Data Mining Approach Using Practical Swarm Optimization (PSO) to Predicting Going Concern: Evidence from Iranian Companies

Mahdi Salehi*, Fezeh Zahedi Fard**

Abstract

Purpose - Going concern is one of fundamental concepts in accounting and auditing and sometimes the assessment of a company's going concern status that is a tough process. Various going concern prediction models' based on statistical and data mining methods help auditors and stakeholders suggested in the previous literature.

Research design - This paper employs a data mining approach to prediction of going concern status of Iranian firms listed in Tehran Stock Exchange using Particle Swarm Optimization. To reach this goal, at the first step, we used the stepwise discriminant analysis it is selected the final variables from among of 42 variables and in the second stage; we applied a grid-search technique using 10-fold cross-validation to find out the optimal model.

Results - The empirical tests show that the particle swarm optimization (PSO) model reached 99.92% and 99.28% accuracy rates for training and holdout data.

Conclusions – The authors conclude that PSO model is applicable for prediction going concern of Iranian listed companies.

Keywords : Data Mining, Going Concern Prediction, Particle Swarm Optimization, Financial Ratios.

JEL Classifications : G11, G33, M41.

1. Introduction

Going concern is one of the fundamental concepts of accounting and auditing. Statement on Auditing Standards (SAS) No.59 requires that on every audit the auditor should evaluate whether considerable doubt exists about the entity ability to continue as a going concern. In particular, the auditor has to evaluate the client's going concern status for a reasonable period, not to exceed one year further than the date of the financial statements was being audited. Relevant information about the continuation of entity's going concern is generally obtained from the application of auditing procedures that are planned and performed in order to attain audit objectives. Examples of conditions and events that cast doubt ability of an entity for survival include negative financial trends, defaults on loans or similar agreements, and non-financial internal and external issues such as work stoppages or substantial dependence on the success of a specific project. When the identified conditions and events taken together lead to substantial doubt about going concern status of an entity, the auditor should detect and assess management's plans to mitigate the effects of adverse conditions or events. If the auditor is convinced that: management's plans have ability to overcome a substantial doubt, a going concern audit report is not required. However, we explore the performance of going concern status using PSO for one year before going concern (t-1) and address a rule-based comprehension if the auditor decides that substantial doubt exists, he should be modified audit report by adding adescriptive paragraph accompanied by the opinion paragraph.

Although the evaluation of an entity's viability is not the main objective of an audit, bankruptcies without a prior going concern report are often observed by the public as audit reporting failures (Geiger & Raghunandan, 2002).

The high frequency of this kind of audit reporting failures is suggestive ofthis fact that the auditor's going concern decision is very complicated and requires a high level of evident judgment. The complexity of the going concern decision has led the development of several models for predicting going concern opinion. These studies is concentrated on development of going concern prediction models, by using a multiple financial and non-financial variables that might be indicative of going concern decision for auditor (Martens et al., 2008). Early studies of going concern predictionwere developed by using statistical techniques such as multiple discriminant analysis, Logit, probit etc. These methods by using historical samples created diagnostic model, in spite of the fact that they cannot inductively learn from new data dynamically, which greatly affects accuracy rate (Sun & Li, 2008). In recent years, data mining, a novel field of intelligent data analysis established, developed ,and began to appear and grow promptly in the background of abundant data and poor information. It also has developed a new approach for the deep research in finance. Based on this principal by using great database or data warehouse which stores a large number of listed companies financial data's, by utilizing data mining technique extract valuable unknown knowledge dynamically, which can be applied to predict going

* Corresponding author, Assistant Professor of Accounting, Ferdowsi University of Mashhad, Mashhad, Iran. Tel. +98-91-2142-5323, E-mail: mehdi.salehi@um.ac.ir

** Department of Accounting and Management, Neyshabur Branch, Islamic Azad University, Neyshabur, Iran. Email: fzahedi_1362@yahoo.com

concern status of companies.

Classification models of going concern that allows easy consultation by auditors and other stakeholders to evaluate viability of companies. Furthermore, in today's litigious economic atmosphere, the number and the magnitude of non-going concern firms filings are increasing considerably. Even skilled auditors, who have a convincing level of concerning firms' situations, may fail to make an accurate judgment on firms going concern status (McKee, 2003). Besides, they can use these models in the final stages of the audit engagement under similar circumstances, as a quality control device or as a benchmark can represent auditor judgment.

2. Literature review

It appears the start of going concern studies in the research literature coincides with the issuance of standards addressing going concern. The first study about going concern prediction published shortly after the issuance of SAS No.2 in 1974 by McKee (1976). Several researches published around the issuance of SAS No.34 in 1981. In addition, just before and within a few years after the issuance of SAS No. 59 in 1988 several studies published. However, SAS 59 included the relevant guidelines criticized because of deeply subjective, general and ambiguous (Koh & Killough, 1988). It is worth to be mentioned, even by creating Public Company Accounting Oversight Board by Sox to oversee audits of public firms, has not issued guideline for evaluation of going concern and to this day, SAS No.59 is authoritative guidance available for investigation of going concern status of an entity (Bellovary et al., 2007). In view of the evaluation of going concern is not possible simply, numerous studies conducted that divided into two categories: statistical methods and data mining techniques.

2.1. Statistical methods

The first studies of going concern used Multivariate discriminant analysis (MDA) to develop models (MCKee, 1976). Further studies on going concern decision making used probit analysis (Dopuch et al., 1987; Koh, 1987; Koh & Brown, 1991). Subsequent research applied logit analysis to test going concern predicting models (e.g. Menon & Schwartz, 1987; Harris, 1989; Bell & Tabor, 1991 McKeown et al., 1991). Among these models, only the model of Menon and Schwartz (1987) has reached to the top accuracy rate of 100%.

2.2. Data mining techniques

Since the early 80s, data mining techniques have been successfully applied to going concern prediction. Data mining respectively include five steps: sample, explore, modify, model and assess (i.e. SEMMA) (SAS Institute, 1998). In 1992, Hansen et al., (1992) was applied neural networks (NN) and Inductive Dichotomizer 3 (ID3) for prediction of going concern statue of firms. Since that time so far, various models based on data mining have been proposed and applied to predicting firms financial status (e.g. Rafiei et al., 2011; Chen, 2012; Li & Miu, 2010; Hsieh et al., 2012; Sun & Li, 2011; Harada & Kageyama, 2011; Xiao et al., 2012; Sun et al., 2011; Tsai & Cheng, 2012). The results of these studies indicate that data mining methodology is an efficient research method. We focus using PSO to predict going concern in this paper for the following reasons:

- It has uncomplicated concepts and sustainable convergence.
- To implement PSO is easy and it includes few parameters for adjusting.
- It requires low memory and CPU resources and comparison with mathematical algorithm and other heuristic optimization techniques has higher computational efficiency (Nalini & Balasubramanie, 2009).

3. Research design

3.1. Data collection

The data set used for this study consists of 146 Iranian manufacturing companies. All of these companies were or still listed in the Tehran Stock Exchange (TSE). 73 companies went bankrupt under paragraph 141 of Iran Trade Law[1]) from 2001 to 2011.The number of going concern opinion companies is equal to number of going concern opinion companies and we used "matched" companies (e.g. Min & Lee, 2005; Etemadi et al., 2009; Andrés et al, 2012; Olson et al., 2012; Chaudhuri & De, 2011; Chen, 2011 Brezigar et al., 2012). Due to small population, we could not match completely two groups in each of industries. Also size of the firms as a potential explanatory variable considered in variable selecting steps.

3.2. Feature selection

Feature selection boosts the prediction performance of the predictors and provides faster and more cost-effective predictors, and prepares a better understanding of the underlying process that is stemmed from the data. In addition, reducing the number of irrelevant or redundant features reduces the running time of a learning algorithm. There are many potential advantages of feature selection such as facilitating data visualization and understandable data, reducing the measurement and storage requirements such as: reducing times of training and utilization (Guyon & Elisseeff, 2003; Ashoori & Mohammadi, 2011). Accordingly, the variables selected in this study are based on a combination all variable selection techniques and experiments. More over there is no reliable published relevant data about cash flow statement before 2008 and that is why our candidate financial ratios do not include indicators directly related with cash flows.

We apply a three stages process for future selection. At the first

1) Under paragraph 141 of Iran Trade Law, a firm is bankrupt when its total value of retained earnings is equal or more than 50% of its listed capital.

<Table 1> Variables used in the research and comparison of means in two groups

	Variables	Means of Group 1		Means of Group 0		Sig level			Variables	Means of Group 1		Means of Group 0		Sig level	
		t	t-1	t	t-1	t	t-1			T	t-1	t	t-1	t	t-1
1	EBIT/TA*	0.16	0.18	-0.07	0.05	0.00	0.00	2	LTD/SE	0.25	0.20	-0.39	0.56	0.12	0.06
3	RE/SC*	0.62	0.65	-1.39	0.02	0.00	0.00	4	MVE/TL	1.16	1.40	0.47	0.66	0.00	0.00
5	MVE/SE	2.10	2.42	2.15	2.57	0.78	0.22	6	MVE/TA	0.67	0.77	0.43	0.48	0.00	0.00
7	Ca/TA	0.05	0.05	0.02	0.03	0.00	0.00	8	Size(logTA)	5.33	5.25	5.23	5.23	0.30	0.83
9	TL/TA**	0.67	0.67	0.97	0.80	0.00	0.00	10	CL/SE*	2.28	2.27	0.76	4.76	0.04	0.00
11	CL/TL	0.85	0.86	0.86	0.85	0.31	0.94	12	(Ca+STI)/CL	0.12	0.11	0.04	0.05	0.00	0.00
13	(R+Inv)/TA	0.55	0.57	0.55	0.57	0.99	0.88	14	R/S	0.55	0.53	0.33	0.40	0.46	0.10
15	R/Inv	1.64	1.18	1.00	1.00	0.22	0.93	16	SE/TL	0.62	0.63	0.06	0.32	0.00	0.00
17	SE/TA	0.35	0.35	0.04	0.22	0.00	0.00	18	CA/CL	1.31	1.31	0.82	1.07	0.00	0.00
19	QA/CL	0.76	0.70	0.43	0.57	0.00	0.00	20	QA/TA	0.40	0.37	0.35	0.36	0.07	0.73
21	FA/(SE+LTD)	0.58	0.60	1.88	0.91	0.00	0.01	22	FA/TA	0.23	0.22	0.25	0.24	0.44	0.63
23	CA/TA	0.70	0.70	0.67	0.68	0.39	0.66	24	Ca/CL	0.10	0.09	0.03	0.04	0.00	0.00
25	IE/GP	-0.01	-0.02	1.86	-1.21	0.70	0.48	26	S/Ca	28.98	35.30	57.50	44.80	0.00	0.11
27	S/TA	0.87	0.93	0.72	0.70	0.02	0.00	28	WC/TA	0.13	0.13	-0.16	0.00	0.00	0.00
29	PIC/SE	0.53	0.53	2.74	0.86	0.03	0.00	30	S/WC	4.05	2.87	1.85	1.73	0.00	0.96
31	RE/TA*'**	0.08	0.08	-0.22	-0.03	0.00	0.00	32	NI/SE	0.36	0.42	-1.28	-0.03	0.00	0.00
33	NI/S	0.19	0.16	-0.35	-0.02	0.00	0.00	34	NI/TA*'**	0.12	0.13	-0.13	0.00	0.0	0.00
35	S/CA	1.25	1.34	1.13	1.07	0.27	0.00	36	OI/S**	0.19	0.20	-0.39	0.06	0.00	0.00
37	OI/TA	0.14	0.17	-0.07	0.03	0.00	0.00	38	EBIT/IE	-1.84	-5.21	0.63	-0.45	0.00	0.05
39	EBIT/S	0.24	0.52	-0.37	0.10	0.00	0.00	40	GP/S*	0.27	0.27	0.11	0.15	0.00	0.00
41	S/SE	3.13	3.32	4.07	4.68	0.57	0.05	42	S/FA	5.95	6.29	5.42	6.44	0.09	0.33

stage, the 42 variables used in this study as shown in Table 1, were selected after reviewing the finance and accounting literature dealing with financial status prediction models in Iran (in order to be consistent with economic conditions in Iran) and finally the complete set of variables that were used by Etemadi et al. (2008) are selected and captured the aforementioned characteristics of financial failure. In the second stage, we applied T-test at a significant level of 0.05 and according to this experiment; variables that potentially had the ability of predicting financial status in the model were selected. In the third stage, stepwise discriminant analysis (SDA) selected final variables. We chose SDA method because it is a dominant method in researches conducted in the accounting field (e.g. Alici, 1996; Altman, 1993; Chen and Shemerda, 1981).

SDA is a prevalent procedure in order to reduce dimensions of a problem and choosing the most significant variables from among of extensive set of variables. SDA select variable on based on a cut-off F value for pre-specified statistical significance level until no significant variables remained.

With significant level set at the 0.05 level, the discriminant stepwise procedure selected 6 and 4variables respectively for t and t-1 from the 42 candidate variables for the models which could best differentiate the going concern firms from the non-going concern firms.

Summary of SDA process is shown in Table 3. During each of these stages, a financial ratio is selected and added to collection of chosen financial ratios. Reduce Wilks' Lambda at each stage meant to boost the diagnostic power of selected variables and Table 2 shows the results. These selected financial ratios for t are: Earnings before interest & taxes to total assets (χ_1), Retained earnings to Stock capital (χ_3), Retained earnings to total assets (χ_{31}), Gross profit to

Sales (χ_{40}), Current liabilities to Shareholders' equity (χ_{10}) and Net income to total assets (χ_{34})and for t-1 are: total liabilities to total assets (χ_9) , Retained earnings to total assets (χ_{31}), Operational income to sales (χ_{36}) and Net income to total assets (χ_{34}).

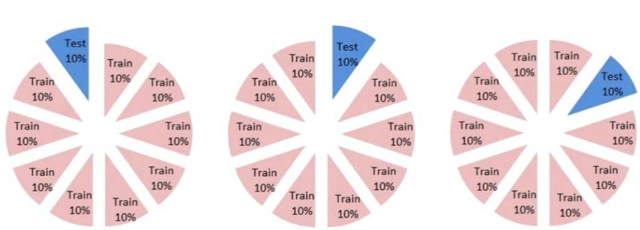
<Table 2> Selected variables in SDA Analysis for t-1

Step	Tolerance	F to Remove	Wilks' Lambda
1. Net income to total assets	1.000	100.772	
2. Net income to total assets	0.938	56.243	0.748
Total liabilities to total assets	0.938	9.068	0.550
3. Net income to total assets	0.513	8.617	0.522
Total liabilities to total assets	0.912	11.103	0.532
Operational income to sales	0.546	6.114	0.512
4. Net income to total assets	0.478	4.749	0.489
Total liabilities to total assets	0.896	8.546	0.503
Operational income to sales	0.539	4.586	0.488
Retained earnings to total assets	0.770	4.369	0.487

3.3. PSO model development

3.3.1. Cross-Validation

The cross-validation is the standard data mining methodology used to evaluate and compare learning algorithms by splitting the data into two main subdivisions: a training set and test set. Quality of the prediction evaluated on the test set. K-fold cross validation is the primal form of cross-validation. In K-fold cross-validation the data is firstly partitioned into K subsets of approximately or exactly the same size. Then, K iterations of training and test are done such that in each iteration a variant fold of the data is held-out for validation while the rest $K-1$ folds are used for learning and K outputs from the folds can be averaged and can produce a single estimation (Figure No.1). The advantage of K-fold cross-validation is that all observation are utilized for both training and test sets (Alpaydin, 2010).In data mining and machine learning K is typically 10 or 30 that in this study $K=10$.



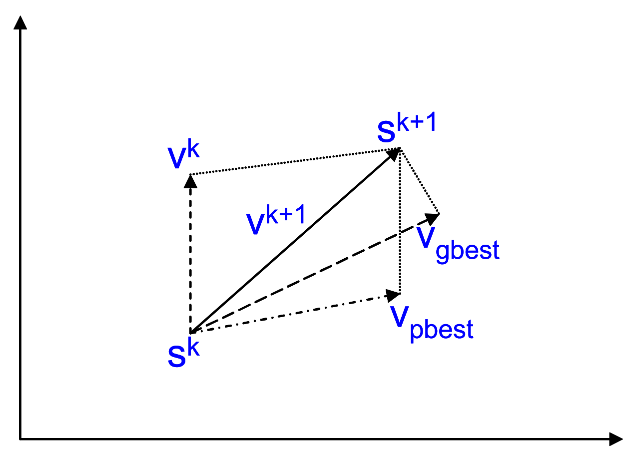
<Figure 1> Three stage of 10-fold cross-validation

3.3.2. Particle Swarm Optimization (PSO)

PSO is a robust stochastic optimization computational technique based on the movementand intelligence of swarms that do not possess any leader in that complement or swarm. It employs the concept of social interaction to problem solving. It was developed and conducted by Kennedy (social psychologist) and Eberhart (electrical engineer) in 1995 pivoting on the social behaviors of birds flocking or fish schooling. It applies a number of agents (particles) that causes a swarm movement around in the search space, seeking for the best alternative or solution (Rini et al., 2011).

Each particle is treated as a point in an N-dimensional space which adjusts its "flying"according to its own flying experience as well as the flying experience of other particles. Each particle keeps track of its coordinates in the solution space which are associated with the best solution (fitness) that has achieved so far by that particle. This value is called personal best, $pbest$. Another best value that is tracked by the PSO is the best value obtained so far by any particle in the neighborhood of that particle. This value is called $qbest$.

The basic concept of PSO lies in accelerating each particle toward it's $pbest$ and the $qbest$ locations, with a random weighted acceleration at each time step as shown in Figure No. 2.



S^k : current searching point
S^{k+1} : modified searching point
V^k : current welocity, Vk^{+1}
modifiedvelocity, V_{pbest}
Velocity based on pbest,
V_{qbest} : Velocity based on qbest

<Figure 2> Concept of modification of a searching point by PSO

Each particle tries to modify its position using the following information: the current positions, the current velocities, the distance between the current position and $pbest$, the distance between the current position and the$pbest$.

The modification of the particle's position can be modeled mathematically according the following equation:

$$V_i^{k+1} = WV_i^k + C_1\, rand_1(...) \times (pbest_i - S_i^k + C_2\, rand_2\,(...) \times (gbest - S_i^k) \;......\; (1)$$

Where,

V_i^k: $velocity\ of\ agent i at eration\ K,$
W: $weighting\ function,$
C_J : $weighting factor,\ rand$
$uniformly\, distributed\, random\ \nu mber\ between$ 0 and 1
S_i^k: $current\ position\ of\ agent\ i\ at\ eration\ K,$
$pbest_i$: $pbest\ of\ agent i,$
$qbest$: $qbest\ of\ the\ qroup$

The following weighting function is usually utilized in (1)

$$W = W_{\max} - [(W_{\max} - W_{\min} \times er]/ Max_{er} \qquad (2)$$

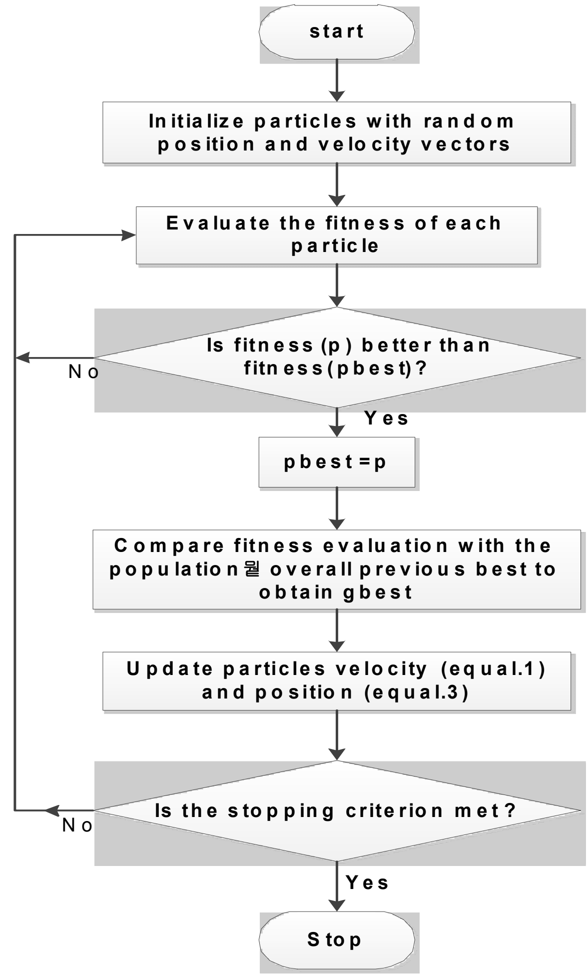
Where : $W_{Max} = initial weight,$
$W_{Min} = final\ weight,$
$Max_{er} = masimum\, eration\ \nu mber,$
$er = current eration\ \nu mber,$

$$S_i^{k+1} = S_i^k + V_i^{k+1} \qquad (3)$$

A large inertia weight (W) facilitates a global search while a small inertia weight facilitates a local search (see Table 3). By linearly reducing the inertia weight from a virtually large value to a small value through the course of the PSO run gives the best PSO performance compared with fixed inertia weight settings. Larger W has greater global search ability and smaller Whas greater local search ability. In Figure 3 we show a flow chart depicting of general PSO algorithm.

<Table 3> Weighting function of PSO

Fold	W1	W2	W3	W4
1	0.67	-1.00	-0.12	-1.00
2	0.63	-1.00	-0.33	-0.75
3	1.00	-0.79	-0.42	-1.00
4	0.61	-1.00	-0.25	-0.68
5	0.43	-0.79	-0.23	-0.92
6	0.50	-1.00	-0.24	-1.00
7	0.77	-1.00	-0.03	-1.00
8	0.33	-0.81	-0.09	-0.70
9	0.80	-1.00	-0.06	-0.11
10	0.83	-1.00	-0.25	-1.00
Average	0.66	-0.94	-0.20	-0.82



<Figure 3> General PSO Algorithm Flow chart depicting

4. Experimental results

The proposed PSO model is implemented using MATLAB 7.6. They are results on the 10 testing data sets. See Table 3. Table 4 shows obtained result from PSO model. This model could classify firms with 99.92% and 99.29% overall accuracy rate in the training and testing sample, respectively as shown as in Table 4 and it can seen the result of error type 1and 2 for each set of data in Table 5. In addition, this model could correct classify for going concern firms with 98% accuracy rate and 100% for non-going concern firms.

<Table 4> Predictive accuracies (%)

Hold-out data	Training data	Fold
100.00	100.00	1
100.00	100.00	2
100.00	99.23	3
100.00	100.00	4

100.00	100.00	5
100.00	100.00	6
92.86	100.00	7
100.00	100.00	8
100.00	100.00	9
100.00	100.00	10
92.86	99.23	Min
100	100	Max
100	100	Median
5.10	0.06	Variance
99.29	99.92	Mean

<Table 5> Errors for training and hold-out data (%)

Error type 2		Error type 1		Fold
Hold-out	Training	Hold-out	Training	
0	0	0	0	1
0	0	0	0	2
0	0	0	1.54	3
0	0	0	0	4
0	0	0	0	5
0	0	0	0	6
0	0	20	0	7
0	0	0	0	8
0	0	0	0	9
0	0	0	0	10
0	0	2	0.154	Average

5. Conclusion

In this paper, we considered a set of features that include 42 variables proposed in prior literature dealing with financial status prediction models in Iran. We applied SDA to identify potential variables for predicting model and finally, we selected 6 and 4 financial ratios in t and $t-1$. We constructed PSO prediction model based on selected features for t-1. Based on the results, the empirical tests show that PSO model achieved 99.92% and 99.28% accuracy rates for training and hold-out data. In summary, obtained results from PSO model from 146 companies of Iran indicate that this model has suitable ability going concern prediction status of firms.

Received: December 19, 2012.
Revised: February 16, 2013.
Accepted: March 18, 2013.

References

Alici, Y. (1996), "Neural network in corporate failure prediction: the UK experience", Proceedings of the Third International Conference on Neural Networks in the Capital Markets,London, October 1995. World Scientific Publishing Co Pte, 11-13.

Altman, E.I. (1993), *Corporate Financial Distress and Bankruptcy: A Complete Guide to Predicting & Avoiding Distress and Profiting from Bankruptcy,* second Edition, New York: John Wiley and Sons.

Alpaydin, E. (2010), *Introduction to machine learning,* 2nd ed., Cambridge, Mass: MIT Press.

Andrés, J., Landajo, M. & Lorca, P. (2012), "Bankruptcy prediction models based on multinorm analysis: An alternative to accounting ratios", *Knowledge-Based Systems,* 30(6), 48-56.

Ashoori, S. & Mohammadi, S. (2011), "Compare failure prediction models based on feature selection technique: empirical case from Iran", *Procedia Computer Science,* 3, 568–573.

Bell, T. & Tabor, R. (1991), "Empirical analysis of audit uncertainty qualifications" , *Journal of Accounting Research,* 29(2), 350-370.

Brezigar-Masten, A. & Masten, I.(2012), "CART-based selection of bankruptcy predictors for the logit model", *Expert Systems with Applications,* 30, 67-77.

Bellovary, J., Giacomino, D. & Akers, M. (2007), "A Review of going concern prediction studies: 1976 To Present", *Journal of Business & Economics Research,* 5, 8-28.

Brezigar-Masten, A. & Masten, I. (2012), "CART-based selection of bankruptcy predictors for the logit model", *Expert Systems with Applications,* 39(11), 10120-10139.

Chaudhuri, A. & De, K. (2011), "Fuzzy Support Vector Machine for bankruptcy prediction", *Applied Soft Computing,* 11, 2472–2486.

Chen, J-H. (2012), "Developing SFNN models to predict financial distress of construction companies", *Expert Systems with Applications,* 39, 823–827.

Chen, K.H. & Shemirda, T.A. (1981), "An empirical analysis of useful financial ratios" , *Journal of Financial Management,* 10(1), 51-60.

Chen, M-Y. (2011), "Predicting corporate financial distress based on integration of decision tree classification and logistic regression", *Expert Systems with Applications,* 38, 11261–11272.

Dopuch, N., Holthausen, R. & Leftwich, R. (1987), "Predicting audit qualifications with financial and market variables", *The Accounting Review,* 63(3), 431-453.

Etemadi, H., Anvary-Rostamy, A.A. & Farajzadeh-Dehkordi, H. (2009), "A genetic programming model for bankruptcy prediction: Empirical evidence from Iran", *Expert Systems with Applications,* 36, 3199–3207.

Geiger, M.A. & Raghunandan, K. (2002), "Auditor tenure and audit reporting failures", *Auditing: A Journal of Practice and Theory,* 21 (1), 67–80.

Guyon, I. & Elisseeff, A. (2003), "An introduction to variable and feature selection", *Journal of Machine Learning Research,* 3, 1157–1182.

Hansen, J., Mc Donald, J. & Stice, J. (1992), "Artificial intelligence and generalized qualitative-response models: An empirical test on two audit decision-making domains", *Decision Sciences* 23(3), 708-723.

Harada, N. & Kageyama, N. (2011), "Bankruptcy dynamics in Japan", *Japan and the World Economy,* 23, 119–128.

Harris, C. (1989), "An expert decision support system for auditor going concern evaluation",Ph.D. dissertation, University of Texas-Arlington.

Hsieh, T-J., Hsiao, H-F. & Yeh, W-Ch. (2012), "Mining financial distress trend data using penalty guided support vector machines based on hybrid of particle swarm optimization and artificial bee colony algorithm", *Neurocomputing,* 82, 196–206.

Koh, H. & Brown,R. (1991), "Probit prediction of going and non-going concerns" *Managerial Auditing Journal,* 6(3), 18-23.

Koh, H. (1987), "Prediction of going-concern status: A probit model for the auditors",Ph.D. dissertation, Virginia Polytechnic Institute and State University.

Leon Li, M-Y. & Miu, P. (2010), "A hybrid bankruptcy prediction model with dynamic loadings on accounting-ratio-based and market-based information: A binary quantile regression approach", *Journal of Empirical Finance,* 17, 818– 833.

Li, H., Sun, J. & Wu, J. (2010), "Predicting business failure using classification and regression tree: An empirical comparison with popular classical statistical methods and top classification mining methods", *Expert Systems with Applications,* 37, 5895 –5904.

Martens, D., Bruyneseels, L., Baesens, B., Willekens, M. & Vanthienen, J. (2008), "Predicting going concern opinion with data mining", *Decision Support Systems,* 45, 765–777.

McKee, T. (1976), "Discriminant prediction of going concern status: A model for auditors",Selected Papers of the AAA Annual Meeting.

McKee, T.E. (2003), "Rough sets bankruptcy prediction models versus auditor signaling rates", *Journal of Forecasting,* 22, 569–586.

McKeown, J., Mutchler, J. & Hopwood, W. (1991), " Towards an explanation of auditor failure to modify the audit opinions of bankrupt companies", *Auditing: A Journal of Practice & Theory,* 10, 1-13.

Menon, K. & Schwartz, K. (1987), "An empirical investigation of audit qualification decisions in the presence of going concern uncertainties", *Contemporary Accounting Research,* 3(2), 302-315.

Min, J. H. & Jeong, C. (2009), "A Binary classification method for bankruptcy prediction", *Expert Systems with Applications,* 36, 5256–5263.

Min, J.H. & Lee, Y.C. (2005), "Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters", *Expert Systems with Applications,* 28, 603–614.

Nalini, C. & Balasubramanie, B.P. (2009), "Performance Analysis of Cooperative PSO Algorithm with ACO and Tabu", *Search, International Journal of Computational Intelligence,* 5(2), 127-150.

Olson, D. L., Delen, D. & Meng, Y. (2012), "Comparative analysis of data mining methods for bankruptcy prediction", *Decision Support Systems,* 52, 464 – 473.

Rini, D.P., Shamsuddin, S.M. &Yuhaniz, S.S. (2011), "Particle Swarm Optimization: Technique" *System and Challenges, International Journal of Computer Applications,* 14(1), 19-27.

SAS Institute (1998), "From Data to Business Advantages: Data Mining", The SEMMA Methodology and SAS Software, SAS Institute, North Carolina.

Sun, J., He, K-Y. & Li, H. (2011), "SFFS-PC-NN optimized by genetic algorithm for dynamic prediction of financial distress with longitudinal data streams", *Knowledge-Based Systems,* 24, 1013–1023.

Sun, J. & Li, H. (2008), "Data mining method for listed companies' financial distress prediction", *Knowledge-Based Systems,* 21, 1-5.

Sun, J. & Li, H. (2011), "Dynamic financial distress prediction using instance selection for the disposal of concept drift", *Expert Systems with Applications,* 38, 2566–2576.

Tsai, C-F. & Cheng, K-C. (2012), "Simple instance selection for bankruptcy prediction" *Knowledge-Based Systems,* 27, 333–342.

Xiao, Z., Yang, X., Pang, Y. & Dang, X. (2012), "The prediction for listed companies'financial distress by using multiple prediction methods with rough set and Dempster-Shafer evidence theory", *Knowledge-Based Systems,* 26, 196–206.