find a suitable linear model compared to non-linear models that works better or close to them. The detailed results of MAE, MSE, and SMAPE cross-validation error in the learning phase are shown in Tables (1), (2), and (3).

Table 1: Details of the cross-validation error MAE in the learning phase

		MAE Criterion	**Fold**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**Average**
Accounting performance criteria	**Current Year**	Linear PINSVR		0.089	0.092	0.016	0.091	0.092	0.093	0.093	0.093	0.090	0.091	**0.092 ± 0.001**
		Nonlinear PINSVR		0.083	0.085	0.014	0.086	0.086	0.086	0.087	0.086	0.084	0.085	**0.085 ± 0.001**
		Linear PLS		0.088	0.092	0.016	0.091	0.092	0.093	0.093	0.094	0.090	0.091	**0.092 ± 0.002**
		Kernel PLS		0.058	0.061	0.007	0.059	0.061	0.060	0.061	0.060	0.058	0.059	**0.06 ± 0.001**
	Next Year	Linear PINSVR		0.088	0.091	0.014	0.088	0.090	0.087	0.088	0.088	0.089	0.087	**0.088 ± 0.001**
		Nonlinear PINSVR		0.080	0.082	0.012	0.080	0.082	0.079	0.080	0.079	0.080	0.079	**0.08 ± 0.001**
		Linear PLS		0.087	0.091	0.014	0.089	0.090	0.087	0.088	0.089	0.089	0.087	**0.088 ± 0.001**
		Kernel PLS		0.057	0.058	0.006	0.057	0.058	0.057	0.056	0.057	0.057	0.057	**0.057 ± 0.001**

Table 2: Details of the cross-validation error MSE in the learning phase

		MSE Criterion	Fold	1	2	3	4	5	6	7	8	9	10	Average
Accounting performance criteria	Current Year	Linear PINSVR		0.015	0.016	0.016	0.016	0.016	0.017	0.017	0.017	0.016	0.016	**0.016 ± 0**
		Nonlinear PINSVR		0.013	0.014	0.014	0.014	0.014	0.014	0.014	0.014	0.013	0.014	**0.014 ± 0**
		Linear PLS		0.015	0.016	0.016	0.016	0.016	0.017	0.017	0.017	0.016	0.016	**0.016 ± 0**
		Kernel PLS		0.006	0.007	0.007	0.006	0.007	0.007	0.007	0.007	0.006	0.007	**0.007 ± 0**
	Next Year	Linear PINSVR		0.015	0.015	0.014	0.015	0.015	0.014	0.015	0.015	0.015	0.014	**0.015 ± 0**
		Nonlinear PINSVR		0.012	0.013	0.012	0.012	0.013	0.012	0.012	0.012	0.012	0.012	**0.012 ± 0**
		Linear PLS		0.015	0.015	0.014	0.015	0.015	0.014	0.015	0.014	0.015	0.014	**0.015 ± 0**
		Kernel PLS		0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	**0.006 ± 0**

Table 3: Details of the cross-validation error SMAPE in the learning phase

		SMAPE Criterion	Fold	1	2	3	4	5	6	7	8	9	10	Average
Accounting performance criteria	Current Year	Linear PINSVR		0.085	0.088	0.088	0.088	0.087	0.088	0.089	0.088	0.086	0.087	**0.087 ± 0.001**
		Nonlinear PINSVR		0.080	0.082	0.081	0.082	0.082	0.082	0.083	0.082	0.080	0.081	**0.081 ± 0.001**
		Linear PLS		0.085	0.088	0.087	0.087	0.087	0.089	0.089	0.089	0.086	0.087	**0.087 ± 0.001**
		Kernel PLS		0.056	0.059	0.057	0.057	0.058	0.058	0.058	0.058	0.056	0.057	**0.058 ± 0.001**
	Next Year	Linear PINSVR		0.079	0.081	0.079	0.079	0.081	0.078	0.079	0.078	0.080	0.078	**0.079 ± 0.001**
		Nonlinear PINSVR		0.072	0.073	0.072	0.071	0.074	0.071	0.072	0.071	0.072	0.071	**0.072 ± 0.001**
		Linear PLS		0.079	0.081	0.078	0.079	0.081	0.078	0.079	0.079	0.080	0.078	**0.079 ± 0.001**
		Kernel PLS		0.052	0.053	0.052	0.051	0.053	0.051	0.051	0.051	0.052	0.052	**0.052 ± 0**

The detailed results of MAE, MSE, and SMAPE cross-validation error in the testing phase are shown in Tables (4), (5), and (6).

Table 4: Details of the cross-validation error MAE in the testing phase

		MAE Criterion	Fold	1	2	3	4	5	6	7	8	9	10	Average
Accounting performance criteria	Current Year	Linear PINSVR		0.115	0.088	0.090	0.092	0.091	0.081	0.079	0.079	0.104	0.095	**0.092 ± 0.012**
		Nonlinear PINSVR		0.103	0.085	0.086	0.082	0.079	0.077	0.070	0.079	0.099	0.091	**0.085 ± 0.01**
		Linear PLS		0.115	0.088	0.090	0.092	0.092	0.082	0.081	0.081	0.104	0.096	**0.092 ± 0.011**
		Kernel PLS		0.073	0.050	0.064	0.067	0.050	0.054	0.051	0.053	0.072	0.062	**0.06 ± 0.009**
	Next Year	Linear PINSVR		0.092	0.068	0.093	0.088	0.071	0.103	0.092	0.093	0.080	0.102	**0.088 ± 0.012**
		Nonlinear PINSVR		0.081	0.066	0.082	0.083	0.062	0.091	0.081	0.086	0.077	0.091	**0.08 ± 0.01**
		Linear PLS		0.093	0.069	0.094	0.090	0.071	0.104	0.094	0.096	0.081	0.103	**0.089 ± 0.012**
		Kernel PLS		0.059	0.048	0.057	0.061	0.049	0.060	0.062	0.058	0.057	0.059	**0.057 ± 0.005**

Table 5: Details of the cross-validation error MSE in the testing phase

		MSE Criterion	Fold	1	2	3	4	5	6	7	8	9	10	Average
Accounting performance criteria	Current Year	Linear PINSVR		0.025	0.017	0.015	0.018	0.017	0.011	0.011	0.011	0.020	0.017	**0.016 ± 0.004**
		Nonlinear PINSVR		0.020	0.014	0.014	0.015	0.012	0.010	0.009	0.011	0.019	0.016	**0.014 ± 0.004**
		Linear PLS		0.025	0.018	0.015	0.018	0.017	0.011	0.011	0.012	0.021	0.018	**0.017 ± 0.004**
		Kernel PLS		0.010	0.004	0.007	0.009	0.005	0.005	0.005	0.005	0.009	0.007	**0.007 ± 0.002**
	Next Year	Linear PINSVR		0.015	0.008	0.017	0.013	0.010	0.021	0.015	0.015	0.012	0.020	**0.015 ± 0.004**
		Nonlinear PINSVR		0.012	0.007	0.013	0.012	0.007	0.016	0.013	0.013	0.011	0.017	**0.012 ± 0.003**
		Linear PLS		0.015	0.008	0.018	0.013	0.010	0.021	0.015	0.016	0.012	0.020	**0.015 ± 0.004**
		Kernel PLS		0.006	0.004	0.007	0.006	0.004	0.006	0.007	0.006	0.005	0.006	**0.006 ± 0.001**

Table 6: Details of the cross-validation error SMAPE in the testing phase

		SMAPE Criterion	Fold	1	2	3	4	5	6	7	8	9	10	Average
Accounting performance criteria	Current Year	Linear PINSVR		0.107	0.082	0.086	0.086	0.089	0.079	0.076	0.078	0.098	0.092	**0.087 ± 0.01**
		Nonlinear PINSVR		0.096	0.079	0.082	0.076	0.077	0.076	0.067	0.077	0.093	0.088	**0.081 ± 0.009**
		Linear PLS		0.107	0.082	0.086	0.086	0.090	0.080	0.078	0.080	0.098	0.093	**0.088 ± 0.009**
		Kernel PLS		0.068	0.047	0.061	0.063	0.050	0.054	0.050	0.053	0.067	0.061	**0.057 ± 0.007**
	Next Year	Linear PINSVR		0.081	0.062	0.082	0.080	0.063	0.088	0.081	0.087	0.072	0.090	**0.079 ± 0.01**
		Nonlinear PINSVR		0.071	0.060	0.072	0.075	0.055	0.077	0.072	0.080	0.069	0.081	**0.071 ± 0.008**
		Linear PLS		0.082	0.063	0.083	0.082	0.063	0.088	0.083	0.089	0.073	0.091	**0.08 ± 0.01**
		Kernel PLS		0.052	0.045	0.050	0.056	0.043	0.052	0.056	0.055	0.051	0.053	**0.051 ± 0.004**

The value shown in each house of the table from left to right (STD±AVG) is the average of the reported errors of the 10-fold cross-validation method and the standard deviation of the 10 reported errors. The standard deviation is expected to be close to zero for different implementations. Because, the further this value is from zero, it means that the learning process is dependent on the input data of the problem.

5-3 Evaluation of Linear and Non-Linear Models

To evaluate linear and non-linear models from three evaluation criteria called Symmetric Mean Absolute Magnitude Error (SMAPE), absolute mean. The mean squared error (MSE) was used to measure the error (MAE). Considering that the model of the current year is independent from the model of the next year, it is expected that the errors reported in the current year will be close to the errors of the next year, and this issue is clear in Tables 1 and 2. In the reported errors, the errors of the current and future years are close to each other, and it indicates good learning of the current and future year models. The next issue that should be investigated the predictive power of the models and the non-occurrence of the phenomenon of overshooting in the process of the learning phase [40].

Table 7: Learning Power of Models Using MAE, MSE and SMAPE Error Evaluation Criteria in the Training Phase

Scale	Training step	
MAE	**Current year**	**Future year**
Linear PINSVR	0.092±0.001	0.088±0.001
Non-Linear PINSVR	0.085±0.001	0.08±0.001
Linear PLS	0.092±0.002	0.088±0.001
Kernel PLS	0.06±0.001	0.057±0.001
MSE	**Current year**	**MSE**
Linear PINSVR	0.016±0	0.015±0
Non-Linear PINSVR	0.014±0	0.012±0
Linear PLS	0.016±0	0.015±0
Kernel PLS	0.007±0	0.006±0
SMAPE	**Current year**	**SMAPE**
Linear PINSVR	0.087±0.001	0.079±0.001
Non-Linear PINSVR	0.081±0.001	0.072±0.001
Linear PLS	0.087±0.001	0.079±0.001
Kernel PLS	0.058±0.001	0.052±0

In the testing phase, for this reason, the test data that were discarded in the 10-fold cross-validation process are entered into the learned models to evaluate their predictive power for samples that have not yet been observed. It is expected that the error difference between the training and test phase is not too extreme. The errors of the test phase may be less or more than the value of the errors of the test phase, this is not an important issue. What is significant is the small difference in reported errors between these two phases. In Table (7), similar to the training phase, the mean and standard deviation of all errors on accounting performance measures are shown in Table (8). As can be seen, the difference between the errors reported in Table (7) and Table (8) is small. Therefore, the phenomenon of over-fitting has not happened and all the topics raised in the training phase are also true in the evaluation phase.

Table 8: Average and Deviation from Error Criteria to Check the Predictive Power of Models in the Test Phase

Scale	Test step	
MAE	**Current year**	**Future year**
Linear PINSVR	0.092±0.012	0.088±0.012
Non-Linear PINSVR	0.085±0.01	0.08±0.01
Linear PLS	0.092±0.011	0.089±0.012
Kernel PLS	0.06±0.009	0.057±0.005
MSE	**Current year**	**Future year**
Linear PINSVR	0.016±0.004	0.015±0.004
Non-Linear PINSVR	0.014±0.004	0.012±0.003
Linear PLS	0.017±0.004	0.015±0.004
Kernel PLS	0.007±0.002	0.006±0.001
SMAPE	**Current year**	**Future year**
Linear PINSVR	0.087±0. 01	0.079±0. 01
Non-Linear PINSVR	0.081±0.009	0.071±0.008
Linear PLS	0.088±0.009	0.08±0. 01
Kernel PLS	0.057±0.007	0.051±0.004

6 Conclusion

Almost the majority of economists, regardless of their school and intellectual point of view, put a lot of emphasis on capital formation and carrying out strategic measures in order to increase investment as the most important factor determining growth and development [41]. However, commercial units are always face with many investment opportunities and need to make a rational decision regarding an optimal investment. In fact, the investment of each commercial unit better done according to the limited resources and its efficiency but the main issue is choosing projects and making decisions about investment opportunities and efficiency through the managers of commercial units, which is foundation on their personal interests. In other words, informational asymmetry and conflict of interests hinders. Therefore, for investing in different projects, commercial units should pay attention to the extent or amount of investment according to the limited resources, in order to get the maximum efficiency from their investment and achieve a double investment [42]. Managers benefit by increasing information asymmetry between the firm and the market more features and opportunities to hide negative news and speed up reviews positive news, as a result of information asymmetry between managers and investors, increases the risk of future stock price falls [26].The optimal portfolio has the ability to provide the final desired wealth of the investor at a 95% confidence level and the portfolio return, including transaction costs, is higher than the return of a single-period portfolio[43]. In neural network models optimized with genetic algorithms, it was found that the neural network optimized with particle swarm algorithm shows a lower error coefficient, which indicates better performance and higher prediction accuracy compared to the neural network optimized with genetic algorithms[44]. The increase in investment is an activity to earn profit, and uncertainty and risk are two factors that affect the expected profit and as a result the decision to invest. Mainly forecasting financial and economic indicators are doing by regression method[45]. Regression is a possible prediction outside a time range that has different types including linear, multinomial, stochastic and transformational regression models [6]. As mentioned, the current research was conducted with the aim of modeling the growth of fixed assets of Tehran Stock Exchange companies. The KPLS (or linear PLS) method has several advantages over previous approaches: (1) The KPLS can reduce feature dimensionality. (2) The KPLS can find a small number of latent variables, e.g., 20, to project thousands of features into a very low-dimensional subspace, which may have great impact on real-time applications. (3) The KPLS regression has an output vector that can contain multiple labels, so that several related problems, can be solve altogether [18]. PINSVR indirectly finds a pair of nonparallel proximal functions with a pair of different parametric-insensitive nonparallel proximal functions by solving two smaller sized quadratic programming problems. By using new parametric-insensitive loss functions, the proposed PINSVR automatically adjusts a flexible parametric-insensitive zone of arbitrary shape and minimal size to include the given data to capture data structure and boundary information more accurately [6]. Compared to two methods, KPLS and PINSVR, there is another method call SVM of support vector machine in solving artificial intelligence methods mentioned in the research background. The SVM algorithm is also a powerful and flexible Machine Learning model. This algorithm supports not only linear and non-linear classification, but also linear and non-linear regression (SVR). The trick to using SVM for regression is to invert the goal, instead of trying to fit the widest two-class binomial path that minimizes the outlier. In SVM regression, the attempt is to place as many samples as possible on this path (edge) while minimizing edge violations. SVM algorithm can also be used for outlier data, but the drawback of this algorithm it is only good for small and medium complex training data, but becomes very slow as the number of training data samples increases [19]. Also, one of the improvements in new studies methods Artificial intelligence is employed, using the feature selection as a pre-stage for It is the main classifier model. In addition, the feature of using the vector machine method is significantly superior to the genetic algorithm. One of the most important disadvantages of the genetic algorithm is their sensitivity to the initial values of the weights, and these methods lack a suitable structure to determine the structural parameters. In this research, according to the importance of investment by Tehran Stock Exchange companies, the factors affecting investment in fixed assets investigated. For this purpose, the information of 101 stock exchange companies during the years 2013-2022 was analyzed using artificial intelligence method and the results of the research are as follows:

How is the impact of the research on the quality of financial accounting on the company's investment in fixed assets? It can be said that among the financial variables, the changes in asset return, stock return, company value compared to Tobin's Q, sales, sales return, operating profit margin, company size, financial leverage, the division ratio of assets with assets are the most effective on fixed assets. The artificial intelligence algorithm predicts the investment in the company's fixed assets and financial accounting standards are effective with the amount of investment in fixed assets and the artificial intelligence algorithm model for predicting investment in fixed assets are obtained through the following. In this step, using the 10-fold cross-validation method the training-validation data set was gived to four algorithms: linear PINSVR, non-linear PINSVR, linear PLS and non-linear PLS, which is call KPLS. These algorithms learn accounting (selected by the Relief-F algorithm as independent variables (features) along with the dependent variable of fixed asset growth to four parameters and hyper-parameters using this data. The set of samples, with performance measures from the algorithm was gived. The training phase of linear and non-linear models after learning, the same training-validation data were again gived to them without the dependent variable to predict the value of the fixed asset growth variable, then by calculating the three error measures MAE, MSE and SMAPE, the learning power and learning error of these models are tested. In all these errors, the closer the error value are to zero, the more powerful the corresponding algorithm is. The power of linear models is extremely low and all three show this error well. Therefore, the phenomenon of over-fitting has not happened and all the topics raised, in the training phase are also true in the test phase. The error difference between linear models and non-linear models is not so great that it can be claim that linear models are ineffective in the problem of predicting asset growth. This conclusion is draw from the complexity of the input space of the problem that for the nature of the problem of fixed asset growth with accounting performance criteria, it is not possible to find a suitable linear model compared to non-linear models that works better or close to them. In the reported errors, the current and future year errors are close to each other, and it is a good indication of the current and future year models. For further research, the following suggestions are provided:

1- Research on the impact of intellectual capital characteristics on the decision to invest capital market actors in fixed assets.

2- Examining the impact of risk management indicators on investment decisions in fixed assets.

3- Due to the artificial optimal performance with PINSVR and KPLS methods, to predict the fixed investment of companies, the genetic algorithm method is also used in asset prediction.

This manuscript is prepared based on PhD thesis of Farzaneh shamsdoost at sanandaj Branch, Islamic. Azad University, sanandaj, Iran.

Refrences

[1] Laurencia, V., and Mulyana, R., Firm Characteristics, Profitability and Corporate Governance Effect on Earnings Management, *International Journal of Business, Economics and Law*, 2022; 26(1): 224-32.

[2] Khushkar, F., Safarikhoh, M., and Ranjbari, S., The effect of ownership structure on the relationship between financial information quality and efficiency, *Journal of Perspectives on accounting and management and investment*, 2020; 3(33): 162-175, (In persian).

[3] Jiménez-Preciado, A. L., Venegas-Martínez, F., and Ramírez-García, A., Stock Portfolio Optimization with Competitive Advantages (MOAT), A Machine Learning Approach, *Journal of Mathematics*, 2022; 10 (23). Doi:10.3390/math10234449

[4] Khalifa Soltani,S. A., Rahimi, A., and Azizollahi, B., The impact of market competition on the investment decisions of stock exchange manufacturing companies, *Journal of accounting studies in Tehran*, 2014; 1(1): 81-96, (In persian).

[5] Gordon, L. A., Loeb, M. P., and Tseng, C. Y., Enterprise risk management and firm performance: A contingency perspective, *Journal of accounting and public policy*, 2009; 28(4): 301-327. Doi:10.1016/j.jaccpubpol.2009.06.006

[6] Sayadi, M., Dastgir, M., and Aliahmadi, S., Increasing the efficiency of investment on the relationship of management ability, *Journal of Research paper Studying the role of company risk management,* 2019; 7(24), 1-37. Doi:10.22051/jfm.2018.20848.1700, (In persian).

[7] McNichols, M. F., and Stubben, S. R., Does earnings management affect firms' investment decisions? *Journal of The accounting review*, 2008; 83(6):1571-1603. Doi:10.2308/accr.2008.83.6.1571

[8] Zhang, X., Hu, Y., Xie, K., Wang, S., Ngai, E. W. T., and Liu, M. A causal feature selection algorithm for stock prediction modeling, *Neurocomputing*, 2014; 142, 48-59. Doi:10.1016/j.neucom.2014.01.057

[8] Zhang, X., Hu, Y., Xie, K., Wang, S., Ngai, E. W. T., and Liu, M., A causal feature selection algorithm for stock prediction modeling, *Journal of Neurocomputing*, 2014; 142, 48-59. Doi:10.1016/j.neucom.2014.01.057

[9] De Vries, S., and Ter Braak, C. J., Prediction error in partial least squares regression, a critique on the deviation used in The Unscrambler, *Chemometrics and intelligent laboratory systems*, 1995; 30(2): 239-245. Doi:10.1016/0169-7439 (95)00030-5

[10] Rännar, S., Lindgren, F., Geladi, P., and Wold, S., A PLS kernel algorithm for data sets with many variables and fewer objects, Theory and algorithm, *Journal of Chemometrics*, 1994; 8(2): 111-125. Doi:10.1002/cem.1180080204

[11] Naim, P., and Condamin, L., Operational risk modeling in financial services, the exposure, occurrence, impact method, *John Wiley and Sons*, 2019; ISBN. 1119508509, 9781119508502

[12] Razia Saif Elahi et al, A comparative study of behavioral factors in the investment of financial assets, *Scientific research quarterly of investment knowledge,* 2016; 4(15):33-52.

[12] Seifollahi, R., Kordloei, H., and Dashti, N., A comparative study of behavioral factors in the investment of financial assets, *Journal of Scientific research quarterly of investment knowledge*, 2016; 4 (15): 33-52, (In persian).

[13] Géron, A., Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow, " O'Reilly Media, Inc", ISBN, 2019; 1498032611, 9781492032618

[14] Porter, M. E., Industry structure and competitive strategy, Keys to profitability, *Financial analyst s journal*,1980; 36(4): 30-41. Doi:10.2469/faj.v36.n4.30

[15] Feng, M., Li, C., McVay, S. E., and Skaife, H., Does ineffective internal control over financial reporting affect a firm's operations? Evidence from firms' inventory management, *Journal of The Accounting Review*, 2015; 90(2): 529-557. Doi:10.2308/accr-50909

[16] Jahanshad, A., and Khalili, A., The relationship between stock returns and the growth rate of fixed assets in different time frames using wavelet transformation, *Journal of Financial engineering and securities management*, 2013; 15: 1-16, (In persian).

[17] Darabi, R., and Karimi, A., the effect of growth rate of fixed assets on stock returns, *Financial accounting and auditing researches*, 2011; 2(8): 130-99, (In persian).

[18] Hajian, N., Anwari Rostami, A., Rahmani, A., and Azar, A., Investigating the effective factors on the information disclosure level of the companies admitted to the Tehran Stock Exchange, *Journal of Financial Accounting Research*, 2015; 7(1), (In persian).

[19] Firmansyah, D., Ahmar, N., and Mulyadi, J. M. V., The Effect of Leverage, Size, Liquidity, Operating Cash Flow on Fixed Asset Revaluation, *Journal of The Indonesian Accounting Review*, 2017; 7(1): 31-43. Doi:10.14414/tiar.v7i1.816

[20] Lopes, A. B., and Walker., MAsset revaluations, future firm performance and firm-level corporate governance arrangements, New evidence from Brazil, *Journal of The British Accounting Review*, 2012 ; 44(2): 53-67. Doi:10.1016/j.bar.2012.03.007

[21] Mirmohammadi, S.E., and Soleimani Amiri.GH., Examining the relationship between debt surplus and current performance (employment rate, investment in fixed assets and financial stress) of small and medium companies in Iran's capital market, *Journal of Accounting and Management Perspective*, 2020; 35(3): 101-121.

[22] Leung, M. T., Daouk, H., and Chen, A. S., Forecasting stock indices, A comparison of classification and level estimation models, *International Journal of Forecasting*, (2000); 16: 173–190. Doi:10.1016/S0169-2070 (99)00048-5

[23] Mansourfar, Gh., the Role of Internal and External Dimensions of Corporate Governance on the Relationship between Information Asymmetry and Investment Efficiency, *Financial Research Journal, 2020;* 2 (22): 227-248. Doi :10.22059/frj.2020.279145.1006855, (In persian).

[24] Taghizadeh, V., Badavar.Y., Motagi. A., and Taghizadeh, H., Validation of investment efficiency models based on representation theory, asymmetry Information, managerial fronting and maximization of company value, *Investing knowledge scientific research quarterly,* 2022; 10(38): 287- 318, (In persian).

[25] Silverstein, B., Managerial Opportunism and Corporate Investment Efficiency, *Kansas State University Department of Finance, Published Social Science Research*, 2020; Doi:10.2139/ssrn.3463419

[26] He, Y., Chen, C., and Hu, Y., Managerial overconfidence, internal financing, and investment efficiency: Evidence from China, *Research in International Business and Finance*, 2019; 47 (1): 501-510.
Doi: 10.1016/j.ribaf.2018.09.010

[27] Chandrinos, S.K., Sakkas, G., and Lagaros, N.D., A risk management tool using machine learning, *Expert Systems with Applications*, 2018; 105: 34-48. Doi:10.1016/j.eswa.2018.03.044

[28] Marchellina, V., and Firnanti, F., Financial ratio and company characteristics effect on earnings management. *In Ninth International Conference on Entrepreneurship and Business Management, Atlantis Press*, 2021; 178-183. Doi. 10.2991/aebmr.k.210507.027

[29] Wickramasinghe, M. B., and Gunawardana, K., Cash flow risk management practices on sustainable financial performance in Sri Lanka, *International Journal of Arts and Commerce*, 2017; 6(8).

[30] Yang, Z. M., Hua, X. Y., Shao, Y. H., & Ye, Y. F., A novel parametric-insensitive nonparallel support vector machine for regression, J*ournal of Neurocomputing*, 2016; 171: 649-663.

[31] Mello-Román, JD., Hernandez, A., KPLS optimization approach using genetic algorithms, *Journal of Procedia Computer Science*, 2020; 170: 1153-1160. Doi:10.3390/math9050506

[32] Huerta, R., Corbacho, F., and Elkan, C., Nonlinear support vector machines can systematically identify stocks with high and low future returns, *Journal of The Algorithmic Finance*, 2013; 2(1): 45-58. Doi. 10.3233/AF-13016

[33] Huang, C. J., Yang, D. X., and Chuang, Y. T., Application of wrapper approach and composite classifier to the stock trend prediction, *Expert Systems with Applications*, 2008; 34(4): 2870-2878. Doi:10.1016/j.eswa.2007.05.035

[34] Lee, M. C., Using support vector machine with a hybrid feature selection method to the stock trend prediction. *Journal of Expert Systems with Applications,* (2009); 36(8): 10896-10904. Doi:10.1016/j.eswa.2009.02.038

[35] Żbikowski, K., Using volume weighted support vector machines with walk forward testing and feature selection for the purpose of creating stock trading strategy, *Journal of Expert Systems with Applications,* 2015; 42(4), 1797-1805. Doi:10.1016/j.eswa.2014.10.001

[36 Manjami, Seyyed Amir Hossein, Abzari, and Raitishawazi., Forecasting stock prices in the stock market using fuzzy neural network and genetic algorithms and comparing it with artificial neural network, *Journal of Quantitative Economics*, 2009; 22(6): 1-26, (In persian).

[37] Dash, M., and Liu, H., Feature selection for classification, *Intelligent data analysis,* 1997; 1(1-4): 131-156. Doi:10.1016/S1088-467X(97)00008-5

[38 Rosipal, R., and Trejo, L. J., Kernel partial least squares regression in reproducing kernel hilbert space, *Journal of machine learning research*, 2001; 2: 97-123.

[39] Robnik-Šikonja, M., & Kononenko, I., An adaptation of Relief for attribute estimation in regression, *In Machine learning, Proceedings of the fourteenth international conference*, 1997; 5: 296-304.

[40] Alpaydin, E., Machine learning and data analytics, *Mit Press*, 2021.

[41] Moez, E. G., & Amina, Z., Overinvestment of free cash flow and manager's overconfidence, *International Business Research,* 2018; 11(3): 48-57.

[42] García Lara, J. M., García Osma, B., and Penalva, F., Accounting conservatism and corporate governance, *Journal of Review of accounting studies,* 2009; 14: 161-201.

[43] Fallahshams, M., and Razmian Moghadam, Z., The Empirical Test of the relationship between information asymmetry, Overvalued Equities and Stock Price Crash Risk, *Advances in Mathematical Finance and Applications,* 2023; 8(3): 769-781. Doi:10.22034/AMFA.2019.1873371.1255.

[44] Nozarpour, Y., Davoodi, S. M. R., and Fadaee, M., Selecting The Optimal Multi-Period Stock Portfolio with Different Time Horizons in the Credibility Theory Framework, *Advances in Mathematical Finance and Applications,* 2023; 8(3): 1043-1056. Doi: 10.22034/AMFA.2022.1953564.1709.

[45] Talebi, B., Abdi, R., Hajiha, Z., and Rezaei, N., Modelling Optimal Predicting Future Cash Flows Using New Data Mining Methods (A Combination of Artificial Intelligence Algorithms), *Advances in Mathematical Finance and Applications*, 2023; 8(3): 1007-1022. Doi: 10.22034/AMFA.2021.1910900.1493