

**PREDICTION OF FIRMS' FINANCIAL DISTRESS USING ADABOOST ALGORITHM AND
COMPARING ITS ACCURACY TO ARTIFICIAL NEURAL NETWORKS.**

(Recibido el 10-07-2017. Aprobado el 21-09-2017)

Mahmoud Hemmatfar

*Associate Professor of Accounting Boroujerd
Branch, Islamic Azad University, Boroujerd, Iran.*

Seyed Alimorad Hosseinipak

*PH.D Scholar of Accounting Boroujerd Branch,
Islamic Azad University, Boroujerd, Iran.*

Abstract.One of the most important topics discussed in the area of financial management is investors' ability to tell favourable investment opportunities from unfavourable ones. One way to help investors is to present firm's financial distress prediction models. So far, different techniques have been used to design firm's financial distress prediction models. Recent studies in the field of financial distress prediction have focused on creation and application of artificial intelligence and machine learning methods, AdaBoost algorithm and artificial neural networks are used in the present study as a comparative model to Companies' financial distress prediction. 660 samples were selected from 112 financially distressed companies and 548 non-financially distressed over a 6-year period from 2007 to 2012 have been selected.

Research Findings suggest that in Companies' financial distress prediction, the model based on AdaBoost algorithm has a higher overall accuracy than the model based on artificial neural network.

Keywords: Prediction, Financial Distress, AdaBoost Algorithm, Artificial Neural

1. NETWORKS

From economic view point, financial distress can be interpreted as operating at a loss, in this case, the company has faced failure; in fact, in this situation, return rate is lesser than cost rate. Another case can be seen when the firm fails to comply with one or some of debt contract clauses as keeping the current ratio or the ratio of equity to total assets under the contract, which is translated to technical default. Other forms of financial distress include, when company cash, is insufficient to repay the debt principal and interest, and also when company's brand equity is a negative number. (Rostami, et al 2011). One of the approaches which can be utilized to exploit investment opportunities and better allocation of resources, is prediction of financial distress. Thus, firstly, with warning, firms can be warned of becoming financially distressed so that they can take proper actions in respect to these warnings. Secondly, investors and creditors, differentiate favorable opportunities from adverse opportunities and invest their resources in appropriate opportunities (Rae and Fallahpour, 2007). On the other hand, the development of new technologies and their application in different sciences has attracted professional accounting and financial management attention to use of these methods in this profession. Technological changes and their application in various disciplines has led the accountants in order to increase efficiency use these new technologies. One of the most important ways to increase the accuracy of prediction of companies' financial distress is using new methods of data mining to predict. So far, at the international level, considerable researches have been done to develop and provide models to identify distressed companies. According to the above, the main problems with this study were to assess and compare the financial distress predictions using neural network and Adaboost algorithm.

1.1. Research goals

1- Prediction of companies' financial distress in research's area using adaboost algorithm compared against neural network and assessing the models' performance.

2- Comparison the power of these models in prediction of financial distress of companies in the research area.

1.2. Research questions and hypothesis:

Research questions

Are artificial networks able to predict companies' financial distress?

Is adaboost algorithm able to predict companies' financial distress?

Is adaboost algorithm's accuracy higher than artificial neural networks in prediction of financial distress?

1.3. Hypothesis

Artificial networks have the ability to predict companies' financial distress.

Adaboost algorithm is able to predict companies' financial distress.

The overall accuracy of companies' financial distress prediction using adaboost algorithm is more than artificial neural networks' .

2. RESEARCH LITERATURE

2.1. Adaboost algorithm

Adaboost is abbreviation of Adaptive Boosting, and is a kind of machine learning algorithm that was developed in 1995 by Yoav Freund and Robert Schapire (Freund and Schapire, 1996). In fact, Adaboost is a meta-algorithm and can be used in association with other learning algorithms, and their enhanced effectiveness. Adaboost algorithm is based on the principle that "it is very difficult to introduce a precise law, while making lots of rules of thumb that have a moderate accuracy is easily done." In other words, in this approach, weak laws that are gathered together, they achieved a strong law. this method differs from boosting due to adaptive omission of errors in every stage of learning.

Adaboost algorithm function is to create a weak classifier in each round ($t = 1, \dots, T$). If we consider weak linear classifier, each of these classifiers draws a line or a page that divides the space into halves. Each weak classifier at every stage, deals with a different series s of learning samples . Precisely speaking, in each round, different weights

are applied on learning samples (Sadeghi, 2008). This means that in each round, wrongly classified samples' weights are escalated and correctly classified weights are reduced. In fact, the examples not correctly recognized as samples now by classifier have more chance to be utilized in the next stage (Zareh Chakook and Moghadam Cherkery, 2010). Once the training process is finished, the weak classifiers are combined and constitute the final classifier, a decisive classifier that is based on training; hence, the final classifier achieves a high level of accuracy in a set of tests. Several authors have demonstrated this fact theoretically and experimentally (Banfield et al., 2004), (Boor and Koohavy, 1999) (Bremen, 1998) (Daytryj, 2000) (Friedman et al., 2000). Artificial neural networks "Neural network is a machine learning system based on a simple model of biological neurons' works" (Angelina et al., 2007). The basis of calculation in the neural network is to model the human brain structure and its function and to use these models to solve complex nonlinear problems or for making computers which can demonstrate most of brain aspects. In fact, artificial neural network is a set of processing units, called neurons, each of these units with a special coefficient, in order to achieve the set's aim, is connected to other processing units (giyupta et al., 2003). Following, a summary of the most important concepts of neural network is presented, and then each of its divisions are expressed in detail. Neural networks are triangles with three concept_ sides: Data analysis system, 2_ neurons or nerve cells, and 3_ network or groups' labor law. In a classic definition Haykyn says "The neural network is a massive collection of parallel processors with innate talent for storing experimental data and its application, and this network is similar to a brain at least in 2 aspects :

(1) the so-called stage of learning (2) and synaptic weights used to store knowledge.

The task of neural networks is learning. Almost like learning in a young child. Learning in neural network are often done "supervised ". Parents, show images of different animals to a child and tell the name of each to the child. Here the focus is on the animal of rabbit. The child sees the pictures of different types of rabbits along with input data (images and sound) for each sample, he is said that the information is either on a "rabbit" or not.

Without having to be told, the brain system analyzes incoming data and reaches findings in the field of each input information such as "color, size,

sound, having toe or ear". After a while, he will be able to recognize a "new type" of a rabbit that has never seen before, since for every animal in the child's learning process, it is said to be a rabbit or not, this type of learning is called supervised. Another type is unsupervised learning which is simulated by neural network but has fewer applications, the rest of the season, each of these systems are described. In some cases, explanations for other systems namely strengthening systems, will be provided.

3. FINANCIAL DISTRESS

In Oxford dictionary, the word «Distress» means pain, grief, lack of funds and being poor is given. Various definitions are given in financial literature of financial distress. Gordon (1971), one of the first academic studies on the theory of financial distress presented it, as "reduction of the profitability in the company which escalates the probability of failure to repay interest and principal debt" (Gordon, 1971). Whitaker (1999) considers the financial distress a situation in which the Cash inflows are less than interest costs related to long-term debt. (Whitaker, 1999). From economic standpoint, financial distress can be interpreted as the detriment of the company, in which case the company suffered heavy and continuous losses.

In fact, in this case, the internal rate of return is less than the cost of capital. Another case of financial distress occurs when the company fails to comply with one or more of the clauses in their contracts, financial liabilities that is known as state of technical default.

Other scenarios include financial distress when the company's cash flow to repay debt principal and interest is insufficient or when the company's equity equals to a negative numeric value (Kopland and Weston, 1992). It should be noted that financial distress will not necessarily lead to bankruptcy, but bankruptcy is one of the consequences of financial distress, the last and most severe stage of financial distress, and bankruptcy is the last resort (Mousavi and Tabarestani, 2008). To determine the exact cause or causes of financial distress is not easily done, often multiple reasons have led to the phenomenon of financial distress. In connection with the causes of financial distress, the researchers offered different views.

(Kaits and Broker 1988) and Newton (2010), believe that the financial distress' determinants,

include external factors (externals) and internal factors (internals).

External factors: factors that are not controlled by the company, but are due to financial problems and alteration in economical structures, changes in public demands and business fluctuations (inflation, prices falling and rising interest rates and ...), the problems associated with financing, events and natural disasters and the intensity of competition in the market.

On the other hand, internal factors include management's mistakes or times when they were incapable if taking necessary decisions in the past, such as creation and development of the size of the credit for the consumer (excessive credit Sales) ; inefficient management (lack of training, experience, ability and initiative in the field of competition, technology and resource management and administrative errors), infidelity and fraud can be noted (Katz and Brakr, 1988) (Newton, 2010). Altman (1983) notes macroeconomic indicators and their impact on companies' financial distress and believes factors such as decline in economic growth, a reduction in the volume of operations in the capital markets, decline in liquidity growth, an increase in the establishment and development of corporations, can influence financial distress in companies (Altman, 1983).

4. RESEARCH BACKGROUND

Researches done in Iran.

In Iran, similar researches have been done in this field and some of the researches conducted as follows.

Financial distress prediction using artificial neural networks by Resa Rai and Saeed Fallah Pour; financial research carried out in spring and summer of 1383, compares the predictive power of neural networks and method of multivariate linear classifier.

They, defined companies that are subject to article 141 of the Commercial Code from the year in which these companies became subject to the code as financially distressed companies, and through financial information of years ago, they predicted these companies.

Overall, they included 40 companies financially distressed from 73 to 80, which reached the requirements, and also 40 companies through

random selection were selected among healthy companies. The average forecast accuracy of linear classifier model generally is 93.4% and forecast accuracy of artificial neural network model is calculated 95.3 %

Their point of discontinuity in linear classifier model is assumed equal to 0.5 Ultimately concluded that "the model of artificial neural networks" in prediction of financial distress, is significantly higher than the "multivariate discriminant model" in forecasting accuracy. However, according to two points, the mentioned conclusion is notable:

Use of 0.5 for the discontinuity point in multivariate discriminant, model, and relatively low level of difference in models' accuracy (95.3 % in comparison with the 93.4% Atrinejad (2005), predicted share price with artificial neural networks.

Neural network used in this study were MLP with two hidden layers that are trained using back-propagation algorithm. The results of his study showed that neural networks perform better than univariate and multivariate time_ series models in prediction of stock price. In a study by Adel Azar and Cyrus Karimi in 1388 under the title of "prediction of stock returns using accounting ratios with artificial neural network approach," the researchers concluded that the use of artificial neural networks leads to substantial reduction in the estimation error in comparison with model of linear least squares Peyravi (2012) in histudy evaluates, the ability of market value added forecast based on the capital structure and profitability using artificial neural network techniques, showed that the perception of two-layer Perception model to predict the added value is better than the linear regression model. The models in this study with achieving a one percent error estimator function, significant relationship between capital structure and market value added is proven.

2. Foreign researches:

The first research which created a model for prediction of bankruptcy were the research of William Beaver in 1966. Beaver selected a collection of 30 financial ratios that he thought to be the best ratios to assess the health of a company. Then, ratios , based on the evaluation of the Companies, were classified into six groups. through his researches, he reached the conclusion that the value of each ratio is in creditability of classifying

bankrupted and non_bankrupted companies, and lesser classifying error points to high value of each ratio. According to this principle, Beaver described the ratios with least classifying error in order of importance: cash flow to total asset, Net revenue to total assets, total debt to total assets, working capital to total assets, current ratio and the ratio of Uncertainty Interval.

Chiang et al (1996) used Genetic Algorithms to predict net assets of investment companies at the end of the year.

They compared data of networks and their results with the results obtained from traditional econometric techniques, they found that genetic algorithm do better than regression methods when there are significantly less data.

In another survey with the aim to include qualitative factors such as political effects with quantitative factors, led to creation of a new type of genetic algorithm, which they examined their models in Taiwan stock market.

Garlyaks (1999) lunched time_ series forecasts using Genetic algorithm associated with kernel function and back-propagation error for the stock market. He concluded that the financial time series prediction by Genetic Algorithm perform better than the classical statistical model and other models.

Chan (2000) predicted financial time series using daily data of Shanghai stock exchanges and genetic algorithm. To speed up and for more convergence, they used gradient descent algorithm and multiple linear regressions to determine the weights.

They concluded that genetic algorithm could satisfactorily predict time series better, and their approach for selecting weights, led to a lower computational cost.

Kim and Han (2000) utilized an adjusted neural network by genetic algorithm to predict stock index, in this case genetic algorithm was used to decrease future complexity of price time series. Olson and musman (2003) using data of 2352 Canadian companies from 1976 to 1993 predicted stock market return; with using three methods of logistic regression, least normal squares and genetic algorithm, they predicted the return.

Current ratio, liquidity ratio, inventory turnover ratio, debt-to-equity ratio, the ratio of return on

equity, ratio of book value to market value and sale price were the studied variables. The result showed superiority of genetic algorithm compared with two other methods. Henry Haglund (2012) in his study entitled "The discovery of earnings management using neural networks" compares the power of regression model with neural network model's.

. His findings showed that the neural network model has more accuracy in detecting earnings management, thus, it is a better tool for detecting earnings management is the result.

5. DATA COLLECTION

Data collection for this study used data provided by the Stock Exchange, library studies, internal and external papers and dissertations..

5.1. statistical population and sampling method

Since the time area of this research was from 2007 to the end of 2012, so all the statistical population are from accept companies in Tehran stock exchange.

6. RESEARCH METHODOLOGY

The research can be categorized in the field of positive research and the purpose is functional. Given that for the test, hypotheses are from historical information, it is Quasi-experimental. The present method is kind of inductive and post-event (using past data) research, and using support vector machine, multiple discriminant analysis and logistic regression models for prediction of companies financial distress. Predictive powers of these models are compared.

6.1. The populations and the samples

Since the time scope of the investigation is from early 1387 until the end of 1392, our statistical society includes all accepted companies in Tehran stock exchange.

The sampling method is systematic elimination. Our observation reached 660 year_companies, from which 112 samples were financially distressed and 548 companies were non_financially distressed. To ensure validity of the model, samples were split into two categories of learning samples and the second were test samples. First samples made up from 540 year _company and the second were made up from 110 year _company.

6.2. Data collection

Data collection for this study, used data provided by the Stock Exchange, library studies, internal and external papers and dissertations.

7. RESEARCH'S FINDINGS

are presented.

As can be seen in Table 1. All variables are significant at 95%. Therefore, we can say there is a significant difference between variables. Significant

Table 1. Variables' mean difference' significance test

difference		healthy		Financially distressed		variable
P-value	t-test	Standard deviation	mean	Standard deviation	Mean	
0.035	2.561	1.012	1.5342	0.928	0.6981	CA/CL
0.017	3.709	1.307	0.2090	0.859	-0.1437	EBIT/TA
0.023	2.917	0.879	0.5732	0.886	0.2981	TS/TA
0.0098	4.128	1.375	0.632	0.809	-1.231	WC/TA
0.11	3.965	0.990	0.3721	1.320	-0.6562	RE/TA

from healthy ones. In fact our aim is to modify the Following

regression function based on five financial ratios from twenty-eight financial ratios.

Table 2 Five selected financial ratios with modified prediction coefficient

Modified coefficient value	Financial ratio		The objective function
-0.965	Earning before interest and taxes to total assets	X ₁	β ₁
-1.045	The ratio of operating profit to sales	X ₄	β ₂
-0.733	The ratio of net working capital to total assets	X ₃	β ₃
-1.462	accumulated Income (loss) to total assets	X ₁₅	β ₄
-0.563	Earnings before interest and taxes to interest expense ratio	X ₁₉	β ₅

models, in prediction of financially distressed companies, populations' means comparison test (f Statistic) was used. The result is presented in table

Table 3. Comparison of the power of predictions of different models in financial distress occurrence year

Neural network	Adaboost algorithm	Group name
29	33	Financially distressed
82.89%	94.28%	
28	31	Without financial distress
77.14%	88.57%	
57	64	Models' overall accuracy
81.42%	91.42%	

First hypothesis: the neural network is capable of prediction of companies' financial distress.

In Table 1, the average and standard deviation of each of these variables and their significance test for the two groups of financially distressed and healthy companies

difference between the two groups confirms the information content of accounting items differs between companies with financial distress and healthy companies.

The second hypothesis: Adaboost algorithm is able to predict companies' financial distress.

At first 28 financial ratios for each of 110 samples were defined as input.

Adaboost algorithm aims to find five financial ratios, which are more capable of differentiating financially, distressed companies

$$\text{Financial Distress}_{it} = \beta_1 X_{A,it} + \beta_2 X_{B,it} + \beta_3 X_{C,it} + \beta_4 X_{D,it} + \beta_5 X_{E,it}$$

7.1.Third hypothesis testing

Third Hypothesis: the overall accuracy of companies' financial distress prediction using Adaboost algorithm is more than artificial neural networks.

In table 4-4 capability of Adaboost Algorithm and neural networks for prediction of financially distressed and healthy companies in the year of financial distress occurrence is presented. In order to comparison the overall (power) accuracy of

According to Table 4, the results of F-test for comparison of mean of prediction accuracy in the models is presented; the result show that at 95% confidence, mean of prediction for the three models are significantly different, because the F statistic in the test (6.121) is more than the acceptable minimum value for the 95% confidence level.

As the result, at the level of acceptable error of 5%, the assumption of significant difference between the means of prediction accuracy in the models cannot be rejected, and the assumption in which accuracy mean of forecast in adaboost

algorithms and neural networks are significant, is confirmed.

Table 4. Mean of models accuracy

neural networks	Adaboost algorithm	method
81.42%	91.42%	Mean of overall accuracy
6.121		F Statistic
0.018		Significance (P-Value)

8. CONCLUSION

Neural network are able to predict companies financial distress. "

According to table (1), each of variables' mean and standard deviation, and test of significance for either of the healthy and financially distressed groups, points to significance at 95 % of each of variables, and this means, neural networks are able to predict companies' financial distress

Adaboost algorithm is capable of prediction of companies' financial distress.

Test results: According to Table 2, adaboost algorithm can predict Company's financial distress."The overall accuracy of companies' inancial distress prediction using adaboost algorithm is more than neural networks."

9. APPLIED PROPOSALS FOR RESEARCH

According to this study's, results, suggestions for their use is presented as follows

1) stock exchange organization is suggested to issue indexes associated with health wise accepted companies' financial distress and help optimal allocation of capital, and transparency of decision making environment.

2) According to the findings of this study, capital market participants, decision-makers, financial analysts, potential and actual investors in stock exchange are recommended to use adaboost algorithm which has a high level of accuracy and prediction power in the analysis of investment projects in: financial assets and securities in order

Friedman, J., Hastie, T. and R. Tibshirani (2000). "Additive logistic regression: a statistical view

to separate financially distressed and not distressed companies, because this model will lead to optimal investment portfolio with minimum risk and maximum return , while doubles the transparency of the decision-making environment and results.

REFERENCE

- Altman, E. (1983). "Why businesses fail". The Journal of Business Strategy, Vol. 3, Pp. 15-20.
- Banfield, R. E., Hall, L. O., Bowyer, K. W., Bhadoria, D., Kegelmeyer, W. P. and S. Eschrich (2004). "A comparison of ensemble creation techniques, in: F. Roli, J. Kittler, T. Windeatt (Eds.), Multiple Classifier Systems".
- Bauer, B. and R. Kohavi (1999). "An empirical comparison of voting classification algorithm: bagging, boosting and variants". Machine Learning, Vol.36, Pp. 105–142.
- Bradstreet, D. (1998). "Bankruptcy Insolvency Accounting Practice and Procedure". Wiley, Vol. 2, Pp. 21-41.
- Breiman, L. (1998). "Arcing classifiers". The Annals of Statistics, Vol. 26, Pp. 801–849.
- Dietterich, T. G. (2000). "Ensemble methods in machine learning, in: J. Kittler, F. Roli (Eds.), Multiple Classifier Systems". Lecture Notes in Computer Science, Vol. 1857, Pp. 1–15.
- Eugene, F. B. and C. G. Louis (1994). "Financial Management: Theory & Practice". Houghton Mifflin Harcourt School. 2nd edition.
- Freund, Y. and R. E. Schapire (1996). "Experiments with a new boosting algorithm". 13th International Conference on Machine Learning, Morgan Kaufmann, San Francisco, Pp. 148–156.
- of boosting". The Annals of Statistics, Vol. 38, Pp. 391–393.

Gerald, E.A. (1998). "The Analysis and Use of Financial Statements". American Business Review, Vol. 16, Pp. 983-1037.

Gitman, L.J. (1998). "Principle of Managerial Finance". Working paper. New York. Harper Collins College.

Keats, B. W. and J. S. Bracker (1988). "Toward a theory of small firm performance: A conceptual model". American Journal of Small Business, Vol. 12, Pp. 41-58.

Lecture Notes in Computer Science. Vol. 3077, Pp. 223-232.

Newton, G. W. (2010). "Bankruptcy and Insolvency Accounting, Practice and Procedure". Vol. 1, Pp. 4-21.

Whitaker, R. (1999). "The Early Stage of Financial Distress". Journal of Economics and Finance, Vol. 23, Pp. 123-133.

Weston, J. F. and T. E. Copeland (1992). Managerial Finance; Dryden Press. 9th Edition, New York: Mc Graw Hill.

Wikipedia (<http://en.wikipedia.org/wiki/AdaBoost>)