IRANIAN JOURNAL OF ACCOUNTING, AUDITING & FINANCE

journal homepage: https://ijaaf.um.ac.ir
DOI: 10.22067/ijaaf.v3i1.82217

Application of Artificial Intelligence Algorithm of Linear and Non-linear Relevance Vector Machine in Predicting the Bankruptcy

Seyed Hesam Vaghfi*
Department of Management,Economics and Accounting,Payame Noor University,Tehran,Iran

Abstract

The purpose of this research is the localization of a method for analyzing and predicting the bankruptcy of companies at three levels (financial health, bankruptcy, and bankrupt). In the first step, using the Relief-F multi-class, AI algorithm among 54 initial independent variables, and using information from 1488 companies-years during the period 2011-2016, the financial risk variables, working capital ratio, long-run debt ratio, asset flow ratio, the economic value added ratio, the ratio of non-executive managers, the ratio of current debts to equity, the ratio of debt to equity, corporate size, earnings management were selected as important variables in the prediction of the three-level bankruptcy situation, respectively. Using the relevance vector machine algorithm, the bankruptcy situation of companies is predicted in the upcoming year and next two years using MATLAB 2017 Software. The results of the research indicate that in general, the predictive power of the relevance algorithm in nonlinear mode is much higher than in linear mode, so that in the nonlinear mode, using the relevance vector machine algorithm, we can determine the company's bankruptcy with accuracy of more than 93% for the upcoming year and more than 86% for the next two years.

Keywords: Bankruptcy, Artificial Intelligence Algorithm, Global and Iranian Models of Bankruptcy, Tehran Stock Exchange

* Corresponding author: Faculty member of Payame Noor University, Email: Vaghfi2018@gmail.com

1. Introduction

The bankruptcy of large corporates and the financial scandals of some companies in recent years point to the need for broader and deeper research in the field of bankruptcy. Bankruptcy is a situation in which the company does not have enough cash to meet its financial obligations (Outecheva, 2007). In the 1990s, researchers such as Altman (1968) Springate (1978), Shirata (1995, etc) investigated the reasons for commercial failure of companies. They used financial ratios of companies to offer models for predicting bankruptcy. In Iran, according to Article 141 of the Commercial Code, if at least half of the company's capital is destroyed in the course of losses, members of the board of directors must immediately call the extraordinary general shareholders to decide on the dissolution or survival of the company. If the company is subject to Article 141, it must exit this situation within two years, otherwise, the acceptance of the company will be canceled (Saeedi & Aghaei, 2009). Owens (2017) argues that with the anticipation of bankruptcy, investor valuation will decline because they consider liquidation and liquidation values more relevant than future economic profits, indicating a decline. The information content is profit.

Bankruptcy literally means helplessness in business and commerce, and the business case is that the business is damaged and its debt is more than its assets. In the financial literature, there are indistinguishable words for bankruptcy. Some of these words are: financial distress, failure, business failure, deterioration, bankruptcy, inability to pay debt, helplessnessFinancial and so on. Brad Street defines the term bankrupt companies as, "Units that suspend their business because of a bankruptcy, bankruptcy, or cessation of credit operations by creditors."(Ghasemi & Ramezanpour, 2018).

2. Literature and hypotheses development

The bankruptcy of companies has always been one of the main concerns of creditors, investors, and governments so that timely detection of companies that are about to be financially distressed can partly offset the potential losses of the stakeholders (Mashayekhi & Ganji, 2014). The previous studies suggest that the returns of companies with a higher risk of bankruptcy are lower than those of other companies, which is due to anomalies in the risk of bankruptcy, because the high risk of bankruptcy or reduced ability to pay off debts at maturity means that the company loses its ability to spread debt with a low risk (George & Hwang, 2010).

2.1. Accounting criteria and its effect on Bankruptcy

The financial health of companies is related to the persistence of activities, bankruptcy, bankruptcy and the qualitative characteristics of accounting information. Therefore, they should assess the performance of firms in the fields of operations, marketing, finance, and accounting (Kritsonis, 2005). Financial ratio indicators which are considered as indicators of profitability and liquidity are those internal factors affecting bankruptcy analysis. Financial ratios are the most widely used tools for calculating the performance and health of bankruptcy of corporates. Profitability ratios measure the full financial status of a company (Mahswara, 2011). Zmijewski M.E using financial, liquidity and performance ratios could provide a model with a precision of 92% for financial health of companies (Zmijewski M.E, 1984). Wing Yu & et al. (2003) and Brewer & et al. (2012) have used financial ratios that include profitability information, liquidity performance to determine financial strength and financial health of companies. According to the researches by Ismail Zadeh & Shakeri (2015), Ahmadi (2016) Pindado & et al. (2008) Chen (2011) Sun & et al. (2011) Hu & Sathye (2015), Zohra & et al. (2015), financial and accounting ratios have been used for analysis of bankruptcy.

2.2. Corporate governance criteria and their impact on Bankruptcy

Financial statements can provide useful information about corporate operations for the public. Investors, in part, also invest on the basis of financial statements. However, when companies manipulate information and do not honestly disclose, the disclosures in the financial statements will be bundled with bias and will be deviated from the truth, but under the appropriate corporate governance mechanism, the board will be able to control the company and will prevent financial data from being distorted by the company's management. An appropriate governance system, first, ensures the minority shareholders that they have received reliable information about the value of the company and have not been abused by corporate executives and other large shareholders. Second, it encourages managers to seek to maximize the value and benefits of the company rather than addressing their personal goals (Sadeghi et al., 2014). Bredart (2014) believes that the proper structure of corporate governance reduces the likelihood of facing Bankruptcy. The results of the research by Alifiah (2014) and Li et al. (2015) indicate the impact of corporate governance variables on bankruptcy. Researchers such as Setayesh & Mansouri (2014), Osmani et al. (2011), Xie et al. (2011) and Li et al. (2015) investigated the relationship between corporate governance criteria and bankruptcy. In this study, we used these criteria to investigate bankruptcy.

2.3. Research Background

Fakhrehosseini, Seyed Fakhreddin, Omid Aghaei Meybodi (2019), forecast and identify companies with a high probability of bankruptcy on the Tehran Stock Exchange (different models of analysis). (Sprint, Altman, Fulmer, Zmijewski and Mckee Genetic models). To achieve this goal, 75 companies that are selected not covered based on 141 of the Commercial law. Required data for the 10-year period (86-95) has been compiled. According to the results in each of the above models, a number of companies were identified as high-risk probability companies and then companies that were identified as most likely to be bankrupt in most of these models. The results also show that, with the exception of Mckee model, in four other models, three companies with high bankruptcy probability were included. Among these four models, Zmijewski model has a higher coefficient of determination, hence, we can say that other models have been more accurately predicted for bankruptcy and have a significant role in corporate bankruptcy among financial ratios, debt ratios, asset turnover, and asset returns.

Ahmadi Amin, E., & Tahriri, A. (2019) Study the impact of bankruptcy transfer in the industry on the information content of corporate profits. Statistical results demonstrate that a firm probability of bankruptcy is negatively associated with the informativeness of good news earnings, whereas it has no negative effect on informativeness of bad news earning. Furthermore, until one bankrupt company exists, earnings informativeness reduction of good news is more than when a number of bankrupt companies exist.

Ghazanfari et al. (2018) predicted corporate bankruptcy based on hybrid intelligent systems. The results show the performance superiority of the combination of backup machine with algorithms to optimize the search for harmony and colonial competition to predict corporates financial loss.

Khajavi and Ghadirian (2018) examined the ability of financial performance managers and the risk of bankruptcy. The results showed that managerial ability has a negative impact on bankruptcy risk and financial performance mediates the effect of managerial ability on bankruptcy risk. In other words, managerial ability reduces the bankruptcy risk through improving financial performance. Therefore, it could be concluded that managerial ability is an important factor in the success of the companies in TSE.

Ahmadi (2016) conducted an analysis on the relationship between the features of the governance system and the systematic risk with the bankruptcy of companies during the years 2010-2014. According to the results of the research, the concentration of ownership and the proportion of institutional shareholders have a reverse and significant effect on bankruptcy. Also, the mixed variable of ownership concentration and systematic risk has a direct and significant effect on bankruptcy.

Asgarnejad Nouri & Soltani (2015) presented their model for predicting bankruptcy in their article entitled, "designing the bankruptcy prediction model, using accounting, market, and macroeconomic variables". The mentioned model was estimated using a complete set of accounting, market, and macroeconomic variables, using logistic regression method. Results show that the accuracy of bankruptcy models based on accounting and market variables is 91.2% and 82.1%, respectively. On the other hand, there was no significant relationship between macroeconomic variables and the probability of bankruptcy of companies.

Shen and Christopher (2017) investigated the relationship between bankruptcy and stock returns. Their research results showed that there is a strong negative relationship between stock returns and bankruptcy of the companies involved in their study.

Fontaine et al. (2017) conducted an analysis of bankruptcy in public joint stock corporates using macroeconomic variables. The projected model was between 2001 and 2014, using the logistic regression technique with panel data, it includes a sample of Brazil's famous business firms. The proposed model, in addition to financial variables, includes macroeconomic variables. Finally, it was said that the relationship between the phenomenon of financial dissatisfaction and distress is better explained by using macroeconomic and financial explanatory variables.

Kingsley et al. (2016) examined the impact of board quality and corporate bankruptcy. They predicted that board quality would reduce the negative effects of bankruptcy. The results indicate that the companies reduce the bankruptcy effects in terms of the quality of the board of directors and the ability of external managers to oversee the CEO on behalf of the shareholders and also they provide advice.

2.4. Research hypotheses

According to the mentioned theoretical foundations and the purpose of the research, the following hypotheses are developed.

1- The artificial intelligence algorithm of the relevance vector machine in nonlinear mode has a higher ability to predict bankruptcy in the current year than in linear mode.

2. Artificial intelligence algorithm of relevance vector machine in nonlinear mode has a higher ability to predict future bankruptcy than in linear mode.

3. Artificial intelligence algorithm of relevance vector machine in nonlinear mode has a higher ability to predict bankruptcy over the next two years than in linear mode.

3. Research design

3.1. Population and statistical sample of research

This research is applied in terms of the purpose and also the present research is fiend-library in terms of the type of study and it is post-event (i.e. the use of past information), using historical information. The statistical population of this research includes all companies listed on Tehran Stock Exchange with the following qualifications:

1. During the period of the study, they should have no change in their financial period.

2. They must not be affiliated with investment companies, financial intermediaries, banks, insurance, and leasing.

3. Their considered data should be available.

Finally, due to the limitations mentioned above, 1488 years - companies have been selected during the years 2011 to 2016 as the statistical population. Due to the availability of information, all companies have been considered as a statistical sample.

3.2. Research Variables

Table 1 shows the initial independent variables of the research.

Table 1: Research variables

initial independent variables		
company value	**Quick ratio**	**audit quality**
Return on sales	Working capital ratio	False self-confidence of manager
Systematic risk	Financial leverage (financial risk)	margin of gross profit
Stock returns	Debt to equity ratio	Margin of operating profit
Unconditional conservatism	Inventories turnover ratio	Asset return
Conditional conservatism	Asset turnover ratio	Return on equity
Cash to asset ratio	Fixed asset turnover ratio	Earnings per share
Operating cash flow to assets ratio	Turnover of receivable accounts	Current ratio
Free cash flow	Cost of capital	Ratio of current assets to assets
Kaplan financial constraint model	Information asymmetry	Ratio of Current Debt to Equity
Dividend profits ratio	Number of board members	Long-term debt ratio to total assets
Intellectual Capital	Gender of board members	Market value added
The proportion of institutional owners	Non-executive managers ratio	Size of company
Concentration of ownership	Dual role of CEO	Stock price
Costs stickiness	Ownership type	Rate of stock liquidity
Earning quality	Lifetime of company	Earning management
Disclosure Quality	Profit of paid stock	Economic value added
Price to income ratio	Liquidity conversion cycle	Tax avoidance

Dependent variable	
In this study, according to Table 2 and No. 3, a new approach was used to identify the financial distress situation with the use of universal repetitive models and Article 141 of the Trade Law in Iran.	Bankruptcy in three levels
Artificial intelligence algorithm of relevance vector machine	Research method
2011-2016 (6 years)	Research period
Two-stage approach, 1- Variable selection test using Relief-F multi-class algorithm. 2- Financial distress prediction with relevance vector machine algorithm for current, future and next two years periods.	Research design
1. Use of all accounting variables including financial ratios, corporate governance criteria as primary input variable (54 variables) 2. Use of artificial intelligence methods to select the best combination of independent input variables among 54 initial independent variables (Multi-class Relief-F algorithm). 3. Localization of Bankruptcy recognition at three levels, using the universal repetitive models and Article 141 of the trade law 4. Prediction of Bankruptcy using linear and nonlinear methods of relevance vector machine algorithm	The aspect of research innovation

Bankruptcy

1. In this study, according to Table 2 and 3, a new approach has been used to identify financially distressed companies. First, according to Table 1, each of global and Iranian bankruptcy methods which has been used in Iran, including (Altman (1968), Springate (1978), Ohlsan (1980), Fulmer (1984), Zemigewski (1984)), modified Altman (2000), Zavgren (1985), Shirata (1995) methods and Article 141 of the trade law were considered as a criterion for bankruptcy.

Table 2: Methods used in 100 recent researches about Bankruptcy

Distress model	number
Altman	22
Altman's modified model	4
Ohlsan	7
Article 141	49
Fulmer	2
Zemigewski	4
Springate	4
Shirata	2
Zavgren	3
Other models	3

According to Table 3 and its analysis, companies are categorized at three levels of financially healthy (companies that are known to be healthy in all models), financially distressed (companies that are not approved to be financially healthy in 1 to 4 models), and financially bankrupt (companies that have not been approved to be financially healthy in 5 to 9 models).

Table 3: The number of companies and the number of models indicating Bankruptcy

Type of company in this study	Number of year-company	Number of models indicating Bankruptcy
Financially healthy	134	0
Financially distressed	408	1
Financially distressed	403	2
Financially distressed	254	3
Financially distressed	150	4
Financially bankrupt	69	5
Financially bankrupt	44	6
Financially bankrupt	26	7
Financially bankrupt	2	8
Financially bankrupt	0	9

4. Empirical results

4.1. Variable Selection of the Relief Algorithm

The Relief method uses a statistical solution to select the feature. This method is an algorithm based on weighting the independent variables, the idea of which is inspired by sample-based algorithms. This algorithm selects a subset of the sample from D educational sample. The user specifies the number of companies (NoSample) in this subset as a predefined value. The algorithm randomly selects a company-year from this subset as a sample, then based on the characteristics (independent variables) of this sample, finds the nearest hit and the nearest miss based on the Euclidean distance estimation function. The nearest hit is the company-year sample which has Euclidean's least distance among other samples with the same level companies as the selected samples. By the same level, we mean if the selected sample is financially distressed, based on the Euclidean distance, it looks for a company-year that has the same characteristics, and that its independent variables are close to the selected company-year in terms of Euclidean distance. The nearest miss is also a company-year that has the least Euclidean's distance among samples that are not at the same level as the selected sample. The main idea behind this algorithm is that as the difference between the size of a feature in the selected company-year and the nearest hit is less, this feature is better, plus a good feature is that the difference between the size of that feature and the nearest miss be greater.

Relief(D, S, *NoSample*, *Threshold*)
(1) $T = \phi$
(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 $j = 1$ to N
    $W_j = W_j - diff(x_j, nearHit_j)^2 + diff(x_j, nearMiss_j)^2$
(4) For $j = 1$ to N
    If $W_j \geqslant Threshold$
      Append feature f_j to T
(5) Return T

Figure 1: Relief algorithm

In this algorithm, each of the independent variables has, at first, a weight of W, which at the start of the algorithm is equal to zero. The algorithm updates the weight of the features after determining the nearest hit the nearest miss. This update is such that the square of the difference between the value of the desired feature in the selected sample and the sample of the nearest hit will be reduced from the weight of the feature and the square of the difference between the value of the feature in the selected sample and the nearest miss will be added to the weight of the feature. The larger the weight, the better the feature can distinguish the companies belonging to a class from other companies. After determining the distance for all companies-years in the sample set, the algorithm eliminates the feature (independent variables) (f) the weight of which is less than or equal to a threshold and negative and restores others as the subset of the response characteristic (T). The threshold level is determined by the user, although it may be determined automatically by a function of the total number of features or determined by trial and error. Relief works well for noisy and correlated features and its time complexity is a linear function of the number of features given and NoSample. This algorithm works well for samples with continuous and nominal features. One of the main constraints of this algorithm is that it does not find features that have redundancy and therefore finds non-optimal sets that have redundancy. This problem can be solved with a subsequent exhaustive search for the subsets selected by the algorithm. In addition, another problem with the algorithm is that it works well with two-class issues. This constraint has also been resolved with the Relief-F algorithm, with the new algorithm, the incomplete data problem (incomplete training samples) has also been solved. There is also another version of this algorithm called RRelief-F for regression issues (Robnik-Šikonja, and Kononenko, 1997). The results of the selection of the independent variables of Relief-F multi-class algorithm are shown in Table 4.

Table 4: Selected independent variables along with weight (significance) to predict Bankruptcy

Independent variables selected by algorithm Relief-F
Financial leverage (financial risk)
Working capital ratio
Long-term debt ratio to total assets
Fixed asset turnover ratio
Economic value added
Non-executive managers ratio
Ratio of Current Debt to Equity
Debt to equity ratio
size of the company
Earnings management

4.2. Relevance Vector Machine (RVM)

Relevance vector machine is a Bayesian Sparse kernel technique for classification and regression issues. In addition, this method leads to the production of a sparse linear model (or a nonlinear model with the help of the kernel trick) and has a high convergence rate. The linear model of RVM for regression is defined as:

$$y(x) = \sum_{i=1}^{M} w_i \phi_i(x) = w^T \phi(x)$$

Where M is the number of independent variables (features) plus one, which actually an intercept term is added to it and $\phi_i(x)$ is in the linear mode of the independent variable ith of the company x. In non-linear mode$\phi_i(x)$is a fixed nonlinear base function mapping the input data from the nonlinear input space of the problem to the linear space of the feature, and in general, the dimension of the space of the feature may be less than or equal to infinity. For example, suppose companies are divided into two categories of bankrupt and healthy, which are shown in Fig. 2 with the signs of the circle and triangle. Also, assume that the coordinate axes in the left-hand figure 2 represent two independent variables of the problem. This space is called the entrance space. It can be seen that it is impossible to separate circles and triangles with a line, that is to say, there is no linear model that can separate healthy and bankrupt firms 100%. For this reason, a non-linear function is used as the mapping function and maps the input data to a new space that has higher dimensions than the input space; this space is called feature space. In this space, the dimensions of space, or, in other words, new independent variables come to existence that no longer define the concept of independent variables in the input space. As you can see in the middle 2nd figure, you can now divide these companies with a single page. The result of this separation is shown in Figure 3-right on the input space, which is no longer a line. For we do not get involved with the computing of the feature space, with obtaining a proper mapping function φ and eventually with losing the concept of independent variables in the new space, we use the kernel trick to bypass this space. Using the kernel trick k (x, y) = <φ (x), φ (y)> k (x, y) = ⟨φ (x), φ (y)⟩, we no longer need to transfer data to the space of the feature. Instead, we use the internal multiplication image of two mapping functions in the input space.

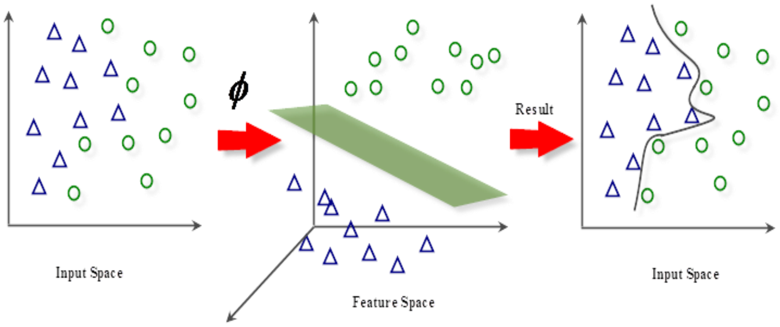

Figure 2. Concept of input space, mapping function and feature space

Suppose that there is a set of N observations with the input vector (independent variables) x, so that all of these observations are represented by an X matrix, so that the nth row is represented by $x_n^T$and represents the independent variables of the nth company, and n = 1,2, ..., N. The corresponding target values (dependent variable) are represented by the vector $t = (t_1, t_2, \dots, t_N)^T$. So, consider the likelihood function as follows.

$$p(t|X, w, \beta) = \prod_{n=1}^{N} p(t_n|x_n, w, \beta^{-1})$$

Where $p(t_n|x_n, w, \beta^{-1})$indicates that the nth company has the dependent variable $t_n$provided that the independent variables x_n, the parameter w and β are known. In this algorithm, it is assumed that the data are independent of each other, therefore, the multiplication of probability is used in equation (2). Now, the prior distribution function is defined as a Gaussian function with mean zero as the relation (3). This relationship shows that the function of the weight distribution has a mean of 0 and the parameter of the variance α is for controlling the sparseness of w. Therefore, one of the goals of RVM is to obtain a weight vector, the values of which are close to zero as much as possible.

$$p(w|\alpha) = \prod_{i=1}^{M} \mathcal{N}(w_i|0, \alpha_i^{-1}) = \mathcal{N}(w|0, A)$$

When $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_M)^T$and $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_M)^T$, in RVM, instead of computing w, we obtain α values that we will see that when the amount of αi goes to infinity, corresponding wi will be zero. In fact, α is the sparseness control parameter of w. The basic functions associated with these parameters do not play any role in the predictions made by the model, therefore, their effect on making the model sparse. The following Gaussian function is used for the posterior distribution with the mean m and Σ of the following covariance.

$$p(w|t, X, \alpha, \beta) = \mathcal{N}(w|m, \Sigma)$$

$$m = \beta \Sigma \Phi^T t$$

$$\Sigma = (A + \beta \Phi^T \Phi)^{-1}$$

Where Φ is the matrix N × M with elements $A = diag(\alpha_i), \Phi_{ni} = \phi_i(x_n)$. The values of α and β are obtained using the type-2 maximum likelihood method known as evidence approximation. For this purpose, the marginal likelihood function (7) is maximized on the w parameter.

$$p(t|X, \alpha, \beta) = \int p(t|X, w, \beta) p(w|\alpha) dw$$

Since the equation (7) shows two Gaussian functions, it can be written in the form of the following log marginal likelihood after doing some simplification steps:

$$ln\, p(t|X,\alpha,\beta) = ln\, \mathcal{N}(t|0,C) = -\frac{1}{2}\{N\, ln(2\pi) + ln|C| + t^T C^{-1} t\}$$

When $t = (t_1, t_2, \ldots, t_N)^T$ and the matrix $C_{N\times N}$ are defined as the relation (9).

$$C = \beta^{-1} I + \Phi A^{-1} \Phi^T$$

At this time, the aim is to maximize the equation (8) toward the parameters α and β. By deriving from the marginal likelihood and putting it equal to zero, and by simplifying, the values of α and β are repeatedly calculated as follows:

$$\alpha_i^{new} = \frac{\gamma_i}{m_i^2}$$

$$(\beta^{new})^{-1} = \frac{\|t - \Phi m\|}{N - \sum_i \gamma_i}$$

Where m_i, is the i the last mean of the equation (5). The quantity γi is defined as follows:

$$\gamma_i = 1 - \alpha_i \Sigma_{ii}$$

Where Σ_{ii} ،is the i th last diagonal component of the covariance of the relation (6).

Solving the RVM optimization problem is solving the problem by making the weight vector w sparse with the aid of the super-parameter α. Therefore, by taking into account the Gaussian distribution function with mean zero for the weight vector w, according to equation (3), we can obtain the value of the weight vector w by obtaining the value of the super-parameter α through the equation (10). For this purpose, in the first step, equation (5) and (6) are firstly calculated by initializing to α and β, and then using the equation (10) and (11), the new α and β values are calculated and the equation (5) and (6) are recalculated. This will continue until the convergence of the algorithm. The convergence condition can be a certain number of repetitions. After obtaining α and β, by the equation (2), the probability of the dependent variable based on the inputs (observations) X is obtained. In the nonlinear model (13), unlike the linear model (1), this nonlinearization is performed with the help of the kernel. The x_n s which are remained in corresponding with nonzero weights, is called relevance vectors (Tipping, 2001).

$$y(x) = \sum_{n=1}^{N} w_n k(x, x_n) + b$$

For the evaluation of linear and nonlinear classification models, the detection rate has been used. The detection rate is the percentage of the correct classification of the dependent variable to the total of the company-years in each dataset and the detection rate is calculated from the following equation.

Detection Rate = the number of correct predictions/ the number of companies-years

4.3. Results of RVM Forecasting for the Bankruptcy of Companies for the Years t, t + 1 and t + 2

Educational and evaluation data were categorized by the 10-Fold Cross-Validation method to RVM classification into linear and nonlinear models. In order to study the degree of learning of linear and nonlinear model, the learning data from the 10-Fold cross-validation method was given to the models, and the mean of learning rate for the learning data is shown in Table 5. It is observed that the linear model is less accurate than the nonlinear model in approximating the bankruptcy of companies, and the linear model for almost every three years offers approximately the same rate. In contrast, the nonlinear model has relatively higher accuracy, compared to the linear model and its prediction power decreases as the years increase.

Table 5: The mean of RVM Linear and Non-Linear Model Detection Rate for predicting the Bankruptcy of companies for training data.

The next-next year		The next year		The current year		
Non-linear model	Linear model	Non-linear model	Linear model	Non-linear model	Linear model	RVM
86.54	81.90	93.19	82.74	95.84	81.68	Detection rate

Now, to test the problem of over-fitting and generality of the model, the test data of the method 10-Fold Cross-Validation is given to the learned models and the detection rates for these models are shown in Table 6. Given the low discrepancy in the detection rate identification of Table 5 and Table 6, it can be said that the models are not over-fitted and also, according to the detection rates obtained in Table 6, it can be claimed that the models are generic.

Table 6: The mean of RVM Linear and Nonlinear Model Detection Rate for Predicting the Bankruptcy of Companies for Test Data

The next-next year		The next year		The current year		
Non-linear model	Linear model	Non-linear model	Linear model	Non-linear model	Linear model	RVM
86.22	81.89	93.16	82.93	95.83	81.46	Detection rate

5. Discussion and conclusion

One of the issues that can help investors make decisions is the availability of appropriate tools and models for assessing financial conditions and corporate status. One of the instruments used to decide about investing is prediction models of bankruptcy. Speeding up activities and economic events have had many positive and negative consequences. One of the most important negative consequences of these changes is the increased competition for obtaining financial resources and limiting access to profits by business units and economic entities. According to Yang and Jiang (2008), one of the ways through which we can make optimal use of investment opportunities and prevent loss of resources is to predict bankruptcy or corporate bankruptcy. The results of the research show that the variables of financial risk, working capital ratio, long-term debt ratio, fixed asset turnover ratio, economic value added, non-executive managers ratio, current debt ratio to equity, debt to equity ratio, the size of the company, earning management have been selected as important variables in the prediction of the threefold bankruptcy condition respectively, and also in the nonlinear mode of relevance vector machine, the company's bankruptcy level can be accurately predicted at more than 93% for the next year and more than 86% for the next two years. These results are consistent with the results of Setayesh and Mansouri (2014); Chang (2009); Meshkai Miyaveghi and Hashemi Sa'adat, (2015); Fich and Slezak (2008); Mollaei and Khazdouzi, (2015); Ozhan A. and Oazkan N. (2004); Kim et al. (1998); Rahimian and Tavakolnia, (2013); Ahmadi, (2016).

According to the primary results of the research, the variables of financial leverage (financial risk), the ratio of non-executive managers, earning management, economic value added, and debt ratios are of the highest importance in predicting the bankruptcy situation. Therefore, managers of active enterprises in the capital market are recommended to consider these variables in order to decide on the bankruptcy of companies.

According to the secondary results of the research, which shows that the non-linear relevance vector machine algorithm has a higher ability to predict the bankruptcy of companies than the linear mode, it is recommended that the owners of capital and company decision-making use the predictive power of artificial intelligence algorithms, especially the nonlinear relevance algorithm in their decisions on investing in the capital market.

Further, the results of this research can draw the attention of managers of the capital market of Iran, practically, so that by predicting the state of bankruptcy in companies and by working on the contributing factors, we would be able to manage the attraction of shareholders' capital, reduce the risk of financial crises, and help investors avoid great losses in the stock market.

References

Ahmadi,S. Sh. (2016). Investigating the Relationship between Corporate Governance and Systematic Risk with Financial Distress in Companies Accepted in Tehran Stock Exchange, Master's Thesis.

Ahmadi Amin, E., and Tahriri, A. (2019). The Effect of Bankruptcy Contagion on Earnings Informativeness. *Journal of Accounting and Auditing Review*, 26 (1), pp. 1-18. (In Persian)

Altman,E.I., (1968). Financial ratios, discriminant analysis and prediction of corporate bankruptcy. *Journal of Finance*, 23, pp. 589-609.

Asgarnezhad Nouri, B., and Soltani, M. (2015). Designing a BankruptcyPrediction Model Based on Account, Market and Macroeconomic Variables (Case Study: Cyprus Stock Exchange), *Iranian journal of management*, 9 (1), winter 2016, pp. 125-147

Mullahi E., and Khidozzi, B. (2015). The Effect of Liquidity Level on the Risk of Financial Distress in Companies Accepted in Tehran Stock Exchange, Industrial Management (Azad Sanandaj), No. 34

Bredart, X. (2014). Financial Distress and Corporate Governance around Lehman Brothers Bankruptcy. International Business Research, 7 (5), pp. 1-8

Brewer, B. E., Wilson, C.A., Featherstone, A. M., Harris, J.M., Erickson, K. and Hallahan, (2012). Measuring the Financial Health of U.S. Production Agriculture.

Chang, C. (2009). The Corporate Governance Characteristics of Financially Distressed Firms: Evidence from Taiwan. *The Journal of American Academy of Business*, 15 (1), pp. 125-132.

Chen, M.Y. (2011). Predicting corporate financial distress based on integrationof decision tree classification and logistic regression. Expert Systems with Applications, 38 (9), pp. 11261-11272.

Eliezer F., Steve S. (2008). Can corporate governance save distressed firms from bankruptcy? An empirical analysis. *Review of Quantitative Finance and Accounting*, 30 (2), pp. 225-251.

Esmaeelzadeh Moghri, A., and Shakeri, H. (2015). Predicting financial distress of companies listed in Tehran Stock Exchange using simple Bayesian network and comparing it with data envelopment analysis. *Journal of Financial Engineering and Portfolio Management*, spring 2015, No. 22, pp. 1-28

Felipe Fontaine Rezende ,. Roberto Marcos da Silva Montezano, Fernando Nascimento de Oliveira, Valdir de Jesus Lameira. (2017). Predicting financial distress in publicly-traded companies, Revista Contabilidade & Finanças, vol.28 no.75 São Paulo Sep./Dec. 2017 Epub July 20, 2017.

Fakhrehosseini, F.,and Aghaei Meybodi, O. (2019). Prediction and Identification of Companies with High Bankruptcy probability in Tehran Stock Exchange (Different analysis of models), Journal of Decisions and operations research, 4 (1), summer 2019, Page 3-15(IN PERSIAN).

Fulmer H Score. (1984) is a bankruptcy classification model, based on the 1984 paper "A Bankruptcy Classification Model for Small Firms.

Ghasemi, M., and Ramezanpour, s. (2018). Prediction of Bankruptcy of Companies Listed on the Securities and Exchange Organization Using Artificial Neural Network, Journal of Investment Knowledge, 7 (27), pp. 227-296(IN PERSION).

Hu, H., and Sathye, M. (2015). Predicting Financial Distress in the Hong Kong Growth Enterprises Market from the Perspective of Financial Sustainability. Sustainability, 7, pp. 1186-1200.

James A. Ohlson (1980). Financial Ratios and the Probabilistic Prediction of BankruptcyJournal of Accounting ResearchVol. 18, No. 1 (spring, 1980), pp. 109-131.

Kingsley Opoku, A., Amon Chizema. (2016). The impact of board quality and nomination committee on corporate bankruptcy, Advances in Accounting, incorporating Advances in International Accounting, journal homepage: www.elsevier.com/locate/adiacpp1-7

Kim, C.-S., Mauer, D.C., and Sherman, A.E. (1998). The determinants of corporateliquidity: theory and evidence. Journal of Financial and Quantitative Analysis, 33, 305–334.

Khajavi, SH., and Ghadirian Arani, M. H. (2018). Journal of accounting knowledge, Article 2, 9 (1), pp. 35-61 (IN PERSAN).

Kritsonis, Alicia. (2005). Assessing a firms future financial health. *International Journal of Scholarly Academic Intellectual Diversity*, California State University vo.8, No.1.

Li, Z., Crook, J., and Andreeva, G. (2015). Corporate Governance and Financial Distress: a Discrete Time Hazard Prediction Model, Retrieved fromhttp://ssrn.com/abstract=2635763.

Maheswara Reddy, D., and Reddy, C. R. (2011). Application of Z score Analysis in EvaluatingThe Financial Health of Pharmaceutical Companies- Acase Study.

Mashayekhi, B., and Ganji, H. (2014). Examining the impact of earnings quality on bankruptcy prediction using artificial neural network. *Financial Accounting and Auditing Research*, Article 6, 6 (22), pp. 147-173.

Mohd Norfian, A., Norhana, S., and Ismail, A. (2013). Prediction of financial distress companies in the consumer products sector in Malaysia. *Journal Sains Humanika*, 64 (14).

Marko Robnik-ŠikonjaIgor, K. (1997). Theoretical and Empirical Analysis of ReliefF and RReliefF Machine Learning Volume 53 (1-2), pp 23–69.

Mohd Norfian, A. (2014). Prediction of financial distress companies in the trading and services sector in Malaysia using macroeconomic variables, Journal.

Meshki Miavaghi, M., and Hashemi, M. (2015). Investigating the Relationship between Corporate Governance with Bankruptcy Probability in Companies Listed in Tehran Stock Exchange,*The Journal of Research Accounting*. 2 (17), pp. 37-58.

Michael E. Tipping (2001) Sparse Bayesian Learning and the Relevance Vector Machine, Journal of Machine Learning Research 1, pp. 211-244.

Outecheva, N. (2007). Corporate Financial Distress: An Empirical Analysis of Distress Risk. University of St. Gallen, Switzerland

Osmani Qasim, M., Javid, D., and Rahimi, Saeed. (2011). Investigating the deterrence impact of corporate governance mechanisms on financial distress in companies listed in Tehran Stock Exchange. *Accounting and Auditing Research*, No. 12, pp. 1-19.

Owens, E. (2017). Asymmetric Effects of Default Probability on Earnings

Informativeness. The University of North Carolina at Chapel Hill. https://ssrn.com/abstract=1869108.

Ozkan, A., and Ozkan, N. (2004). Corporate cash holdings: an empiricalinvestigation of UK companies. *Journal of Banking and Finance*, 28, pp. 2103–2134.

Pindado, J., Rodrigues, L., and de la Torre, C. (2008). Estimating financial distresslikelihood. *Journal of Business Research*, 61, pp. 995-1003.

Rahimian, N., Tavakolnia, E. (2013). Financial leverage and its relationship with financial distress and growth opportunities in companies listed in Tehran Stock Exchange (linear and curvature relationships); *Quarterly of Financial Accounting*, 5 (20), pp. 108-129.

Returns Pengjie Gao Christopher A. Parsons Jianfeng Shen (2017). Global Relation between Financial Distress and Equity. *The Review of Financial Studies*, hhx060, DOI: 10.1093/rfs/hhx060 Published: 14 June 2017.

Sadeghi, H., Rahimi, P., and Salmani, Y. (2013),The Effect of Macroeconomic and Governance Factors on Financial Distress i Manufacture Firms Listed in Tehran Stock Exchange, *Financial monetary economics*, 21 (8), pp. 107-127.

Saeedi, Ali. Aghaei, Arezoo. (2009). A Review of Bankruptcy Prediction Methods and Models, Accounting Knowledge and Research, No 16

Setayesh ,M. H., and Mansouri., Sh. (2014). The comparative investigation of corporate governance mechanisms in financial distressed and nonfinancial distressed listed companies of Tehran Stock Exchange. *Journal oh financial Research*, Volume 16, Issue 1 (Spring and Summer), Page 99-112

Shirata, C.Y. (1995). Read the Sign of Business Failure. Journal of Risk and Management, Vol. 23, pp.117-138.

Springate, G. L. V. (1978). Predicting the possibility of failure in a Canadian firm (Unpublished master’s thesis). Simon Fraser University, Canada

Sun, J., Jia, M-Y., and Li, H. (2011). AdaBoost ensemble for financial distress prediction: An empirical comparison with data from Chinese listed companies.Expert Systems with Applications, 38 (8), pp. 9305-9312.

Thomas J.George,Chuan-Yang Hwang (2010). A resolution of the distress risk and leverage puzzles in the cross section of stock returns" .Journal of Financial Economics (96), pp. 56-79.

Wing Yu, i., Ho, F., Law, E. and Fung, L. (2003). An Analysis of the Financial Health of Hong Kong Corporations.

Xie, C. and C. Luo, and X. Yu. (2011). Financial distress prediction based on SVM and MDA methodes: the case of Chinese listed companies. *Quality and quantity*, 45 (3), pp. 671-686.

Yang, J.; and Y. Jiang. (2008). Accounting information quality, free cash flow and Overinvestment: A Chinese study. *The Business Review*, Vol. 11, pp. 159-166

Zmijewski M.E. (1984). Method logical Issues Relate to the Estimation of Financial Distress prediction Models. *Journal of Accounting Research,* Vol 22 supplement.

Zavgren, C. (1985). Assessing the vulnerability to failure of American industrial firms: A logistic analysis. Journal of Business Finance & Accounting, 12 (1), pp. 19–45.

Zohra, K. F., Mohamed, B., Elhamoud, T., Garaibeh, M., Ilhem, A., and Naimi,H. (2015). Using Financial Ratios to Predict Financial Distress of Jordanian Industrial Firms, "Empirical Study Using Logistic Regression". *Academic Journal of Interdisciplinary Studies*, 4 (2), pp. 137-142.