Application of Fuzzy Classification in Bankruptcy Prediction

Zijiang Yang[1] and Guojun Gan[2]

[1] York University
zyang@mathstat.yorku.ca
[2] York University
gjgan@mathstat.yorku.ca

Abstract. Classification refers to a set of methods that predict the class of an object from attributes or features describing the object. In this paper we present a fuzzy classification algorithm to predict bankruptcy. Our classification algorithm is modified from a subspace clustering algorithm called fuzzy subspace clustering (FSC). As our algorithm associates each feature of a class with a fuzzy membership, feature selection is not necessary. Our experiments show that the classification results produced by our algorithm can translate into large financial and other benefits to organizations through such activities as credit approval, and loan portfolio and security management.

Keywords: Bankruptcy prediction, Fuzzy subspace clustering, Data mining, Classification.

1 Introduction

Bankruptcy prediction has always attracted significant global attention. Financial institutions, in particular, are interested in an effort to reduce the level of risk in their investments. Given the growing importance of the bankruptcy prediction, there have been many attempts to model the prediction of business failure.

The first publication appeared in 1966 and was authored by W. Beaver [2]who created a univariate discriminant model using financial ratios selected by a dichotomous classification test. Since then bankruptcy prediction models have evolved to use both statistical analysis and data mining techniques to refine the decision support tools and improve decision making. In addition to discriminant analysis, traditional statistical methods include regression, logistic models, factor analysis, etc. More recent data mining techniques include decision trees, neural networks (NNs), fuzzy logic, genetic algorithms (GA) and support vector machines (SVM) among others. The statistical applications, although enhanced over time, were restricted by the rigorous assumptions of traditional statistics such as: linearity, normality, independence among predictor variables and preexisting functional form relating the criterion variable and the predictor variable.

Balcean and Ooghe [1] presented an overview of classic statistical methods for predicting business failure developed thus far and provided a detailed analysis

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 921–928, 2008.

of four types: (1) univariate analysis, (2) risk index models, (3) multivariate discriminant analysis, and (4) conditional probability model (logit, probit, linear probability models).

Min and Lee [9] were among the first to apply SVM (support vector machines) to bankruptcy prediction problem. Another attempt at hybrid intelligent systems for bankruptcy prediction was made by Tsakonas, Dounias, Doumpos and Zopounidis [10] who developed a model employing neural logic networks through genetic programming. The goal was to obtain the optimal topology of neural networks using genetic programming. A comprehensive survey of research work published between 1968 and 2005 has been compiled by Kumar and Ravi [8]. The paper analyzed a variety of statistical and intelligent methodologies applied to bankruptcy prediction. Organized by technique category, the study concluded that the stand-alone statistical methods were no longer used and neural networks are the most commonly intelligent technique used in the stand-alone mode. However, the author emphasized the potential of hybrid intelligent systems and identified it as the current trend an! d a direction for the future.

The above sampling of the recently published literature on predicting corporate failure shows a vast number of approaches taken to the subject in an attempt to refine the classification model. Most forecasts achieve accuracy between 65% to 85%. Although in the last decade the standard statistical techniques including clustering/classification methods have been largely replaced by more advanced intelligent techniques our study will take a new attempt at subspace clustering algorithm and alter it and adapt for the purpose of accurate prediction.

This paper modifies a fuzzy subspace clustering (FSC) algorithm to a classification algorithm and applies the resulted classification algorithm to bankruptcy prediction. Data clustering provides a number of methods to analyze data sets and extract relevant information from data sets. Technically, data clustering refers to an unsupervised process that divides a given data set into homogeneous groups called clusters such that points within the same cluster are more similar than points across different clusters. An overview of the topic can be found in [5,4].

Unlike data clustering, classification refers to a set of methods that predict the class of an object from attributes or features describing the object [6,7]. In other words, classification relates to the problem of predicting the unknown class of an object. In classification, an object is classified into a pre-defined class using the features that distinguish it from other classes. In general, a classification system consists of two major stages: training and prediction. In the training stage, a training data set is used to train the system; In the prediction stage, the trained system is used to predict the unknown class of an object.

In clustering and classification, a data point or object can be any real item such as a company or a bank branch. An object is characterized by a set of features which are numerical variables such as total assets and total debt. Therefore, an object corresponds to a point in the d-dimensional feature space. A data set is usually described by an $n \times d$ matrix which contains a row for each of the n objects and a column for each of the d features [5,4].

In general, the application of classification for prediction consists of the following steps:

1. Collect data for known objects, i.e. the training set;
2. Train the clustering algorithm using the training set;
3. Use the trained algorithm to classify new objects and to predict their properties.

The remaining part of this paper is organized as follows. In Section 2, we briefly review the FSC method. In Section 3, we present experimental evaluation of the modified FSC algorithm. In Section 4, some conclusions are given.

2 The FSC Method

The FSC algorithm was proposed by Gan et. al. [3] to cluster high-dimensional data sets. In the FSC algorithm, each dimension of the original data is associated with each cluster by a weight. The higher density of a cluster in a dimension, the more weight will be assigned to that dimension.

Before introducing the FSC algorithm, we first define some notations. Let $D = \{\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n\} \subset \Re^d$ be a finite data set in the Euclidean space $\Re^d$ and k be an integer $2 \leq k < n$. A fuzzy k-partition of D can be represented by a real $k \times n$ matrix $U = (u_{ji})$ which satisfies

$$u_{ji} \in [0, 1], \quad 1 \leq j \leq k, \, 1 \leq i \leq n, \tag{1a}$$

$$\sum_{j=1}^{k} u_{ji} = 1, \quad 1 \leq i \leq n, \tag{1b}$$

$$\sum_{i=1}^{n} u_{ji} > 0, \quad 1 \leq j \leq k. \tag{1c}$$

A $k \times d$ matrix $W = (w_{jh})$ is said to be a fuzzy dimension weight matrix if W satisfies the following conditions

$$0 \leq w_{jh} \leq 1, \quad 1 \leq j \leq k, \quad 1 \leq h \leq d, \tag{2a}$$

$$\sum_{h=1}^{d} w_{jh} = 1, \quad 1 \leq j \leq k. \tag{2b}$$

The element w_{jh} specifies the probability of the dimension h belonging to the set of cluster dimensions of the cluster j.

Mathematically, the objective function of the FSC algorithm is defined as

$$E_{m,\alpha,\epsilon}(W, Z, U) = \epsilon \sum_{j=1}^{k} \sum_{h=1}^{d} w_{jh}^{\alpha} + \sum_{j=1}^{k} \sum_{i=1}^{n} u_{ji}^{m} \sum_{h=1}^{d} w_{jh}^{\alpha} (x_{ih} - z_{jh})^2, \tag{3}$$

where W, Z and U are the fuzzy dimension weight matrix, the center and the fuzzy k-partition of D, respectively, $m \in (1, \infty)$, $\alpha \in (1, \infty)$ is a weight component or fuzzier, and ϵ is a very small positive real number. It should be noted that any one of W, Z and U can be determined from the other two.

It can be shown that (W^*, Z^*, U^*) is a local minimum of $E_{m,\alpha,\epsilon}$ if and only if, for any $m > 1$, $\alpha > 1$ and $\epsilon > 0$, there holds

$$w_{jh}^* = \frac{1}{\sum\limits_{l=1}^{d} \left[\frac{\sum_{i=1}^{n} (u_{ji}^*)^m (x_{ih} - z_{jh}^*)^2 + \epsilon}{\sum_{i=1}^{n} (u_{ji}^*)^m (x_{il} - z_{jl}^*)^2 + \epsilon} \right]^{\frac{1}{\alpha - 1}}}, \tag{4a}$$

for $1 \leq j \leq k$ and $1 \leq h \leq d$,

$$z_{jh}^* = \frac{\sum\limits_{i=1}^{n} (u_{ji}^*)^m x_{ih}}{\sum\limits_{i=1}^{n} (u_{ji}^*)^m}, \tag{4b}$$

for $1 \leq j \leq k$ and $1 \leq h \leq d$, and

$$u_{ji}^* = \frac{1}{\sum\limits_{l=1}^{k} \left[\frac{d_{ji}}{d_{li}} \right]^{\frac{1}{m-1}}}, 1 \leq j \leq k, 1 \leq i \leq n, \tag{4c}$$

if $d_{li} = \sum\limits_{h=1}^{d} (w_{lh}^*)^\alpha (x_{ih} - z_{lh}^*)^2 > 0$ for all l, i. If $d_{li} = 0$ for some l, i, then u_{ji}^* can be any nonnegative real numbers satisfying

$$\sum_{j=1}^{k} u_{ji}^* = 1, \text{ and } u_{ji}^* = 0 \text{ if } d_{ji} \neq 0. \tag{4d}$$

3 Bankruptcy Prediction by the FSC Method

3.1 Modified FSC

Let $D = \{\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n\}$ and $D^* = \{\mathbf{x}_{n+1}, \mathbf{x}_{n+2}, ..., \mathbf{x}_{n+t}\}$ denote the training data set and the test data set, respectively. Objects in the training set are classified into two groups: bankrupted and non-bankrupted. Suppose the first group contains all bankrupted objects and the second group contains all non-bankrupted objects, we then specify the $u_1(\mathbf{x})$ and $u_2(\mathbf{x})$ of the object $\mathbf{x}$ in the training set as

$$u_1(\mathbf{x}) = \begin{cases} 1, \text{If } \mathbf{x} \text{ is bankrupted;} \\ 0, \text{Otherwise,} \end{cases} \tag{5}$$

$$u_2(\mathbf{x}) = 1 - u_{1i}. \tag{6}$$

Then we can train the the fuzzy dimension weight matrix and the centers w_{jh} and z_{jh} using the u_{ji}s specified above.

In details, in order to predict whether $\mathbf{y} \in D^*$ is bankrupted, we first need to calculate z_{jh}'s and w_{jh}'s according to Equation (4b) and (4a), respectively. That is,

$$z_{jh} = \frac{\sum\limits_{\mathbf{x} \in D} u_j(\mathbf{x})^m x_h}{\sum\limits_{i=1}^{n} u_j(\mathbf{x})^m}, \tag{7a}$$

for $1 \leq j \leq 2$ and $1 \leq h \leq d$, where $u_j(\mathbf{x})$'s are defined in Equation (5) and Equation (6), and

$$w_{jh} = \frac{1}{\sum\limits_{l=1}^{d} \left[\frac{\sum\limits_{\mathbf{x} \in D} u_j(\mathbf{x})^m (x_h - z_{jh})^2 + \epsilon}{\sum\limits_{\mathbf{x} \in D} u_j(\mathbf{x})^m (x_l - z_{jl})^2 + \epsilon} \right]^{\frac{1}{\alpha - 1}}}, \tag{7b}$$

for $1 \leq j \leq 2$ and $1 \leq h \leq d$, where u_{ji}'s are defined in Equation (5) and Equation (6), and z_{jh}'s are calculated as in Equation (7a). Using the z_{jh}'s in Equation (7a) and w_{jh}'s in Equation (7b), we calculate the memberships of $\mathbf{y}$ as

$$u_1(\mathbf{y}) = \frac{1}{1 + \left[\frac{d_1(\mathbf{y})}{d_2(\mathbf{y})} \right]^{\frac{1}{m-1}}}, \tag{8}$$
$$u_2(\mathbf{y}) = 1 - u_1(\mathbf{y}),$$

where

$$d_l(\mathbf{y}) = \sum_{h=1}^{d} w_{lh}^{\alpha} (y_h - z_{lh})^2, \quad l = 1, 2,$$

with w_{lh}'s and z_{lh}'s being defined in Equation (7b) and Equation (7a), respectively. Here we assume that $d_1(\mathbf{y}) > 0$ and $d_2(\mathbf{y}) > 0$. For real data set, this assumption is reasonable. Based on $u_1(\mathbf{y})$ calculated above, $\mathbf{y}$ is classified as follows. If $u_1(\mathbf{y}) > 0.5$ (equivalently $d_1(\mathbf{y}) < d_2(\mathbf{y})$), $\mathbf{y}$ is classified into the bankrupted group; if $u_1(\mathbf{y}) < 0.5$ (equivalently $d_1(\mathbf{y}) > d_2(\mathbf{y})$), $\mathbf{y}$ is classified as the non-bankrupted group; otherwise, the classification is inconclusive.

Once we classified an object in D^*, we move this object from the test data set to the training data set D and and update w_{jh}'s and z_{jh} according to Equation (7b) and Equation (7a), respectively. Then we repeat the aforementioned process to classify another object in D^* until all objects in the test data set D^* have been classified. The pseudo code of the algorithm is described in Algorithm 1.

3.2 The Data Set

The data set includes 303 companies and 28 of them went bankrupted one year later. Each company is described by 10 attributes, which includes total assets (TA), working capital (WC), earnings before income, tax, depreciation

Algorithm 1. The pseudo code of modified FSC for prediction of bankruptcy

```
Require: $D$ – the training data set, $D^*$ – the test data set;
repeat
    Calculate $z_{1h}$'s and $z_{2h}$'s according to Equation (7a);
    Calculate $w_{1h}$'s and $w_{2h}$'s according to Equation (7b);
    Calculate $u_1(\mathbf{y})$ according to Equation (8);
    Classify $\mathbf{y}$;
    if $u_1(\mathbf{y}) \neq 0.5$ then
        Move $\mathbf{y}$ from $D^*$ to $D$;
    else
        Remove $\mathbf{y}$ from $D^*$;
    end if
until $D^*$ is empty
```

and amortization (EBITDA), retained earnings (RE), shareholders equity (EQ), total current liabilities (CL), interest expense (IN), cash flow from operations (CF), stability of earnings (SE) and total liabilities (TL).

3.3 Experimental Evaluation

The modified FSC algorithm is coded in the MATLAB script language. Our experiments are conducted on a PC with 1.7G CPU and 512M RAM. In our experiments, we specify $\alpha = 2$, $m = 2$ and $\epsilon = 0.0001$.

For the given data set, we randomly select part of the data as training data and the remaining data as test data. Since the training data is selected randomly from the whole data set, we run the algorithm 100 times and calculate the average accuracy. The average accuracy is calculated as follows. Let p be the percentage of training data and n the number of objects in the training data set, then we have $n = [275p] + [28p]$, where $[a]$ denotes the largest integer less than or equal to a. Therefore, the number of objects in the test data set is $n_A = 303 - n$. Let $n_B = 28 - [28p]$ denote the number of bankrupted objects in the test data set, n_C be the number of objects classified correctly, and n_D the number of bankrupted objects classified correctly. The total accuracy and accuracy (Bankrupted) are defined as

$$R = \frac{n_C}{n_A},$$

and

$$r = \frac{n_D}{n_B},$$

respectively. The average total accuracy and average accuracy (bankrupted) of 100 runs are defined as

$$\bar{R} = \sum_{i=1}^{100} \frac{R_i}{100},$$

and

$$\bar{r} = \sum_{i=1}^{100} \frac{r_i}{100},$$

Table 1. Average prediction accuracy of 100 runs of the modified FSC algorithm with various percentages of training data

Training Data	Avg. Accuracy (Total)	Avg. Accuracy (Bankrupted)	Avg. Accuracy (Non-bankrupted)
95%	91.19%	35.00%	99.21%
90%	93.58%	40.33%	99.29%
85%	92.40%	38.20%	98.86%
80%	92.66%	34.33%	99.02%
75%	93.00%	34.71%	98.91%
70%	92.79%	34.11%	99.16%
50%	92.95%	33.36%	98.99%
20%	89.20%	26.09%	95.80%
10%	78.75%	21.62%	84.74%

respectively, where R_i and r_i are total accuracy and accuracy (Bankrupted) for the ith run.

The classification results of the algorithm are given in Table 1. We can easily observe that the proposed method provides an impressively high prediction accuracy for the non-bankrupted companies. However, the prediction accuracy rate for the bankrupted companies is farily low. This is not a surprising result since the data set only includes less than 10% bankrupted companies and the training process can not accurately recognize the features of the bankruptcy. If a balanced data set is used, the proposed algorithm will produce a very high prediction rate for both bankrupted and non-bankrupted companies. Nevertheless, the current result still provides significant insights to the industry. Our models' high levels of non-bankruptcy prediction accuracy can translate into large financial and other benefits to organizations through such activities as credit approval, and loan portfolio and security management.

4 Conclusion

In this paper we presented a fuzzy classification algorithm and used it for the purpose of bankruptcy prediction. The presented algorithm was modified from a subspace-clustering algorithm and adapted for object categorization. The most important advantage of our algorithm is that this it is able to classify an object without the need for feature selection. The implementation (empirical testing) of our algorithm demonstrates an exceptional and consistent predictability ratio of non-bankrupted objects averaging at about 99%. In terms of total accuracy, reaching 93%, the results produced by our algorithm are above average. Due to the small percentage of bankrupted companies (28 bankrupted objects among 303 objects), the results produced by our algorithm are not high in accuracy for bankrupted objects. We claim that given a more balanced data set with an appropriate level of bankrupted objects, an empirical test performed on this new data using modified FSC would yield an eq! uivalently high performance.

Our FSC algorithm affords opportunities for future research. The proposed methodology can be extended and used with other classification algorithm as a hybrid system to amplify the advantages of the algorithm and further improve its classification performance.

References

1. Balcean, S., Ooghe, H.: 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 (2006)
2. Beaver, W.: Financial ratios predictors of failure. Empirical research in accounting: selected studies 1966. Journal of Accounting Research (Suppl. 4), 71–111 (1967)
3. Gan, G., Wu, J., Yang, Z.: A fuzzy subspace algorithm for clustering high dimensional data. In: Li, X., Zaïane, O.R., Li, Z. (eds.) ADMA 2006. LNCS (LNAI), vol. 4093, pp. 271–278. Springer, Heidelberg (2006)
4. Gordon, A.: A review of hierarchical classification. Journal of the Royal Statistical Society. Series A (General) 150(2), 119–137 (1987)
5. Jain, A., Murty, M., Flynn, P.: Data clustering: A review. ACM Computing Surveys 31(3), 264–323 (1999)
6. Jain, A., Duin, R.P.W., Mao, J.: Statistical pattern recognition: a review. IEEE Transactions on Pattern Analysis and Machine Intelligence 322(1), 4–37 (2000)
7. Kulkarni, S.R., Lugosi, G., Venkatesh, S.S.: Learning pattern classification-a survey. IEEE Transactions on Information Theory 44(6), 2178–2206 (1998)
8. Ravi Kumar, P., Ravi, V.: Bankruptcy prediction in banks and firms via statistical and intelligent techniques - A review. European Journal of Operational Research 180(1), 1–28 (2007)
9. Min, J.H., Lee, Y.C.: Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters. Expert Systems With Applications 28(4), 603–614 (2005)
10. Tsakonas, A., Dounias, G., Doumpos, M., Zopounidis, C.: Bankruptcy prediction with neural logic networks by means of grammar-guided genetic programming. Expert Systems With Applications 30(3), 449–461 (2006)