Expert Systems with Applications 38 (2011) 12747–12756

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Predicting business failure using support vector machines with straightforward wrapper: A re-sampling study

Hui Li *, Jie Sun

School of Economics and Management, Zhejiang Normal University, Jinhua City 321004, Zhejiang Province, PR China

ARTICLE INFO

Keywords:
Business failure prediction (BFP)
Support vector machine (SVM)
Wrapper
Re-sampling study
Feature selection
Feature ranking

ABSTRACT

Business failure prediction (BFP) is an effective tool to help financial institutions and relevant people to make the right decision in investments, especially in the current competitive environment. This topic belongs to a classification-type task, one of whose aims is to generate more accurate hit ratios. Support vector machine (SVM) is a statistical learning technique, whose advantage is its high generalization performance. The objective of this context is threefold. Firstly, SVM is used to predict business failure by utilizing a straightforward wrapper approach to help the model produce more accurate prediction. The wrapper approach is fulfilled by employing a forward feature selection method, composed of feature ranking and feature selection. Meanwhile, this work attempts to investigate the feasibility of using linear SVMs to select features for all SVMs in the wrapper since non-linear SVMs yield to over-fit the data. Finally, a robust re-sampling approach is used to evaluate model performances for the task of BFP in China. In the empirical research, performances of linear SVM, polynomial SVM, Gaussian SVM, and sigmoid SVM with the best filter of stepwise MDA, and wrappers respectively using linear SVM and non-linear SVMs as evaluating functions are to be compared. The results indicate that the non-linear SVM with radial basis function kernel and features selected by linear SVM compare significantly superiorly to all the other SVMs. Meanwhile, all SVMs with features selected by linear SVM produce at least as good performances as SVMs with other optimal features.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

During the period of 1986–1991, bank failures averaged 175 per year (Barr & Siems, 1994). Many of these failures can be avoided if some people, such as: managers and auditors, recognized poor performance of these companies sufficiently early (Whalen, 1991). Business failure prediction (BFP) is an effective tool to help financial institutions and relevant people to make a decision on business performance of companies. The tool of BFP provides auditors and managers a chance to identify problems early. Thus, relevant people have an opportunity to intervene early into affairs of problems to reduce the expected cost of business failures (Sarkar & Sriram, 2001). The primary objective is to reduce as much as possible the loss of relevant people, such as: investors, managers and employees.

Prediction of business failure has received abundant interest from lots of researchers in finance, business and engineering. The research of predictive models for business failure started in the late 1960s and still be vital to this day. One of the early work of predicting business failure is conducted by Beaver (1966). He investigated the predictive abilities of 14 financial ratios by using 158 samples of failed and non-failed companies, with the conclusion that the single financial ratio of cash flow/total debt was the best predictor for distinguishing healthy companies from failed ones. Beaver's research belongs to the use of univariate discriminant approach for BFP. Altman (1968) attempted to use multivariate discriminant analysis (MDA) for BFP. According to the result provided by Altman, business failure can be explained effectively by using just five financial ratios out of 22 initial ratios. These five financial ratios have been widely used as input variables of various models since then. However, the early researches of using classical statistical models for BFP need some strict assumptions, such as normal distribution of data and normal inputs. In the late 1970s and early 1980s, Martin (1977) and Ohlson (1980) respectively attempted to use the model of logistic regression (logit) to predict business failure without needing strict assumptions. The model has good probabilistic interpretation, since the output of logit is between 1 and 0. These types of researches for BFP using classical statistical models are chiefly conducted from the aspect of finance and accounting. Nowadays, classical statistical models for BFP have been standard benchmarks in the area. However, assumptions of classical statistical models, such as: linearity, normality and variable independence, restrict their applications in real world. A common drawback of classical statistical models in BFP is the relative lower predictive accuracy rate than some newly developed models.

* Corresponding author. Tel.: +86 579 8228 8602.
E-mail address: li.1999@osu.edu (H. Li).

0957-4174/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2011.04.064

In order to exploit more accurate predictive models without assumptions of classical statistical models, some researches attempt to employ intelligent techniques, which are chiefly from the view of business and engineering. Frydman, Altman, and Kao (1985) proposed the application of recursive partitioning algorithm for BFP and compare it with MDA, with the conclusion that tree algorithms outperformed MDA. Tam (1991) used back propagation neural networks (BPNN) to predict business failure with data collected from Texas banks. The result is that BPNN offered better predictive accuracy than all the compared models, such as: MDA, logic, k-nearest neighbor and the tree algorithm of ID3. Becerra, Galvao, and Abou-Seada (2005) employed wavelet network models for BFP of British companies, with the conclusion that wavelet network model was a valid alternative to MDA. Li and Sun (2008, 2009) constructed two different hybrid case-based reasoning (CBR) methods to solve the task of BFP with data collected from Chinese listed companies. The results indicate that CBR was a valid alternative to predict business failure, though the early application of CBR in the problem did not achieve satisfactory results (Jo, Han, & Lee, 1997). Dimitras, Slowinski, Susmaga, & Zopounidis, (1999) used rough set theory to solve the task of BFP with data collected in Greece. The result indicates that the approach based on rough set outperformed MDA and logit. These researches are followed by lots of researchers (Ahn & Kim, 2009; Gepp, Kumar, & Bhattacharya, 2010; Hardle, Lee, Schafer, & Yeh, 2009; Hu & Ansell, 2009; Lee, 2007; Lin, Wang, Wu, & Chuang, 2009; Min & Lee, 2008; Psillaki, Tsolas, & Margaritis, 2010; Yeh, Chi, & Hsu, 2010; Yoon & Kwon, 2010), especially the issue of BFP based on neural networks (NN) (Chauhan, Ravi, & Chandra, 2009; Chen & Du, 2009; Chen, Huang, & Lin, 2009; Kim & Kang, 2010; Ravisankar & Ravi, 2010; Tsai & Wu, 2008; Tseng & Hu, 2010; Wilson & Sharda, 1994; Yang, Platt, & Platt, 1999).

In general, intelligent models yield to produce more accurate predictive performance than classical statistical models. Recently, support vector machine (SVM) is gaining popularity due to its attractive features and excellent generalization performance. The foundation of SVM is structural risk minimization principle, which is superior to traditional empirical risk minimization principle employed by NN (Bloch, Lauer, Colin, & Chamaillard, 2008; Cao, Xu, Chen, & Qiao, 2009; Celikyılmaz & Türkşen, 2007; Jayadeva, Khemchandani, & Chandra, 2008; Lingras & Butz, 2007; Maldonado & Weber, 2009; Vapnik, 1998; Yu & Li, 2008; Zhang & Wang, 2008). The former principle minimizes an upper bound of generalization error, which guarantees the global optimum. In contrast, the latter principle may fall into local optimum. Meanwhile, the problem of over-fitting of many machine learning algorithms is unlikely to occur with SVM (Cristianini & Shawe-Taylor, 2000). To fulfil the purpose of finding more accurate predictive models for BFP, SVM is an effective alternative. Wrapper approach is useful in helping classifiers produce accurate rate.

This work devotes to the investigation of predicting business failure by using SVM with a straightforward wrapper. Meanwhile, this work also attempts to investigate performances of using linear SVM and non-linear SVMs to select optimal feature subsets in the wrapper. When selecting optimal features, the wrapper approach is implemented by using SVM as the assessing function. The approach is composed of feature ranking under SVM and feature selection on the base of ranking information. Whether the linear SVM is more suitable to be employed in the wrapper than non-linear SVMs is also of interest. The four kernels of SVM, that is, linear kernel, polynomial kernel, radial basis function (RBF) kernel and sigmoid kernel are employed to make an insight research. The paper is organized as follows. Section 2 presents the review on related work of SVM-based BFP and the significance of this work. Section 3 specifies the use of SVM with straightforward wrapper, followed by the design of empirical study in Section 4. Section 5 presents the re-sampling results and discussion. Concluding remarks are provided in Section 6.

2. Related work and analysis

SVM is a valid alternative to predict business failure. An early application of SVM to predict business failure is conducted by Shin et al. (2005). They compared predictive performance of SVM with that of BPNN under the assessment of a hold-out method, with the conclusion that accuracy and generalization performance of SVM were better than those of BPNN. The data set, collected from Korea Credit Guarantee Fund, was split into two parts. One part, occupying 80%, was used for training and validating. The other, occupying 20%, was used as testing data set. Both SVM based on RBF kernel and BPNN were trained using the training and validating data set. For variable selection, they employed a two-stage feature selection process, which was composed of *t*-test and stepwise MDA in consecutive sequence. Min and Lee (2005) used grid-search technique and fivefold cross-validation to find optimal parameter values of SVM when solving the problem of BFP. The two classical statistical models of MDA and logit and the intelligent model of BPNN were employed to make comparisons. The results show that SVM outperformed all the three compared models with the data collected from the largest credit guarantee organization. For variable selection of SVM, they used principal components analysis. For comparative models, stepwise logit was used to select optimal features. The assessment of models was also under a hold-out method, where 80% of the whole data was used for training and 20% of the whole data was used for testing. For SVM, they tried all of the four kernel functions, that is, linear kernel, RBF kernel, polynomial kernel and sigmoid kernel, and found that RBF kernel was the optimal. Min, Lee, and Han (2006) constructed a hybrid model for BFP by integrating genetic algorithm with SVM. Genetic algorithm was used to improve predictive performance of SVM in two aspects: one for feature subset selection and the other for parameter optimization of SVM. The RBF kernel was used for SVM. Three filters, that is, *t*-test, stepwise MDA and stepwise logit, were used to make comparisons. The whole data set was spilt into two parts, that is, 30% for testing and 70% for training and validating. The results indicate that SVM with genetic algorithm outperformed NN, logit and SVM with filters.

Hui and Sun (2006) firstly attempted to investigate the feasibility of applying SVM to predict business failure of Chinese listed companies. The contribution was the use of accuracy rate produced by cross-validation as the assessment. Grid-search technique was also employed to search optimal values of SVM with RBF kernel as the heart. Optimal features were picked out by using the filter of stepwise MDA. The results indicate that SVM was superior to NN, MDA and logit. Wu, Tzeng, and Goo (2007) investigated the predictive performance of employing genetic algorithm only to optimize parameter values of SVM with RBF kernel. Nineteen features that were used in previous researches as significant in predicting business failure were picked out to construct the SVM model. Hold-out method was again utilized as the assessment. The findings show that SVM with genetic algorithm was superior to logit, MDA, NN and pure SVM. Hua, Wang, Xu, Zhang, and Liang (2007) also attempted to apply SVM to predict business failure of China. SVM was integrated with logit to generate a forecasting. Optimal features were selected by using a two-stage approach, which firstly used *t*-test and then used univariate discriminant approach to choose variables. The findings, supported by cross-validation, mean that the presented SVM model outperformed initial SVM models. In very recent, Ding, Song, and Zeng (2008) re-applied SVM for BFP in China, which again provided the evidence from hold-out method that SVM with RBF as the heart outperformed

Download English Version:

https://daneshyari.com/en/article/388541

Download Persian Version:

https://daneshyari.com/article/388541

Daneshyari.com