Predicting Bankruptcy of Companies using Data Mining Models and Comparing the Results with Z Altman Model

Somaye Fathi
Department of Accounting, Boroujerd Girls' Technical University, Lorestan, Iran
(Corresponding author)
somayeh.fathi@yahoo.com

Samira Saif
Department of Accounting, Payame Noor University, Nahavand, Hamadan, Iran

Zohre Heydari
Department of Accounting, Kosar University of Bojnord, Bojnord, Iran

ABSTRACT

One of the issues helping make investment decisions is appropriate tools and models to evaluate financial situation 0f the organization. By means of these tools, investors can analyze financial situation of the organization and identify financial distress or an ideal condition, they become aware of making decisions to invest in appropriate conditions. The main objective of this study is to evaluate the power of using data mining models which are among new tools of prediction. This tool was used to predict the bankruptcy of companies listed in Tehran stock exchange and comparison the results with the Altman model as one of the prevalent methods of prediction the bankruptcy of a company. The research data includes information of all companies listed in Tehran stock exchange during the years 2013 to 2018 subjected to Title 141 of the law of trade and were bankrupt. Variables used in both models were five financial ratios. The data mining models on the average in the base year had a predictive ability of 92.4 percent and the Altman model had a predictive ability of 82.41 percent. Considering the results, it was shown that the data mining model has more power to predict bankruptcy.

Keywords:

Altman Model, Bankruptcy, Data Mining Models.

1. Introduction

One of the most important threats to the national economy is its companies' bankruptcy. Evaluating bankruptcy provide valuable information that governments, investors, shareholders can use them to make their financial and investment decision in order to prevent possible loss. In one of the first academic studies on the theory of bankruptcy, company's inability to profitability has been defined as it increases the probability of a lack of ability to repay interest and debt. In other definition, consider the situation that the cash flows of the company are less than the cost of the interest related to long-term debt (Witker, 1999). A company's financial position and its stability against bankruptcy has effect on stakeholders, i.e. shareholders, creditors, employees. Therefore, bankruptcy is one of the important and challengeable subject in scientific studies. The most important topics in financial management are investment and investment confidence. One of the things helping make the investment decisions is appropriate tools and models to assess the financial and organizational situations. For investors and many private firms affected by bankruptcy, quickly and easily using tool is important scientifically because they often have to make decision about their capital faster and may not have analytical understanding of firms performance (Permacahndra, 2009) Since the capital market in Iran is young and is not regarded as an efficient market, scientific study can be theoretical expression to establish the information published by Tehran Stock Exchange and it is useful in making decision by capital owners and potential investors. It is tried to provide useful tools to make decision about information for users of the capital markets. One of these tools is Bankruptcy prediction models that may be helpful.

According to inverse effects of bankruptcy on capital markets and economy, researchers tried to develop predicting models by using various procedures so that they can reduce losses and its effects. Firm's bankruptcy is usually determined by various related factors, so determining the exact cause or causes of bankruptcy and financial problems in any particular case is not easy. Generally, the bankruptcy factors include external factors (external) and internal factors (internal) .The external factors which are not controllable by the company and cause financial problems are including economical system properties, economical structures changes, trade and improvements changes, public demand, trade fluctuations(incapability between production and consumption, lack of employment, reduction in sales, inflation, prices falling and higher interest rates, etc.), problems associated with the financing, natural events and disasters, and the intensity of competition in the market. On the other hand, internal factors include things in which managers make mistake or they were failed to do what is necessary in management decisions of the past such as development of excessive credit for customers, sales credit), inefficient management (lack of training, experience, ability and initiative in the field of competition and technology management and resource, management errors), inability to manage capital effectively referred treason and fraud.

The issue of bankruptcy can be studied by observing direct effects of internal and external factors on bankrupt companies or by critically examining the financial ratio of bankrupt firms. The financial ratios reflect interactive effects of external factors and analysis of traditional risk on the company's poor financial situation. Using the financial ratios give signals about company for financial distress or bankruptcy before the company declare bankruptcy specialties.

Bankruptcy prediction models are tools to predict future situation of a company that estimate firm bankruptcy by combination a group of financial ratios, they can be classified to three statistical models and theoretical and artificial intelligence. The statistical models are divided into two groups: single, multivariate. Multivariate statistical models are made of multiple discriminant analysis models, linear probability, logit , probit , cumulative total and partial adjustment process artificial neural networks, genetic algorithms, software hard set, support machine vector, the argument based on the findings and fuzzy logic forming techniques of artificial intelligence and theoretical models including criteria analysis of balance sheet / theory of entropy theory, bankruptcy speculator theory, cash management and credit risk theory. In addition, the accuracy of the model provided in Iran environmental conditions according to the study with an overall accuracy 98.3% in comparison to Wallace research with an overall accuracy of 94% predictive variables in both studies is higher. According to the research conducted in this area until now, research on data mining model has not carried out according to the power of data mining and artificial intelligence in processing large databases and finding complex patterns and non-linear in that, a lot of research by using data mining has been done in various fields. Meanwhile, the financial decisions and vocation judgment in accounting and auditing because of the nature of the turbulent influencing variables have been suitable for using tools for discovery and data mining , so that this modeling by data mining technique in the world is as one of the most versatile items in the financial sphere . Therefore, this study aimed to determine the effectiveness of the Altman model data mining to predict the bankruptcy companies listed on the Stock Exchange of Tehran in order to prevent capital losses through with timely alerts to investors, shareholders and companies. The research question in this case, "Which one of two models of data mining models to prediction bankruptcy Altman is more credible.

2. Literature Review

2.1. Background Research

Mehrani et al. (2005) also sought to design new models on the basis of Zimsky's and Shirata's bankruptcy prediction models in accordance with Iran's environmental conditions. Their research results showed that both models have the ability to identify companies in two groups bankrupt and non-bankrupt, and independent variables don't have the same effect on the prediction of corporate bankruptcy. In this study, the independent variables of Shirana and Rimsky's model include the accumulated profits to total assets, net profit to total assets, total debts to total assets, liabilities and equity of previous years to debts and salaries stockholders in the current year.

GHadirimgadam et al (2009) also compared two predictive models of bankruptcy by Altman and Olson in accordance with Iran's environmental conditions and presented a statistical model suitable for predicting the exact proportion of firms in one, two, and three years before the financial crisis, companies so that they can study the financial status of the companies and the issue of their continued activity using the model and promote qualitative decision making by stakeholders and stakeholders. The research results indicate that the model presented by Olson (1980), and the model extracted by the logistic regression method has a higher accuracy in predicting corporate bankruptcy. In this research, they used the variables of the total index, the ratio of working capital to total assets, the ratio of accumulated profits to total assets, the profit before interest and taxes on total assets, the ratio of the book value of the company's shares to the book value of the total debt, the profit before interest and the tax on total assets.

Godrati and Mogadam (2010)also in their research studied the accuracy of bankruptcy prediction models(Altman, Shirana, Olson Zemiski, Springeit, C.I.Score, Fulmer, Farajzadeh Genetic models and McKay genetic)in Tehran Stock Exchange. Their results showed that the predictive models of the financial crisis presented by Zemiski, Springeit, C.I.Score, Fulmer, Farajzadeh Genetic models and McKay genetic have ability to predict continuity of activities in the companies listed in the Tehran Stock Exchange. Also, the models had been modeled using artificial intelligence (genetic algorithm) compared to models that had been modeled using statistical techniques (Classic models) were more likely to be in bankruptcy ahead.

In this study, they used the variables of the ratio of working capital to total assets, the profit ratio accumulated to total assets, profit before interest and taxes on total assets, value ratio the equity market and the total book value of the debt.

Makian et al (2010) compared neural network modeling with two statistical methods of logistic regression and audit analysis to predict bankruptcy, and in addition to introducing neural network models designed a neural network model for prediction the bankruptcy of manufacturing companies. The information used is related to Kerman province between 1985-1997. In this study they used the ratios of current assets to current liabilities, profit before interest expense and total taxes assets, equity to debt, working capital, profits before interest expense and sales tax. Their research results showed that designed neural network model has higher predictability compared to two other statistical methods.

Firoozian et al (2011) examined the application of genetic algorithm in prediction bankruptcy and compared it with Z-Altman model in listed companies in Tehran Stock Exchange Their results showed that the general accuracy of the genetic algorithm model is more than the Z-Altman model,Hence, the genetic algorithm model is a more appropriate tool to predict corporate bankruptcy. In this research, they used current proportions, capital, profitability, and debt.

Karami and Seyyed Hosseini (2012) examined the usefulness of accounting information to market information in the prediction of bankruptcy. In this research, they used variables including profitability, liquidity, leverage and market. Their research results showed that accounting patterns have a precision of 91.1%, market patterns of 70.3%, and the combination patterns has a precision of 79/44% in prediction of bankruptcy.

Hung Wu et al. (2007) used Genetic Algorithm Model to optimize the support vector machine to predict bankruptcy. They are current ratio, immediate ratio, and gross profit ratio to sales, operating income to sale .Their results showed that using genetic algorithm improves the performance of the support vector machine and increases its predictive ability.

Boyakiaglo et al (2009) used the combination of neural network with support vector machine and statistical models to predict bankruptcy. In this research, they used equity to total assets, equity to total debt, total debt to total assets, net profit to average assets, current assets to total assets. The result of their research showed that training and setting valid data and achieve a unique technique for solving problems and anticipating bankruptcy by neural network are difficult alone. Therefore, it is proposed to combine it with other patterns.

Yoon and Yang (2010) also investigated the bankruptcy of small firms using sales credit card data. In this research, they developed a support vector machine and used the average sales in every 8 months, the average sales per month, and the average sales per 6 months, the lowest sales and the most sales. The evidence of this research showed a better performance of this model to the neural network, multivariate analysis and regression analysis.

Premachandra et al (2011) in a research with the development of data envelopment analysis model suggested an indicator for predicting bankruptcy or organizational success. Their example was a spectrum of different industries. In this research, they used current cash ratios to total assets, net profit to total assets, profit before interest and tax on total assets. The results of the study showed that data envelopment analysis model is not suitable for prediction alone and with DEA development, weaknesses can be improved.

Jardin and Séverin (2012) using the internal pattern of the organization predicted the bankruptcy of companies that provided financial reports each year. In this research, to extract the internal model of the organization, the variables including equity to total Assets, total debt equity ratio, cash to current debts, cash flows to total debts, and cash flows to total assets are used. The result of this study was that this model is appropriate to predict bankruptcy in Short-term periods (one year) before bankruptcy.

Jeong et al. (2012) in a research combined the artificial neural network with collective patterns and the genetic algorithm. They used the combined algorithm to predict the bankruptcy of company and collective models to determine the input variables. In this model they used the sales cost ratio to goods sold, current debt to total assets, interest expense to sales, current debt to total assets to predict the bankruptcy of companies. Comparing this algorithm with decision tree, collective models, multivariate analysis and various linear models, shows the superiority of the combined algorithm.

Kim and Kang (2012) in a research used a combination of genetic algorithms with optimization patterns to predict corporate bankruptcy. Variables used in this study include net profit to total assets, financial expenses to total assets, net profit to sales, current assets to total assets, current assets to 9current debts and ongoing debt to sales. The result of their research is expressive that combining the genetic algorithm with optimization patterns resulted in better prediction of bankruptcy.

Angga Pertapan et al. (2018) in his research titled Bankruptcy Prediction in PT Blue Bird, Tbk 2011-2016 Using Altman Z-Score, Springate, and Zmijewski Model .The results of the analysis using Springate model shows that companies in the category good condition. The analysis result using Zmijewski model also shows companies in the category good condition.

Błazej Prusak (2018) in his research titled Review of Research into Enterprise Bankruptcy Prediction in Selected Central and Eastern European Countries during the period Q4 2016–Q3 2017. Their results showed regarding the development of bankruptcy forecasts, it is also worth mentioning the differences that may occur when applying different accounting standards. Differing results are also obtained when financial statements are prepared in accordance with national or international accounting standards

Given the many studies that have been done so far, the importance of the bankruptcy companies can be more known, but each of these studies have considered one or several models of prediction models and this issue has been investigated:

Can it be identified in companies before the bankruptcy and financial crisis? Are these patterns able to predict corporate bankruptcy? In these studies, each one of the researchers has examined factors in bankruptcy prediction models according to the taste or limitations of the data, or taking into account previous research. If we can use them in prediction patterns considering the importance of factors affecting bankruptcy, certainly we will come up with better answers, which this study will identify and rank important factors in corporate bankruptcy.

Considering that so far any research has not used Data Mining Models to predict bankruptcy and comparison with Altman model. nowadays the modeling by means of the techniques given in the world is as one of the most widely items in financial state, so this research is to predict the bankruptcy using Data Mining Models and comparison it Altman's famous model.

2.2. Theoretical Framework

Data mining

Data mining is an inter field activity and is the product of encounter and synergism between various sciences such as statistics, machine learning, artificial intelligence, data base technology and visualizing (description and indexing) (Nakhaeizadeh, 1998). At the same time, this common method has been distinguished from other methods in the field. One of the applications of data mining in the context of bonds and helping investors is acquisition of best economic decisions.

Data mining is applied in many branches such as marketing, financial affairs, banking, production, medicine, management of customer relations, tracking, predicting damages, organizational teaching, etc. We can refer to the following cases:

1) Existence of large data base from which valuable trade information can be extracted from this data base.
2) It is not possible to use past traditional methods to support decisions and analyses.
3) Human analyses are influenced by data dimensions and volume.
4) Traditional statistical ranking methods are not able to prioritize and there is need for important and significant specialists and analysts.

Therefore, we can use various data mining methods to predict bankruptcy classified into two main groups as follows:

- First group, parametric models:
  1) Linear probability model
  2) Logit and probit model
  3) Differentiating analysis models

- Second group, nonparametric models:
  1) Linear programming
  2) Neural networks
  3) Decision trees
  4) Closest neighboring models
  5) Hierarchal analysis strategy
  6) Expert systems
  7) Genetic algorithm

Parametric method requires a mathematical function that dependent variable is estimated based on it by using independent variables. This method is used to estimate parameters of a function of the experimental observed data. In fact, in this method, first, a particular form is considered for the production function, and then the coefficients of the unknowns (parameters) are estimated by using one of the estimating methods of functions, which is common in the statistics and econometrics. Since in this methods, the parameters of the assumed function are estimated, these methods are called parametric methods. Non-parametric methods generally study the performance of a firm or making decision unit with the best performance of businesses within the industry. Non-parametric methods can be considered as the simplest methods of observation and the efficiency estimation, because in these methods the specified form is not considered for the production function and we use directly the observed data, and because this method is not statistical, so we cannot use statistical tests on it. Parametric methods with respect to the production function design, and use of more complicated mathematic formula are more difficult than Non-parametric methods. Moreover, in order to determine the efficiency of comparative evaluation, Non-parametric methods are more appropriate. At the level of Non-parametric patterns, there are different methods to view inefficiency:

1) Observations method
2) Step method
3) Average linear method
4) Plug –point's method
5) Data analysis method in the decision making trees are as one of the techniques of classification of methods of Non-parametric methods in validation. The trees can identify the characteristics of customers and separate them to the group of good and bad and validate them. These features describe the profile of bank customers and specify their class in the validation. From the other side, knowing the appropriate model in the classification of bank customers in the much volume of the data requires pre-processing of the data. There are a variety of methods for pre-processing the data that two cases of them are applied in this article including cluster and selecting features. With the pre-processing of the data, an appropriate pattern can be created to identify and validation of the banks customers.

Concepts of Data Mining

A- Stages (process) of data mining

Stages of data mining can be expressed in more detail as follow:

1) Identification of purpose: At this stage, what the user wants and what extent of information is to be acquired from the data base is identified.
2) Data Selection: At this stage, data should be selected based on specific criteria.
3) Data preparation: Usable form of data and identification of excessive variables is the purpose of this stage of the process.
4) Data evaluation: The overall framework of this step includes criteria such as kind of distribution of data, characteristics and structure of the data bank, overall condition of the data, etc.
5) Putting the response into a frame: The output of this section is presentation of the format in the form of a picture, diagram, neural network and other types of diagrams.
6) Instrument selection: At this stage, appropriate tools for data mining are selected.
7) Modeling: The data mining process mainly begins at this stage which includes search of patterns in the data collection, classification and evaluation of the data and etc.
8) Finding validity: This stage involves testing the models.
9) Presentation of the results: The result of this section is final report for the user.
10) Use of the results: The main purpose of data mining is use of discovered results for decision making, policy setting and predicting in order to create a better and new situation.

Neural networks

Neural networks are computing methods based on connection of multiple processing units. The network under consideration is formed of an optional number of cells or nodes, 41 or 42neurons relating the input set to the output set. In fact, synthetic neural networks are inspired to extent from natural learning systems in which a complex set of connected neurons are involved in the learning task. Along with each connection, there is a numerical value which is called weight 43. Each neuron receives 44signals from connected neurons. If sum of input signals exceeds a threshold 45, the neuron will lose functionality. The neurons are layered in groups. A layered network at least includes an input layer 46and and output layer 47. Between these input and output layers, wide layers 48 may exist. When a signal or pulse is transmitted to one layer, the signal begins activity from the upper layer and is evaluated and revised by the neurons in that layer. In fact, each neuron increases the signal strength and transfers the pulse to the next layer (saffai, 2006). Various kinds of neural networks have alternating layers. Self-organized maps have only one input and output layer. On the other hand, post error transmission neural network 50 has one or multiple extra discrete wide layers. To give the neural network ability to learn, after the signal is transferred from the first layer of the network to the lower network, each neuron should exert the effect on the signal. Therefore, the main application of artificial networks can be accurate estimate nonlinear functions appropriately.

Decision making tree

The decision making tree is an event reevaluation to express an alternating classification process. Each leaf of the decision making tree represents a different class. The decision making tree has the ability to break the complex decision making process into a set of simpler decisions that are easily interpretable. Areas of overall complex decision making (particularly in spaces with numerous dimensions) can be estimated by union of simpler regional decision making areas at various levels of the tree. Unlike prevalent single stage classifications where each sample is tested on all classes, in a tree classifier, a sample is tested on special subsets of classes and unnecessary calculations are omitted.

In the single step classifier, only subsets of adjectives are used for the method of between classes which is generally selected based on the best overall criteria. In the tree classifier, there is flexibility in choosing various subsets of adjectives in the different internal groups of the tree such that the selected subset differentiates between the classes of this group in the best way. This flexibility can lead to improvement in efficiency compared to single stage classifiers.

Supporting Vector Machine (SVM)

In 1965, a Russian researcher, Vladimir Vepnick, took a very important step in designing classifiers. He formulated the statistical theory of learning in stronger ways and presented supporting vector machines on this basis. Supporting vector machines have the following characteristics:

1) Classifying design with maximum generality,
2) Reaching the best overall point of the function,
3) Automatic determination of the best structure and topology for the classifier,
4) Modeling nonlinear discrimination functions using nonlinear nuclei and the concept of internal multiplication in Hilbert spaces.

SVM is an algorithm that finds a special kind of linear models leads to maximum margin for the cloud plain. Maximization of the cloud plain margin leads to maximization of discrimination between classes. The closest teaching points to the maximum cloud plain margin are called supporting vectors. Only from these vectors, points for identification of the boundaries between classes are used. SVM has better performance in model identification, regression estimation, prediction of financial time-series, marketing, production estimated efficiency, text classification, facial recognition using pictures, identification of hand writing and medical diagnoses in comparison to other learning techniques.

The Altman Model

In 1968, Edward Altman for the first time evaluated the effect of various hybrids of financial ratios to predict financial distress of companies. In this study, Altman used the method of discriminate analysis A model that he is famous for and is still used called “Z-score”. The Z-score is an index for financial health of companies.

The main theory of Altman was that his bankruptcy prediction model consisted of five financial ratios can be used to identify bankrupt companies from non-bankrupt. He suggested that his model is used to evaluate business loan applications, internal control processes and investment choices. In his model, Altman selected five ratios that appeared to have the most importance in predicting bankruptcy. Altman’s “Z-score” model is as follows:

$$Z = 1.2x1 + 1.4x2 + 3.3x3 + 0.6x4 + 0.999x5$$

x1. Circulating capital to total assets
x2. Accumulated profit or loss to total assets
x3. Profit before interest and tax to total assets
x4. Value of stock owner income to total value of debts
x5. Total sales to total assets

In this model, companies are classified as bankrupt that their Z values are lower as 1.8 and as non-bankrupt with values higher than 2.99 and there is no special classification for values between 1.8 and 2.99.

Main hypothesis 1: Data mining models are appropriate tools for prediction bankruptcy.

Sub-hypothesis 1: Supporting Vector Machine (SVM) model is a suitable tool for f prediction bankruptcy.

Sub-hypothesis 2: Decision tree model is an appropriate tool for prediction bankruptcy.

Sub-hypothesis 3: Neural network model is an appropriate tool for prediction bankruptcy.

Principal hypothesis 2: The Altman Z model is a suitable tool for prediction bankruptcy.

Hypothesis 4: for prediction bankruptcy, data mining models are more appropriate than the Altman model.

3. Methodology

Statistical Society and Sample

All research on bankruptcy prediction have divided sample companies into two bankrupt and non-bankrupt, to determine the authenticity of integrity prediction models. In this study, we also selected 64 companies as samples in 5 years, ie, the 320-Year-Company as following.

The statistical population of this research consisted of all companies accepted in the Tehran stock exchange during 5 years from 2013 to 2018. Based on information in the Rahavard Novin software, the number of accepted companies in the stock exchange was 545. Therefore, all companies that were the members of the statistical population and met the following criteria were included in the statistical sample:

- were manufacturing companies
- were accepted in the stock market before 2013.
- Their financial reports had been presented to the stock market during the time period from 2013 to 2018.

Since the criteria of bankruptcy and financial distress in this research was subject hood to Title 141 of the law of trade, initially a list of companies accepted in Tehran’s stock exchange that had become bankrupt in this period was prepared. This 6-year period was selected with attention to limitation in information. The number of these companies was 51. Among them, 38 companies where there was possibility of access to their information were selected. It should be noted that considering that for each company information from one and two years prior to bankruptcy was used, in total information of companies between the years 2013 to 2018 was accessed.

After selecting 38 financially distressed companies as above, 38 healthy companies were also needed as the second group selected from healthy companies. Therefore, the evaluated sample included 76 companies selected to conform regarding financial calendar year. Additionally, for each company, information of 3 financial years was used. The base year for bankrupt companies is the year when the company became subjected to Title 141 of trade law and for healthy companies; it is the year that the companies had the highest compiled profits.

Research Variables

The Altman model variables are as follows:
In 1983, to address criticism to his early model, Altman presented below model as version z to compare two models, data mining model and Altman's model Z. This variable group has been used for two models.

$$Z = 1.2x1 + 1.4x2 + 3.3x3 + 0.6x4 + 0.999x5$$

x1. Circulating capital to total assets
x2. Accumulated profit or loss to total assets
x3. Profit before interest and tax to total assets
x4. Book value of stock owner income to total book value of debts
x5. Total sales to total assets

Considering the results, this research was a fundamental research, because it was tried to identify criteria for predicting companies' bankruptcy. On the other hand, according to the application of this research in predicting companies' bankruptcy, it was an applied study. The research method with consideration of the nature of the research in the area of financial sciences was a survey study. The implementing stages and main steps in performing this research are summarized as follows:

1) Determining the financial ratios
2) Collecting data from the existing data base (organization for stock exchange bonds)
3) Identifying 38 companies as bankrupt and 38 as non-bankrupt in each year based on financial ratios.
4) Dividing sample data to two sets of educational and test data.
5) Evaluating specificity and validity of the models in predicting bankruptcy.
6) Providing the best model to predict bankruptcy of companies based on the performed research.

The conceptual model of research can be presented as figure1.

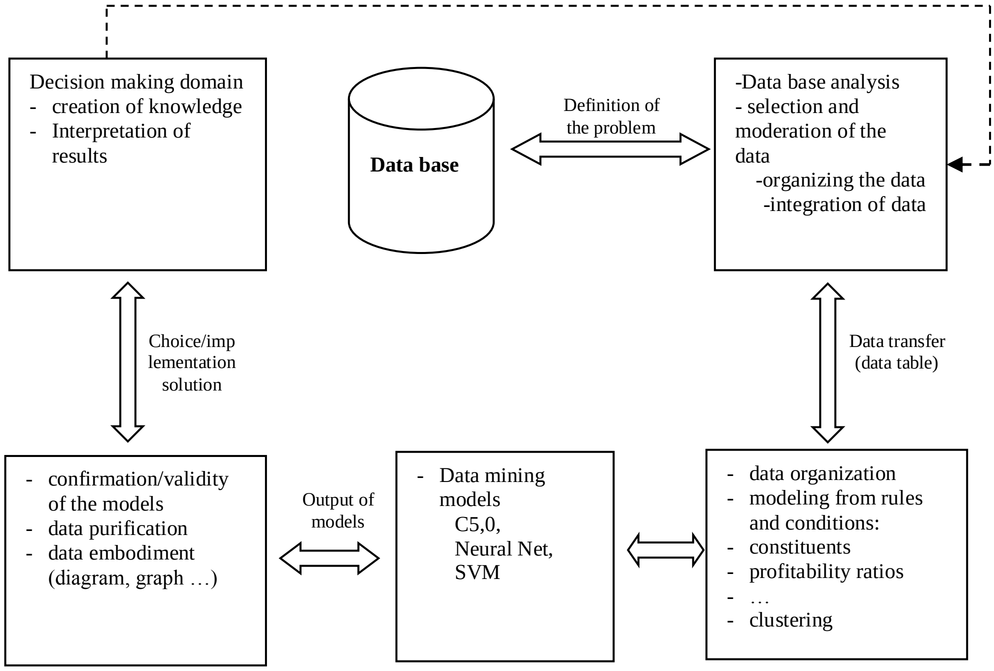

Figure 1: Conceptual Model of Research

4. Results

To test the hypotheses, the SPSS Clementine software was used such that the raw data was first collected and classified and entered into the software in the frame of SQL.

This software involves multiple algorithms and software capabilities such as svm, C5, ANNs. According to the basic concepts of data mining models, the data of research should be divided into two clusters %25 test and%75 teaching model, test data is given to teaching model.

In this research, 25 percent of the data were test data and 75 percent educational data. In this study, 3 models of support vector machine, decision making tree and neural network were used. The result of implementing these models is as follows.

Results of Implementing the Data Mining Models

- **SVM Model**

The SVM model was able to correctly predict 88.5 percent of bankrupt and 85.5 percent of non-bankrupt companies two years before the base year (t-2) and on the average, this model in year t-2 had a specificity of 87 percent. The results of this evaluation are presented in Table 1.

In year t-1 (one year before the base year), it had a specificity of 91.4 percent for bankrupt and 85.6 percent for non-bankrupt companies and on the average, the total specificity of this model has been 88.5 percent and the results of evaluation are presented in Table 2.

In year t (the base year), it had a specificity of 92.5 percent for bankrupt and 88.5 percent for non-bankrupt companies and on the average, the total specificity of this model has been 90. 5 percent and the results of evaluation are presented in Table 3.

Considering the results obtained, it can be concluded that the SVM model has a higher specificity of prediction in classifying companies into two groups of bankrupt and non-bankrupt. Therefore, the first minor hypothesis is confirmed

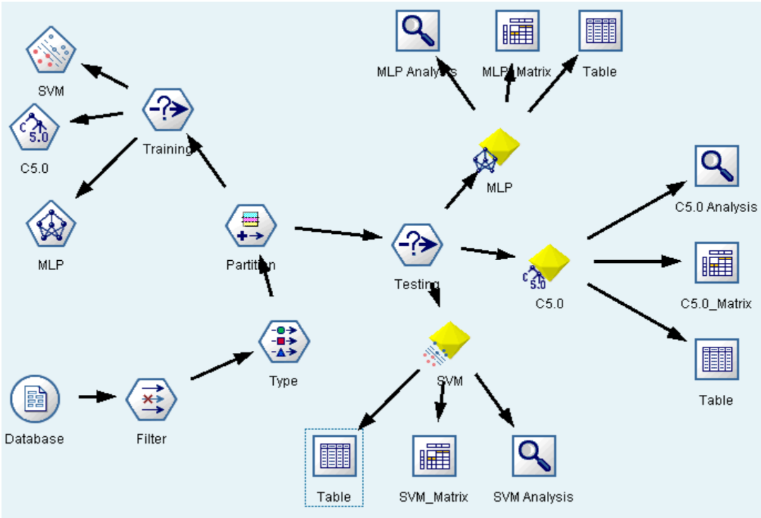

Figure 2: How to build and implement the algorithm in the Clementine SPSS software

Table 1. Results of classification of companies using the SVM model in year t-2

Kind of company	Total number	Number bankrupt	Number nonbankrupt	Percent bankrupt	Percent nonbankrupt
Bankrupt	38	30	8	0/885	0/115
Nonbankrupt	38	9	29	0/145	0/855

Table 2. Results of classification of companies using the SVM model in year t-1

Kind of company	Total number	Number bankrupt	Number nonbankrupt	Percent bankrupt	Percent nonbankrupt
Bankrupt	38	32	6	0/914	0/086
Nonbankrupt	38	9	29	0/144	0/856

Table 3. Results of classification of companies using the SVM model in year t

Kind of company	Total number	Number bankrupt	Number nonbankrupt	Percent bankrupt	Percent nonbankrupt
Bankrupt	38	34	4	0/925	0/075
Nonbankrupt	38	8	30	0/115	0/885

- **Model C5, 0**

Model C5, 0 was able to correctly predict 93.8 percent of bankrupt and 87.5 percent of non-bankrupt companies two years prior to the base year (t-2) and on the average, this model in year t-2 had a sensitivity of 90 percent. The result of the evaluation is shown in Table 4.

In year t-1 (one year before the base year), it had a specificity of 91.5 percent for bankrupt and 93.5 percent for non-bankrupt companies and on the average, the overall specificity of this model was 92.5 percent. The result of the evaluation is shown in Table 5.

In year t (the base year), it had a specificity of 89.5 percent for bankrupt and 93 percent for non-bankrupt companies and on the average, the overall specificity of this model was 91.25 percent. The result of the evaluation is shown in Table 6.

Considering the results obtained, it can be concluded that model C5, 0 has a higher predicting specificity in classification of companies into two groups of bankrupt and non-bankrupt. Therefore, the second minor hypothesis is confirmed

Table 4. Results of classification of companies using the C5,0 model in year t-2

Kind of company	Total number	Number bankrupt	Number nonbankrupt	Percent bankrupt	Percent nonbankrupt
Bankrupt	38	30	8	0/938	0/062
Nonbankrupt	38	8	30	0/062	0/875

Table 5. Results of classification of companies using the C5,0 model in year t-1

Kind of company	Total number	Number bankrupt	Number nonbankrupt	Percent bankrupt	Percent nonbankrupt
Bankrupt	38	33	5	0/915	0/085
Nonbankrupt	38	7	31	0/175	0/935

Table 6. Results of classification of companies using the C5,0 model in year t

Kind of company	Total number	Number bankrupt	Number nonbankrupt	Percent bankrupt	Percent nonbankrupt
Bankrupt	38	35	3	0/895	0/105
Nonbankrupt	38	6	32	0/07	0/93

- **The Neural Net Model**

The neural net model was able to correctly predict 94.5 percent of bankrupt and 90 percent of non-bankrupt companies two years prior to the base year (t-2) and on the average, this model had a sensitivity of 92.25 percent in year t-2. The result of the evaluation is shown in Table 7.

In year t-1 (one year before the base year), it had a specificity of 96.9 percent for bankrupt and 87.5 percent for non-bankrupt companies and on the average, the overall specificity of this model was 92.2 percent. The result of the evaluation is shown in Table 8.

In year t (the base year), it had a specificity of 95.7 percent for bankrupt and 96.9 percent for non-bankrupt companies and on the average, the overall specificity of this model was 96.3 percent. The result of the evaluation is shown in Table 9.

Table 7. Results of classification of companies using the Neural Net model in year t-2

Kind of company	Total number	Number bankrupt	Number nonbankrupt	Percent bankrupt	Percent nonbankrupt
Bankrupt	38	36	2	0/945	0/055
Nonbankrupt	38	3	35	0/096	0/904

Table 8. Results of classification of companies using the Neural Net model in year t-1

Kind of company	Total number	Number bankrupt	Number nonbankrupt	Percent bankrupt	Percent nonbankrupt
Bankrupt	38	37	1	0/969	0/031
Nonbankrupt	38	4	34	0/125	0/875

Table 9. Results of classification of companies using the Neural Net model in year t

Kind of company	Total number	Number bankrupt	Number nonbankrupt	Percent bankrupt	Percent nonbankrupt
Bankrupt	38	37	1	0/957	0/043
Nonbankrupt	38	2	36	0/031	0/969

- **Analysis of the Altman Z Model**

To analyze the Altman Z model, the values related to the variables of this model were extracted from the Rahavard Novin software. Subsequently, this information was stored in Excel software. Using the Excel software, values for the variables were calculated. Next, in one stage, information related to 2 years prior to the base year was extracted from company financial reports and the Z index was calculated for them. In the second stage, information related to one year prior to the base year was extracted from financial reports and after calculating the ratios and placing them into the model, the Z index for each of the companies was calculated. In the final step, information related to the base year was extracted from financial reports and after calculating the ratios and placing them into the model, the Z index for each of the companies was calculated.

Results of Implementing the Altman Z Model

After computing the Altman Z value in this model, companies for values less than 1.8 were considered bankrupt and for values higher than 2.99 were considered non-bankrupt and for values between 1.8 and 2.99, special classification was not available. Therefore, according to the results, it can be concluded that the Altman model at year t-2 correctly predicts 79.1 percent of bankrupt and 82.2 percent of healthy companies and on the average, this model has specificity equal to 80.6 percent in year t-2as shown in Table 10.

In year t-1, it has had a sensitivity of 87.5 percent for bankrupt companies and 76.5 percent for healthy ones and on the average, this model in year t-1 has had a specificity equal to 82 percent as shown in Table 11.

In year t, it has had a sensitivity of 90.4 percent for bankrupt companies and 78.9 percent for healthy ones and on the average, this model in year t has had a specificity equal to 84.65 percent as shown in Table 12.

Table 10. Results of classification of companies using the Altman Z model in year t-2

Kind of company	Total number	Number bankrupt	Number nonbankrupt	Percent bankrupt	Percent nonbankrupt
Bankrupt	38	28	10	0/791	0/209
Nonbankrupt	38	11	27	0/178	0/822

Table 11. Results of classification of companies using the Altman Z model in year t-1

Kind of company	Total number	Number bankrupt	Number nonbankrupt	Percent bankrupt	Percent nonbankrupt
Bankrupt	38	28	10	0/875	0/125
Nonbankrupt	32	9	29	0/235	0/765

Table 12. Results of classification of companies using the Altman Z model in year t

Kind of company	Total number	Number bankrupt	Number nonbankrupt	Percent bankrupt	Percent nonbankrupt
Bankrupt	38	30	8	0/904	0/096
Nonbankrupt	38	9	29	0/211	0/789

5. Discussion and Conclusions

Prediction bankruptcy of companies is among interesting and important studies in the financial sector. With prediction of bankruptcy and after that, finding the roots of the problem and solving it can lead to very satisfactory results With increasing corporate and their capital structure complexity and the emergence of severe financial crises in terms of macro and micro economic, owners and stockholders of entities are looking for a way to protect themselves against such risks and that make sensitive them use of prediction tools and models to assess the financial power of companies. Bankruptcy prediction can be used as an effective tool to help them. Prediction methods are constantly evolving and today's data mining models have a special place among these methods.

In this research, the models of support vector machine, decision making tree and neural network which are data mining methods were used to predict the production company bankruptcy. Two years prior to the base year, on the average, models created by data mining were able to predict bankruptcy with a specificity of 89.35 percent; while, the Altman Z model had an average specificity of 80.6 percent. One year prior to the base year, data mining models had an average specificity of 90.4 percent; while, the Altman Z model had an average specificity of 82 percent. In the base year, data mining models had an average specify of 91.8 percent and the Altman Z model had an average specificity of 84.65 percent. Therefore, considering the results, the specificity of the data mining models is better compared to the Altman Z model. As a result, in general, data mining models are more appropriate tools to predict bankruptcy of companies. Managers are recommended to prevent company from bankruptcy by full understanding of market situation, preventing from increasing the ratio of total debt to total assets, having experience, expertise, spending related educational courses. It is recommended to the shareholders, creditors, lenders and all those who care about company's performance to use data mining models to predict.

References

1) Angga Pertapan, F, Sri Hartono, Dyah Purnomo WulanA. (2018). Bankruptcy Prediction in PT Blue Bird, Tbk 2011-2016 Using Altman Z-Score, Springate, and Zmijewski Model, The 2nd International Conference on Technology, Education, and Social Science. (The 2nd ICTESS 2018)
2) Altman ,E.L,.(1968). Financail Ratios, Disarmament Analysis and the prediction of corporate Bankruptcy. the Journal of finance, 23(5) ,598-609
3) Boyacioglu, M., Yakup ,K., Mer Kaan ,B. (2009).Predicting bank financial failures using neural networks, supportvector machines and multivariate statistical methods: A comparative analysis inthe sample of savings deposit insurance fund (SDIF) transferred banks inTurkey. Expert Systems with Applications, 36,3355–3366
4) Firouzian .,M.,. Javid., D .,Najmodin, N. (2011). Evaluation of application of models of predicting bankruptcy of Altman and the genetic algorithm in companies accepted in the Tehran stock exchange. Accounting and Audit Reviews., 18(65),99-114.
5) Hung, chilhli. & Chen, jing-hong. (2007). A selective ensemble based on expected probabilities for bankruptcy prediction, Expert Systems with Applications, 36(3), 5297-5303
6) GHodrati, H, Moghadam, A. (2010).The Accuracy of Bankruptcy Forecasting Models (Altman, Shirata, Fulmer, CI, Scheringite, Olson, Chemically, Genetic Frajzadeh and McKay Genetics) in Tehran Stock Exchange . Accounting and Audit Reviews, 16 (58) ,153-12.
7) GHadirimoghadam, A; Gholampourfard, M; F Nasirzadeh, F. (2009). Review the ability of Altman and Olson bankruptcy prediction models to predict bankruptcy Companies listed on the Stock Exchange. Journal of Science and Development. 16(28), 100-125.
   Karami, G; Seyyed Hosseini, S . (2012).The usefulness of accounting information to market information to predict bankruptcy. Accounting Knowledge, 3 (11), 93-116
8) Kim, Myoung-Jong & Dae-Ki Kang. (2012) . Classifiers selection in ensembles using genetic algorithms for bankruptcy prediction. Expert Systems with Applications, 15, 1-7.
9) Jardin, Philippe Du & Eric Séverin. (2011). Predicting corporatebankruptcy using a self-organizing map: An empirical study to improvethe forecasting horizon of a financial failure model. Decision SupportSystems, 51: pp. 701–711
10) Jeong ,Chulwoo, Jae ، Min , H., Myung, S. (2012). A tuning method for the architecture of neural network models incorporating GAM and GA as applied to bankruptcy prediction. Expert Systems with Applications, 39,3650–3658.
11) Makiyan, N; Al-Mudarresi,S; Salim K. (2009). Comparison of artificial neural network with logistic regression and audit analysis in prediction companies bankruptcy. Quarterly Journal of Economic Research, (10) 10 , 141-161.
12) Mehrani, S; Mehrani, K, Monasfi, Y; Karami , Gh. (2005) Applied research On Zimiski and Shirana 's prediction Pattern of bankruptcy in

companies accepted in Tehran Stock Exchange. Accounting and auditing reviews. (12) 41, 105-131.

13) Nakhaiezade, G. (1998). Data Mining, Theoretische Aspekte and Anwendungen Beiträge zur Wirtschaftsinformatik. Physica-Verlag.
14) Premachandra, I. M., Bhabra, G. S. & Sueyoshi, T. (2009). DEA as a tool for bankruptcy assessment: A comparative study with logistic regression technique. European Journal of Operational Research, 193, 412-424
15) Prusak, B.(2018). Review of Research into Enterprise BankruptcyPrediction in Selected Central and Eastern European Countries International Journal of Financial Studies Review ,(6) 60,20-48
16) Saffai, A. (2006). How to build a smart neural network? Example of object writing in artificial neural networks and artificial network. Network Monthly ,6 (71),9-87.
17) Yoon, J,.Young, K. (2010). A practical approach tobankruptcy prediction for small businesses: Substitutingthe unavailablefinancial data for credit card sales information. Expert Systems with applicationWhitaker, R. (1999). The Early Stage of Financial Distress. Journal of Economics and Finance, 23 (2), 123-133