COMPANY FINANCIAL PATH ANALYSIS USING FUZZY C-MEANS AND ITS APPLICATION IN FINANCIAL FAILURE PREDICTION

Jiaming LIU, Chong WU*

Harbin Institute of Technology, School of Management, Harbin, Hei Longjiang, China

Abstract. This study investigates the dynamic process of the company's financial status over years and proposes to design financial path using fuzzy c-means (FCM) approach. FCM is firstly used to quantize company financial paths of behavior over consecutive years. Financial paths are deployed to depict different patterns of failure and to understand the dynamics of financial failure. Then financial failure prediction model is built based on the proposed financial path (FP) approach. Empirical experiment is carried out with data samples of Chinese listed companies. Through analyzing financial path, it is found that there are mainly four patterns of process to terminal failure. Besides, in order to validate the prediction performance of the financial path prediction model, four prevalent financial failure prediction models, logistic regression (LR), support vector machine (SVM), decision tree (DT), and neural network (NN), are deployed to compare with the proposed model respectively. Experimental results show that FP has significantly better financial failure performance than other four models in terms of accuracy, type I error, and type II error. Therefore, the financial path monitors the change of financial status and acts as a prediction tool for financial failure prediction, which is a significant determinant of the financial success. Managers can recognize the financial failure signal in advance and understand their evolution trend on the future. In addition, the financial path prediction model is also an effective supplement to the research field of financial failure analysis and prediction.

Keywords: financial failure, financial dynamics, financial path analysis, fuzzy c-means, forecasting, prediction model.

JEL Classification: C53, G11.

Introduction

Financial failure prediction (FFP) has long been a well studied topic in both academic world and industrial world. The increase in financial failure leads to economic decline and gives rise to a lot of social problems. It becomes crucial important to discover the potential financial failure behaviors of firms and investigate the implicit patterns from the perspective of early

---

*Corresponding author. E-mail: *wuchong1971@126.com*

warning. There have a myriad of research on financial failure prediction by using different kinds of methods, such as univariate discriminant analysis (Beaver 1966), multivariate discriminant analysis (Altman 1968), logistic regression (Ohlson 1980), neural network (Odom, Sharda 1990), decision tree (Frydman *et al.* 1985), and support vector machine (Min, Lee 2005). Due to the financial failure prediction models are designed to predict a firm's ability to cover its financial obligations, so the prediction ability of the model is a great concern for the stakeholders, such as investor, creditor, financial institution, and company itself.

Because no general theory or theoretical model is available to describe the evolution of a company towards terminal failure, so failure prediction models were usually designed empirically by means of quantative methods. Statistical models were firstly used for financial failure prediction, but data mining techniques were the most frequently used methods to build prediction models. The companies' financial profiles are mainly reflected by the financial statement, in addition, empirical research has proved that the companies' future status can also be predicted according to their historical behaviors. So financial data, mainly annual accounting data, was used as the training sample to construct prediction models. The constructed prediction model is used to predict whether a firm will "fail" within a given period. To assess the prediction ability, models usually calculate the boundary between financial healthy companies and financial failure companies based on the characteristics of financial data. The testing sample is compared with the boundary in order to obtain the prediction result (the company will/will not fail within certain years). This particular way of forecasting is commonly used by most data mining techniques, a prediction relies on the estimation of a value that measures a distance to financial failure.

However, through summarizing previous studies, most of the financial failure prediction models are constructed on a static snapshot of financial situation. That is, traditional models are mono-period models which rely on a snapshot of a firm's financial situation measured at year t to predict if it will fall into failure at year $t+1$, $t+2$, or $t+3$. These models do not take a time dimension into account when modelling. However, research has shown that financial failure is a dynamic process (Laitinen 1991). A firm's history performance or financial behavior trail has critical influence on its future behavior (Balcaen, Ooghe 2006). Hence, the financial status of a company measured at a given time depends heavily on its history performance (Hambrick, D'Aveni 1988). Actually, financial failure is often the result of a sequence of cascading events rather than a sudden event. So, models that able to assess the behavior of a firm over time based on the firm's historical financial data should be developed. From the firm's point of view, mangers would expect that they could take corrective actions to prevent financial failure according to the behaviors of the company. However, financial failure process which characterizes the dynamic changes of financial situation receives little attention, only a few works attempted to analyze the temporal sequence of financial situations.

To the best of our knowledge, Argenti (1976) pioneered to use the concept of financial paths to explain the behaviors of firms before going bankrupt. He found that there are three categories of firms and they failed in different ways. Miller and Friesen (1977) investigated the relationship between financial variables and archetypal that best describe healthy firms and failure firms. They found ten different archetypes using factor analysis, six of them describing sound firms, the rest of them describing unsound firms. Firms within the same

archetypes have some common characteristics and behave in the same manner. These archetypes are likely to share a similar changes of dynamics. That is, some quantization methods, especially clustering methods can be used to group firms based on their financial characteristics, then help to understand the dynamics of firms behavior. Therefore, financial situations can be represented by a specific firm samples which share the common behaviors, and these behaviors are reflected through financial indicators using clustering method. Following these studies, Martín-del-Brío and Serrano-Cinca (1993) have designed models using Kohonen map to analyze patterns among healthy and failure firms that traditional dichotomous models are unable to assess. Neophytou and Molinero (2004) have used multidimensional scaling approach to assess the behaviors of firms over time. Jardin and Séverin (2011) have used a Kohonen map to design corporate bankruptcy prediction model through analyzing financial trajectories.

Based on the above analysis, there is no specific method that could be used to describe and analyze the dynamics of financial failure. It is for this reason that we study this issue –financial path analysis –from a quantitative point of view. The aim of this research is to provide a general framework that makes it possible to quantize the paths of firms may take over the several years of their existence. D'Aveni (1989) has confirmed the existence of different patterns of decline. That is, the healthy firms and decline firms have experienced very different financial paths, the paths experienced by healthy and failure firms are discriminating. Hence, based on this ground assumption, we design a financial failure prediction model based on financial path analysis using fuzzy c-means clustering method. Therefore, our study is to assess how to use a fuzzy c-means to quantize typical failure paths, to analyze how such quantification may contribute to understanding of failure process and how it can be used to predict a company's financial status in the future.

Specially, we used a fuzzy c-means to design these financial paths. First, the fuzzy c-means clustering method was used to delimit boundaries between spaces representing various company financial status. Second, we analyzed how companies moved over time within these clusters to estimate a typology of behavior we called "financial paths". Third, we used these financial paths to forecast firm financial status. Finally, we compared the results achieved using the paths and results estimated using the most common methods of designing financial failure models: logistic regression, support vector machine, decision tree, and neural network.

The remainder of this paper is organized as follows. In Section 1, we present a literature review that explains our research question. In Section 2, we describe the samples and variables used in our study. In Section 3, we design the prediction model of our study. In Section 4, we discuss our results. In the last Section we summarize our main findings.

1. Literature review

Most financial failure models are single-period models. These models are designed using financial variables collected at time t, and their prediction performance is measure at time $t+1$, $t+2$, or $t+3$. Since the seminal work proposed by Altman in 1968, a myriad number of financial failure prediction models have been investigated in such a way. However, almost all classic financial failure prediction models are subjected to the problem related to using only

the transverse data samples (i.e. one annual account) to build the prediction models. The use of transverse financial data is because they neglect the important factor of time dimension in building models. Obviously, instead of a sudden and unexpected event, financial failure is a dynamic process which consists of several phases (Luoma, Laitinen 1991). So financial failure prediction models can be divided into two categories based on the idea of modelling techniques, they are snapshot model and historical model respectively. Snapshot model ignore the time factor when designing models, whereas historical model introduces time as a dimension when building prediction models.

1.1. Financial failure prediction based on snapshot models

The corporate failure prediction model was pioneered by Beaver (1966). He developed prediction model using univariate discriminant analysis model, and built cutoff threshold values for financial ratio variables to classify firms into two groups. Altman (1968) introduced multivariate discriminant analysis (MDA) to predict the company failure. Altman *et al.* (2016) extended Altman's Z-score Model to an international context, and assessed the prediction performance to test the model's usefulness for all parties. Univariate discriminant analysis model and MDA have been replaced by less restrictions statistical methods such as logit analysis (LA), probit analysis (PA). Ohlson (1980) was the pioneer in applying LA in company failure prediction, whereas Zmijewski (1984) pioneered using PA. Over the past few decades, artificial intelligent techniques have played a crucial role in firm failure prediction and have been reported to be quite effective in some cases. Neural networks are firstly used in bankruptcy prediction literature in 1990 (Odom, Sharda 1990). The use of neural networks has become popular due to the ability to learn and capture the nonlinear relationships between variables (e.g. Chauhan *et al.* 2009; Chen, Du 2009). Iturriage and Sanz (2015) combined multi-layer perceptron and self-organizing maps to study the bankruptcy of US banks, and provided a visualization tool to assess bank risk. Zhao *et al.* (2017) employed kernel extreme learning machine, which is a variation of neural network model, to predict bankruptcy. Moreover, some other artificial intelligence methods such as decision tree (Sun, Li 2008; Delen *et al.* 2013), support vector machine (Shin *et al.* 2005; Hsieh *et al.* 2012), and genetic algorithm (Shin, Lee 2002; Kim, Han 2003) were also useful for financial failure prediction. Due to the superior generalization capability, ensemble method has gained in popularity for firm failure prediction. Tsai *et al.* (2014) presented a comprehensive study of comparing ensemble classifiers including bagging and Adaboost. Wang *et al.* (2014) proposed an improved boosting model to predict bankruptcy and demonstrated the effectiveness of proposed method. Finally, Geng *et al.* (2015) conducted a forecast competition of four prevalent classification models, such as decision tree, neural network, support vector machine, and majority voting ensemble, to handle with the task of predicting financial distress of listed Chinese companies. These classic statistical and intelligent methods have achieved numerous static failure prediction models. However, this kind of modelling uses financial data fragments to build models. So we call this modelling method a snapshot model, because the static failure prediction models ignore the fact that companies change over time. The

financial data over consecutive years are not fully utilized, thus the changes of company's financial status can not be reflected by the static models.

1.2. Financial failure prediction based on historical models

Different from the traditional static failure prediction model, historical-based model takes time dimension or historical dimension into account in modeling. The historical-based models are more closer to the reality of the financial process, because firm failure is a series of cascading events which can be characterized by specific behavior of certain variables or specific symptoms of failure (Laitinen 1991). These processes can be used as an effective prior knowledge to predict the financial status of the company in the future. Sung (1999) considered the effective of bankruptcy prediction model over different economic phases. They found that the prediction ability under normal conditions and crisis conditions is very variable, so the need for a different model in crisis economic conditions is justified. Sun and Li (2011) investigated the dynamic prediction of firm failure using instance selection method. They pointed out that the prediction ability of the model is decreased with time passing on, so the prediction model is constantly rebuilt with latest samples using instance selection method. Beside, Sun *et al.* (2013), Sousa *et al.* (2016) also explored different dynamic failure prediction models based on the idea of instance selection. Liu and Wu (2017) also took time dimension into consideration when designing firm failure prediction model. They proposed incremental bagging based on selective ensemble to achieve dynamic failure prediction. Jardin (2017) depicted the change of firm financial health over time using quantization method to improve the forecast horizon in bankruptcy prediction.

Many studies have pointed out that time dimension plays a critical role in the establishment of forecasting model (Edmister 1972; Dambolena, Khoury 1980; Betts, Belhoul 1987; Balcaen, Ooghe 2006). For this reason, we investigate the prediction model which is not in the same way as previous research. Instead of using financial indicators measured over several years to design a model, we introduce time dimension into the modelling and build financial paths of firms as a prediction tool. Specifically, fuzzy c-means were used to design paths, and historical-based model was then compared to that of traditional models in terms of accuracy, type I error, and type II error.

2. Samples and variables

2.1. Sample selection

The financial data used in this study are collected from the China Center for Economic Research (CCER) database. This is an internet-based data service platform in China. More specifically, this database contains balance sheets and income statements of Chinese listed firms over a period of 10 years. We extracted from this database a set of samples of companies using criteria that are presented below, and we computed financial ratios that were used as explanatory variables. Only financial ratios are used as explanatory variables in this study.

Commonly, Chinese listed companies will be specially treated (ST) if a company has had negative net profit in consecutive two years or a company's net capital per share is lower

than its face value. The ST companies are labeled as financial failed companies, and non-ST firms are identified as financial healthy firms. In order to avoid class imbalance problem for firm failure prediction, firms are prepared to find a matching healthy listed firm to form a pairwise data sample for classification. We selected paired firms with an asset structure as homogeneous as possible to control for the size effect and to allow comparisons of ratios, besides, the paired firms should belong to the same industry as the ST firm.

Unlike traditional prediction classifiers only using the snapshot character of the annual account, that is, the boundary used to classify healthy and failure firms is constructed by using the transverse financial data. The historical-based model proposed in this study uses a certain length of financial history data (longitudinal financial data) to establish a financial path. A financial path is a trajectory which a firm moves from one class of risk to another over time. These classes of risk can be considered the hierarchies of financial profiles that best summarize all kinds of company financial situations. Table 1 presents how the longitudinal financial data were collected over time to design the financial path of historical models. From Table 1 we can see that a span of four years' financial data are used to design financial path. For instance, to build the firms' financial paths of year 2014, financial data from 2009 to 2012 are collected, that is, there are four financial statuses to describe the financial path of a firm. We take firm's ST year as the benchmark year (t–0). Then, our empirical experiment is done based on year (t–2), which represents two years before ST occurs. It means that financial failure is tried to be predicted according to financial information two years in advance. So there is no need to use the financial data of year 2013 to establish the company's financial path of year 2014.

Table 1. Financial data used to design the financial paths of historical model

Data span used to design paths				Target year
2009	2010	2011	2012	2014
2008	2009	2010	2011	2013
2007	2008	2009	2010	2012
2006	2007	2008	2009	2011
2005	2006	2007	2008	2010
2004	2005	2006	2007	2009

Table 2 lists the training sets and test sets of each year to be predicted. From Table 2 can be seen that three years' firm samples are collected as training dataset. If the financial performance in 2014 is to be predicted, financial data from 2011 to 2013 are selected as training dataset. The financial data used to design financial path is also presented. Table 2 also lists the samples and paths in the experiment. Take year 2014 as an example, the number of firms in 2014 is 50, so each company corresponds to one of its own financial paths, the number of financial paths in 2014 is also 50. Because of data span used to design financial path is 4 years, they are 2009, 2010, 2011 and 2012 respectively, so there are 200 (50*4) samples to design the 4-year paths in 2014. Thus, the initial datasets are built according to the model framework.

Table 2. Training samples and test samples

	Test set	Training set		
Year of samples	2014	2013	2012	2011
Data span used to design paths	2009–2012	2008–2011	2007–2010	2006–2009
Number of paths	50	34	42	30
Number of samples	200	136	168	120
Year of samples	2013	2012	2011	2010
Data span used to design paths	2008–2011	2007–2010	2006–2009	2005–2008
Number of paths	34	42	30	72
Number of samples	136	168	120	288
Year of samples	2012	2011	2010	2009
Data span used to design paths	2007–2010	2006–2009	2005–2008	2004–2007
Number of paths	42	30	72	48
Number of samples	168	120	288	192

2.2. Financial variable

We selected a set of 30 financial indicators as inputs based on the qualitative analysis of China Accounting Standards. These indicators were not chosen arbitrarily, but based on the main financial dimensions that represent financial situation. We choose the five dimensions that are: per share ratio (reflect the profit-gaining ability and development level of listed company), solvency (reflect the ability of a company to pay its debts), profitability (reflect the efficiency of a firm transforms business activities into profits), growth ability (reflect the future development of a company), and turnover (reflect the ability of a firm to turn assets, liabilities and capital accounts into cash or sales). Prediction models were then designed using the variable selection technique. The variables selection process are presented in Section 5.1. The initial financial indicators are listed in Table 3.

3. Modeling method

We selected two types of modelling techniques to design financial failure prediction models. First, with a procedure presented below, we used a fuzzy c-means to design paths of financial situation over time, then predicted the firm's financial situation according to patterns of financial paths. Second, for comparative purpose, we chose four of the most commonly used modeling techniques in the financial literature: logistic regression (LR); neural network (NN), especially a multi-layer perceptron (MLP); decision tree (DT); support vector machine (SVM). The four aforementioned methods, unlike the historical-based model, rely not on data that measure changes of a firm's financial status over several consecutive years but on a snapshot of a company's financial profile taken at a particular point in time.

Table 3. Initial financial indicators

Category	Variable	Name of variable	Formula
Per share ratio	V_1	Basic earnings per share	Net profit / Total shares
	V_2	Net asset value per share	(Total assets-Total liabilities) / Total shares
	V_3	Operating income per share	Sales revenue / Total shares
	V_4	Undistributed profit per share	Total undistributed profit / Total shares
	V_5	Provident fund per share	Provident fund / Total shares
Profitability	V_6	Gross profit margin	Gross sales margin / Sales revenue
	V_7	Net profit margin	Net profit / Sales revenue
	V_8	Cash sales rate	Net cash flow / Sales revenue
	V_9	Sales expense rate	Sales expense / Sales revenue
	V_{10}	Operating profit ratio	Operating profit / Sales revenue
	V_{11}	Return on assets	Net profit / Total assets
Solvency and Liquidity	V_{12}	Assets liabilities ratio	Total liabilities / Total assets
	V_{13}	Tangible assets asset ratio	Intangible assets / Total assets
	V_{14}	Current assets ratio	Current assets / Total assets
	V_{15}	Tangible assets debt ratio	Intangible assets debt / Total assets
	V_{16}	Current debt ratio	Current liabilities / Total assets
	V_{17}	Current ratio	Current assets / Current liabilities
	V_{18}	Quick ratio	Quick assets / Current liabilities
	V_{19}	Cash liabilities ratio	Net cash flow / Total liabilities
Turnover	V_{20}	Inventory turnover	Sales cost / Average balance of inventory
	V_{21}	Current assets turnover	Sales revenue / Average balance of current assets
	V_{22}	Total assets turnover	Sales revenue / Total assets
	V_{23}	Accounts receivable turnover	Net credit income / Average balance of accounts receivable
Growth ability	V_{24}	Net cash flow growth rate	Net cash flow of this year / Net cash flow of last year
	V_{25}	Operating income growth rate	Main business income of this year / Main business income of last year
	V_{26}	Operating profit growth rate	Operating profit of this year / Operating profit of last year
	V_{27}	Total profit growth rate	Total profit of this year / Total profit of last year
	V_{28}	Net assets growth rate	Net assets of this year / Net assets of last year
	V_{29}	Total assets growth rate	Total assets of this year / Total assets of last year
	V_{30}	Net profit growth rate	Net profit of this year / Net profit of last year

3.1. Historical model development

In this section, instead of designing snapshot models, we designed financial path models based on consecutive financial data. The method and procedure used to design this kind of model is presented below.

3.1.1. Framework of historical-based model for financial failure prediction

This study tries to investigate on financial path method for financial failure prediction, whose framework is designed as Figure 1. In order to make full use of financial data information, consecutive financial data over several years are deployed to design financial path by using fuzzy c-means approach. Then, various financial paths are designed corresponding to the training firm samples. After that, two patterns of financial path, financial path of healthy firms and financial path of failure firms respectively, are formed based on the label information of the firms. Due to the trajectories experienced by healthy and failure companies are disparate, so these two patterns of financial paths are employed to predict the firms' financial status in the future through calculating the Euclidean distance between path positions. The components of financial path method, such as fuzzy c-means approach, financial path design mechanism, and historical model design mechanism, are to be stated respectively in the following subsections.
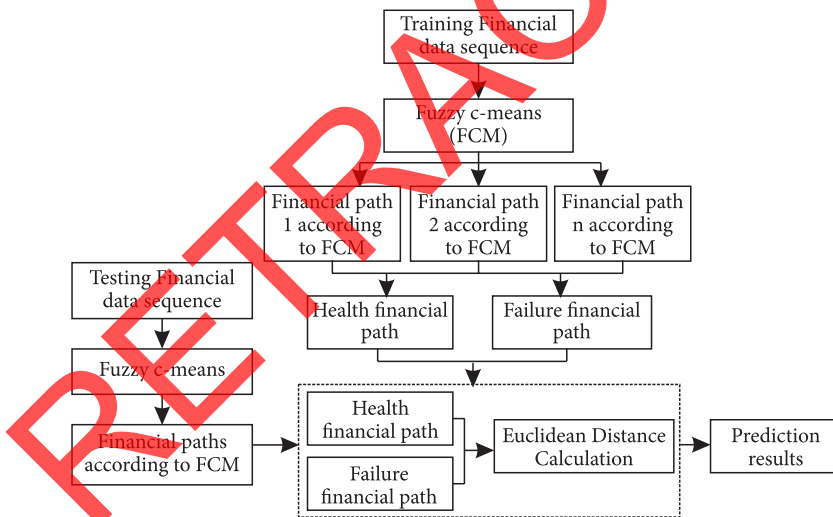

Figure 1. Framework of historical-based model for financial failure prediction

3.1.2. Fuzzy C-means

Fuzzy c-means (FCM) is a method of clustering which allows one piece of data to belong to two or more clusters. This method (developed by Dunn in 1973 and improved by Bezdek in 1981) is frequently used in pattern recognition. It is based on minimization of the following objective function:

$$J_m = \sum_{i=1}^{N} \sum_{j=1}^{C} u_{ij}^m \left\| x_i - c_j \right\|^2, \ 1 \le m \le \infty, \tag{1}$$

where m is any real number greater than 1, u_{ij} is the degree of membership of x_i in the cluster j, x_i is the ith of d-dimensional measured data, c_j is the d-dimension center of the cluster, and $\|*\|$ is any norm expressing the similarity between any measured data and the center.

Fuzzy partitioning is carried out through an iterative optimization of the objective function shown above, with the update of membership u_{ij} and the cluster centers c_j by:

$$u_{ij} = \frac{1}{\sum_{k=1}^{C}\left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|}\right)^{\frac{2}{m-1}}}; \tag{2}$$

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}^m \cdot x_i}{\sum_{i=1}^{N} u_{ij}^m}. \tag{3}$$

This iteration will stop when $\max_{ij}\left\{\left|u_{ij}^{(k+1)} - u_{ij}^{(k)}\right|\right\} < \varepsilon$, where ε is a termination criterion between 0 and 1, whereas k are the iteration steps. This procedure converges to a local minimum or a saddle point of J_m. The algorithm is composed of the following steps:

1. Initialize $U=[Uij]$ matrix, $U^{(0)}$.
2. At k-step: calculate the centers vectors $C^{(k)}=[c_j]$ with $U^{(k)}$

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}^m \cdot x_i}{\sum_{i=1}^{N} u_{ij}^m}.$$

3. Update $U^{(k)}$, $U^{(k+1)}$

$$u_{ij} = \frac{1}{\sum_{k=1}^{C}\left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|}\right)^{\frac{2}{m-1}}}.$$

4. If $\|U^{(k+1)} - U^{(k)}\| \leq \varepsilon$ then STOP; otherwise return to step 2.

3.1.3. Financial path design

A financial path represents the way some firms move forward over a period of several years. Costea *et al.* (2017) demonstrated that fuzzy c-means can be used to group the firms into similar financial performance classes according to their financial indicators. Indeed, the fuzzy c-means approach can handle multi-dimensional financial data and tackle non-linear relationships among input variables. The result has shown that the clustering of firm samples represent different financial statuses. So fuzzy c-means made it possible to show different levels of financial status, each of these levels contributing to a particular financial profile. Some groups consist of profitable, good structure, and sound companies, others are composed of poor profitability, very unsound companies, still others represent firms in intermediate financial situations. Accordingly, fuzzy c-means can be used to delimit boundaries between groups at risk of failure and other groups at low risk of failure or without any risk of failure,

each group accounts for a given financial profile associated with a probability of failure. As a consequence, fuzzy c-means can be used to design the change in financial situation of a company over time. A financial path shows the way companies move on the clusters over several consecutive years in different levels at risk of failure. Therefore, a financial path is presented by a sequence of positions on the clusters over a given period.

In this study, to depict the path of a firm, a fuzzy c-means was used to reorganized the firm samples. Since a cluster represents firms with similar financial situation and profiles the state that firms may embody at a given point in time, the clusters obtained by fuzzy c-means represents different risks of failure of these firm samples, then the financial process corresponds to the way a firm moves into the situation of risk over time. Therefore, the financial situation was quantized using fuzzy c-means. More specifically, we first used fuzzy c-means to organize firm samples for each of the 4 years. Once the clustering phase was completed, we computed firm financial paths, that is, the positions of companies on the clusters over the four-year period. To calculate the different positions of a company on the cluster, we recorded the cluster number of the company in each year. There are four positions of a company on the cluster over the period analyzed here. A financial path is then a sequence of four positions. Since the number of firm samples determines the number of financial paths, it is impossible to analyze all possible financial paths, so we used fuzzy c-means method again to reorganize the financial paths into several prototype groups. We then calculated the percentage of healthy and failed firms within each group. And we labeled each prototype financial path with the class (healthy or failure) whose percentage was higher. Finally, we used these financial paths to achieve the forecast based on Euclidean distance metric. The following subsection presents some specifications of the prediction model achieved with financial path.

3.1.4. Historical model design

Different groups of financial paths represent the way distinct groups of firms change from one possible financial situation to another with some common behaviors over time. We assume that there exist m financial paths representing the trajectory of the healthy company, and there exist n financial paths representing the trajectory of the failure company. Let $P = \{p_1, p_2, ..., p_j; l\}$ represents the financial statuses (cluster positions) over t years, where $p_j = (p_{j1}, p_{j2}, ..., p_{jt})$ denotes the j-*th* common financial path of firm samples with similar financial situation over years, and $l = (0, 1)$ denotes the class label of the financial path, "0" represents financial failure firm and "1" refers to non-failure firm. That is, if one financial path is labeled as "0", the firms in the same group have similar financial situation over time, and finally fall into financial failure, and vice versa. Let $X = \{x_1, x_2, ..., x_t\}$ represents the financial path of testing sample, we calculate the Euclidean distance between the sequence of financial states of the test sample and the sequence of existing financial states, then the test samples' financial statuses can be predicted according to the distance metric. The prediction model can be expressed as follows:

$$Distance(X, p_j) = \underset{m+n}{\arg\min} \sqrt{(x_1 - p_{j1})^2 + (x_2 - p_{j2})^2 + ... + (x_t - p_{jt})^2}. \tag{4}$$

The fundamental assumption of historical model is that the financial paths experienced by firms with the similar financial status have some common characteristics in history, so

the financial paths between healthy firms and fail firms are separable. Therefore, the model is thus developed to predict financial failure.

3.2. Snapshot models used as benchmarks

A set of snapshot models was constructed using logistic regression, decision tree, support vector machine and neural network. These four models are traditionally mono-period models which rely on a snapshot of a company's financial situation taken at a particular point time. The range of training data for snapshot models is same with the historical model, that is, three years of financial data are selected to build the prediction models.

3.3. Evaluation of model performance

The evaluation criteria of the each models' classification performance were adopted from the established standard measures in the fields of financial failure prediction. In this study, we used accuracy, type I error, and type II error to measure the performance of each method. A type I error is made when a ST firm is misclassified as a norm firm, whereas a type II error is made when a norm firm is wrongly assigned to the ST group. The arithmetic definition of the measures can be explained with respect to a confusion matrix as shown in Table 4.

Table 4. Confusion matrix

		Actual condition	
		Positive(Norm company)	Negative(ST company)
Test result	Positive(Norm company)	True Positive(TP)	False Positive(FP)
	Negative(ST company)	False Negative(FN)	True Negative(TN)

The aforementioned performance measures are defined as follows:

$$accuracy = \frac{TP + TN}{TP + FP + FN + TN};\tag{5}$$

$$type\ I\ error = \frac{FN}{TP + FN};\tag{6}$$

$$type\ II\ error = \frac{FP}{FP + TN}.\tag{7}$$

4. Results and discussion

4.1. Variables used to design historical models

The training dataset was used to select the variables. Their characteristics are presented in Tables 5 and 6. Statistical results in Table 5 were calculated with data from the training

sample of year 2014, and data has been normalized from 0 to 1 beforehand. Financial variable selection is first conducted by ranking each indicator based on the strength of its relation to the dependent variable. A one-way ANOVA F test is performed to rank the financial variables. Variables are sorted according to their *p*-value in an ascending order and the score of each variable is used as ranking method to evaluate the predictive ability of each financial indicator. The first 50% financial indicators are selected for future use. From Table 5 we see that a number of 16 financial indicators are selected from the initial set by the ANOVA F test. The quartile of each variable show the discrepancy of the deviations in and between the two groups of firms. Table 5 also indicates the *p*-values of a Kolmogorov-Smirnov normality test. As the Kolmogorov-Smirnov test shows that none of the variables are normally distributed, the Mann-Whitney U test is more reliable than Student's t test. This test underscores that all variables present significant differences between the two groups. Then the variables are used to conduct Mann-Whitney U test. The statistical results obtained from Mann-Whitney U test are listed in Table 6.

Table 5. Characteristics of variables

Variables	Healthy companies percentiles			Fail companies percentiles			*K–S* test
	25%	50%	75%	25%	50%	75%	
V_1	0.640	0.699	0.755	0.530	0.594	0.618	.000
V_2	0.328	0.404	0.508	0.162	0.255	0.334	.015
V_4	0.509	0.567	0.643	0.360	0.440	0.496	.081
V_5	0.092	0.168	0.242	0.044	0.095	0.159	.000
V_{11}	0.777	0.799	0.824	0.723	0.754	0.768	.000
V_{12}	0.265	0.364	0.491	0.380	0.553	0.645	.073
V_{13}	0.899	0.948	0.973	0.895	0.949	0.980	.000
V_{15}	0.020	0.032	0.045	0.013	0.018	0.030	.000
V_{17}	0.041	0.052	0.066	0.035	0.040	0.054	.000
V_{18}	0.033	0.042	0.054	0.029	0.033	0.043	.000
V_{19}	0.183	0.207	0.236	0.174	0.190	0.204	.000
V_{20}	0.016	0.026	0.043	0.014	0.024	0.045	.000
V_{23}	0.002	0.003	0.008	0.002	0.003	0.006	.000
V_{24}	0.797	0.798	0.799	0.797	0.798	0.799	.000
V_{28}	0.194	0.200	0.214	0.181	0.192	0.198	.000
V_{29}	0.133	0.145	0.167	0.123	0.135	0.147	.000

As can be seen from Table 6, V_1, V_2, V_4, V_5, V_{11}, V_{12}, V_{13}, V_{17}, V_{18}, V_{19}, V_{28}, V_{29} show significant differences at the 1% confidence level. The remaining variables show no significant difference between two groups. Therefore, a total of twelve variables were chosen as input to build prediction models.

Table 6. Mann-Whitney U test

	V_1	V_2	V_4	V_5	V_{11}	V_{12}	V_{13}	V_{15}
Z	−13.01	−10.58	−11.84	−6.91	−12.93	−7.87	−7.63	**−0.97**
Sig.	.000	.000	.000	.000	.000	.000	.000	**.331**
	V_{17}	V_{18}	V_{19}	V_{20}	V_{23}	V_{24}	V_{28}	V_{29}
Z	−5.62	−5.69	−5.64	**−1.30**	**−1.22**	**−0.32**	−7.78	−5.39
Sig.	.000	.000	.000	**.195**	**.221**	**.747**	.000	.000

4.2. Firm financial behaviors analysis

In order to design financial path, the financial data was organized into several groups using FCM, Table 7 presents the clustering profile statistics. From Table 7 can be seen, financial data was grouped into 6 clusters, and there are big differences between the number of healthy samples and the number of fail samples within each cluster. Cluster 1, 2, and 3 are labeled as failure status due to the failure samples percentage is high, cluster 4, 5, and 6 are labeled as healthy status because of the healthy samples percentage is high.

Table 7. Clustering profile

	1	2	3	4	5	6
Number of Samples	81	73	75	76	46	73
Number of Healthy samples	3	20	29	58	38	64
Number of Fail samples	78	53	46	18	8	9

Statistical result about the number of samples within each cluster and proportion between failure and healthy firms is also depicted in Figure 2. The blue curve indicates that the proportion between failure firms and healthy firms decreases from cluster 1 to cluster 6. That is, the financial situation of samples between cluster 6 to cluster 1 is progressively deteriorated.
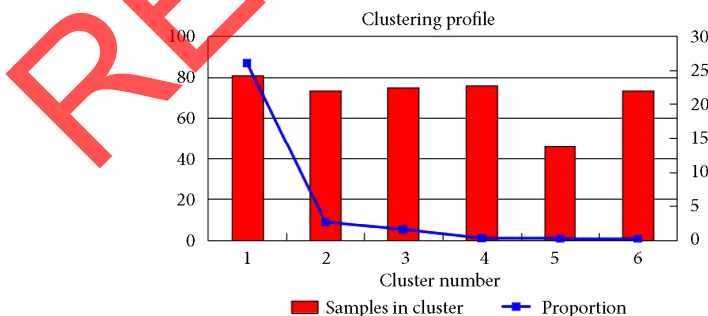

Figure 2. Sample number within each cluster and proportion between failure and healthy firms

The firm samples were grouped into 6 cluster by fuzzy c-means. Within each cluster, we calculated the means of all variables to rank the cluster by financial health. These statistics are shown in Table 8. Table 8 shows that group 1 is made up of very healthy, profitable, and

liquid companies, as opposed to group 6, which is made up of unsound firms with the lowest profitability and solvency. Take V_1 as an example, V_1 represents basic earning per share, the higher value the earnings per share, the better the profitability and growth ability of the company. From Table 8 we can see that the value of V_1 is increased from cluster 1 to cluster 6. The remaining variables have the similar tendency in Table 8. That is to say, the firms within the same group have similar financial behavior, besides, the indicators of healthy firms which representing profitability, solvency and growth ability are stronger than firms with unsound financial status, due to the values of such indicators are larger than that of unsound firms, and the liabilities are lower than that of unsound firms.

Table 8. Characteristics of the variables within each prototype cluster

	Failure firms			Healthy firms		
	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
V_1	0.516	0.600	0.608	0.681	0.699	0.738
V_2	0.203	0.296	0.257	0.395	0.428	0.513
V_4	0.371	0.471	0.456	0.556	0.583	0.637
V_5	0.103	0.133	0.088	0.155	0.200	0.241
V_{11}	0.715	0.754	0.761	0.795	0.792	0.800
V_{12}	0.621	0.557	0.351	0.324	0.424	0.411
V_{13}	0.019	0.023	0.043	0.052	0.039	0.040
V_{17}	0.044	0.046	0.059	0.068	0.058	0.060
V_{18}	0.037	0.039	0.048	0.057	0.049	0.050
V_{19}	0.189	0.193	0.196	0.226	0.216	0.221
V_{28}	0.185	0.197	0.199	0.206	0.208	0.210
V_{29}	0.136	0.147	0.140	0.146	0.152	0.155

4.3. Failure firm paths analysis

In this section, we analyzed the firms financial paths of each pattern. First, the paths of healthy firms and failure firms are analyzed respectively. Based on the above description, we know that the financial status is declined from cluster 6 to cluster 1. The mean value of positions from healthy firms and the mean value of positions from failure firms are calculated to describe the variation of financial process. On the graph, the scale of the abscissa corresponds to the six years, and the scale of the ordinate to the six clusters. Figure 2 displays the two types of firm financial paths, from Figure 3 we can see that the discrimination between two types of firms is obvious, and these firms can be identified by the positions of the company's financial status. From the red curve, it can be seen that the financial positions of the healthy company are relatively stable and the positions are relatively high over the past four year. Conversely, the financial status of failed companies has declined year by year, and failed companies locate in a relatively low position compared with healthy companies. This shows that the company's financial status can be classified through the financial path.

In addition to analyze the panorama of the financial path of two firm types, we have also examined the financial path patterns of companies before failure. Based on the position

trajectory, we have classified the financial paths of these failed companies into four categories. The mean values of four patterns of failure process are calculated, and the four curves are depicted in Figure 3.

From Figure 4 we can see that failure firms have four major patterns of process in common. The first refers to a sudden deterioration in the last year before failure. These firms were initially in a healthy state, but fell into financial failure suddenly. We call this kind of failure process is sudden deterioration. Sudden deterioration is characterized by the rapid collapse of a firm, financial failure is occurred consequently. In Figure 3, the sudden deterioration curve shows that the firms financial status are at good position initially but suddenly declined in the forth year without any symbols in previous three years.

The second reveals slower and sustained deterioration, the recession in financial is inexorable. The financial profiles of these types of firms are steadily weaken several years before financial failure. We call this kind of failure process is gradual decline. Gradual decline is exemplified by the slow, incremental deterioration of a firm, financial failure is occurred immediately after the decline is complete. In Figure 4, the gradual decline curve shows that financial status is gradually declined over past four years, so these companies' financial failure are traceable and prescient.
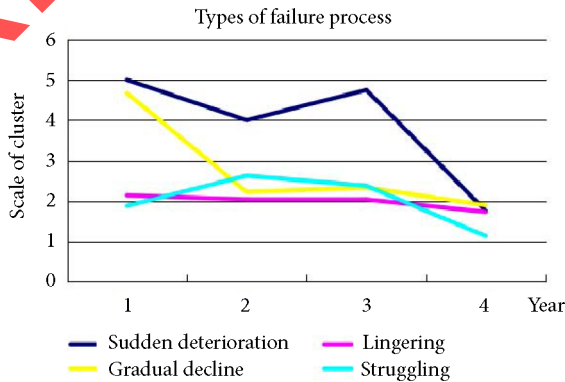

Figure 3. Two types of financial paths

Figure 4. Patterns of failure process

The third pattern has to do with firms that were unsound for many years but never had the opportunity to get better even though they managed to survive for a long time. We call this kind of failure process is lingering. Lingering is characterized by the changes of firms situation, which are neither rapidly nor gradually. Although these firms financial status are rather stable, they have fell in financial failure for several years. The lingering curve is also depicted in Figure 4.

The forth pattern refers to these firms that were unsound for years, they try to get better but failed finally. We call this kind of failure process is struggling. Struggling is exemplified by a period of remission, in which companies shifted from a very difficult financial situation to a better place then get worse again. The firms in this group have certainly tried to prevent failure, but they were unable to adapt the environment to survive. From Figure 4 we can see that the struggling curve is volatile and finally declined.

4.4. Prediction results and comparative analysis

Based on the financial path analysis, we design financial path prediction model and conduct empirical experiments. Forecasting results were estimated using the test sample, we selected 3 years to test the prediction models, the test sample were collected from 2012, 2013, and 2014, respectively. Table 9 shows the prediction results using the five models, they are logistic regression (LR), support vector machine (SVM), decision tree (DT), neural network (NN), and financial path (FP), respectively. As mentioned in the earlier section, three measures of prediction performance are reported in Table 9 and they are accuracy, type I error, and type II error respectively. We have made 20 repeated experiments based on each model. As shown in Table 9, FP gets the highest average accuracy in each test year, and also obtains the higher overall accuracy, 85.14%. The overall accuracy performance of NN ranked second, 83.11%, close to but not exceeding the accuracy obtained using FP significantly. SVM also gets the higher average accuracy, 82.32%. For the type I and II errors, FP reduces the two indicators into 11.36% and 14.79%. Compared with other four models, FP reduces the type I error greatly. Overall, FP performs relatively superior and stable than other models in terms of both error rate.

In order to determine the best performing model, we not only compared the average prediction performance of the five models but also conducted two-sample t test between the performance of FP and that of the other four models. The null hypothesis of each t test was that the prediction accuracy of FP was higher than that of LR, SVM, DT, and NN. The statistical results indicated that the prediction accuracy of FP was significantly higher than LR, SVM, DT, and NN. In addition, the null hypothesis of each t test was that the error rate of FP was lower than that of LR, SVM, DT, and NN. The statistical results indicated that the error rate of FP was significantly lower than LR, SVM, DT, and NN in terms of type I error. However, the null hypothesis is not rejected when comparing the type II error of FP with SVM. This indicated that SVM performed as well as FP in these experiments. Therefore, FP enhances the prediction accuracy of firm failure and reduce the error rate in some extend. It can be served as an alternative diagnosis tool for financial failure prediction.

Table 9. Accuracy, Type I error and Type II error of the prediction results

		2012	2013	2014	Overall
Accuracy	LR	80.48%	79.56%	81.20%	80.41%
	SVM	82.50%	81.77%	82.70%	82.32%
	DT	79.88%	80.00%	81.40%	80.43%
	NN	83.93%	81.91%	83.50%	83.11%
	FP	86.79%	83.24%	85.40%	85.14%
Type I error	LR	18.33%	19.12%	17.60%	18.35%
	SVM	17.38%	22.65%	16.00%	18.68%
	DT	21.19%	21.18%	15.80%	19.39%
	NN	13.33%	16.18%	17.60%	15.70%
	FP	11.67%	10.60%	11.80%	11.36%
Type II error	LR	20.71%	21.77%	20.00%	20.83%
	SVM	17.62%	13.82%	18.60%	16.68%
	DT	19.05%	18.82%	21.40%	19.76%
	NN	18.81%	20.00%	15.40%	18.07%
	FP	14.76%	12.20%	17.40%	14.79%

Table 10. Significant test

			LR	SVM	DT	NN
FP	Accuracy	t-value	4.539	2.957	5.320	2.277
		p-value	0.000***	0.008***	0.000***	0.034**
	Type I error	t-value	−6.640	−8.400	−7.011	−5.028
		p-value	0.000***	0.000***	0.000***	0.000***
	Type II error	t-value	−4.056	−1.137	−3.174	−2.287
		p-value	0.001***	0.269	0.005***	0.033**

Conclusions

Financial failure prediction is one area in which data mining and forecasting tools have expanded greatly over the past few years. In economies, failure prediction models have become essential to credit decision-making frameworks. Different from traditional dichotomous models, to obtain a better understanding of financial failure process and assess the behaviors of firm over time, a new framework of assessing a firm's financial situation over years is proposed by using fuzzy c-means. Financial path is design to trace the behavior of a firm and is used as a diagnosis tool for financial failure prediction.

In this study, we have investigated the abilities of fuzzy c-means for representation of financial data and prediction of financial status. Chinese real financial data within consecutive several years are collected to conduct the empirical experiments. The financial paths between healthy firms and failure firms are compared, from the results can be seen that there

is a discrepancy between two types of firms, financial paths of healthy firms are relatively stable over past several years, whereas financial paths of failure firms are declined year by year, which provides a deterioration sign with a visible manner. Furthermore, the patterns of failure paths are also analyzed, the financial paths show that there exist different paths to terminal failure. Therefore, financial path tool not only help managers observe the financial trend of a firm, but also provide manager early warning signals to take timely activities.

From the prediction performance view of point, taking the three assessment of total accuracy, type I error, and type II error, the conclusion of this study is that the proposed financial path model compares favourably with the four models (i.e. LR, DT, NN, and SVM). In terms of total accuracy, FP obtained significant better performance than other models compared in statistic. In terms of type I error, which is more important than type II error in the assessment of model's performance, FP also outperforms other models significantly. In terms of type II error, SVM wins the best performance. FP outperforms other three models statistically. Thus, FP significantly outperforms other four models from both metrics of accuracy and type I error, and is not significantly outperformed by other four models (except for SVM) from the view of type II error.

In conclusion, financial path model designed with fuzzy c-means presents some merits in the application of financial analysis, including financial data description and financial failure prediction. However, this study also has certain limitations due to the sample data used for experimental study are only collect from Chinese listed firms, and some other countries' real world financial data samples should also included in our future study. The other limitation is that the forecasting horizon of our proposed model is 3 year before the failure, but the prediction horizon of 4, 5 or long-term year are not tested. Lastly, fuzzy c-means is an effective tool when grouping firms based on the parsimonious financial variables. As we know, there exist many financial variables, and the relationship between these variables are complex, so identification of intrinsic characteristics and topology of high-dimensional financial data with other clustering methods is another important task for financial path analysis. Because of the limitations of the study, some questions deserve a further study, such as how many years can the financial path model detect the symbol of financial failure in advance. Further, if there have any other clustering methods to better analyze the financial path when handling with high-dimensional financial data, then enhancing the prediction performance of financial failure according to the complex interactions among the financial variables.

References

Altman, E. I. 1968. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, *The Journal of Finance* 23(4): 589–609. https://doi.org/10.1111/j.1540-6261.1968.tb00843.x

Altman, E. I.; Iwanicz-Drozdowska, M.; Laitinen, E. K.; Suvas, A. 2017. Financial distress prediction in an international context: a review and empirical analysis of Altman's Z-Score model, *Journal of International Financial Management and Accounting* 28(2): 131–171. https://doi.org/10.1111/jifm.12053

Argenti, J. 1976. Corporate planning and corporate collapse, *Long Range Planning* 9(6): 12–17. https://doi.org/10.1016/0024-6301(76)90006-6

Balcaen, S.; Ooghe, H. 2006. 35 years of studies on business failure: an overview of the classic statistical methodologies and their related problems, *The British Accounting Review 38*(1): 63–93. https://doi.org/10.1016/j.bar.2005.09.001

Beaver, W. H. 1966. Financial ratios as predictors of failure, *Journal of Accounting Research* 4: 71–111. https://doi.org/10.2307/2490171

Betts, J.; Belhoul, D. 1987. The effectiveness of incorporating stability measures in company failure models, *Journal of Business Finance and Accounting* 14(3): 323–334. https://doi.org/10.1111/j.1468-5957.1987.tb00098.x

Bezdek, J. C. 2013. *Pattern recognition with fuzzy objective function algorithms*. Springer Science & Business Media.

Chauhan, N.; Ravi, V.; Chandra, D. K. 2009. Differential evolution trained wavelet neural networks: application to bankruptcy prediction in banks, *Expert Systems with Applications* 36(4): 7659–7665. https://doi.org/10.1016/j.eswa.2008.09.019

Chen, W. S.; Du, Y. K. 2009. Using neural networks and data mining techniques for the financial distress prediction model, *Expert Systems with Applications* 36(2): 4075–4086. https://doi.org/10.1016/j.eswa.2008.03.020

Costea, A.; Ferrara, M.; Serban, F. 2017. An integrated two-stage methodology for optimising the accuracy of performance classification models, *Technological and Economic Development of Economy* 23(1): 111–139. https://doi.org/10.3846/20294913.2016.1213196

Dambolena, I. G.; Khoury, S. J. 1980. Ratio stability and corporate failure, *The Journal of Finance* 35(4): 1017–1026. https://doi.org/10.1111/j.1540-6261.1980.tb03517.x

D'Aveni, R. A. 1989. The aftermath of organizational decline: a longitudinal study of the strategic and managerial characteristics of declining firms, *Academy of Management Journal* 32(3): 577–605. https://doi.org/10.2307/256435

Delen, D.; Kuzey, C.; Uyar, A. 2013. Measuring firm performance using financial ratios: a decision tree approach, *Expert Systems with Applications* 40(10): 3970–3983. https://doi.org/10.1016/j.eswa.2013.01.012

Du Jardin, P. 2017. Dynamics of firm financial evolution and bankruptcy prediction, *Expert Systems with Applications* 75(1): 25–43. https://doi.org/10.1016/j.eswa.2017.01.016

Du Jardin, P.; Séverin, E. 2011. Predicting corporate bankruptcy using a self-organizing map: an empirical study to improve the forecasting horizon of a financial failure model, *Decision Support Systems* 51(3): 701–711. https://doi.org/10.1016/j.dss.2011.04.001

Dunn, Joseph C. 1973. A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters, *Journal of Cybernetics* 3(3): 32–57. https://doi.org/10.1080/01969727308546046

Edmister, R. O. 1972. An empirical test of financial ratio analysis for small business failure prediction, *Journal of Financial and Quantitative Analysis* 7(2): 1477–1493. https://doi.org/10.2307/2329929

Frydman, H.; Altman, E. I.; Kao, D. L. 1985. Introducing recursive partitioning for financial classification: the case of financial distress, *The Journal of Finance* 40(1): 269–291. https://doi.org/10.1111/j.1540-6261.1985.tb04949.x

Geng, R.; Bose, I.; Chen, X. 2015. Prediction of financial distress: an empirical study of listed Chinese companies using data mining, *European Journal of Operational Research* 241(1): 236–247. https://doi.org/10.1016/j.ejor.2014.08.016

Hambrick, D. C.; D'Aveni, R. A. 1988. Large corporate failures as downward spirals, *Administrative Science Quarterly* 33(1): 1–23. https://doi.org/10.2307/2392853

Hsieh, T. J.; Hsiao, H. F.; Yeh, W. C. 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(1): 196–206. https://doi.org/10.1016/j.neucom.2011.11.020

Iturriaga, F. J. L.; Sanz, I. P. 2015. Bankruptcy visualization and prediction using neural networks: a study of US commercial banks, *Expert Systems with Applications* 42(6): 2857–2869. https://doi.org/10.1016/j.eswa.2014.11.025

Kim, M. J.; Han, I. 2003. The discovery of experts' decision rules from qualitative bankruptcy data using genetic algorithms, *Expert Systems with Applications* 25(4): 637–646. https://doi.org/10.1016/S0957-4174(03)00102-7

Laitinen, E. K. 1991. Financial ratios and different failure processes, *Journal of Business Finance and Accounting* 18(5): 649–673. https://doi.org/10.1111/j.1468-5957.1991.tb00231.x

Liu, J.; Wu, C. 2017. Dynamic forecasting of financial distress: the hybrid use of incremental bagging and genetic algorithm – empirical study of Chinese listed corporations, *Risk Management* 19(1): 32–52. https://doi.org/10.1057/s41283-016-0012-6

Luoma, M.; Laitinen, E. K. 1991. Survival analysis as a tool for company failure prediction, *Omega* 19(6): 673–678. https://doi.org/10.1016/0305-0483(91)90015-L

Martín-del-Brío, B.; Serrano-Cinca, C. 1993. Self-organizing neural networks for the analysis and representation of data: some financial cases, *Neural Computing and Applications* 1(3): 193–206. https://doi.org/10.1007/BF01414948

Miller, D.; Friesen, P. H. 1977. Strategy-making in context: ten empirical archetypes, *Journal of Management Studies* 14(3): 253–280. https://doi.org/10.1111/j.1467-6486.1977.tb00365.x

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(4): 603–614. https://doi.org/10.1016/j.eswa.2004.12.008

Neophytou, E.; Molinero, C. M. 2004. Predicting corporate failure in the UK: a multidimensional scaling approach, *Journal of Business Finance and Accounting* 31(5): 677–710. https://doi.org/10.1111/j.0306-686X.2004.00553.x

Odom, M. D.; Sharda, R. 1990. A neural network model for bankruptcy prediction, in *Neural Networks IJCNN International Joint Conference*, 17–21 June 1990, San Diego, CA, USA, 163–168. https://doi.org/10.1109/IJCNN.1990.137710

Ohlson, J. A. 1980. Financial ratios and the probabilistic prediction of bankruptcy, *Journal of Accounting Research* 18(1): 109–131. https://doi.org/10.2307/2490395

Shin, K. S.; Lee, T. S.; Kim, H. J. 2005. An application of support vector machines in bankruptcy prediction model, *Expert Systems with Applications* 28(1): 127–135. https://doi.org/10.1016/j.eswa.2004.08.009

Shin, K. S.; Lee, Y. J. 2002. A genetic algorithm application in bankruptcy prediction modeling, *Expert Systems with Applications* 23(3): 321–328. https://doi.org/10.1016/S0957-4174(02)00051-9

Sousa, M. R.; Joao, G.; Elisio, B. 2016. Dynamic credit score modeling with short-term and long-term memories: the case of Freddie Mac's database, *Journal of Risk Model Validation* 10(1): 59–80. https://doi.org/10.21314/JRMV.2016.147

Sun, J.; Li, H. 2008. Data mining method for listed companies' financial distress prediction, *Knowledge-Based Systems* 21(1): 1–5. https://doi.org/10.1016/j.knosys.2006.11.003

Sun, J.; Li, H. 2011. Dynamic financial distress prediction using instance selection for the disposal of concept drift, *Expert Systems with Applications* 38(3): 2566–2576. https://doi.org/10.1016/j.eswa.2010.08.046

Sun, J.; Li, H.; Adeli, H. 2013. Concept drift-oriented adaptive and dynamic support vector machine ensemble with time window in corporate financial risk prediction, *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 43(4): 801–813. https://doi.org/10.1109/TSMCA.2012.2224338

Sung, T. K.; Chang, N.; Lee, G. 1999. Dynamics of modeling in d ata mining: interpretive approach to bankruptcy prediction, *Journal of Management Information Systems* 16(1): 63–85. https://doi.org/10.1080/07421222.1999.11518234

Tsai, C. F.; Hsu, Y. F.; Yen, D. C. 2014. A comparative study of classifier ensembles for bankruptcy prediction, *Applied Soft Computing* 24: 977–984. https://doi.org/10.1016/j.asoc.2014.08.047

Wang, G.; Ma, J.; Yang, S. 2014. An improved boosting based on feature selection for corporate bankruptcy prediction, *Expert Systems with Applications* 41(5): 2353–2361. https://doi.org/10.1016/j.eswa.2013.09.033

Zhao, D.; Huang, C.; Wei, Y.; Yu, F.; Wang, M.; Chen, H. 2017. An effective computational model for bankruptcy prediction using kernel extreme learning machine approach, *Computational Economics* 49(2): 325–341. https://doi.org/10.1007/s10614-016-9562-7