Bankruptcy Prediction, Accounting Variables and Economic Development: Empirical Evidence from Iran

Gholamreza Karami[1], Navid Attaran[1], Seyed Mostafa Seyed Hosseini[1] & Seyed Mojtaba Seyed Hossein[2]

[1] Faculty of Management, University of Tehran, Tehran, Iran

[2] Faculty of Engineering, Azad University, Branch of Mashhad, Mashhad, Iran

Correspondence: Seyed Mostafa Seyed Hosseini, Faculty of Management, University of Tehran, Pol-e-Nasr, Tehran, P.O. Box 14155-6311, Iran. Tel: 98-935-473-7464. E-mail: m.hosseini.ut@gmail.com

Abstract

One of the key factors in reaching economic development and accelerating its growth is increasing investment opportunities. However investment resources are scarce especially in developing countries, thus its efficient allocation between those activities with higher efficiency is noteworthy. Financial reports extracted from accounting systems are somehow viewed as rich financial media facilitating economic decision making. In this regard, assessing firm's financial condition via accounting indicators can in fact help investors in decision making process, prevent resource wastage and expedite economic development. The purpose of this research is to study the usefulness of accounting information in predicting financial condition of companies listed in Tehran Stock Exchange between 2004 and 2010 using a Genetic Algorithm model, further to improve results generalization, cross validation is executed. The GA model predicted financial condition of sample firms one year ahead with 90.3% of accuracy and 1.7% standard deviation.

Keywords: economic development, accounting information, bankruptcy, Genetic Algorithm, cross-validation

1. Introduction

Many economists agree upon the notion that while increasing investment may not be sufficient for economic development and growth, it's necessary. These investment resources are scarce especially in developing areas so one can claim the most valuable element in performing an efficient investment is possession of relevant information and on time prediction. Decision making has always some level of uncertainty which is caused by information asymmetry. Business environment can be seen from two perspectives: Ideal and uncertain conditions. In ideal realm there is complete information symmetry between market participants, in other words there is no difference between organizations inside and outside information. But in uncertain condition there is no such thing due to information asymmetry which has various levels depending on information quality, quantity and transparency. Ideal condition resembles a utopia in business so economists have focused on uncertain environment. Scott (2009) identifies two symptoms of information asymmetry, namely Moral hazard and adverse selection and insists on the importance of information resources in such condition. Allen (1999) has also cited the vital role of information resources in economic development. Financial reports as an information resource can reduce the asymmetry between market players and provide a reasonable basis for prediction and decision making process. Poor financial condition is not rare in market but inability in predicting it can result in resource wastage which deteriorates investor's confidence leaving them no choice but to abandoning the market. Albeit market performance is not predictable due to its random walk nature but we can learn from the past andform someexpectations relevant to it. Conducting a model for assessing financial condition of firms can surely foster economic development.

2. Literature Review

Three main categories in predicting corporate failure include statistical models, Artificial Intelligence Expert System Model (AIES) and theoretical models (Aziz & Dar, 2006). Statistical bankruptcy classification methods include Univariate and Multivariate analysis (Altman, 1986). Non-parametric statistical methods are designed based on machine learning techniques such as Artificial Neural Networks. Odom and Sharda (1990) are among the first researches who introduced ANN to the bankruptcy prediction realm, by investigating 128 firms with

various proportions of bankrupt and healthy firms in the training set, they achieved Type I correct classification accuracy in the range of 77.8% to 81.5% and type II accuracy in the range of 78.6% to 85.7% depending on the training set. Numerous studies regarding bankruptcy prediction have been executed; however the variables used in these surveys are somehow heterogeneous. In a classic study by Beaver (1966), the condition of 158 and 117 firms one to five years before the bankruptcy were assessed respectively. The variables were all extracted from financial statements; the results are depicted in table 1.

Table 1. Bankruptcy prediction result-Beaver 1966

Years before bankruptcy	1 year	2 years	3 years	4 years	5 years
Accuracy	%87	%79	%77	%76	%78

He argues cash and leverage ratios have the most predict ability. Altman (1968) and Ohlson (1980) models are technically linear models that classify healthy and bankrupt firms using financial ratios as inputs. Altman uses the classical multivariate discriminant analysis technique (MDA) which is based on applying the Bayes classification procedure; the following financial ratios were used as inputs:

1) working capital/total assets;

2) retained earnings/total assets;

3) earnings before interest and taxes/total assets;

4) market capitalization/total debt;

5) sales/total assets.

Shumway (1999) introduced the Hazard model which needs three market and two accounting ratios of firm to estimate related bankruptcy risk. By comparing prediction results of ANN and Linear Regression, Zhang et al. (1999) showed the superior performance of ANN with 88.2% accuracy against LR with 78.6%, he also executes five-fold cross-validation procedure to enhance the generalizability of the results. Chava and Jarrow (2004) compared the prediction results of Hazard, Altman and Ziminski models and concluded that Hazard model has more accurate predictions than logit and probit models. Bencivenga (1991) addresses the ability of capital markets in directing investment capital into manufacture and profitable sector which leads to economic development through diversification in investor's portfolio and reducing liquidity risk. Mulford and Comiskey (2005) have cited the importance of cash flow information regarding bankruptcy prediction. Lambert (2007) demonstrated that accounting information would improve the efficiency of investments via reducing adverse selection and liquidity risk. Capital is scarce in developing countries and investment risk is high, thus providing useful information can reduce the mentioned risk. Pritchard (2002) has found a significant correlation between profit calculated by accounting system and market gains across Baltic countries. Beuselinck (2005) acknowledges the vital role of accounting information in proper working of a dynamic economy. However the more demand regarding accounting information, the higher the quality of thisinformation (Ball & shivakumar, 2005). Chen (2010) has verified a positive relationship between the quality of accounting information and investment efficiency. Although Ball (2000) has found limited use of accounting information in business environments structured by institutional investors combined with low protection for individual investors.Undoubtedly investor's participation depends on safe environment for investment which would be feasible partly by rendering useful information. Financial reporting which is generated by accounting system has this characteristic and will aid investors in predicting financial condition of firms. The main reason of financial reporting is to produce useful information for decision makers; However chava and Jarrow (Ibid) claim accounting variables add little predictive power when market variables are included in bankruptcy prediction models. McKee and Greenstein (2000) developed a prediction method based on decision trees and applied it to a number of US firms one year ahead bankruptcy. Their method obtains better results than NN and multivariate discriminant analysis (MDA) for Type II error, but less accurate predictions for Type I error. Kerling and Poddig (1994) compared NN with MDA for a database of French firms for a three-year-ahead forecast. The NN achieved prediction accuracy in the range of 85.3–87.7% compared to 85.7% for MDA.

3. Research Hypothesis

Following the arguments made on the literature the question is related to usefulness of GA and accounting variables in failure prediction. Authenticating poor financial condition by accounting variables can facilitate efficient resource allocation and accelerate economic development. To this end the research hypothesis seeks tocomparethe means of two samplegroups, model predictions and reality, hence,

H1: There is no significant difference between model predictions and reality.

4. Research Method

The main purpose of this research is to assess whether proposed GA model along accounting variablesaresuitable in predicting financial condition of firms. In this study those companies which are not subject to article 141 of commercial codes of Iran are classified as healthy firms. This article states "had the cumulative losses of a firm reaches at least halve of the firms legal capital, the board is responsible for holding an extraordinary general meeting of shareholders, deciding between liquidation and survival". However due to non-existence of such trade law in other countries, researchers can perform their surveys based on other assumptions. Data are derived one year prior to financial assessment. The model used in this research is based on Genetic Algorithm where statistical constraints such as normality and independence of predicting variables are not subject to controversy. Statistical population in this survey contains listed firms in TSE between 2004 and 2010. There are totally 2958 firm/years in this period in which 74 firm/years were subject to article 141 of commercial codes of Iran. Based on the principals of sampling in limited population, the minimum required sample data is 37 firm/years. We randomly selected 70 firm/years, consisting of 35 bankrupted and 35 non-bankrupted firms/years.

Table 2. Number of bankrupt and non-bankrupt firms

Year	Firms	Bankrupt Firms	Non-bankrupt firms
2004	406	14	392
2005	416	18	398
2006	424	7	417
2007	426	6	420
2008	428	9	419
2009	429	13	416
2010	429	7	421
Firm/Year	2958	74	2884

4.1 Model and Variables Presentation

The model presented in this research belongs to Artificial Intelligence group in which we used Genetic Algorithm technique. GA along with simulation of natural evolution of environment within which an element called "gene" transfers physiological attributes which seeks optimization. The GA model provides an if-then rule; each rule is associated with a chromosome consisting of n genes. Each gene has three fields:

$$C_i(gene1(X1,Le,V1,), gene2(X2,Le2,V2,),\ldots\ldots, geneN(Xn,Len,Vn,))$$

Elements of each gene are:

X_i: variable used for prediction,

Le_i: logical equation of cut-off points for bankrupt firms,

V_i: cut-off point for each variable.

Logical equation is simply "<" or ">" related to cut-off point, V_i is the cut-off point determined by the rule which can be considered as a benchmark for each variable set. The fitness function below is set to maximize the accuracy. Assigning scores to model predictions is the same for true and false predictions. Carvalho (2002) used the function below to compute fitness,

$$Accuracy = (TP+TN) / N \text{ (size of population)}$$

"TP" represents the total number of non-bankrupt firms which model predicted correctly and "TN" represents the total number of bankrupt firms that model predicted correctly. "N" is the total number of predictions. The GA model is supposed to explore the cut-off point related to each variablein order to maximize the accuracy function, which in turn can be reached via maximizing the above numerator.

The variables in this research are extracted from financial statements one year prior to authenticating financial condition. Selecting accounting variables were based on conceptual framework of prior studies in bankruptcy field, first 18 of them were chosen and subsequently t-test for mean comparison between two groups was executed, then 10 variables with most significant difference were elected. Variables and t-test results are shown in table 3.

Table 3. Variables and t test results

row	variable	Mean in bankrupt firms	Mean in health firms	t-statistics	Significance of t-statistics
1	ROS	0.015	0.269	4.227	0.000
2	Operating income over total assets	0.019	0.185	6.146	0.000
3	retained earnings over total assets	-0.005	0.163	7.295	0.000
4	ROE	-0.834	0.627	2.124	0.037
5	ROA	-0.002	0.17	7.723	0.000
6	Current assets over current liabilities	0.938	1.117	2.248	0.028
7	cash over total liabilities	0.267	0.06	3.416	0.001
8	Current liabilities over total assets	0.728	0.576	-4.669	0.000
9	Total liabilities over total assets	0.821	0.596	-6.187	0.000
10	Total liabilities over owners equity	6.156	2.773	-3.986	0.000

4.2 Model Validation

In order to improve the generalizability of the GA results in determining model accuracy, cross validation is used Cross-Validation is a statistical method of evaluating and comparing learning algorithms by dividing data into two segments: one used to learn or train a model and the other used to validate the model. In typical cross-validation, the training and validation sets must cross-over in successive rounds so that each data point has a chance of being validated against. The basic form of cross-validation is k-fold cross-validation. Other forms of cross-validation are special cases of k-fold cross-validation or involve repeated rounds of k-fold cross-validation. Kohavi (1995) obtained good results with 10-fold cross-validation with empirical decision trees (C4.5). Also Kerling (1995) tested several cross-validation procedures for corporate distress diagnosis. Values of k as small as 5 or even 2 may work even several different random k-way splits of the data are analyzed which in turn reduces the variability of the cross-validation estimate. In this paper we have used 5 fold cross validation:

1. Dividing total population into 5 classifiers each containing 14 samples (7 bankrupt and 7 non bankrupt).

2. Introducing 4 classifiers to the algorithm as training set and the one left out is to be validation set. Due to random nature of extracting cut-off points by GA we repeat the procedure 30times for each step so at the end of this step we have 30 rules each having its own accuracy and type I and II errors.

3. Substituting one of the training sets with previous validation set. This procedure is continued until all classifiers are once used as validation set.

4. At the end we have 150 rules each having different accuracy. In this step we select and run 30 rules with the highest accuracy over the total population and introduce the mean of accuracy and type I and II errors as the final results of the GA algorithm.

5. Empirical Results

The model results which are cross validated 30 times are depicted in table 4. Vividly GA could predict the financial condition of firms one year ahead with 90.3% accuracy using accounting information. A low standard deviation as 1.7% represents high model stability. After determining the model accuracy it's necessary to compare the prediction results with realty in order to accept or reject research hypothesis, hence Mann-Whitney test is used to test the difference between averages of two non parametric groups. In another words it assesses whether observations in one sample tend to be larger than observations in the other sample (predictions and reality). A Mann-Whitney's U test is also known as Wilcoxon Rank sum test, and basically a non-parametric version of t test.

Had the statistical significance reaches above 5% the research hypothesis is accepted. The results are available in table 6. The results of spearman correlation test are also shown in table 5 which signals high correlation between predictions and reality.

Table 4. Model results

Average accuracy	Type I error	Type II error	Average accuracy Std.
% 91.42	% 5.71	% 11.42	% 1.70

The results of Mann-Whitney test show 0.53 significance, thus verifying acceptance of research hypothesis; subsequently accounting information is verified as a useful tool in bankruptcy prediction.

Table 5. Spearman correlation coefficient

Sample No.	Spearman correlation coefficient	Significance
70	0.83	0.000

Table 6. Mann-Whitney Test results

Z	Significance
- 0.675	0.53

6. Conclusion

The failure prediction research has suffered from lack of any unified theory since the 1930's when first empirical studies on this subject were published (Back et al., 1996). However accounting information is one the most rich information resources in economic decision making realm. Inabilities in discerning healthy from bankrupt firms have undesirable social and economical impacts. So designing and presenting a model in order to predict the financial condition of firms can accelerate economic development and improve efficient resource allocation in developing countries. The purpose of this study is to build a GA in order to assess the usefulness of accounting information in predicting financial condition of firms. However complementary information resources such as market information may generate desirable predictions as well. The samples are a mixture of different industries and variables used are obtained from firm's accounts. The model used in this survey is based on Genetic Algorithm and is set to maximize the predictions accuracy one year ahead. In order to improve the generalization of results 5 fold cross validation is executed while extracting 30 rules from each fold. The empirical results of GA model have achieved 90.3% accuracy along with 1.7% standard deviation. To test the research hypothesis non parametric Mann-Whitney test is used which shows 0.53 significance, thus null hypothesis is rejected. In addition achieving 0.8 correlations between model predictions and reality further verifies research results. This survey has some constraints in practice since determining a benchmark in order to segregate bankrupt firms form non-bankrupt ones is related to Article 141 of commercial codes of Iran. However the usefulness of accounting information in efficient resource allocation and accelerating economic development is a proven fact.

References

Allen, F. (1999). The Role of Information in Financial and Capital Markets.Wharton School. University of Pennsylvania. Prepared for Symposium on Financial Market Evolution Its Perspectives in Tokyo.

Altman, E. (1968). Financial ratios, discriminant analysis and the classification of corporate bankruptcy. *The Journal of Finance, 23*(4), 589-609. http://dx.doi.org/10.2307/2978933

Back, B., Laitinen, T., Sere, K., & Van Wezel, M. (1996). Choosing bankruptcy predictors using discriminant analysis, logit analysis, and genetic algorithms. *Turku Centre for Computer Science*, 40.

Ball, R., & Shivakumar, L. (2005). Earnings quality in U.K. private firms: comparative loss Recognition timeliness. *Journal of Accounting & Economics, 39*, 83-128. http://dx.doi.org/10.1016/j.jacceco.2004.04.001

Ball, R., Kothari, S. P., & Robin, A. (2000). The Effect of International Institutional Factors on Properties of Accounting Earnings. *Journal of Accounting and Economics, 29*, 1-51. http://dx.doi.org/10.1016/S0165-4101(00)00012-4

Bencivenga, V. R., & Smith, B. D. (1991). Financial Intermediation and Endogenous Growth. *Review of Economic Studies, 58*, 195-209. http://dx.doi.org/10.2307/2297964

Beuselinck, C. (2005). Essays on Financial Reporting Quality, Earning Management, and corporate Disclosure. University of Ghent, Ghent, Belgium.

Carvalho, D. R., & Freitas, A. A. (2004). A Hybrid Decision Tree Genetic Algorithm Method for Data Mining. *Information Sciences, 163*(1-3), 13-35. http://dx.doi.org/10.1016/j.ins.2003.03.013

Chava, S., & Jarrow, R. (2004). Bankruptcy prediction with industry effects. *Review of Finance, 8*, 537-569. http://dx.doi.org/10.1007/s10679-004-6279-6

Chen, F., Hope, O-K., Qingyuan, L., & Wang, X. (2011). Financial Reporting Quality and Investment Efficiency of Private Firms in Emerging Markets. *The Accounting Review, 86*, 1255-1288. http://dx.doi.org/10.2308/accr-10040

Fama, E. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance, 25*, 383-417. http://dx.doi.org/10.1111/j.1540-6261.1970.tb00518.x

Kerling, M., & Poddig, T. (1994). Klassifikation von Unternehmenmittels KNN. In Ökonomie, H. Rehkugler and H. G. Zimmermann (Eds.), *Neuronale Netze in der onomie*. München, Germany.

Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. *14th International Joint Conference on Artificial Intelligence (IJCAI-95)*, 1137-43.

Lambert, R., Leuz, C., & Verrecchia, R. E. (2007). Accounting Information, Disclosure, and the Cost of Capital. *Journal of Accounting Research, 45*, 385-420. http://dx.doi.org/10.1111/j.1475-679X.2007.00238.x

McKee, T. E., & Greenstein, M. (2000). Predicting bankruptcy using recursivepartitioning and a realistically proportioned data set. *J. Forecasting, 19*, 219-230. http://dx.doi.org/10.1002/(SICI)1099-131X(200004)19:3<219::AID-FOR752>3.3.CO;2-A

MokhatabRafiei, F., Manzari, S. M., & Bostanian, S. (2011). Financial health prediction models using artificial neural networks, genetic algorithm and multivariate discriminant analysis: Iranian evidence. *Expert System withApplication, 38*(8), 10210-10217.

Mulford, C. W., & Comiskey, E. E. (2005). Creative cash flow reporting: Uncovering sustainable financial performance. Hoboken, NJ: *John Wiley and Sons.*

Ohlson, J. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research, 18*, 109-131. http://dx.doi.org/10.2307/2490395

Pritchard, N. J. (2002). The Relationship between Accounting Numbers and Returns in the Baltic Stock Markets. *CERT Discussion Papers 2002/06*, Centre for Economic Reform and Transformation, Heriot Watt University.

Scott, W. R. (2009). *Financial Accounting Theory* (6th ed.). Prentice Hall.

Sen, A. (1983). Development: Which Way Now? *Economic Journal, 93*, 745-762. http://dx.doi.org/10.2307/2232744

Shumway, T. (1999). Forecasting bankruptcy more accurately: A Simple Hazard Model. *Journal of Business, 74*, 101-124. http://dx.doi.org/10.1086/209665

Zhang, G., Hu, M., & Patuwo, B. (1999). Artificial neural networks inbankruptcy prediction: General framework and cross-validation analysis. *European J. Oper. Res., 116*, 16-32. http://dx.doi.org/10.1016/S0377-2217(98)00051-4