Adv. Math. Fin. App., 2022, **7(**1), P. 245-260

Advances in Mathematical Finance & Applications
www.amfa.iau-arak.ac.ir
Print ISSN: 2538-5569
Online ISSN: 2645-4610
Doi: 10.22034/AMFA.2019.579662.1145

Applied-Research Paper

Evaluation of Intelligent and Statistical Prediction Models for Overconfidence of Managers in the Iranian Capital Market Companies

Shokoufeh Etebar[a], Roya Darabi[b, *], Mohsen Hamidian[b], Seiyedeh Mahbobeh Jafari[b]

[a] *College of Skills and Entrepreneurship, Karaj Branch, Islamic Azad University, Karaj, Iran*
[b] *Department of Economics and Accounting, Islamic Azad University, South Tehran Branch, Tehran, Iran*

ARTICLE INFO
Article history:
Received 2018-11-21
Accepted 2019-08-04

Keywords:
managerial overconfidence
machine learning Adaboost
Algorithm
Probit Regression

ABSTRACT

Behavioural characteristic managerial overconfidence of managers effects on the investment and financing decisions and company performance in the long run. the purpose of the present study was to validate the Adaboost machine learning and probit regression in the prediction of Management's overconfidence at present and in the future. It also compares the predicted models obtained during the years 2012 to 2017. The samples of the research were the companies admitted to the Tehran Stock Exchange, Data collection in the theoretical section of this study uses content analysis of international scientific articles in the library method and calculating the data used by Excel software using Matlab 2017 and Eviews10.0 to test the research hypothesis. The empirical findings demonstrate that The Adaboost's algorithm nonlinear prediction model represents the highest power in learning and prediction (performance of this model) the managerial over-confidence for this year and the next year, proved to be better than the probit regression prediction model.

1 Introduction

One of the Managers goals is to utilize scarce resources for better rendering services and providing competitive benefits in the manufacturing companies by improved in the (technology, employee efficiency, industry practices, macroeconomic status, and investment) in enterprise projects [8]. In psychology science and financial management, characteristic behavior manager's is one of the known possible factors of overconfidence of management. Managerial overconfidence is one of the insights judgments and affected in decision making. The accounting information system, as a sub-system of management information, is one of the critical tools for making information relevant for decision-making. The work of this system, on the one hand, depends on the quality of the information provided, and on the other hand, on the functioning of management [12]. More importantly, this is a well-documented measurable

* Corresponding author. Tel.: +98 9123194494
E-mail address: *: royadarabi110@yahoo.com*

administrative function and a significant explanation for the company's financial policies [27]. Empirical evidence in various social studies shows the overconfident managers are too optimistic. The market considers the value of corporate projects to be less realistic. Optimistic managers often predict cash flow, and the company's investment opportunities will be more valuable. The result also increases the likelihood of exchange of shares, which results in long-term debt rather than short-term debt. Overconfident managers have a wrong belief that the stock market values the value of the company's projects less than real they are. [10]. Optimistic managers often predict, exceed cash flows, and the company's investment opportunities will be even most valued [18].

International studies in the field of financial management in the past two years have been using a combination of machine learning methods, artificial intelligence algorithms, such as researches by Marucci-Wellman et al. [29] that used from neural networks to make short-term predictions in their research [29]. Also, Baig et al. [6] performed the AdaBoost algorithm for learning to use artificial neural networks [6]. The importance of research because of the request by various stakeholder groups and other users of the company's financial information on the lay down of rules for protecting against Conflict of interest, the information asymmetry and managerial overconfidence that will be lead to manipulation of accounting figures and the detriment the company's stakeholders in the future [7]. This research has three perspectives on innovation. First: providing a model for managerial overconfidence in the capital market of Iran and Comparison of methods machine learning Adaboost Algorithm and Probit Regression for predicting managerial overconfidence. Two: The use of the most significant number of independent variables in the present research, which has been carried out in researches both inside and abroad of Iran. Three: Use of LVF filter algorithm to select variables of the present research. Also, considering that so far no research inside Iran has not compared the validation and the predictive power of artificial intelligence machine learning Adaboost algorithm and Probit Regression. This research is important. The goals of this research are: validating and predicting the managerial overconfidence with the AdaBoost algorithm machine learning and probit regression and compare the Adaboost algorithm model with a probit regression for the predictive power of managerial overconfidence in Iran's capital market. Therefore, this research attempted to render the prediction models regarding recent literature and applied research abroad and theoretical literature of Iran about managerial overconfidence, in companies admitted to the capital market of Iran. This paper organized as follows. Section1 is the introduction, and section 2 presents the relevant literature and develops research hypotheses. In Section 3, describe the sample, data, and choice of variables by LVF filter algorithm, research design. Section 4 discusses the empirical results, while Section 4 conducts the robustness and performance models test. In part 4.2, provide additional analysis tests. Section 5 draws conclusions based on the findings.

2 Literature Review and Development of Hypotheses

What is the effect of Managerial overconfidence in Long-term performance of firms? Research in finance and economics so far has given little consideration to this question. Theoretical research suggests a reason: over-confidence can benefit shareholders by increasing investment in risky projects. Findings suggest that over-confidence helps CEOs exploit innovative growth opportunities [18]. Many studies show that overconfidence affects corporate investment, financing, and dividend policies such as [26-11-13-18]. Several authors have researched about managerial overconfidence in Iran such as [14] that Evidence from the experimental results of their research showed that the behavioral variables studied in the research has a significant and inverse effect on the stock return of the companies. In the other studies also [5-6-21-28] managerial overconfidence subject with company value, capital structure and is investigated. Also, the impact of the management performance evaluation methods on the information quality in accounting had studied in [30]. Recent work in accounting examines, the impact of overconfidence such as [34] in 2012 and the possibility of issuing a management forecast by [19, 25, 2] an investigation on financial reporting quality, showed that overconfidence effects on the financial reporting. In research on overconfident managers and internal controls, the findings of [10] suggest that threshold for cost-effective internal controls will differ across firms based on the characteristics of their management team [34-19-25-2-10]. In the results of research [29] indicate that, in terms of accounting interpretation, the rate of earnings response coefficient does not affect the relationship between overconfident of management and conditional conservatism. In research [11], there was a significant relationship between the managerial overconfidence index and the financial performance and the quality of financial reporting [11].

Bharati et al. [9] provide evidence of the critical impact that Sarbanes–Oxley Act has on the relationship between CEO overconfidence and firm policies, and the role of the financial reporting environment in selecting a new CEO from within versus outside the organization [9]. Weak reporting controls allow the CEO to misreport performance information, which reduces the board's ability to detect and replace poorly-performing CEOs as well as aggravating incentive contracting [28]. The stock market is affected by news and information. If the stock market is not efficient, the reaction of stock price to news and information will place the stock the behavioral finance approach modeled the market in overreaction and under-reaction states in the study [31], the reaction of the stock price in the stock market. Managers often express their behavioral characteristics when making long-term decisions, and manage the investment of financial resources and respond to changes in the business environment. View in this research is using accounting variables and financial ratios. Can predict managerial overconfidence? In this research will be reviewed the recent researches (Ability to predict Managerial overconfidence and Applying Artificial Intelligence Algorithm) by using financial data. Therefore, the central question of the research is: whether, by accounting variables and financial ratios, in the methods of probit binary regression and machine learning Adaboost algorithm can predict managerial overconfidence? So far, numerous studies on modeling in various topics have been done using Artificial intelligence algorithm and Artificial Neural Networks. Among the many issues related to data stream applications, those involved in predictive tasks such as classiifcation and regression play a signi. cant role in Machine Learning (ML). Junior and Nicoletti [23] in their research, using a new classification in the Boosting algorithm, were able to play an essential role in machine learning (ML) among many issues related to data flow programs (easy and flexible updates) and those who play predictive activities such as classification and regression by suggesting a new competitive model. In the research will be reviewed the Ability to predict Managerial overconfidence and Applying Artificial Intelligence Algorithm by using financial data [23].

The probit model is a famous model for fitting the binary response variables, which have only two outcomes that can found in the research variable, managerial overconfidence. Antunes et al. [5] in their paper," Forecasting banking crises with dynamic panel probit models"; forecast performances of several (dynamic) probit models to develop common vulnerability indicators with early warning properties [5]. The paper of Han and Vytlacil in 2017 provides identification results for a class of models specified by a triangular system of two equations with binary endogenous variables [17]. Martinetti and Geniaux present a new estimation method for spatial binary probit models in 2017 both spatial autoregressive (SAR) and spatial error (SEM) models considered. Whether a firm can attract foreign capital and whether it may participate in the export market depends on whether the fixed costs associated with doing so at least covered by the incremental operating profits [28]. These findings established through the estimation of a spatial bivariate probit model.

H1: Probit binary regression model has the ability to predict overconfidence of managers in the companies admitted to Tehran Stock Exchange for the current and the next year.
Jun et al. [22] in their research at 2018, in their research proposed a model for the tourism industry by using the forecasting algorithm model. It compared with three other ANN-based models and the most popular ARIMA model using three non-linear, non-stationary tourist arrivals data series. Popular ARIMA model using three non-linear, non-stationary tourist arrivals data series. Studies on experimental cases demonstrated that the proposed combination method consistently outperformed the other related methods. Their results indicated that using ANN and ARIMA models with three, time series (combined method) improved the results of other related methods. A boosting-based method of learning a feed-forward artificial neural network (ANN) with a single layer of hidden neurons and a single output neuron was presented [22]. Baig et al. [6] in a study titled "Adaboost-based artificial neural network learning" in the year 2018 presented a detailed performance comparison of various neural network models by training the proposed methods. The proposed method uses series representation to approximate non-linearity of activation functions, by training the coefficients of nonlinear terms by AdaBoost [6]. With considering such research, decided to test the machine learning Adaboost algorithm method for the prediction of management overconfident at the level of detection rate (Predictive Power) 90% hypotheses one mentioned and tested as follows:
H2: Machine learning Adaboost Algorithm model has the ability to predict overconfidence of managers in the companies admitted to Tehran Stock Exchange for the current and the next year.
In applied research, Kang et al. in 2017, the results from simulation and empirical analyses support the model's predictions. Thus, while managers' cognitive biases, when considered separately, negatively impact firm performance, they can be beneficial when considered jointly [24]. Other research such as [1-29-27] that compared predictive models with each other created hypothesis 3 in our minds.
H3: There is a significant difference (At the level of Predictive Power or detection rate 90%) between the overconfidence of management based on Adaboost algorithm and probit regression prediction models for managers of companies admitted to Tehran Stock Exchange for the current and the next year.

3 Research Design

3.1 Data and Samples Selection

This applied research is prediction research based on quantitative and post-event data that uses documentary data and financial information through referring to financial statements of companies admitted to Iran Stock Exchange and databases and Sites related to Tehran Stock Exchange. The statistical population of this research includes all companies accepted in Tehran Stock Exchange from 2012 to 2017. Due to limitations: High Tolerance Filtering (Upper and Lower Limits) the range of variations of research variables, in other words, Homogeneous data to be used. Other constraints were: The financial statements data should also be available, and they are not the investment, bank, and leasing companies. They did not change the activity.
Did not change the financial period (The financial period ended March 19 per year) and for comparability of the information, their fiscal year ends were March 19., As for Limitations on companies' activation in Tehran Stock Exchange (Their stocks traded on the market during the research period), year-company observations (sample), do not require a time-series feature in conducting research. The final samples were 784 data (observations of the year-company) that samples were available from 193 companies in the period 2012-2017 and therefore selected as a statistical sample for this research.

3.2 Model and Method of Selecting Research Variables

In the present research, the central question of the research is: whether, by accounting variables and financial ratios, in the methods of probit binary regression and machine learning Adaboost algorithm can predict managerial overconfidence? Therefore, three questions were raised in connection with the central research question:

1. Whether Machine learning Adaboost Algorithm model has the ability to predict overconfidence of managers in the companies admitted to Tehran Stock Exchange for the current and the next year?
2. Whether probit binary regression model has the ability to predict overconfidence of managers in the companies admitted to Tehran Stock Exchange for the current and the next year?
3. Whether There is a significant difference (At the level of Predictive Power or detection rate 90%) between the overconfidence of management based on the Adaboost algorithm and probit regression prediction models for managers of companies admitted to Tehran Stock Exchange for the current and the next year? Due to the Literature review and research questions, the following hypotheses were presented in this study.

H1: Probit binary regression model has the ability to predict the overconfidence of managers in the companies admitted to Tehran Stock Exchange for the current and the next year.
H2: Machine learning Adaboost Algorithm model has the ability to predict the overconfidence of managers in the companies admitted to Tehran Stock Exchange for the current and the next year.
H3: There is a significant difference (At the level of Predictive Power or detection rate 90%) between the overconfidence of management based on Adaboost algorithm and probit regression prediction models for managers of companies admitted to Tehran Stock Exchange for the current and the next year.
Overconfidence of management dependent variable is the difference between objective management precision and subjective certainty in management decisions [15].
The most consistent operational definition of the dependent variable was provided by [19-31-27-34], who stated that the first criterion is the remainder regression equation for capital expenditure. This research measure is based on the Earnings per share prediction error criterion. This criterion calculates by the difference between the forecast earnings per share and actual profit. If the expected profit is higher than real profit, it will get 1, in which case, the manager is overconfident; otherwise, it will be zero, in which case, the manager is not overconfident.

3.2.1 Independent Variables

The independent variables of the study include accounting variables and financial ratios extracted from financial statements in Iran's capital market. Definition of independent variables will be presented in the following Table 1 [35].

3.2.2 Conceptual Model of Research

The method of testing the hypotheses based on the research model was firstly tests of WIFE, and Dickey fuller was done to examine the status of the research variables. (1) Study of the Multicollinearity between independent variables, and (2) Evaluation of the false (or correct) of the estimated regression model. Then the method of testing the hypnosis 1 (the probit regression) is run. In the second section, Hypothesis 2 is tested, using the implementation process of the artificial intelligence algorithm (machine learning Adaboost algorithm). The method of testing the hypotheses one based on the variable-selection method by the step by step regression, that it is a statistical approach for selecting independent variables.

Table 1: Variable Definitions

Variables	definition
Value of company	Logarithm of the market value of the company
Margin of net profit	Sales minus net profit divided to net sales
Return on assets	Net profit divided to total assets
Earnings per share	Net profit divided to the number of shares
Current ratio	Current assets divided to current debt
Quick ratio	(Current assets minus inventory minus prepayment) divided to current debt
Working capital ratio	(Current assets minus current debt) Divided to assets
Financial leverage	Debts divided to assets
The proportion of sales to total inventories	sales divided to inventories
The proportion of sales to total assets	sales divided to total assets
The proportion of sales to fixed assets	sales divided to fixed assets
The proportion of sales to accounts receivable	sales divided to fixed accounts receivable
Net profit to sales ratio	Net profit divided to net sales
Stock return	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 dividend approved minus the increase of capital from the place of cash and claims divided to the market value of the company at the beginning of the year
Cash to assets ratio	cash dividend to total assets
Operating Cash to assets ratio	Operating cash dividend to total assets
Current asset to assets ratio	Current asset dividend to total assets
Firm size	Logarithm of the total assets
Stock price	Stock price in the end of year
Dividend ratio	Dividends per share dividend to earnings per share
Operating profit ratio	Operating profit dividend to earnings per share
Price to earnings per share ratio	Price of per Stock dividend to earnings per share

Steps to Implementing the Machine Learning Algorithm are as follows:
Collecting data: Proper design of the variables needed to hold the problem data.

- **Primary data analysis**

Statistical analysis of data in order to achieve a better understanding of the data structure

- **Select variable or extract suitable features**

The decision, about which variables can help reach a right answer, is considered a significant challenge in pattern recognition. Various methods have been developed to evaluate only a small number of sub-regional regression models by adding or removing single to single regression variables. Stepwise regression is a modification of the forward selection method, in which at each step, all of the regression variables already entered into the model with their incomplete statistics re-evaluated. In this research, in order to implement the second part of the step-by-step regression, the choice of research variables is used.

- **Clustering or categorization without data monitors (data split using 10-fold cross-validation method)**

Because this classification is done unambiguously at this stage, this stage considered as an exploratory analysis of the data. Due to its automatic nature and the lack of the need for supervisory intervention, this can result in useful results. One of the criteria used to evaluate a predictor is the error rate. Usually, algorithms tend to approach their actual error rate to the estimated error rate, which is possible by repeatedly executing the learning and evaluation process; So when a dataset will be available, algorithm can it left out part of it for final evaluation and used the rest for learning, and again changed two sets and re-evaluate the model. One standard method for this purpose is K-Fold Cross-Validation. In this way, the datasets are divided into equal parts randomly. In the first part of the K section, the part of the remainder is used to training. In the second installment of the second part of the K section, in order to evaluate, the K-1 remainder part is used for learning. K rank the order of the algorithm is executed in the same way. The learning and evaluation datasets should be large enough to bring the estimated error closer to the actual value. However, learning and evaluation data with the learning data and evaluation of other repetitions should have the least overlap so that all data is involved in the learning and evaluation process.

- **Implementation (training and evaluation process) of machine learning Adaboost algorithm**

First, the data of the independent variables selected using the ten methods of mutual validation to training and evaluation by the algorithm are continuously divided (up to the repetition order of the algorithm) to be able to estimate Bayesian algorithm with using educational and test data. Estimated Error (Estimation) and evaluation of the performance of the model reported in ten cross-validations (ten-layer validation), and the average of these detection rates (error and predictive power) provide in the outputs of the algorithm's matrix.

- **Interpreting the results**

By using from the ten-cross validation method, there are always many companies to learn algorithms and many companies to test or evaluate the efficiency of the algorithm, the algorithm executed. According to instructions and through the average prediction percentages by algorithm, finally shows that the predictive power of the Bayesian algorithm will be several percent's of the 100% or the actual value. If the percentage of prediction or detection rate by the evaluation data that represents the performance or applicability of the model is closer to 100%, and the error of prediction is lower, the prediction of AdaBoost algorithm is closer to reality in the matlab software.

3.3 Research AdaBoost Machine Learning Algorithm

Adaboost, a concise version of adaptive boosting, is an automated learning algorithm under monitoring. Adaboost can combine a large number of learning algorithms to improve performance and improve applicability. The basic classification used for the Adaboost algorithm only needs to be better than random classification and, in this way, the performance of the algorithm is improved with more repetitions. Even higher-than-randomized classifications improve overall performance by obtaining a negative coefficient. AdaBoost is sensitive to Noise data with distinct sections and, in over fitting issues; it is less sensitive to other learning algorithms. At first, the weight of all samples is the same, but at each repetition, the poorly trained structure offers the classification and the weight of the samples classified using this incorrect classification increases. Thus, the focus of the algorithm is on hardly categorized samples. The final classification is made by majority voting on the classifier, where classifiers that have less error have more weight.

Algorithm: AdaBoost type alpha: Inputs are X, T, D. Definitions of the inputs are included, X is: Training data set and T: is algorithm the number of iteration, D: is Initial weight of samples. Definitions of the Outputs are included, ω is: The final weight of the classifiers h is: Final classifiers the machine learning. Adaboost algorithm runs in accordance with this command.

Table 2: Descriptive Statistics of Independent Variable

Variables	Mean	median	Minimum	Maximum	Standard deviation	Skewness	Kurtosis
Value of company	13.908	13.717	10.133	18.863	1.566	.706	.758
Margin of net profit	0.849	.865	.333	1.681	.153	-.014	1.888
Return on assets	0.119	.106	-.230	.564	.111	.556	.994
Earning per share	810.714	514.233	-925/799	4659.799	892.743	1.413	1.979
Current ratio	1.368	1.236	.223	5.552	.674	2.148	7.237
Quick ratio	.791	.723	.058	3.235	.453	1.790	5.735
Working capital ratio	.123	.129	-.660	.658	.210	-.139	.172
Financial leverage	.601	.615	.131	1.935	.179	.648	5.211
The proportion of sales to total inventories	4.735	3.812	.437	28.692	3.497	2.662	10.784
The proportion of sales to total assets	.911	.785	.076	3.795	.509	1.933	5.673
The proportion of sales to fixed assets	5.258	4.064	.147	20.839	4.242	1.280	1.251
The proportion of sales to accounts receivable	6.788	3.380	.159	86.694	11.360	4.406	22.453
Net profit to sales ratio	.151	.135	-.681	.667	.153	.014	1.888
Stock return	46.562	13.970	-64/490	739.090	98.030	2.803	11.053
Cash to assets ratio	.042	.027	.0001	.367	.046	2.581	9.081
Operating Cash to assets ratio	.131	.113	-.283	.557	.118	.533	.748
Current asset to assets ratio	.644	.681	.138	.948	.190	-.607	-.472
Firm size	14.114	13.903	10.617	19.152	1.513	.854	1.285
Stock price	6308.983	4094.000	482.000	41323.000	6182.604	2.295	6.422
Dividend ratio	.061	.034	.000	.529	.078	2.442	7.985
Operating profit ratio	.190	.174	-.475	.602	.150	-.036	.562
Price to earnings per share ratio	8.647	6.796	-99.077	57.759	12.478	-.569	10.717

I. Primary weighing: (This step is executed once in an algorithm)

Uniforms Primary Distribution for Trial Samples at the Start of the Algorithm: $D_1(i) = 1/n$

T: Determination the number of repetitions and selecting weight or confidence measure $\omega_t \in R$ in the stage.

II. Repeat it for t = 1, ..., T: (this step repeats for 1,…, n data)

Applying the classifiers, to the samples and Calculating for the classifiers error $h_t(\epsilon_t)$ is the step. In the steps the amount of weight (ω_t)to be determined in each classifier, In accordance with this formula: $\omega_t = \frac{1}{2}\ln\left(\frac{1-\epsilon_t}{\epsilon_t}\right)$ then Optimize the distribution of the training Collection with this formula:$D_{t+1}(i) = \frac{D_t(i)e^{-\omega_t y_i h_t(x_i)}}{Z_t}$ is the step of runs Adaboost algorithm which is repeated for n degree. When Z_t is the normalization factor of weights $\omega = \omega_t,\ h = h_t, Z = Z_t, D = D_t$ and y_i is the target output and will be ended Repeat the algorithm command.

The AdaBoost algorithm has many different types, including the alpha and beta types, which have undergone various improvements. Boosting action (Reinforcement) can be considered as minimizing a convex function on a convex set of functions. Specifically, a function that is minimized is the Exponential function:

$$G = \sum_i e^{y_i f(x_i)}$$

The final form of the AdaBoost algorithm is to find the following model.

$$f(x) = \sum_t \omega_t h_t(x)$$

4 Empirical Results

4.1 Descriptive Statistics

Table 2 presents the descriptive statistics for variables. Since the dataset is a pooling data, Descriptive indexes of the variables selected including mean, the range of variation, standard deviation, and skewness and kurtosis, which are used in this study to predict the overconfidence of management, are presented in Table 2. For the quality variable of the managerial overconfidence in use, the number and percentage of descriptive statistics have been used, as presented in Table 3.

Table 3: Descriptive Statistics of Depended Variable

Managerial overconfidence	Number	percentage
no	451	57.5
yes	333	42.5

Table 4: VIF Test Results

Variables	Tolerance	Variance inflation factor (VIF)
Value of company	.064	15.573
Return on assets	0.078	12.824
Earning per share	0.143	6.974
Current ratio	0.144	6.956
Quick ratio	0.231	4.333
Working capital ratio	0.083	12.019
Financial leverage	0.183	5.476
The proportion of sales to total inventories	0.514	1.946
The proportion of sales to total assets	0.368	2.716
The proportion of sales to fixed assets	0.427	2.339
The proportion of sales to accounts receivable	.666	1.502
Net profit to sales ratio	.134	7.471
Stock return	.769	1.301
Cash to assets ratio	.781	1.280
Operating Cash to assets ratio	.477	2.094
Current asset to assets ratio	.160	6.251
Firm size	.068	14.783
Stock price	.202	4.956
Dividend ratio	.476	2.099
Operating profit ratio	.197	5.071
Price to earnings per share ratio	.892	1.121

4.2 Summary Statistics

WIFE test did for financial data, before the implementation of probit regression modeling in Eviews software. Table 3 shows the WIFE test (Intensity of co-linearity between independent variables) was used to examine the existence of a co-linearity between the independent variables that despite the high coefficient of determination, the validity of the model goes under the question. In order to investigate the co-linear severity of the independent variables of the research, the inflation factor variance has used. The Detection threshold amount of variance inflation factor (VIF) is less than 5. According to the results of variance inflation factor independent variables provided in the Table 4, co-linearity between the independent variables (Quick ratio, The proportion of sales to total inventories, The proportion of sales to total assets, The proportion of sales to fixed assets, The proportion of sales to accounts receivable, Stock return, Cash to assets ratio, Operating Cash to assets ratio, Stock price, Dividend ratio, Price to earnings per share ratio) is not proved because their variance inflation factor (VIF) was less than 5, so these variables does not co-linearity, and in the following to test the probit regression, they are tested.

Table 5: The Dickey-Fuller Test Results

Variables	The probability obtained in the probability level of <0.001 is considered	Statistic	The Dickey-Fuller result test
Quick ratio	<0.0001	-13.458	Variable with other variables haven't unit roots.
The proportion of sales to total inventories	<0.0001	-15.422	Variable with other variables haven't unit roots.
The proportion of sales to total assets	<0.0001	-12.653	Variable with other variables haven't unit roots.
The proportion of sales to fixed assets	<0.0001	-14.580	Variable with other variables haven't unit roots.
The proportion of sales to accounts receivable	<0.0001	-16.050	Variable with other variables haven't unit roots.
Stock return	<0.0001	-32.069	Variable with other variables haven't unit roots.
Cash to assets ratio	<0.0001	-20.720	Variable with other variables haven't unit roots.
Operating Cash to assets ratio	<0.0001	-19.423	Variable with other variables haven't unit roots.
Stock price	<0.0001	-16.250	Variable with other variables haven't unit roots.
Dividend ratio	<0.0001	-15.524	Variable with other variables haven't unit roots.
Price to earnings per share ratio	<0.0001	-19.890	Variable with other variables haven't unit roots.

Table 5 presents the results of the Dickey-Fuller Test. The validity of regression estimation examined in different ways. Usually, non-durable variables lead to a false evaluation of the regression. Before estimating the model, it is necessary that the durability of all variables used in the estimation tested. In this research, before the probit regression was used to investigate the unit root test of Dickey-Fuller, The results indicate that independent variables in the research are at a significant level (p <0.0001) So, in short, we can say that based on Dickey Fuller's method, the null hypothesis of the test concerning the existence of a single root was rejected and, accordingly, the variables used in this research have not The root of the unit.

4.3 Assessment of Probit Regression Model

In Table 6 we report the Binary probit regression results on total independent variables for predicting managerial overconfidence. The resulting regression model is as follows:

$$CEO_{i,t+1} = \alpha_0 + \alpha_1 QR_{i,t} + \alpha_2 STI_{i,t} + \alpha_3 AT_{i,t} + \alpha_4 FAT_{i,t} + \alpha_5 TAR_{i,t} + \alpha_6 SR_{i,t} + \alpha_7 CTA_{i,t} + \alpha_8 OCFTA_{i,t} + \alpha_9 SP_{i,t} + \alpha_{10} DPS_{i,t} + \alpha_{11} P/E_{i,t} + \varepsilon_{i,t}$$

Independent variables were associated with statistically significant results obtained in Table 6 are expressed. As shown in Table 6, the findings support the prediction in the first hypothesis that the probi.. binary regression model has the ability to predict the overconfidence of managers in the Companies admitted to Tehran Stock Exchange for present and the next year. The Probit regression model with three variables (The proportion of sales to fixed assets with The significance level 0.046, Cash to assets ratio with The significance level 0/017, Stock price with The significance level 0/010) create the maximum coefficient of determination for the probit regression model in the significance of probability less than <0.0001. The determination coefficient is 26% that can say that independent variables of the prediction model only explain 26% of the variations of the dependent variable of overconfidence of management. The variable coefficient coefficients are shown in Table 6.

Table 6: Probit Regression Prediction Model Results

Variables in the probit regression model	Variable coefficient (beta)	The standard deviation (std error)	T Statistic	Sig
Quick ratio	0.015	0.261	0.003	0.956
The proportion of sales to total inventories	-0.020	0.036	.307	0.579
The proportion of sales to total assets	-.277	0.278	.993	0.319
The proportion of sales to fixed assets	-0.060	0.030	3.969	0.046
The proportion of sales to accounts receivable	-.003	0.010	.086	0.770
Stock return	0.0001	0.001	0.002	0.964
Cash to assets ratio	-6.588	2.772	5.647	0.017
Operating Cash to assets ratio	-0.446	1.083	0.169	0.681
Stock price	0.0001	<0.001	6.583	0.010
Dividend ratio	0.887	1. 588	.312	0.576
Price to earnings per share ratio	0.007	0.010	0.572	0.450
Constant amount of model	0.153	0.298	.265	0.607
determination coefficient	0.260			
Likelihood Ratio Statistic	17.522			
the amount of obtained probability	<0.001			

4.4 Artificial Intelligence Algorithm

4.4.1 Variable Selection with Step-By-Step Regression Method

In order to predict managerial overconfidence using AdaBoost algorithm used in this research, using the stepwise method, we selected the significant variables and then used for analysis in the AdaBoost algorithm. According to stepwise regression results, the proportion of sales to fixed assets, Stock return, Cash to assets ratio, Operating Cash to assets ratio, Stock price, and variables from between initial variables (Twenty-two independent variable) selected for predicting managerial overconfidence.

Table 7: Stepwise Regression Results

Variables	Variable coefficient (beta)	The standard deviation (std error)	Statistic	Sig
The proportion of sales to fixed assets	-0.081	0.019	17.590	<0.001
Stock return	-0.003	0.001	13.613	<0.001
Cash to assets ratio	-6.543	1.982	10.895	0.001
Operating Cash to assets ratio	-1.586	0.691	5.270	0.022
Stock price	0.0001	<0.001	8.715	0.003
Constant amount of model	0.469	0.164	8.154	0.004
determination coefficient	0.27			
Wald statistic	17.625			
the amount of obtained probability	<0.001			

Table 8 reports the empirical results of the percentage of learning, Adaboost machine learning algorithms, by the use of the training database. To evaluate the reliability managerial overconfidence prediction models, which are based on the Adaboost machine learning algorithm, and which are the non-linear model. Also, in order to ensure fairness and to investigate the over-fitting phenomenon has been used from 10 Cross-Validation methods were used. The company's database divided into two groups, namely training, and test data, and then the command to execute the algorithm is given with the using 10 Cross-Validation methods. Learning data is given to the AdaBoost algorithm with the primary decision tree classes based on the Gini coefficient. After running the learning algorithms process, in order to investigate how the AdaBoost model has completed the learning process, then, the test data were given to the algorithm again. The average recognition rate of the AdaBoost algorithm model is aimed at evaluating the percentage of learning algorithm to predict the overconfidence of the management. The closeness of learning errors to zero or the amount of learning to 100% is an indication of better learning of the AdaBoost algorithm.

Scrutiny of Non-Occurrence: The Over Fitting Phenomenon (Applicability of Adaboost Machine Learning Algorithm for Predicting Managerial Overconfidence)

The companies test data (2012 to 2017) that is not seen by algorithm yet give to the Adaboost machine learning algorithm model. The AdaBoost algorithm predicts the overconfidence of management for all of these companies-years (test data). Comparing the average of estimated detection rate in the current and next year by ten cross-validation method, be determined the value of the applicability of prediction model. Adaboost algorithm model for unseen company-evaluation year's data (2012 to 2017) has predictive accuracy close to the company- training year's data (2012 to 2017) for the predictive model. So the overfitting phenomenon has not happened for the prediction of the Adaboost model. According to the average of the total results at 90% detection rate, the finding in Table 7 provide support for hypothesis 2, because, the average rate of detection equal 92/65 and the average of efficiency (applicability) equal 91/85 prediction model, provides stronger results (detection rates more than 90%) for the current year. The average rate of detection equal 91/85 and the average of efficiency (applicability) equal 89/43 prediction model, that is provided (detection rates very closer to 90%) for the next year.

Table 8: Results of the Prediction Model of the Adaboost Machine Learning Algorithm

Description	Results of learning (training) algorithm		Results of performance evaluation (applicability)	
Ten cross-validation method	Current year	Next year	Current year	Next year
1	91.92	93.67	90.19	90.49
2	92.13	89.36	94.11	92.40
3	94.00	94.87	92.59	90.22
4	94.11	91.20	92.57	89.87
5	91.70	90.63	98.03	86.07
6	92.35	91.92	94.11	84.61
7	93.44	91.07	88.23	91.02
8	93.23	91.78	90.19	88.46
9	92.13	91.02	86.27	90.65
10	91.48	94.87	92.15	90.50
Mean	92.65	92.04	91.85	89.43

4.4.2 AdaBoost Machine Learning Algorithm Prediction Model

In the AdaBoost algorithm to predict the dependent variable $f(x)$ Objective is to reach the parameter ω_t and Decisions tree learning $h_t(x)$ in the form of the following relationship. Each tree presents the prediction decision for the dependent variable that weighted collection of total predictions are (average) from the decision making.

$$f\left(x\right)=\sum_{t}\omega_{t}h_{t}\left(x\right)$$

Given that the purpose of the AdaBoost learning machine learning algorithm is combination linearly the number of t weak learning and in this research, the number t is 50 decision trees because the machine's learning error from amount of given weighted to the per path achieve to below the threshold of less than 0.5 or in other words modulus w_t-w_0 to be smaller than 0.5 forced to be selected 50 floors. Therefore, due to the massive volume and complexity of the nonlinear model, this artificial intelligence algorithm (learning AdaBoost machine) cannot show a linear model, and for better understand Only one tree from the next year is presented (one block) as a sample of a whole nonlinear model.

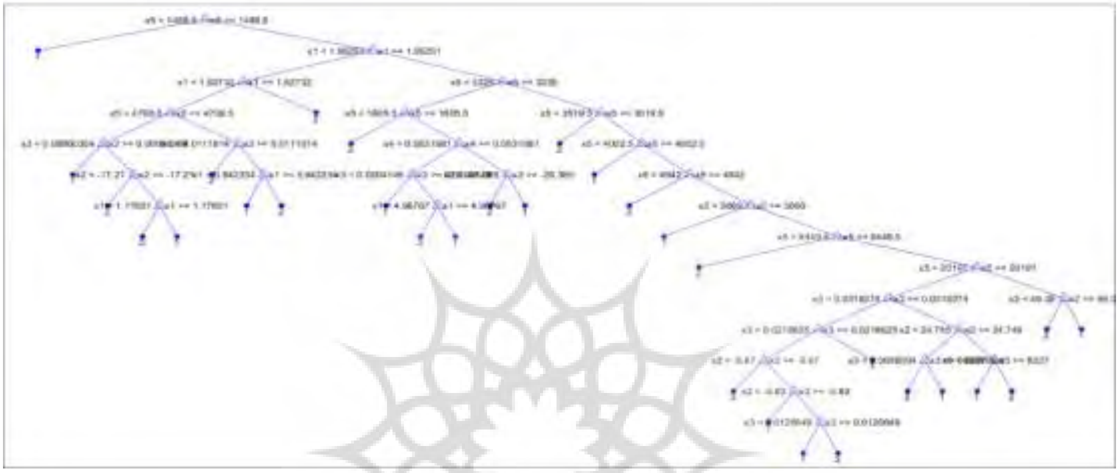

Fig. 1: The First Tree, Part of the Model Relating to Predicting Managerial Overconfidence for the Next Year

5 Conclusion

Considering the objectives of this research, we have validated AdaBoost machine learning algorithm as well as the probit regression to predict managerial overconfidence in accepted companies in Tehran Stock Exchange. Then, we compared the apparent and hidden (nonlinear) predictability models for the current and future management of the research period from 2012 to 2017. The remarkable point is that in the present study, the variables that have been examined Compared to other studies conducted so far in Iran such as [3-4-12-14-28-30]; all accounting variables used in the theoretical foundation's models of the Iranian studies covered and evaluated many more variables. Comparison of prediction models concerning the average rate of recognition (rate of learning by Adaboost model) and Average performance or Efficiency detection rate of algorithmic model with the obtained average of determination coefficient (that is the predictive power of the Probit regression model), It is possible.

The AdaBoost machine learning algorithm model is capable of predicting management overconfident in Matlab software, but AdaBoost machine learning algorithm, compared with the probit regression prediction model, has better results in predicting the overconfidence of management of this year and the next year in Tehran SEC. The probit regression prediction model, the results of which can be seen in the output of Eviews software as a regression model, have the lower ability to predict the validity of this year and the future years in Tehran SEC. Finally, we can claim that there is a significant difference between the predictive powers of the two models presented in the research for predicting managerial overconfidence. Results in the research in case of effect managerial overconfidence in the market value and component of capital structure are Compatible with the result of Ali Nejad Saroklaeiand and Sobhi in the year 2016. Also, Hribar and Yang research in the year 2016 results are consistent with the predictive behavior of managers. The results of research Bharati, Doellman, and Fu in the year 2016, as well as Chen et al. in the 2014 year, had to assess managerial overconfidence and stock returns. They obtained a significant and positive relationship. The results of their research are not consistent with the results of this research. The current research is consistent with the results of Kang et al. in the year 2017, which claimed that overconfident of management harms the financial performance of the company. The present research is consistent with the use of research variables with the research of Fonseca Costa et al. in the year 2017. Also, the results of this study are consistent with the use of combined algorithms for making models with the results of John et al., research in the year 2018.

Based on the results from accepted companies in Tehran SEC, the following suggestions for future research can be offered: investors and Financial market analysts and stockbrokers must have understood the required training and expertise in full recognition of various types of overconfidence and its consequences since the long-term impacts of overconfidence of managers will have adverse consequences for firms in the capital market, While in Iran these outcomes will be evaluated at a later time and with delay. As for suggestion can propose use from the other artificial intelligence algorithms and artificial neural networks and Fuzzy Logic and other alternative variable selection methods, and other Statistical and economic variables for future research such as macroeconomic, GDP and, etc. Also using variables that not used in the regression modeling of this research can provide newer prediction models for predicting overconfidence of management in a world because managers play a vital role in advancing the goals of public and private organizations.

References

[1] Abhinav, R., Pindoriya, N. M., Wu, J., Long, C., *Short-term wind power forecasting using wavelet-based neural network*, Energy Procedia, 2017, **142**(1), P. 455- 460. Doi: 10.1016/j.egypro.2017.12.071

[2] Ahmed, A. S., Duellman, S., *Managerial overconfidence and accounting conservatism*, Journal of Accounting Research, 2013, **51**(1), P. 1- 30. Doi:10.2139/ssrn.887301

[3] Aliahmadi, S., Jamshidi, A., Mousavi, R., *The Managerial Ability and Value of Cash: Evidence from Iran*, Advances in mathematical finance & applications, 2016, **1**(1), P. 19-31. Doi: 10.22034/AMFA.2016.526241

[4] Ali Nejad Saroklaei, M., Sobhi, M., *The Effect of More Trusted Managers on Capital Structure*, Financial Accounting and Audit Research, 2016, **8**(31), P. 93-109. Doi:10.30699/ijf.2020.121887

[5] Antunes, A., Bonfim, D., Monteiro, N., Rodrigues, P. M., *Forecasting banking crises with dynamic panel probit models,* International Journal of Forecasting, 2018, **34**(2), P. 249-275.
Doi: 10.1016/j.ijforecast.2017.12.003

[6] Baig, M. M., Awais, M. M., El-Alfy, E.-S. M., *AdaBoost-based artificial neural network learning, Neurocomputing*, 2017, Doi: 10.1016/j.neucom.2017.02.077

[7] Bamber, M., McMeeking, K., *An examination of international accounting standard-setting due process and the implications for legitimacy*, The British Accounting Review, 2016, **48**(1), P. 59-73.
Doi: 10.1016/j.bar.2015.03.003

[8] Berk, J., Stanton, R., *Managerial ability, compensation, and the closed-end fund discount*, Journal of Finance, 2007, **62**(2), P. 529-556. Doi: 10.1111/j.1540-6261.2007.01216.x.

[9] Bharati, R., Doellman, T., Fu, X., *CEO confidence and stock returns*, Journal of Contemporary Accounting & Economics, 2016, **12**(1), P. 89-110. Doi: 10.1016/j.jcae

[10] Chen, S. S., Lai, S.-M., Liu, C.-L., McVay, S., *Overconfident managers and internal controls*, working paper, National Taiwan University and University of Washington, 2014, Available at SSRN. Doi: 10.2139/ssrn.2510137 .

[11] Cordeiro, L., *Managerial overconfidence and dividend policy*, 2009, Available at SSRN. Doi: 10.2139/ssrn.1343805

[12] Demerjian, P., Lev, B. McVay, S., *Quantifying managerial ability: A new measure and validity tests*, Management Science, 2012, **58**(7), P. 1229-1248. Doi:10.1287/mnsc.1110.1487

[13] Deshmukh, S., Goel, A. M., Howe, K. M., *CEO overconfidence and dividend policy*, Journal of Financial Intermediation, 2013, **22**(3), P. 440-463. Doi: 10.1016/j.jfi.2013.02.003

[14] Abbasi, I., Vakilifard, H., Marouf, M., *Investigating the Impact of Managerial Uncertainty on Financial Reporting Quality and Conditional Conservatism in Tehran Stock Exchange*, Journal of Accounting and Auditing Management, 2018, **7**(25), P. 193-205

[15] Fonseca Costa, D., Carvalho, de Mel F., de Melo Moreira, B. C., do Prado, J. W., *Bibliometric analysis on the association between behavioral finance and decision making with cognitive biases such as overconfidence, anchoring effect and confirmation bias*, Scientometrics, 2017, **111**(3), P.1775-1799. Doi: 10.1007/s11192-017-2371-5

[16] Foroughi, D., Moazzeni, N., *The Effect of Excessive confident of Management on Company Value*, Financial Accounting Knowledge, 2016, **4**(4), P. 65-84

[17] Han, S., Vytlacil, E. J., *Identification in a generalization of bivariate probit models with dummy endogenous repressors*, Journal of Econometrics, 2017, **199**(1), P. 63-73. Doi: 10.1016/j.jeconom.2017.04.001

[18] Hirshleifer, D., Low, A., Teoh, S. H., *Are overconfident CEOs better innovators*, the journal of Finance, 2012, **67**(4), P. 1457-1498. Doi:10.1111/j.1540-6261.2012.01753.x

[19] Hribar, P., Yang, H., *CEO Overconfidence and Management Forecasting*, Contemporary Accounting Research, 2016, **33**(1), P. 204-227, Doi: 10.1111/1911-3846.12144.

[20] Jokar, H., Shamsaddinia, K., Daneshi, V., *Investigating the Effect of Investors' Behavior and Management on the Stock Returns: Evidence from Iran*, Advances in mathematical finance & applications, 2018, **3**(3), P. 41-52, Doi:10.22034/AMFA.2018.544948

[21] Jongjaroenkamol, P., Laux, V., *Insider versus outsider CEOs, executive compensation, and accounting manipulation*, Journal of Accounting and Economics, 2017, **63**(2-3), P. 253-261. Doi: 10.1016/j.jacceco.2017.01.002

[22] Jun, W., Yuyan, L., Lingyu, T., Peng, G., *Modeling a combined forecast algorithm based on sequence patterns and near characteristics: An application for tourism demand forecasting*, Chaos, Solutions & Fractals, 2018, **108**, P. 136-147. Doi: 10.1016/j.chaos.2018.01.028

[23] Junior, B., do Carmo Nicoletti, M., *An iterative boosting-based ensemble for streaming data classification*, Information Fusion, 2019, 45, P. 66-78. Doi: 10.1016/j.inffus.2018.01.003

[24] Kang, J., Kang, J.-K., Kang, M., Kim, J., *Curbing Managerial Myopia: The Role of Managerial Overconfidence in Owner-Managed Firms and Professionally Managed Firms*, Social Science Network of America, 2018, Doi: 10.2139/ssrn.2944998.

[25] Libby, R., Rennekamp, K., *Self-serving attribution bias, overconfidence, and the issuance of management forecasts*, Journal of Accounting Research, 2012, **50**(1), P.197-231. Doi: 10.1111/j.1475-679X.2011.00430.x

[26] Malmendier, U., Tate, G., *Who makes acquisitions? CEO overconfidence and the market's reaction*, *Journal of financial Economics*, 2008, **89**(1), P. 20-43. Doi: 10.1016/j.jfineco.2007.07.002

[27] Malmendier, U., Tate, G., Yan, J., *Overconfidence and early life experiences: the effect of managerial traits on corporate financial policies*, 2011, the journal of Finance, **66**(5), P. 1687-1733. Doi:10.1111/j.1540-6261.2011.01685.x

[28] Martinetti, D., Geniaux, G., *Approximate likelihood estimation of spatial probit models*, Regional Science and Urban Economics, 2017, 64, P.30-45. Doi: 10.1016/j.regsciurbeco.2017.02.002

[29] Marucci-Wellman, H. R., Corns, H. L., Lehto, M. R., *Classifying injury narratives of large administrative databases for surveillance- a practical approach is combining machine learning ensembles and human review*, Accident Analysis & Prevention, 2017, 98, P. 359-371. Doi: 10.1016/j.aap.2016.10.014

[30] Nasirzadeh, F., Abbaszadeh, M. R., Zolfaghararani, M. H., *Investigating the effect of managers' optimism and information asymmetry on stock price fall risk*, Journal of financial accounting, 2017, **9**(34), P. 35-70

[31] Panayiotis C, A., Doukas, J. A., Koursaros, D., Louca, C., *CEO Overconfidence and the Valuation Effects of Corporate Diversification & Refocusing Decisions*, Social Science Network of America, 2017. Doi: 10.1016/j.jbankfin.2019.01.009

[32] Parsamehr,H., Kasravi ,A., and Fazli, M., *Impact of the Management Performance Evaluation Methods on the Data Quality in Accounting*, Advances in mathematical finance & applications, 2017, **2**(1), P. 41-53,

[33] Rezaei, N., Elmi, Z., *Behavioral Finance Models and Behavioral Biases in Stock Price Forecasting*, Advances in mathematical finance & applications, 2018, **3**(4), P. 67-82

[34] Schrand, C. M., Zechman, S. L., *Executive overconfidence and the slippery slope to financial misreporting*, Journal of Accounting and Economics, 2012, **53**(1), P. 311-329. Doi: 10.1016/j.jacceco.2011.09.001

[35] Soleimani, M., Taherabadi,A., Karimipouya,M., *Accounting Ratios in the Analysis of Financial Statements: Financial Ratios of Companies Listed on the Tehran Stock Exchange*, The Secret Secret, 2015, **1**, P.51-58(in Persian).

[36] Zanjardar, M., Rafiei, Z., *Effect of the Earnings Reaction Coefficient on Extreme Confidence Relationships and Contingent Conservatism*, Financial Accounting and Audit Research, 2017, **9** (34), P.113-135