ARTICLE IN PRESS

Tourism Management xxx (2011) 1–13

Contents lists available at ScienceDirect

Tourism Management

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

Forecasting business failure: The use of nearest-neighbour support vectors and correcting imbalanced samples – Evidence from the Chinese hotel industry

Hui Li [a,b,*], Jie Sun [a]

[a] School of Economics and Management, Zhejiang Normal University, P.O. Box 62, 688 YingBinDaDao, Jinhua, Zhejiang 321004, PR China
[b] College of Engineering, The Ohio State University, 470 Hitchcock Hall, 2070 Neil Avenue, Columbus, OH 43210, USA

ARTICLE INFO

Article history:
Received 23 September 2010
Accepted 8 July 2011

Keywords:
Firm failure prediction
Tourism imbalanced dataset
Nearest-neighbour support vector machine
Bagging ensemble
Tourism risk forecasting

ABSTRACT

Previous studies on firm failure prediction (FFP) have chiefly addressed predictions based on balanced datasets without considering that the real-world target population consists of imbalanced data. The current study investigates tourism FFP based on the imbalanced data of Chinese listed companies in the hotel industry. The imbalanced dataset was collected and represented in terms of significant financial ratios, and a new up-sampling approach and forecasting method were proposed to correct imbalanced samples. To balance the imbalanced dataset, the up-sampling method generates new minority samples according to random percentage distances from each minority sample to its nearest neighbour (NN). The NNs of unlabelled samples are retrieved from the balanced dataset to produce a knowledge base of nearest-neighbour support vectors, from which base support vector machines (SVMs) are generated and assembled. Empirical results indicate that the proposed sampling approach helped models produce more accurate performance on minority samples, with accuracy rates in excess of 90 per cent. This method of using nearest-neighbour support vectors and correcting imbalanced samples is useful in controlling risk in tourism management.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

An increasing number of hotels in the US and China face the possibility of business failure. Particularly in the years immediately following the global financial crisis, the US hotel industry braced itself for failure. For example, Ilikai Hotel and Suites, the first luxury high-rise hotel in Hawaii, faced foreclosure and bankruptcy in 2009, and the owner of the Riviera Hotel and Casino on the Las Vegas Strip filed for bankruptcy protection in 2010. Other US hotels that have faced foreclosure in the past two years include the Renaissance Grand and Suites Hotel in St. Louis, Missouri, the W Hotel and InterContinental Montelucia Resort and Spa in Scottsdale, Arizona, the Ritz-Carlton Lake and Tropicana in Las Vegas, Nevada, and the Sheraton Downtown in Orlando, Florida, among others.

A similar situation has occurred in the hotel industry in China. China has approximately 15,000 star-level hotels out of more than 360,000 lodging service providers. For the past two years, their average occupancy rate was about 50%, although this rate improved to 60.22% in the second season of 2010 because of the Shanghai World Expo. The occupancy rate of hotels in China was 60.5% during the depression in the US hotel industry in 2001, when the 9/11 tragedy occurred, and it fell to 60% between 1998 and 2004 due to the oversupply in the Korean lodging industry (Youn & Gu, 2010a). Oversupply resulted in hotel failures in China. For example, XiaoShan International Hotel, the top hotel in XiaoShan, filed for bankruptcy in 2010. Many of these failures could have been avoided if the management of these firms had been aware of their poor performances at an earlier stage (Sarkar & Sriram, 2001). The development of firm failure prediction (FFP) models for the tourism industry benefits managers, customers, investors, and government officials by reducing loss among hotel-related businesses.

FFP refers to forecasting whether a firm will fail prior to one year ($t-1$), two years ($t-2$), or three years ($t-3$) by modelling financial or/and non-financial information. Two key factors of FFP include data preparation and model construction. Although more hotels face the prospect of failure, the number of failed firms is far smaller than the number of firms that do not fail. In the US data source used by Ryu and Yue (2005), only 2% of the firms in the sample failed. A dataset is imbalanced if the classification categories are not represented approximately equally (Chawla et al. 2002). Thus, the real-world tourism population of FFP consists of an imbalanced dataset. Previous research on FFP has primarily focused on balanced datasets by

* Corresponding author. School of Economics and Management, Zhejiang Normal University, P.O. Box 62, 688 YingBinDaDao, Jinhua, Zhejiang 321004, PR China. Tel.: +86 579 8229 8602; fax: +86 579 8229 8567.
E-mail addresses: lihuihit@gmail.com (H. Li), sunjiehit@gmail.com (J. Sun).

0261-5177/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tourman.2011.07.004

Please cite this article in press as: Li, H., & Sun, J., Forecasting business failure: The use of nearest-neighbour support vectors and correcting imbalanced samples – Evidence from the Chinese hotel industry, *Tourism Management* (2011), doi:10.1016/j.tourman.2011.07.004

ARTICLE IN PRESS

matching each failed sample with a corresponding non-failed sample. The current study investigates tourism FFP based on the imbalanced data of listed Chinese hotel companies, including almost the entire real-world population of FFP. This study describes a new up-sampling approach, a minority-samples generating approach based on a random percentage distance to the nearest neighbour (MSGA-RPD-NN), and a forecasting method, the nearest-neighbour support vector machine (NNsSVM). These contributions address imbalanced dataset processing and business forecasting modelling. The current study contributes to controlling risks in tourism-related business decisions by providing more accurate FFP performances. Furthermore, this study adds to the existing tourism literature by providing an approach to imbalanced dataset processing, a predictive model for FFP, and an analysis of the failures of listed Chinese hotel businesses.

This paper is organised as follows. Section 2 reviews previous studies on FFP based on balanced and imbalanced datasets within and outside the tourism industry. Section 3 presents the imbalanced data and variable information on FFP in listed Chinese tourism companies with hotels. The new up-sampling approach for imbalanced dataset processing and the modelling method are discussed in Section 4. Section 5 presents the results and discussion on the performance of the method based on nearest-neighbour support vectors and imbalanced samples corrected for $t-2$ and $t-3$ FFPs. SVM, bagged SVM (BSVM), NNs, bagged NNs (BNNs), logistic regression (logit), and multivariate discriminate analysis (MDA) are benchmarks. Section 6 presents the conclusions and the limitations of the study.

2. Literature review

FFP in the tourism industry is a specific type of business forecasting. Various models have been constructed based on business data to facilitate risk control in tourism-related decisions. Failed samples are initially identified when collecting datasets. Non-failed samples are matched with failed samples to generate a balanced dataset, in which positive/non-failed and negative/failed samples are approximately equally represented. Negative samples are not always matched with positive samples. This approach to data handling dominates research and applications of FFP both within and outside the tourism industry. A dataset can be classified as balanced or imbalanced. In our review of the datasets of Letter, Glass, Image, Vehicle, Wine, Iris, Waveform, and Pima provided by the Machine Learning Repository of the University of California, Irvine (UCI), we found that minority samples constituted 3.95%, 7.94%, 14.29%, 23.52%, 26.97%, 33.33%, 33.33%, and 34.77% of the datasets, respectively. Thus, we formulated a definition of an imbalanced dataset: if minority samples constitute less than 35% of a dataset, the dataset is classified as imbalanced. If minority samples constitute less than 10% of a dataset, the dataset is classified as highly imbalanced.

2.1. FFP with balanced datasets outside of the tourism industry

1) **US FFP**. To determine FFP in the US outside of the tourism industry, Altman (1968) collected data on 33 bankrupt manufacturers for 1945–1965 and matched them with 33 non-bankrupt firms according to industry and assets. An MDA model was employed in addition to the 33 pairs of samples to predict $t-1$ bankruptcy with 95% in-sample accuracy and 70% out-of-sample accuracy. The paired samples were used by another early investigation on FFP, Beaver (1966), in which 79 failed and 79 non-failed samples were collected. Balanced datasets have frequently been used in recent FFP studies. To test the applicability of isotonic separation in US FFP, Ryu and Yue (2005) identified 88, 109, and 104 failed firms from a US database and matched them with equal numbers of non-failed firms. The results of a 10-fold cross-validation indicated that their approach produced accuracy levels of 85.8% $t-1$, 84.86% $t-2$, and 81.25% $t-3$, which were better than the levels produced by MDA, logit, probit, neural network, decision tree, and rough set. Pendharkar (2005) identified 100 bankrupt and 100 non-bankrupt samples and constructed a threshold-varying neural network for FFP. Click-and-mortar firms have contributed significantly to the current economy. Bose and Pal (2006) collected 120 failed and 120 non-failed samples from Wharton Research Data Services and applied MDA, neural network, and SVM for the predictions. Their models generated about 75% testing accuracy on the balanced dataset. The same dataset was used by Ravisankar, Ravi, and Bose (2010), Chandra, Ravi, and Bose (2009), and Bose (2006) to investigate the applicability of a neural network enhanced by GA, hybrid intelligent techniques, and rough sets. McKee (2002) and McKee and Lensberg (2002) investigated FFP with a rough set on a balanced US dataset consisting of 146 bankrupt and 145 non-bankrupt samples.
2) **European FFP**. Tsakonas et al. (2006) used a dataset of 59 failed and 59 non-failed samples for Greek FFP. The balanced dataset was divided into 80 samples for training and 38 samples for testing according to the years when they were bankrupt. A neural logit network enhanced by genetic programming (GP) was used for modelling and produced 93.3% in-sample accuracy and 76.3% out-of-sample accuracy. This dataset was used in Dimitras et al. (1999) to verify Greek FFP with rough sets. Alfaro et al. (2008) collected 590 failed and 590 non-failed firm samples for 2000–2003 from the SABI database of Bureau Van Dijk and used Adaboost for FFP with 92.4% $t-1$ training accuracy and 91.1% $t-1$ testing accuracy. Of the balanced dataset, 80% of it was used as a training set, and the other 20% was used as a testing set. Tseng and Hu (2010) compared the performance of 35 failed and 45 non-failed samples using logit, quadratic interval logit, neural, and fuzzy neural networks for the UK FFP. In-sample and validating accuracies indicated that fuzzy neural networks outperformed the benchmarks.
3) **Asian FFP**. Using a dataset of 1160 bankrupt and 1160 non-bankrupt firms from Korea, Shin, Lee, and Kim (2005) investigated FFP with SVM. The balanced dataset was subjectively divided into two parts: 80% as a training set and 20% as a testing set. Empirical results indicated that SVM outperformed the neural network by optimising the kernel parameters of SVM with a grid search. Min and Lee (2005) forecasted Korean firm failure with a dataset of 944 failed and 944 non-failed samples. The results showed that SVM outperformed MDA, logit, and neural networks. A similar study was conducted by Hua et al. (2007) with a Chinese balanced dataset consisting of 60 negative and 60 positive samples. Min, Lee, and Han (2006) attempted to improve the performance of SVM-based FFP by GA with a Korean balanced dataset consisting of 307 failed and 307 non-failed samples. The results showed that SVM with optimised features and samples outperformed MDA, logit, and basic SVM. Similarly, Wu, Tzeng, and Goo (2007) investigated the use of GA to optimise kernel parameters for Taiwan FFP with a balanced dataset. The results indicated that optimised SVM outperformed MDA, logit, probit, neural network, and basic SVM. Tsai (2008) compared the performances of neural network and SVM. The results from a balanced FFP dataset consisting of 112 failed and 128 non-failed samples indicated that SVM outperformed neural network. Sun and Li (2008) investigated the combination of SVM with MDA, logit, and neural network for $t-2$ FFP with a balanced dataset and showed that the combination could improve the performance of a single model. Besides the popularity of SVM in Asian FFP, Li and Sun (2008, 2009, 2010) and Li et al. (2011) investigated the use of case-based reasoning with NNs to predict Chinese firm failure with three balanced datasets. For further discussion on

Please cite this article in press as: Li, H., & Sun, J., Forecasting business failure: The use of nearest-neighbour support vectors and correcting imbalanced samples – Evidence from the Chinese hotel industry, *Tourism Management* (2011), doi:10.1016/j.tourman.2011.07.004

ARTICLE IN PRESS

FFPs with balanced datasets, refer to the reviewed articles in Kumar and Ravi (2007).

2.2. FFP with balanced datasets in the tourism industry

Previous studies of FFP in the tourism industry have employed balanced samples, with the following results.

1) **US tourism FFP**: Olsen, Bellas, and Kish (1983) identified 7 failed and 12 non-failed restaurant firms to generate a balanced dataset in which the minority samples accounted for 36.84% of the total data. In recent years, FFP in the tourism industry has attracted more attention. Gu (2002) matched 18 bankrupt restaurant firms from Standard & Poor's Compustat for 1986–1998 with 18 non-failed restaurant firms in terms of assets. They achieved 92% $t-1$ in-sample accuracy and 80% $t-1$ out-of-sample accuracy with an MDA model. With the same balanced dataset, Kim and Gu (2006a) employed a logit model that produced 94% in-sample accuracy and 93% out-of-sample accuracy. With a balanced dataset consisting of 16 pairs of bankrupt and non-bankrupt samples, Kim and Gu (2006b) investigated hospitality FFP with logit, which produced 91% $t-1$ accuracy and 84% $t-2$ accuracy. Youn and Gu (2010b) collected 31 failed bankrupt and 31 non-failed restaurant firms between 1996 and 2008 from New Generation Research Inc. for US restaurant FFP. Logit and neural networks produced (88.1%, 76.19%) and (88.1%, 78.57%) in-sample accuracies, respectively, and (95%, 85%) and (95%, 80%) out-of-sample accuracies prior to one and two years, respectively, for the balanced datasets.
2) **Asian tourism FFP**: For the Korean lodging FFP, Youn and Gu (2010a) identified 102 failed lodging firms from a financial supervisory service database for 2000–2005 and used one-by-one matching to collect 102 non-failed samples. They achieved 83.33% in-sample accuracy and 77.27% out-of-sample accuracy with a logit model, and they achieved 87.75% in-sample accuracy and 81.82% out-of-sample accuracy with a neural network model for $t-1$ FFP.

2.3. FFP modelling on imbalanced datasets outside the tourism industry

Of the studies that have used imbalanced datasets for FFP, only a few have employed highly imbalanced samples with a maximum of 10% minority samples. All of these studies have focused on FFP outside the tourism industry. However, the term "imbalanced" has rarely appeared in these studies, and none of them has treated available samples as imbalanced datasets by using an up-sampling approach to generate new minority samples. Instead, previous studies have commonly employed probabilistic models in which the cut-offs were used as the parameters to turn predictions rather than setting them at the default value of 0.5. The evidence is as follows.

1) **US FFP**: Sarkar and Sriram (2001) used imbalanced data with 1139 samples for bank FFP, of which approximately 89% of the samples were positive. Their objective was to apply the two probabilistic models, naïve Bayes and a composite attributes model, in forecasting bank failures. The models change output predictions by using different thresholds of probability. All data were divided into a training set and a testing set in subjectively pre-determined proportions. Hwang, Cheng, and Lee (2007) collected 79 failed US samples and matched them with 156 non-failed samples in 1994–2002. They generated about 95% $t-1$ total accuracy with a straightforward use of a semi-parametric logit model on the dataset. Gepp, Kumar, and Bhattacharya (2010) used an imbalanced dataset consisting of 58 failed and 142 non-failed samples during 1971–1981 for US FFP. The results from a five-fold cross-validation indicated that simpler models had better predictive powers.
2) **Oceania FFP**: Jones and Hensher (2004) collected 2838 non-failed firm samples, 78 samples with solvency difficulties, and 116 samples in failure for 1996–2000 for Australian FFP. These samples were used as a training set. Furthermore, 4980 non-failed samples, 119 samples with solvency difficulties, and 110 samples in failure were obtained as a testing set for 2001–2003. Their results proved that a multi-nominal logit model performed poorly in FFP (only 5% in-sample accuracy and a maximum of 6.4% out-of-sample accuracy) and in forecasting firms with solvency difficulties (24% in-sample accuracy and a maximum of 29% out-of-sample accuracy). In comparison, the mixed logit model generated a maximum of 95% in-sample and out-of-sample accuracies in forecasting failed samples and over 90% in-sample and out-of-sample accuracies in forecasting firms with solvency difficulties.
3) **Asian FFP**: Ding, Song, and Zeng (2008) collected 56 failed and 194 non-failed samples from Chinese listed companies and subjectively divided the dataset into a training set and a testing set. The experiment results indicated that SVM outperformed MDA, logit, and neural network by generating 83.2% out-of-sample total accuracy.
4) **European FFP**: Focussing on FFP with rule-extraction techniques, Martens et al. (2010) collected 74 failed and 348 solvent samples for 1989–1997 and modelled these with AntMiner+. Although AntMiner+ performed poorly compared with SVM, AntMiner+ can generate predictive rules. Without processing the imbalanced dataset, the predictive model yielded low prediction on minority samples if the threshold of probability was not properly set. Gestel et al. (2006) collected 74 bankrupt firms and 348 solvent firms from a major Benelux financial institution to create an imbalanced dataset. The receiver operating characteristic curve was used to assess the performance of the hybrid of the Bayesian evidence framework with least squares SVM, namely, kernel Fisher MDA, for FFP, with an average accuracy of 90%. Different threshold values of probability were used to adjust the predictions. Similar studies were conducted by Gaganis, Pasiouras, and Doumpos (2007) and Hu and Ansell (2007, 2009). Gaganis and his colleagues applied the probabilistic neural networks in FFP to a UK dataset consisting of 175 negative and 2215 positive samples as a training set for 1997–2002 and 89 negative and 1029 positive samples as a testing set for 2003–2004. Hu and Ansell applied sequential minimal optimisation, naïve Bayes, logit, recursive partitioning, and neural network in the US FFP with a dataset consisting of 51 failed and 195 non-failed samples.

The limitation of using the models directly for imbalanced datasets is that the cut-off of probability should be determined before prediction. Without pre-determining the threshold of probability, the FFP tool based on these models cannot produce predictions, and without properly setting the cut-off as an optimal value, a predictor's performance is low. The determination of the threshold value is difficult in real-world applications. McKee and Greenstein (2000) collected a huge imbalanced dataset based on the real proportions of failed firms. Five pairs of training and testing datasets were used, consisting of [(54, 288) (12, 1154); (66, 336) (16, 2683); (59, 240) (19, 2468); (72, 292) (19, 2643); and (76, 309) (18, 3396)] failed and non-failed samples. The results indicated that ID3, neural network, and logit with default values of thresholds all produced poor performances in terms of forecasting failed firms compared with predicting non-failed samples. By searching the optimal values of thresholds of probability, ID3 could generate about 80% Type I and Type II accuracies.

Please cite this article in press as: Li, H., & Sun, J., Forecasting business failure: The use of nearest-neighbour support vectors and correcting imbalanced samples – Evidence from the Chinese hotel industry, *Tourism Management* (2011), doi:10.1016/j.tourman.2011.07.004

2.4. Significance of the current study

Most previous studies have used balanced samples for FFPs outside the tourism industry, and all previous studies have used balanced datasets for FFPs in the tourism industry. Among the few studies that have used imbalanced datasets, imbalanced samples have seldom been treated as imbalanced. However, the target population of FFP consists of an imbalanced dataset in which failed samples are the minority and non-failed samples are the majority. A balanced sample set is a subset of the entire population. FFP models constructed on samples may not be effective for the entire population. If the value of the (volume of samples)/(volume of population) is large, the estimated performance of a predictor will be more accurate. Thus, the imbalanced sample set should be used, and its volume should be approximate to the volume of the population. When assessing the performance of models, previous studies have used subjective divisions of entire datasets into one training set and one testing set. Because the proportion of training sets to testing sets was determined subjectively, these studies were unable to estimate the performance of a predictor based on the availability of training samples.

To increase the effectiveness of FFP, the current study seeks to conduct a timely investigation on forecasting tourism firm failure for nearly the entire real-world imbalanced population of Chinese listed companies in the hotel industry. This study proposes a new up-sampling approach to manage the real-world imbalanced dataset and a new model to enable the tool to produce more accurate predictions. The rationale of the current study is as follows.

1) The estimation of the performance of a predictor is more accurate because nearly the entire population of Chinese listed companies in the hotel industry will be modelled and assessed.
2) The use of an up-sampling approach to the imbalanced dataset before modelling is more realistic than the use of balanced samples because the real-world target population is imbalanced. This approach can be integrated in a straightforward manner with other datasets, assessments, and models to forecast firm failure.
3) Based on the literature review, SVM is superior to previously constructed models for FFP. Only support vectors are useful in modelling SVM. The use of the NN technique to select effective support vectors can enable SVM to perform better.
4) The use of 20%, 35%, 50%, 65%, and 80% of the entire available dataset in modelling provides a more thorough and more accurate assessment of the performance of a predictor. The use of the Wilcoxon significance test in analysing the model performance renders the estimation of the performance of a predictor more statistically significant.

A recent study (Hardle et al., 2009) that treated a real-world imbalanced dataset as imbalanced collected a dataset of 811 bankrupt firms and 10,468 non-bankrupt firms for 1997–2002. Smooth SVM (Lee & Mangasarian, 2001) with adjustments was applied to the German FFP dataset under hold-out methods and was repeated five times *according to year*; it generated a 75% $t-1$ hit ratio on total, Type I, and Type II accuracies. There are several differences between previous studies and the current study. First, previous studies have adopted an oversampling approach to duplicate minority samples. Duplication of minority samples in an imbalanced dataset is not applicable when the multiple cross-validation or multiple hold-out method is used on *panel data without considering the year* of the samples and the number of minority samples is too small. The duplication approach will cause training and testing datasets to employ the same samples. For example, if a sample is (1,1,1,1,1), duplication of the minority samples will generate a sample (1,1,1,1,1). If the dataset is divided into training and testing datasets, the sample (1,1,1,1,1) may appear in the training and testing datasets. Only support vectors are useful in modelling SVM. The duplicated samples seldom contribute to SVM modelling if no further approaches are integrated. Second, the current study uses a new forecasting model, NNsSVM in terms of SVM-based hotel FFP. Third, we focus on FFP with the current real-world imbalanced dataset from Chinese listed companies in the hotel industry. With the proposed model and approach to processing an imbalanced dataset, our paper proposes a pioneering investigation of tourism FFP. These differences make our study valuable to the current literature.

3. Imbalanced data and variables for FFP in the tourism industry

3.1. Imbalanced datasets

The data sources for FFP in the Chinese tourism industry are the Shanghai Stock Exchange, the ShenZhen Stock Exchange, and the China Stock Market and Accounting Research Database. Currently, there are 30 listed companies in the tourism and hotel sector of the two stock exchanges. By deleting 7 companies not engaged in the hotel industry, we obtained 23 listed companies, which are presented in Table 1. The listed year refers to the year the company was listed on the Shanghai Stock Exchange or the ShenZhen Stock Exchange. The failed year refers to the year a company was assessed as failed by the China Securities Supervision and Management Committee (CSSMC). Tourism firm failure refers to three different concepts: economic failure, technical insolvency, and bankruptcy (Gu, 2002; Youn & Gu, 2010a, 2010b). The economic failure of a tourism firm is declared by the CSSMC if the firm has a negative net income for two consecutive years. Non-failed firms refer to tourism companies that have not been declared failures. The period of the study was 1998–2010 because there were too many missing values in the statements of listed tourism companies prior to 1998 and too few tourism companies were listed prior to 1998.

Of the sample, 7 failed and 16 non-failed tourism firms with hotel businesses were identified. Non-failed firms with hotel businesses were represented annually from 1998 to 2010 (as long as they were listed) because the operating information of a non-failed firm in each year denotes a different non-failed sample. Traditionally, studies have extracted one non-failed sample from

Table 1
Chinese tourism listed companies with hotel business in the year span of 1998–2010.

Number	Firm name	Listed year	Failed year
1	Yellow Mountain Tourism	1997	–
2	*Tibet Tourism*	*1996*	*2003*
3	Emei Mountain A	1997	–
4	Guilin Tourism	2000	–
5	Lijiang Tourism	2004	–
6	*Zhang Stock*	*1996*	*2001; 2007; 2010*
7	Dalian ShengYa	2002	–
8	City of Overseas Chinese A	1997	–
9	Yunnan Tourism	2006	–
10	United National Tourism	2000	–
11	*Zero Seven*	*1992*	*2001; 2007*
12	*Wan Hao Wan Jia*	*2003*	*2006*
13	Jinjiang Stock	1996	–
14	Jinling Hotel	2007	–
15	*New City Hotel*	*1994*	*2001*
16	China Sky Hotel	1996	–
17	Eastern Hotel	1993	–
18	Xian Tourism	1996	–
19	*Eastern Sea A*	*1997*	*2001; 2005*
20	Xian Yin Shi	1997	–
21	China Young Tourism	1997	–
22	Capital Tourism Stock	2000	–
23	*Peking Tourism*	*1998*	*2004*

Please cite this article in press as: Li, H., & Sun, J., Forecasting business failure: The use of nearest-neighbour support vectors and correcting imbalanced samples – Evidence from the Chinese hotel industry, *Tourism Management* (2011), doi:10.1016/j.tourman.2011.07.004

one non-failed tourism firm to create a balanced dataset. The number of non-failed tourism firms is much larger than the number of failed tourism firms. Traditional studies have adopted this method to include as many non-failed tourism firms as possible in modelling. The current study extracted as many samples as possible from a firm. From the failed and non-failed tourism firms, 11 failed samples and 159 non-failed samples were extracted for $t-1$ FFP, and 16 and 32 non-failed samples representing information from 2009 and 2008–2009, respectively, were deleted for $t-2$ and $t-3$ FFP modelling. Of the available tourism samples, the minority accounted for 6.47%, 7.14%, and 7.97%. These samples consisted of the current real-world imbalanced data for Chinese listed tourism companies in the hotel business.

3.2. Variables

The financial ratios of failed and non-failed tourism samples were used to represent the dataset. To date, the most popular and widely used variable set that effectively distinguishes $t-1$, $t-2$, and $t-3$ failed samples from non-failed samples is the five-variable set suggested by Altman and Hotchkiss (2005), although subsequent studies have indicated that other financial or non-financial ratios might be useful in specific tourism FFP applications. These five significant variables are listed in Table 2. Atiya (2001) offers several reasons why the five variables should be adopted. The total assets of a tourism firm consist of current and long-term assets. Total assets indicate the size of a firm, so they can be used as a normalising factor. The liabilities of a tourism firm consist of current liabilities and long-term debts. Current assets minus current liabilities are equal to working capital, which indicates a firm's ability to pay its short-term obligations. The ratio of retained earnings indicates a tourism firm's accumulated earnings, which, together with market capitalisation, is a useful indicator to assess the health of a firm. The ratio of market capitalisation to total assets indicates the ability of a tourism firm to issue and sell new shares in the market to repay its debts. Sales indicate whether a tourism firm is healthy in terms of its business. These ratios are helpful in identifying tourism firm failure in theory.

Table 2 presents the statistical details of the current real-world imbalanced samples on the significant financial variables for non-failed firms, $t-3$ failed firms, $t-2$ failed firms, and $t-1$ failed firms in the Chinese tourism industry. A failing, non-failed firm will become a $t-3$ failed firm, $t-2$ failed firm, and $t-1$ failed firm before becoming a failed firm. Variation trends of variable means from non-failed samples to failed samples are illustrated in Fig. 2. The Shapiro–Wilk test was used to ascertain whether the five variables distribute normally. The results indicate that they

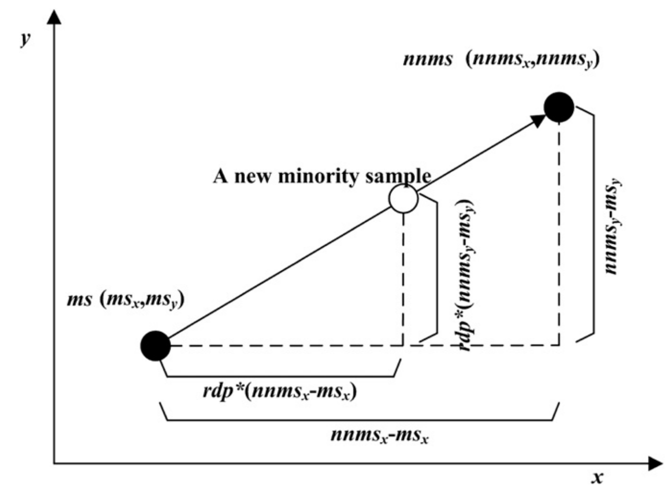


Fig. 1. Mechanism of generating a new minority sample in the up-sampling approach.

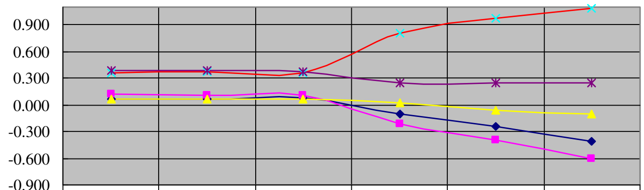
distribute non-normally except for the first variable for $t-3$, $t-2$, and $t-1$ tourism FFP and the third variable for $t-3$ FFP. Ascertaining the high correlation of these selected variables is important because these variables are used in modelling. Spearman's correlation coefficient was used to test the high correlation of each pair of variables with Altman's five variables. If the absolute value of Spearman's correlation coefficient is less than 0.8, the pair of variables is assumed to be not highly correlated. The results indicate that all absolute values of Spearman's correlation coefficient are less than 0.613; thus, these ratios are not highly correlated. The non-parametric Wilcoxon significance test was used to determine whether the five variables can significantly distinguish failed tourism from non-failed tourism samples, except for the *t*-test that was used on *V*1 for $t-2$ and $t-3$ and on *V*3 for $t-3$. Levene's test was used to test the homogeneity of variances of the variables, and the results rejected the null hypothesis. The results of the significance test are provided in Table 2. The findings are as follows.

1) Statistical information on the five variables indicates the behaviour of the failed tourism firms with hotel businesses. The $t-3$ and $t-2$ failed samples in the Chinese tourism industry have smaller values on *V*1, *V*2, *V*3, and *V*5 and a larger value on *V*4 than non-failed samples. When the failed samples are closer to the failed years, they behave worse on the first four variables and are stable in terms of sales to total assets. If a tourism sample is predicted as failed, the corresponding firm should

Table 2
Statistical details of the imbalanced samples for FFP of Chinese tourism industry.

	Working capital to total assets (*V*1)	Retained earnings to total assets (*V*2)	Earnings before interest and tax to total assets (*V*3)	Total debts-to-market capitalisation (*V*4)	Sales to total assets (*V*5)
Mean ± SD of $t-1$, $t-2$, and $t-3$ non-failed samples	0.059 ± 0.215 0.064 ± 0.219 0.075 ± 0.219	0.116 ± 0.074 0.112 ± 0.071 0.108 ± 0.069	0.065 ± 0.049 0.064 ± 0.049 0.064 ± 0.050	0.363 ± 0.394 0.369 ± 0.410 0.355 ± 0.422	0.387 ± 0.226 0.383 ± 0.220 0.379 ± 0.215
Mean ± SD of $t-3$ failed samples	−0.100 ± 0.325	−0.213 ± 0.671	0.023 ± 0.069	0.800 ± 0.841	0.243 ± 0.155
Sig.	*1.76*[a] *(0.107*[b]*)*	2.81 (0.005)***	*1.93 (0.081)**	2.89 (0.004)***	2.38 (0.018)**
Failing sample indicator	−	−	−	+	−
Mean ± SD of $t-2$ failed samples	−0.238 ± 0.390	−0.394 ± 0.935	−0.057 ± 0.098	0.977 ± 0.698	0.252 ± 0.206
Sig.	*2.55 (0.028)***	4.78 (0.000)***	4.80 (0.000)***	3.37 (0.001)***	2.38 (0.018)**
Failing sample indicator	−−	−−	−−	++	−+
Mean ± SD of $t-1$ failed samples	−0.410 ± 0.339	−0.606 ± 0.968	−0.096 ± 0.074	1.090 ± 0.892	0.247 ± 0.250
Sig.	*4.42 (0.000)****	5.51 (0.000)***	5.33 (0.000)***	3.39 (0.001)***	2.88 (0.004)***
Failing sample indicator	−−−	−−−	−−−	+++	−+−

+: The mean value of failed samples is lager, −: the mean value of failed samples is smaller, ***: significant at 1% level, **: significant at 5% level, *: significant at 10% level.
[a] *t* Value or *Z* value.
[b] *p* Value.



	t-1 non-failed samples	t-2 non-failed samples	t-3 non-failed samples	t-3 failed samples	t-2 failed samples	t-1 failed samples
V1	0.059	0.064	0.075	-0.100	-0.238	-0.410
V2	0.116	0.112	0.108	-0.213	-0.394	-0.606
V3	0.065	0.064	0.064	0.023	-0.057	-0.096
V4	0.363	0.369	0.355	0.800	0.977	1.090
V5	0.387	0.383	0.379	0.243	0.252	0.247

Fig. 2. Trends of variable means from non-failed samples to failed samples.

attempt to improve its performance in terms of *V*1, *V*2, *V*3, *V*4, and *V*5 by maintaining more working capital, more earnings and retained earnings, more sales, and fewer debts.

2) Compared with non-failed samples, failed samples with hotel businesses in the Chinese tourism industry have a large proportion of debt, as indicated in *V*4. The variable becomes increasingly larger in changes in non-failed samples → $t-3$ failed samples → $t-2$ failed samples → $t-1$ failed samples. If the stock price of a tourism firm with hotel businesses decreases significantly and the amount of debt increases significantly, the firm must decide whether it is failing.
3) Improvement in sales is one of the objectives of tourism firms. If a tourism firm is failing, it should design a thorough and well-planned development strategy to improve its sales and earnings. As indicated by *V*5 in Table 2, failing tourism firms with poor sales performance in the $t-3$ year are probably failed firms in the $t-2$ year because earnings and retained earnings deteriorate even though the firm manages to prevent its sales performance from deteriorating.
4) The average total assets of non-failed tourism firms (9.81 billion RMB after deleting outliers) are at least twice those of failed firms (4.23 billion RMB). However, non-failed tourism samples have smaller proportions of debts in terms of *V*4. Failed tourism firms place themselves at a significant business risk by borrowing more from banks without significantly improving their sales (as indicated by *V*5) to cover financial and operating costs. If a tourism firm with hotel businesses can maintain approximately 9.81 billion RMB in total assets and maintains its debt-to-market capitalisation at approximately 36%, the firm can effectively keep itself from failing.
5) It is more difficult to distinguish $t-2$ and $t-3$ failed samples from non-failed samples than to distinguish $t-1$ samples from non-failed samples. Thus, in the following sections, we focus on $t-2$ and $t-3$ FFPs.

4. Methodology

4.1. Up-sampling approach to correcting imbalanced samples: MSGA-RPD-NN

The current real-world imbalanced dataset for $t-1$ FFP of Chinese listed tourism companies in the hotel business is composed of 6.47% negative and 93.53% positive samples. A predictive model can obtain 93.53% accuracy by simply forecasting all samples as non-failed. However, this model is not useful in practice. The same is true for $t-2$ and $t-3$ FFPs. For these imbalanced tourism datasets, a model should predict Type I accuracy as accurately as possible and should perform well in forecasting Type II accuracy. This type of prediction is useful in practice. Two types of approaches are useful for forecasting with imbalanced tourism datasets: the solution at the algorithm level and the solution at the data level (Chawla, Japkowicz, & Kotcz, 2004). The former is specific algorithm-dependent and involves adjusting the costs of minority samples, the probabilistic estimate, and the decision threshold. Previous FFP studies, which used probabilistic models on some datasets consisting of less than 10% minority samples, belong to this type, although the word “imbalanced” did not appear in these papers. The limitation of this type of solution is that the adjustment cannot be used in a straightforward manner with other models. The latter is specific algorithm-independent, so the solution at the data level can be integrated with existing models. Commonly used solutions include random or directed oversampling with replacement, random or directed undersampling, and oversampling with informed generation of new samples. Undersampling deletes some majority samples, which reduces the quantity of knowledge in modelling. Oversampling with replacement duplicates existing minority samples, which makes the training set and the testing set consist of the same samples. The assessment suffers from an overfitting problem. Therefore, we focus on up-sampling with informed generation of new samples.

The MSGA-RPD-NN uses a random point between a minority sample and its NN in collecting minority samples as the informed information to generate new tourism samples. The MSGA-RPD-NN approach is presented as follows. Input information is the imbalanced tourism dataset.

Step 1: Separate the imbalanced tourism dataset, *ibd*, into majority and minority samples. Set the number of minority samples as *mim* and the number of majority samples as *mam*.
Step 2: Randomly retrieve a minority sample, *ms,* from the minority sample set.
Step 3: Retrieve the NN, *nnms*, of the current *ms* from minority samples by computing the Manhattan distance between *nnms* and *ms*, namely, $nnms - ms$. The Manhattan distance of two samples, *sa* from *sb*, can be calculated by the following formula, where w_i is the weight of the *i*th variable, and x_{sai} and x_{sbi} represent the values of the two samples on the *i*th variable:

$$MHDD(sa, sb) = \sum_{i=1}^{n} w_i \times (x_{sai} - x_{sbi}). \qquad (1)$$

By employing the Manhattan distance, the coordinate values of new tourism samples can be directly obtained by integrating the coordinate values of a minority sample with a random number and the Manhattan distance between the minority sample and its NN with the same label. The weight of each feature can be set equally after feature selection. Further approaches, such as the Analytic Hierarchy Process, can also be employed.

Step 4: Generate a random distance parameter: $rdp \in (0, 1)$.
Step 5: Generate a new minority tourism sample, *nms*, by randomly selecting a point between *ms* and *nnms*, namely, $nms = ms + rdp \times (nnms - ms)$.
Step 6: Add *nms* into the minority sample set.
Step 7: Delete the duplicated minority samples from the minority sample set and compute the new number of the minority samples, *min*.
Step 8: Repeat Steps 2–7 until $min < (mam - min/2)$. Thus, *ibd* is transferred as a balanced tourism dataset, namely, *bd*.

Thus, the function of MSGA-RPD-NN is illustrated as the following formula:

Please cite this article in press as: Li, H., & Sun, J., Forecasting business failure: The use of nearest-neighbour support vectors and correcting imbalanced samples – Evidence from the Chinese hotel industry, *Tourism Management* (2011), doi:10.1016/j.tourman.2011.07.004

ARTICLE IN PRESS

$$bd = \text{MSGA} - \text{RPD} - \text{NN}\,(ibd). \quad (2)$$

The basic concept of generating a new minority sample in the up-sampling approach is illustrated in Fig. 1 in a two-dimensional space, where ms_x and ms_y refer to the coordinate values of a minority sample, $nnms_x$ and $nnms_y$ refer to the coordinate values of its NN with the same label, $(nnms_x - ms_x)$ and $(nnms_y - ms_y)$ refer to differences between the coordinate values of the two samples, and $rdp \times (nnms_x - ms_x)$ and $rdp \times (nnms_y - ms_y)$ refer to the coordinate values of the new minority sample. If the minority samples, which are a smaller category, account for 10% of the total samples, the difference between the number of the final minority samples and the number of the final majority samples is less than 5% after using the up-sampling approach. The up-sampled dataset is a balanced tourism dataset. These procedures will not generate duplicated minority samples because the process will not generate a new (1,1,1,1,1) sample in the example. For non-failed samples, there are no identical samples, even though different samples are collected from the same firm. The financial ratios of a company in 2004, 2005, and 2006 are not the same. Each year, a company behaves differently in terms of financial ratios as long as it is operating.

4.2. Forecasting models based on the use of nearest-neighbour support vectors: NNsSVM and its ensemble

The SVM is a newly developed technique among the commonly used predictive models in FFP, and it has been demonstrated to be superior to many previous models. The SVM-based FFP is chiefly conducted outside the tourism industry. The review showed that current studies on SVM-based FFP have focused on optimising the single SVM model. Only support vectors are useful in modelling SVM, so selecting effective support vectors is crucial for improving the performance of SVM. The NN of each unlabelled sample is used as the training set of SVM. The NNsSVM approach is presented as follows. Let *trs* express the tourism training set, and let *ts* express the tourism testing set. If *trs* and *ts* are generated from *bd*, then $bd = [trs;\ ts]$.

Step 1: For each unlabelled tourism sample in *ts*, its *k* NNs are retrieved from *trs*, that is, *sts_nns_bd*. Thus, the entire NN set can be expressed as *ts_nns_bd*. Euclidean distance is used to calculate which sample from *trs* is the NN to a sample in *ts*. The Euclidean distance of two samples, that is, *sa* and *sb*, can be calculated by the following formula:

$$ED(sa, sb) = \sqrt{\sum_{i=1}^{n} w_i \times (x_{sai} - x_{sbi})^2}. \quad (3)$$

Step 2: SVM modelling on the NN set of *ts_nns_bd*, which is the set of nearest-neighbour support vectors. The model generated is called NNsSVM.
Step 3: Make a prediction using NNsSVM.

The group decision of various predictive models can improve the performance of FFP models (West, Dellana, & Qian, 2005). A bagging algorithm is integrated with NNsSVM to generate bagged NNsSVM (BNNsSVM) for tourism FFP. Input information includes *trs*, *ts*, the number of base models (i.e., *t*) and the number of nearest neighbours for NNsSVM (i.e., *k*). To facilitate the use of the model, we used the default value of the number of base models, namely, $t = 10$.

Step 1: Randomly generate 10 base tourism training sets, that is, *bts-j*, with replacements from *trs*, which has the same number of samples as *trs*, $j \in [1, 10]$.
Step 2: Generate 10 NNsSVM base models from the 10 *bts-j*, which are called NNsSVM-BTS-*j*.
Step 3: Make a group decision among the 10 models of NNsSVM-BTS-*j* with majority voting on labels of unlabelled samples, that is, *ts*. If no consensus among the 10 base models is generated, the result is set as -1.
Step 4: Generate predictions of *ts* according to the results of the group decision.

5. Results and analysis of FFP with the current imbalanced datasets in the Chinese tourism industry

The SVM, NNs, BNNsSVM, BSVM, BNNs, logit, and MDA were used as benchmarks in the FFP of the Chinese tourism industry with NNsSVM. Almost the entire population was divided into two parts: one for training and the other for testing. The division was repeated 100 times each for $t-2$ and $t-3$ FFPs with the balanced and imbalanced tourism datasets. The proportions of the division were respectively set as (20%, 80%), (35%, 65%), (50%, 50%), (65%, 35%), and (80%, 20%). Thus, 2000 rounds of experiments for each model were conducted. The SVM parameters, kernel and (C, v), were respectively set as RBF, (2^0, 2^0) and RBF, (2^8, 2^0) for the $t-2$ and $t-3$ FFPs of the Chinese listed tourism companies. The number of nearest neighbours was set as 3 for $t-2$ FFP and 5 for $t-3$ FFP to save computing costs and avoid sensitivity to outliers.

5.1. Forecasting results on imbalanced datasets

We investigated whether employing MSGA-RPD-NN to balance the imbalanced tourism datasets was necessary. The mean predictive performances of the eight models on the $t-2$ and $t-3$ initial imbalanced tourism datasets, 154 and 138 samples, are presented in Table 3. All eight models performed poorly in terms of Type I accuracy, which forecast failed tourism samples as failed. They produced approximately 90% $t-2$ and $t-3$ total accuracy by simply forecasting as many positive samples as non-failed ones. This type of forecasting model is not useful in practice. This result proves that imbalanced tourism datasets should be processed using a different approach; otherwise, these predictive models will perform poorly on minority samples. Generally, performance using the 20%/80% ratio for training/testing is worse than performance using the 80%/20% ratio. The former performance is sometimes better than the latter performance in terms of Type I accuracy because the number of failed samples is lower than that of non-failed samples. Thus, the employment of MSGA-RPD-NN to process the two imbalanced datasets for the FFP of Chinese listed tourism companies in the hotel business is necessary and valuable.

5.2. Forecasting results on balanced tourism datasets from MSGA-RPD-NN

After balancing each imbalanced tourism dataset, we obtained 281 samples consisting of 143 non-failed and 138 failed samples for $t-2$ FFP modelling and 249 samples consisting of 127 non-failed and 122 failed samples for $t-3$ FFP modelling. The results of the eight models of the two balanced FFP datasets from these companies on 100 splits from 5 proportions are presented in Figs. 3 and 4.

Please cite this article in press as: Li, H., & Sun, J., Forecasting business failure: The use of nearest-neighbour support vectors and correcting imbalanced samples – Evidence from the Chinese hotel industry, *Tourism Management* (2011), doi:10.1016/j.tourman.2011.07.004

ARTICLE IN PRESS

Table 3
Mean accuracies of the 8 models on the two initial imbalanced datasets for FFPs of Chinese tourism industry.

	$t-2$ Tourism FFP					$t-3$ Tourism FFP				
Total	20–80%	35–65%	50–50%	65–35%	80–20%	20–80%	35–65%	50–50%	65–35%	80–20%
SVM	92.65	92.65	93.19	93.04	93.07	87.82	87.87	88.48	88.69	89.74
NNs	92.66	92.61	92.94	92.32	92.20	91.74	91.72	91.32	91.13	91.85
NNsSVM	92.41	92.40	93.00	92.96	93.17	86.44	85.19	86.06	87.00	87.81
BSVM	92.65	92.65	93.23	93.13	93.10	89.13	88.84	89.06	88.92	89.33
BNNs	92.46	92.15	92.64	92.32	92.57	91.64	91.60	91.30	90.88	91.52
BNNsSVM	92.24	92.39	93.04	93.04	93.13	88.91	88.55	88.64	88.44	88.74
MDA	89.74	90.37	91.88	92.17	92.80	85.06	83.72	85.00	85.42	87.56
Logit	91.01	91.01	92.69	93.04	93.57	87.36	89.37	90.68	91.35	91.74
Type I	20–80%	35–65%	50–50%	65–35%	80–20%	20–80%	35–65%	50–50%	65–35%	80–20%
SVM	0.48	0.76	0.92	0.00	0.00	11.79	10.66	6.88	4.89	3.80
NNs	7.38	14.01	18.23	15.81	16.43	0.54	0.87	0.00	0.00	0.00
NNsSVM	18.30	13.70	7.55	3.89	3.30	14.75	15.94	9.42	8.30	1.76
BSVM	0.48	0.88	0.98	2.10	0.55	7.33	8.59	4.92	4.76	3.33
BNNs	8.69	14.23	18.96	19.27	25.95	0.98	0.63	0.87	0.34	0.00
BNNsSVM	16.13	14.74	8.05	3.74	1.10	9.15	10.54	7.72	5.10	1.48
MDA	42.15	52.38	60.58	58.60	65.41	24.00	35.07	33.92	31.88	27.09
Logit	35.92	41.38	46.04	41.97	46.42	19.51	20.43	15.83	15.16	12.52
Type II	20–80%	35–65%	50–50%	65–35%	80–20%	20–80%	35–65%	50–50%	65–35%	80–20%
SVM	99.99	99.99	99.96	100.00	100.00	94.65	94.61	95.78	96.43	97.07
NNs	99.56	99.02	98.57	98.16	98.16	99.83	99.53	99.40	99.54	99.63
NNsSVM	98.50	98.83	99.36	99.68	99.97	92.92	91.29	92.93	94.24	95.12
BSVM	99.99	99.98	99.99	100.00	100.00	96.44	95.83	96.53	96.74	96.65
BNNs	99.24	98.52	98.16	97.90	98.06	99.69	99.40	99.32	99.25	99.26
BNNsSVM	98.46	98.74	99.35	99.78	100.00	96.06	95.35	95.89	96.17	96.10
MDA	93.73	93.55	94.40	94.91	95.05	90.56	87.92	89.57	90.38	92.54
Logit	95.60	95.14	96.30	96.99	97.29	93.46	95.34	97.30	98.44	98.45

A performance indicator that integrated the total, Type I, and Type II accuracies was used to represent the predictive performance in the experiment, as illustrated in the following formula:

$$PI = 1 - (\text{Type I error} \times \text{Type I cost} + \text{Type II error} \times \text{Type II cost} + \text{total error} \times \text{total cost}), \tag{4}$$

where Type I cost + Type II cost + total cost = 1. In this performance indicator, we integrated three assessing principles: 1) if a predictor produces a smaller Type I error rate, the predictor is better; 2) if a predictor produces a smaller Type II error rate, the predictor is better; and 3) if a predictor produces a smaller total error rate, the predictor is better. The three assessing principles are commonly used when analysing the performance of a predictor. The *PI* integrates them using a weighted mechanism. The indicator benefits from these three assessing principles, although the total error, Type I error, and Type II error are correlated to some extent.

There are different opinions on whether the importance of Type I and II errors should be valued equally (Ooghe & Spaenjers, 2010). The performance indicator was calculated under two assumptions: they should be valued equally and they should be valued differently. If these two types of errors are valued equally, they should have equal weight. If they are valued differently, Type I errors should be given a higher value. There is no widely accepted definition of the cost difference between Type I and Type II errors. We assumed that Type I errors cost three times as much as Type II errors, and the total error rate costs twice that of Type II errors. The results of the performance indicators are provided in Table 4. The Shapiro–Wilk method was used to test whether the predictive performance of each predictor distributes normally, and the result shows that they do not. The Wilcoxon test was used to ascertain whether there are significant differences between NNsSVM and the seven benchmarks. The results are presented in Table 5. Blue indicates that NNsSVM outperforms the benchmark, and black indicates that NNsSVM and the benchmark are similarly ranked.

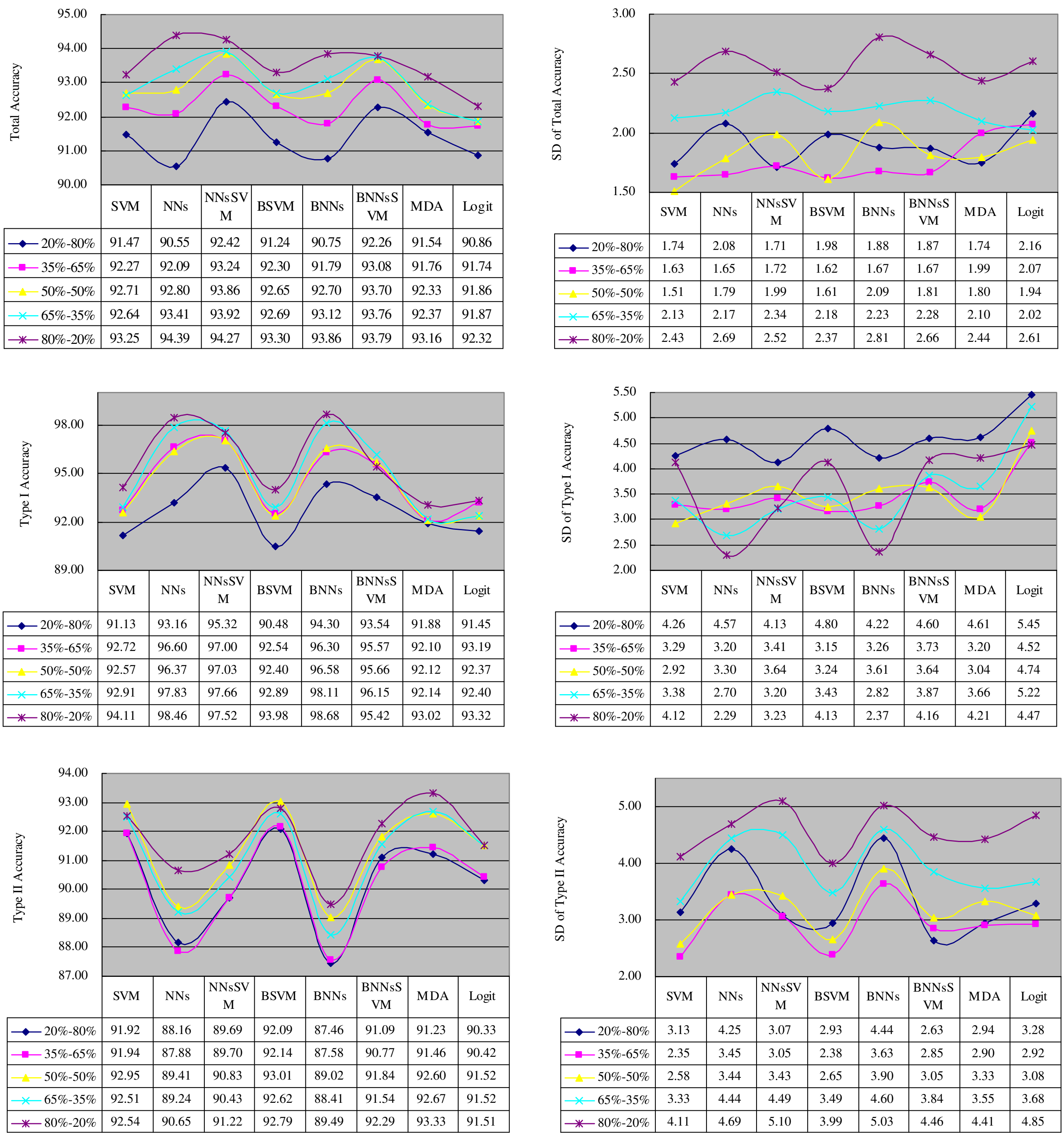
5.3. Discussion and implications

5.3.1. Results discussion

In terms of average out-of-sample performance, NNsSVM/BNNsSVM generated 93.54%/93.32% total accuracy, 96.90%/95.27% Type I accuracy, and 90.37%/91.51% Type II accuracy for $t-2$ FFP, and it generated 86.73%/87.03% total accuracy, 93.29%/91.93% Type I accuracy, and 80.59%/82.51% Type II accuracy for $t-3$ FFP. Table 4 shows that NNsSVM outperformed all seven benchmarks, except for the $t-3$ BNNsSVM model, with equal importance of total, Type I, and Type II accuracies for $t-2$ and $t-3$ FFPs of Chinese tourism companies in the hotel business, whether Type I and Type II accuracies were valued equally or differently. The NNsSVM and BNNsSVM models produced similar performance in the exceptional situation. The NNsSVM reduced the predictive errors of SVM, NNs, BSVM, BNNs, BNNsSVM, MDA, and logit by 14.63%, 12.02%, 15.03%, 14.26%, 3.66%, 17.52%, and 22.23%, respectively, for $t-2$ tourism FFP and by 6.50%, 6.03%, 6.42%, 9.73%, −2.21%, 41.63%, and 36.27%, respectively, for $t-3$ tourism FFP when the costs of Type I and Type II errors were valued equally. The performances of NNsSVM on SVM, NNs, BSVM, BNNs, BNNsSVM, MDA, and logit improved by 28.65%, 12.03%, 29.61%, 12.41%, 11.73%, 31.57%, and 33.47%, respectively, for $t-2$ tourism FFP and by 9.14%, 3.97%, 12.47%, 5.85%, 2.34%, 53.04%, and 46.98%, respectively, for $t-3$ tourism FFP when the costs of Type I and Type II errors were valued differently. Table 5 indicates that the difference between NNsSVM and each benchmark is generally significant at the 1% level. However, there is no significance with BNNsSVM for the $t-3$ tourism FFP under the assumption of equal importance of total, Type I, and Type II accuracies. This result shows that NNsSVM is the preferred method in FFP, followed by the BNNsSVM model with the dataset from MSGA-RPD-NN, which is effective in balancing current real-world imbalanced samples from the Chinese tourism industry.

The MSGA-RPD-NN generates new minority samples to balance the minority and majority. Without the up-sampling approach, all

Please cite this article in press as: Li, H., & Sun, J., Forecasting business failure: The use of nearest-neighbour support vectors and correcting imbalanced samples – Evidence from the Chinese hotel industry, *Tourism Management* (2011), doi:10.1016/j.tourman.2011.07.004



Total Accuracy

	SVM	NNs	NNsSVM	BSVM	BNNs	BNNsSVM	MDA	Logit
20%-80%	91.47	90.55	92.42	91.24	90.75	92.26	91.54	90.86
35%-65%	92.27	92.09	93.24	92.30	91.79	93.08	91.76	91.74
50%-50%	92.71	92.80	93.86	92.65	92.70	93.70	92.33	91.86
65%-35%	92.64	93.41	93.92	92.69	93.12	93.76	92.37	91.87
80%-20%	93.25	94.39	94.27	93.30	93.86	93.79	93.16	92.32

SD of Total Accuracy

	SVM	NNs	NNsSVM	BSVM	BNNs	BNNsSVM	MDA	Logit
20%-80%	1.74	2.08	1.71	1.98	1.88	1.87	1.74	2.16
35%-65%	1.63	1.65	1.72	1.62	1.67	1.67	1.99	2.07
50%-50%	1.51	1.79	1.99	1.61	2.09	1.81	1.80	1.94
65%-35%	2.13	2.17	2.34	2.18	2.23	2.28	2.10	2.02
80%-20%	2.43	2.69	2.52	2.37	2.81	2.66	2.44	2.61

Type I Accuracy

	SVM	NNs	NNsSVM	BSVM	BNNs	BNNsSVM	MDA	Logit
20%-80%	91.13	93.16	95.32	90.48	94.30	93.54	91.88	91.45
35%-65%	92.72	96.60	97.00	92.54	96.30	95.57	92.10	93.19
50%-50%	92.57	96.37	97.03	92.40	96.58	95.66	92.12	92.37
65%-35%	92.91	97.83	97.66	92.89	98.11	96.15	92.14	92.40
80%-20%	94.11	98.46	97.52	93.98	98.68	95.42	93.02	93.32

SD of Type I Accuracy

	SVM	NNs	NNsSVM	BSVM	BNNs	BNNsSVM	MDA	Logit
20%-80%	4.26	4.57	4.13	4.80	4.22	4.60	4.61	5.45
35%-65%	3.29	3.20	3.41	3.15	3.26	3.73	3.20	4.52
50%-50%	2.92	3.30	3.64	3.24	3.61	3.64	3.04	4.74
65%-35%	3.38	2.70	3.20	3.43	2.82	3.87	3.66	5.22
80%-20%	4.12	2.29	3.23	4.13	2.37	4.16	4.21	4.47

Type II Accuracy

	SVM	NNs	NNsSVM	BSVM	BNNs	BNNsSVM	MDA	Logit
20%-80%	91.92	88.16	89.69	92.09	87.46	91.09	91.23	90.33
35%-65%	91.94	87.88	89.70	92.14	87.58	90.77	91.46	90.42
50%-50%	92.95	89.41	90.83	93.01	89.02	91.84	92.60	91.52
65%-35%	92.51	89.24	90.43	92.62	88.41	91.54	92.67	91.52
80%-20%	92.54	90.65	91.22	92.79	89.49	92.29	93.33	91.51

SD of Type II Accuracy

	SVM	NNs	NNsSVM	BSVM	BNNs	BNNsSVM	MDA	Logit
20%-80%	3.13	4.25	3.07	2.93	4.44	2.63	2.94	3.28
35%-65%	2.35	3.45	3.05	2.38	3.63	2.85	2.90	2.92
50%-50%	2.58	3.44	3.43	2.65	3.90	3.05	3.33	3.08
65%-35%	3.33	4.44	4.49	3.49	4.60	3.84	3.55	3.68
80%-20%	4.11	4.69	5.10	3.99	5.03	4.46	4.41	4.85

Fig. 3. Results of the 8 models on $t-2$ FFP of Chinese tourism industry.

eight models, SVM, NNs, NNsSVM, BSVM, BNNs, BNNsSVM, logit, and MDA, performed poorly in predicting failed samples. They simply achieved high total accuracy by forecasting as many positive samples as non-failed ones because the number of positive samples was over 92% of the total. An imbalanced dataset valued the majority samples more than the minority samples in modelling. The real-world dataset for tourism FFP belongs to the imbalanced dataset, which makes MSGA-RPD-NN useful in tourism FFP. The performances of all models significantly improved with the up-sampling approach. The results from $t-2$ and $t-3$ FFPs prove that the method based on nearest-neighbour support vectors and correcting imbalanced samples effectively improves the predictive performance of FFP in the Chinese tourism industry.

5.3.2. *Management implications*

The method based on the use of nearest-neighbour support vectors and correcting imbalanced samples, which was developed in the current study, has important management implications for the Chinese tourism industry.

1) *Variable*. Altman's five variables are useful in the FFP of the Chinese tourism industry. Chinese listed tourism firms in the

Total Accuracy

	SVM	NNs	NNsSVM	BSVM	BNNs	BNNsSVM	MDA	Logit
20%-80%	81.08	78.23	80.53	81.53	77.77	81.27	75.36	76.64
35%-65%	84.39	83.79	84.94	83.99	82.80	85.25	76.64	78.07
50%-50%	86.66	87.56	87.94	86.35	87.02	88.02	78.48	79.61
65%-35%	87.97	88.68	89.40	87.85	88.16	89.40	78.62	80.89
80%-20%	88.88	90.96	90.86	89.35	90.47	91.18	78.37	81.10

SD of Total Accuracy

	SVM	NNs	NNsSVM	BSVM	BNNs	BNNsSVM	MDA	Logit
20%-80%	4.27	3.93	4.35	4.41	4.60	4.50	4.92	5.16
35%-65%	3.70	3.24	3.40	3.59	3.79	3.55	4.01	4.26
50%-50%	2.73	3.06	3.23	2.91	3.33	3.31	3.70	3.94
65%-35%	3.39	3.22	3.16	3.37	3.45	3.11	3.97	4.24
80%-20%	4.11	4.34	4.07	3.68	4.63	3.56	5.97	5.89

Type I Accuracy

	SVM	NNs	NNsSVM	BSVM	BNNs	BNNsSVM	MDA	Logit
20%-80%	84.08	85.13	84.10	83.20	86.22	83.07	72.47	76.05
35%-65%	89.06	91.91	90.48	86.77	92.17	88.96	73.99	77.72
50%-50%	93.33	96.03	95.36	91.20	96.17	93.10	75.88	79.03
65%-35%	96.14	97.64	97.89	94.83	97.89	96.46	76.45	81.50
80%-20%	96.44	97.47	98.64	96.00	98.11	98.07	74.44	80.52

SD of Type I Accuracy

	SVM	NNs	NNsSVM	BSVM	BNNs	BNNsSVM	MDA	Logit
20%-80%	10.11	8.29	8.89	10.40	8.82	9.27	11.30	11.52
35%-65%	7.99	5.71	6.70	8.46	5.75	6.59	8.20	9.43
50%-50%	5.35	3.48	4.55	5.68	3.06	5.51	6.80	7.93
65%-35%	5.17	2.28	2.95	5.95	2.31	3.83	7.84	7.63
80%-20%	5.62	3.02	2.89	5.26	2.71	3.40	10.49	10.34

Type II Accuracy

	SVM	NNs	NNsSVM	BSVM	BNNs	BNNsSVM	MDA	Logit
20%-80%	78.38	71.89	77.20	80.12	69.95	79.68	78.25	77.50
35%-65%	80.25	76.36	79.94	81.71	74.18	82.02	79.33	78.88
50%-50%	80.55	79.67	80.94	81.96	78.50	83.30	81.17	80.63
65%-35%	80.33	80.18	81.34	81.40	78.99	82.82	80.77	80.63
80%-20%	81.86	84.94	83.51	83.30	83.49	84.72	82.43	82.23

SD of Type II Accuracy

	SVM	NNs	NNsSVM	BSVM	BNNs	BNNsSVM	MDA	Logit
20%-80%	7.25	9.53	6.53	6.07	10.85	6.42	8.54	9.00
35%-65%	5.85	6.45	5.60	6.10	7.04	5.70	7.53	7.60
50%-50%	5.18	5.30	4.53	5.51	5.64	4.24	6.64	6.50
65%-35%	5.66	5.41	5.23	6.05	5.71	5.25	6.62	6.73
80%-20%	7.64	7.80	7.21	7.18	8.30	6.48	8.63	8.76

Fig. 4. Results of the 8 models on $t-3$ FFP of Chinese tourism industry.

hotel business may fail if the firms' value of working capital to total assets, retained earnings to total assets, earnings before interest and tax to total assets, and sales to total assets decrease significantly and the value of total debts-to-market capitalisation of firms increases significantly. Our finding that failed tourism firms behave worse than non-failed ones in earnings is consistent with the findings of Gu (2002) and Kim and Gu (2006a). Failing firms may incur more expenses when they try to improve sales, and they may use loans to cover these expenses. Because oversupply is a problem in the Chinese hotel industry, it is highly probable that failing tourism firms will not achieve significant improvements in sales. If their sales do not improve as significantly as they expect, failing tourism firms will be more likely to fail.

2) *Dataset.* The target population of FFP consists of imbalanced samples, in which failed samples constitute the minority and non-failed samples constitute the majority. The imbalanced dataset should be used in the FFP in the Chinese tourism industry to reflect the situation in the real world. Predictive models without appropriate pre-processing techniques usually

Please cite this article in press as: Li, H., & Sun, J., Forecasting business failure: The use of nearest-neighbour support vectors and correcting imbalanced samples – Evidence from the Chinese hotel industry, *Tourism Management* (2011), doi:10.1016/j.tourman.2011.07.004

ARTICLE IN PRESS

Table 4
Mean performance indicators from the integration of mean total, Type I and II accuracies of the 8 models on the 5 types of splits for $t-2$ and $t-3$ FFPs of Chinese tourism industry.

	Equal importance of total, Type I and Type II errors (costs of total, Type I and Type II errors are assumed as 1/3, 1/3, and 1/3)		Different importance of total, Type I and Type II errors (costs of total, Type I and Type II errors are assumed as 2/6, 3/6, and 1/6)	
	$t-2$ Tourism FFP	$t-3$ Tourism FFP	$t-2$ Tourism FFP	$t-3$ Tourism FFP
SVM	92.51 (5)	85.96 (5)	92.56 (5)	87.88 (5)
NNs	92.73 (3)	86.03 (3)	93.97 (3)	88.54 (3)
NNsSVM	**93.61 (1)**	*86.87 (2)*	**94.70 (1)**	**88.99 (1)**
BSVM	92.48 (6)	85.97 (4)	92.46 (6)	87.42 (6)
BNNs	92.54 (4)	85.46 (6)	93.94 (4)	88.31 (4)
BNNsSVM	*93.36 (2)*	**87.16 (1)**	*93.99 (2)*	*88.73 (2)*
MDA	92.25 (7)	77.51 (8)	92.25 (7)	76.55 (8)
Logit	91.78 (8)	79.40 (7)	92.03 (8)	79.23 (7)

The number in () indicates the ranking sequence of a model in the same column.

perform poorly on imbalanced datasets. The models will generate high performances by simply concentrating on the prediction of the majority samples, which is not useful in practice. The up-sampling approach of MSGA-RPD-NN is useful in helping predictive models perform well for both majority samples and minority samples. In the FFP of tourism companies on an imbalanced dataset, a processing approach such as MSGA-RPD-NN can be used before constructing predictive models. This usage effectively improves the predictive performance of FFP models in the tourism industry.

3) *Model*. Another way to improve the performance of tourism FFP is the use of nearest-neighbour support vectors in constructing predictive models. NNsSVM outperforms the two traditional parametric models, MDA and logit, and the other non-parametric models, SVM, NNs, BSVM, and BNNs, regardless of whether the costs of Type I and Type II errors are valued equally. BNNsSVM performs the second best.
4) *Practice*. Creditors and managers may apply the model based on the use of nearest-neighbour support vectors and correcting imbalanced samples to forecast the possibility that a tourism firm in the hotel business will fail. They make decisions according to the performance of the predictive model, which directly affects the level of risk in their decisions. With NNsSVM, creditors and managers can estimate the possibility more accurately, initiate corrective actions, and turn their firms around. For example, many clients did not recover their deposits when the XiaoShan International Hotel went bankrupt. By using the predictive model to forecast whether a hotel will fail before booking rooms and services in the long term, customers could reduce such risks when making decisions.

5.3.3. *Different findings*

In contrast with the findings of previous research on FFP in the tourism industry, the results of our study demonstrated that MSGA-RPD-NN improves the predictive performances of the models in imbalanced datasets and that NNsSVM, followed by BNNsSVM, performs better than the other models in $t-2$ and $t-3$ FFPs regardless of whether the costs of Type I and Type II errors are valued equally. The differences between the current research and previous studies are provided below.

1) Previous studies on FFP in the tourism industry have used balanced datasets and have recommended the use of parametric techniques such as logit and MDA. The difficulty of tourism FFP on an imbalanced dataset and the use of non-parametric techniques of NNsSVM, SVM, NNs, BSVM, BNNs, and BNNsSVM have not been addressed. The findings of the present study indicate that non-parametric techniques could improve the performance of FFP. For example, the average improvement of NNsSVM on logit and MDA for $t-2$ and $t-3$ tourism FFPs of Chinese listed companies in the hotel business is approximately 35.34%. This is a significant improvement in predictive performance.
2) Another conflicting finding of the current study pertains to the relationship between non-failed and failed tourism samples on the variable of sales to total assets. The findings of Gu (2002) and Kim and Gu (2006a) on the US tourism FFP indicate that in terms of this ratio, non-failed tourism samples have a smaller mean value than failed samples do, at 1.553 and 1.627, respectively. This finding implies that failed samples behave better than non-failed samples in terms of achieving sales with

Table 5
Significant test on performance difference between NNsSVM and each benchmark on FFP of Chinese tourism industry.

NNsSVM	Equal importance of total, Type I and Type II accuracies (costs of total, Type I and Type II accuracies are assumed as 1/3, 1/3, and 1/3)		Different importance of total, Type I and Type II accuracies (costs of total, Type I and Type II accuracies are assumed as 2/6, 3/6, and 1/6)	
	$t-2$ Tourism FFP	$t-3$ Tourism FFP	$t-2$ Tourism FFP	$t-3$ Tourism FFP
SVM	8.08[a] (0.000[b])***	3.53 (0.000)***	14.03 (0.000)***	5.60 (0.000)***
NNs	6.33 (0.000)***	2.38 (0.018)**	4.92 (0.000)***	3.52 (0.000)***
BSVM	8.04 (0.000)***	3.62 (0.000)***	14.10 (0.000)***	6.89 (0.000)***
BNNs	7.63 (0.000)***	3.87 (0.000)***	5.36 (0.000)***	4.23 (0.000)***
BNNsSVM	1.72 (0.086)*	−0.73 (0.463)	4.44 (0.000)***	2.61 (0.000)***
MDA	9.52 (0.000)***	22.06 (0.000)***	15.45 (0.000)***	23.13 (0.000)***
Logit	12.06 (0.000)***	19.26 (0.000)***	14.92 (0.000)***	20.98 (0.000)***

***: Significance at 1% level, **: significance at 5% level, *: significance at 10% level.
[a] *Z* Value.
[b] *p* Value.

Please cite this article in press as: Li, H., & Sun, J., Forecasting business failure: The use of nearest-neighbour support vectors and correcting imbalanced samples – Evidence from the Chinese hotel industry, *Tourism Management* (2011), doi:10.1016/j.tourman.2011.07.004

ARTICLE IN PRESS

available total assets. Thus, failed tourism firms may expand too rapidly, with high operating and financial costs (Gu, 2002). In contrast, Chinese failed tourism samples behave worse than non-failed samples in this ratio; the mean value of this variable is 0.247 for $t-1$ failing samples, 0.252 for $t-2$ failing samples, 0.243 for $t-3$ failing samples, and 0.387 for non-failed samples. Because oversupply is a major problem in the Chinese hotel industry, non-failed firms behave better than failed firms in terms of sales.

6. Conclusion

We successfully forecast the failure of Chinese listed tourism firms in the hotel business by constructing a new up-sampling MSGA-RPD-NN approach and a model based on the use of nearest-neighbour support vectors with current real-world imbalanced datasets. A series of experiments were conducted to ensure that the results were not biased and that they had statistical significance. By employing SVM, NNs, BSVM, BNNs, BNNsSVM, logit, and MDA as benchmarks, we proved that NNsSVM outperforms all of the models used in the tourism FFPs of Chinese listed companies in the hotel business, followed by BNNsSVM. Because many hotels face the risk of failure, our research contributes to helping to reduce risks in hotel-related businesses.

The limitations of the current study are as follows. 1) The optimisation of parameters in the models based on the use of nearest-neighbour support vectors can improve the performance of FFP in the tourism industry. This issue is not addressed in this paper and requires further study. 2) The imbalanced datasets for the Chinese tourism FFP were collected from listed companies in the hotel business. Data for non-listed tourism companies were not used because they were not available at the time of the study. 3) Some managers believe that non-parametric techniques are more complex than parametric techniques in FFP. It may be useful to investigate how to make non-parametric techniques more interpretable by, for example, integrating rule-extraction techniques with non-parametric techniques. 4) The current study focused on Chinese tourism FFP in imbalanced datasets. Further studies should be conducted on tourism FFP using a method based on nearest-neighbour support vectors and correcting imbalanced samples with data from other countries.

Acknowledgements

The authors gratefully thank Prof. Chris Ryan (the Editor-In-Chief), editors, and the three anonymous referees for their valuable comments and recommendations. This research is partially supported by the National Natural Science Foundation of China (No. 71171179) and the Zhejiang Provincial Natural Science Foundation of China (No. Y7100008).

References

Alfaro, E., Garcia, N., Gamez, M., & Elizondo, D. (2008). Bankruptcy forecasting: an empirical comparison of AdaBoost and neural networks. *Decision Support Systems, 45*(1), 110–122.

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance, 23*(4), 589–609.

Altman, E. I., & Hotchkiss, E. (2005). *Corporate financial distress and bankruptcy: Predict and avoid bankruptcy, analyze and invest in distressed debt* (3rd ed.). Wiley.

Atiya, A. F. (2001). Bankruptcy prediction for credit risk using neural network: a survey and new results. *IEEE Transactions on Neural Networks, 12*(4), 929–935.

Beaver, W. (1966). Financial ratios as predictors of failure. *Journal of Accounting Research, 4*(Suppl.), 71–111.

Bose, I. (2006). Deciding the financial health of dot-coms using rough sets. *Information & Management, 43*(7), 835–846.

Bose, I., & Pal, R. (2006). Predicting the survival or failure of click-and-mortar corporations: a knowledge discovery approach. *European Journal of Operational Research, 174*(2), 959–982.

Chandra, D., Ravi, V., & Bose, I. (2009). Failure prediction of dotcom companies using hybrid intelligent techniques. *Expert Systems with Applications, 36*(3), 4830–4837.

Chawla, N., Hall, L., Bowyer, K., & Kegelmeyer, W. (2002). SMOTE: synthetic minority oversampling technique. *Journal of Artificial Intelligence Research, 16*(1), 321–357.

Chawla, N., Japkowicz, N., & Kotcz, A. (2004). Editorial: special issue on learning from imbalanced data sets. *ACM SIGKDD Explorations Newsletter, 6*(1), 1–6.

Dimitras, A., Slowinski, R., Susmaga, R., & Zopounidis, C. (1999). Business failure prediction using rough sets. *European Journal of Operational Research, 114*(2), 263–280.

Ding, Y., Song, X., & Zeng, Y. (2008). Forecasting financial condition of Chinese listed companies based on support vector machine. *Expert Systems with Applications, 34*(4), 3081–3089.

Gaganis, C., Pasiouras, F., & Doumpos, M. (2007). Probabilistic neural networks for the identification of qualified audit opinions. *Expert Systems with Applications, 32*(1), 114–124.

Gepp, A., Kumar, K., & Bhattacharya, S. (2010). Business failure prediction using decision trees. *Journal of Forecasting, 29*(6), 536–555.

Gestel, T., Baesens, B., Suykens, J., Van den Poel, D., Baestaens, D.-E., & Willekens, M. (2006). Bayesian kernel based classification for financial distress detection. *European Journal of Operational Research, 172*(3), 979–1003.

Gu, Z. (2002). Analyzing bankruptcy in the restaurant industry: a multiple discriminant model. *Intentional Journal of Hospitality Management, 21*(1), 25–42.

Hardle, W., Lee, Y., Schafer, D., & Yeh, Y. (2009). Variable selection and oversampling in the use of smooth support vector machines for predicting the default risk of companies. *Journal of Forecasting, 28*(6), 512–534.

Hu, Y., & Ansell, J. (2007). Measuring retail company performance using credit scoring techniques. *European Journal of Operational Research, 183*(3), 1595–1606.

Hu, Y., & Ansell, J. (2009). Retail default prediction by using sequential minimal optimization technique. *Journal of Forecasting, 28*(8), 651–666.

Hua, Z., Wang, Y., Xu, X., Zhang, B., & Liang, L. (2007). Predicting corporate financial distress based on integration of support vector machine and logistic regression. *Expert Systems with Applications, 33*(2), 434–440.

Hwang, R., Cheng, K., & Lee, J. (2007). A semi-parametric method for predicting bankruptcy. *Journal of Forecasting, 26*(5), 317–342.

Jones, S., & Hensher, D. (2004). Predicting firm financial distress: a mixed logit model. *Accounting Review, 79*(4), 1011–1038.

Kim, H., & Gu, Z. (2006a). Predicting restaurant bankruptcy: a logit model in comparison with a discriminant model. *Journal of Hospitality & Tourism Research, 14*(1), 474–493.

Kim, H., & Gu, Z. (2006b). A logistic regression analysis for predicting bankruptcy in the hospitality industry. *The Journal of Hospitality Financial Management, 14*(1), 17–34.

Kumar, P. R., & Ravi, V. (2007). Bankruptcy prediction in banks and firms via statistical and intelligent techniques – a review. *European Journal of Operational Research, 180*(1), 1–28.

Lee, Y.-J., & Mangasarian, O. L. (2001). SSVM: a smooth support vector machine for classification. *Computational Optimization and Applications, 20*(1), 5–22.

Li, H., Adeli, H., Sun, J., & Han, J.-G. (2011). Hybridizing principles of TOPSIS with case-based reasoning for business failure prediction. *Computers & Operations Research, 38*(2), 409–419.

Li, H., & Sun, J. (2008). Ranking-order case-based reasoning for financial distress prediction. *Knowledge-Based Systems, 21*(8), 868–878.

Li, H., & Sun, J. (2009). Hybridizing principles of the Electre method with case-based reasoning for data mining: Electre-CBR-I and Electre-CBR-II. *European Journal of Operational Research, 197*(1), 214–224.

Li, H., & Sun, J. (2010). Forecasting business failure in China using case-based reasoning with hybrid representation. *Journal of Forecasting, 29*(5), 486–501.

Martens, D., Gestel, T. V., Backer, M., Haesen, R., Vanthienen, J., Mues, C., & Baesens, B. (2010). Credit rating prediction using ant colony optimization. *Journal of the Operational Research Society, 61*(4), 561–573.

McKee, T. (2002). Rough sets bankruptcy prediction models versus auditor signaling rates. *Journal of Forecasting, 22*(8), 569–586.

McKee, T., & Greenstein, M. (2000). Predicting bankruptcy using recursive partitioning and a realistically proportioned dataset. *Journal of Forecasting, 19*(3), 219–230.

McKee, T., & Lensberg, T. (2002). Genetic programming and rough sets: a hybrid approach to bankruptcy classification. *European Journal of Operational Research, 138*(2), 436–451.

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.

Min, S.-H., Lee, J.-M., & Han, I. (2006). Hybrid genetic algorithms and support vector machines for bankruptcy prediction. *Expert Systems with Applications, 31*(3), 652–660.

Olsen, M., Bellas, C., & Kish, L. V. (1983). Improving the prediction of restaurant failure through ratio analysis. *International Journal of Hospitality Management, 2*(4), 187–193.

Ooghe, H., & Spaenjers, C. (2010). A note on performance measures for business failure prediction models. *Applied Economic Letters, 17*(1), 67–70.

Please cite this article in press as: Li, H., & Sun, J., Forecasting business failure: The use of nearest-neighbour support vectors and correcting imbalanced samples – Evidence from the Chinese hotel industry, *Tourism Management* (2011), doi:10.1016/j.tourman.2011.07.004

ARTICLE IN PRESS

Pendharkar, P. (2005). A threshold-varying artificial neural network approach for classification and its application to bankruptcy prediction problem. *Computers & Operations Research, 32*(10), 2561–2582.

Ravisankar, P., Ravi, V., & Bose, I. (2010). Failure prediction of dotcom companies using neural network-genetic programming hybrids. *Information Sciences, 180*(8), 1257–1267.

Ryu, Y., & Yue, W. (2005). Firm bankruptcy prediction: experimental comparison of isotonic separation and other classification approaches. *IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans, 35*(5), 727–737.

Sarkar, S., & Sriram, R. S. (2001). Bayesian models for early warning of bank failures. *Management Science, 47*(11), 1457–1475.

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.

Sun, J., & Li, H. (2008). Listed companies' financial distress prediction based on weighted majority voting combination of multiple classifiers. *Expert Systems with Applications, 35*(3), 818–827.

Tsai, C.-F. (2008). Financial decision support using neural networks and support vector machines. *Expert Systems, 25*(4), 380–393.

Tsakonas, A., Dounias, G., Doumpos, M., & Zopounidis, C. (2006). Bankruptcy prediction with neural logit networks by means of grammar-guided genetic programming. *Expert Systems with Applications, 30*(3), 449–461.

Tseng, F., & Hu, Y. (2010). Comparing four bankruptcy prediction models: logit, quadratic interval logit, neural and fuzzy neural networks. *Expert Systems with Applications, 37*(3), 1846–1853.

West, D., Dellana, S., & Qian, J. (2005). Neural network ensemble strategies for financial decision applications. *Computers & Operations Research, 32*(10), 2543–2559.

Wu, C., Tzeng, G., & Goo, Y. (2007). A real-valued genetic algorithm to optimize the parameters of support vector machine for predicting bankruptcy. *Expert Systems with Applications, 32*(2), 397–408.

Youn, H., & Gu, Z. (2010a). Predicting Korean lodging firm failures: an artificial neural network model along with a logistic regression model. *International Journal of Hospitality Management, 29*(1), 120–127.

Youn, H., & Gu, Z. (2010b). Predict US restaurant firm failures: the artificial neural network model versus logistic regression model. *Tourism and Hospitality Research, 10*(3), 1–17.

Please cite this article in press as: Li, H., & Sun, J., Forecasting business failure: The use of nearest-neighbour support vectors and correcting imbalanced samples – Evidence from the Chinese hotel industry, *Tourism Management* (2011), doi:10.1016/j.tourman.2011.07.004