International Journal of Forecasting 34 (2018) 440–455

Contents lists available at ScienceDirect

International Journal of Forecasting

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

Forecasting bank failures and stress testing: A machine learning approach

Periklis Gogas, Theophilos Papadimitriou, Anna Agrapetidou *

Democritus University of Thrace, Department of Economics, Komotini 69100, Greece

A R T I C L E   I N F O

Keywords:
Machine learning
Bank failures
Stress testing
Forecasting

A B S T R A C T

This paper presents a forecasting model of bank failures based on machine-learning. The proposed methodology defines a linear decision boundary that separates the solvent banks from those that failed. This setup generates a novel alternative stress-testing tool. Our sample of 1443 U.S. banks includes all 481 banks that failed during the period 2007–2013. The set of explanatory variables is selected using a two-step feature selection procedure. The selected variables were then fed to a support vector machines forecasting model, through a training–testing learning process. The model exhibits a 99.22% overall forecasting accuracy and outperforms the well-established Ohlson's score.

1. Introduction

Historically, the banks' role of intermediation between surplus and deficit agents has been crucial for economic activity. This has remained true over the last few decades, despite the increased significance of capital markets and direct financing. Financial institutions are highly interconnected operationally, and the complex channels of interconnection increase the associated systemic risk. As a result, the issue of their financial health is always topical, and a prerequisite to maintaining stability in the economy. Usually, a banking crisis is transmitted swiftly to other sectors within the originating country or to other economies, triggering financial distress on an international scale. The financial crisis of 2007 is an example of the importance of this interconnection. Only 29 banks out of a total of more than 6000 failed in the U.S. during the seven years from 2000 to 2006. During the next seven years (2007–2013), though, bankruptcies increased by 17 times, reaching a total of 492 failed banks. At the same time, the crisis started spreading internationally. The global financial crisis that followed highlighted the need for a stricter and more efficient supervision of financial institutions, in addition to raising macro-prudential concerns. Stress-testing has proved to be a useful and popular tool for regulators internationally. In 2009, the Federal Reserve implemented the Supervisory Capital Assessment Program (SCAP), known as stress-testing, on the 19 largest bank-holding companies. Since 2011, stress tests have been being conducted as part of the Comprehensive Capital Assessment Review (CCAR) and the Dodd-Frank Act.

Various types of modeling techniques have been applied in the literature in an attempt to forecast bankruptcies. Among the most prominent techniques are: linear probability (Meyer & Pifer, 1970), multivariate discriminant analysis (MDA) (Altman, Haldeman, & Narayanan, 1977; Cox & Wang, 2014; Sinkey, 1975; Stuhr & Van Wicklen, 1974), probit and logit (Cole & Gunther, 1998; Cole & White, 2012; Espahbodi, 1991; Estrella, Park, & Peristiani, 2000; Hanweck, 1977; Martin, 1977; Ohlson, 1980; Thomson, 1991), and Cox proportional hazards models

* Corresponding author.
E-mail address: aagrapet@econ.duth.gr (A. Agrapetidou).

https://doi.org/10.1016/j.ijforecast.2018.01.009

(Lane, Looney, & Wansley, 1986; Shumway, 2001; Whalen, 1991; Wheelock & Wilson, 2000). Recently, several studies have also investigated the use of machine learning based techniques for this purpose.[1]

Boyacioglu, Kara, and Baykan (2009) compared logistic regression, MDA, *k*-means cluster analysis (CA), support vector machine (SVM) and four neural network (NN) architectures in forecasting Turkish bank failures for the period 1997–2003. They used a very small sample with 21 failed banks, 44 solvent banks and 20 explanatory variables. Although NN yielded the best results, SVM outperformed the majority of the other techniques, with a forecasting accuracy of 90.90%. In a similar setup, Ecer (2013) also compared the performances of NN and SVM for forecasting bank failures for the period 1994–2001. The dataset consisted of a small sample of 34 Turkish banks and 36 financial ratios. Again, NN yielded the optimum out-of-sample result (97.06%). Along the same lines, Erdogan (2013) applied SVM to the forecasting of Turkish bank failures in the period 1997–2003. The dataset consisted of 42 Turkish commercial banks, of which 18 had failed and 24 were solvent. He achieved a forecasting accuracy of approximately 95% using 19 financial ratios.

In the relevant literature, most studies rely on accounting data that are augmented in some cases with macroeconomic and market-based variables (Agarwal & Taffler, 2008; Berger & Bouwman, 2013; Cole & Gunther, 1998; Curry, Elmer, & Fissel, 2007; Espahbodi, 1991; Kolari, Glennon, Shin, & Caputo, 2002; Männasoo & Mayes, 2009; Martin, 1977; Meyer & Pifer, 1970; Thomson, 1991). The most commonly used variables are those based on CAMELS[2] indicators. These include measures of capital adequacy, asset quality, management, earnings, liquidity and sensitivity to market risk. Regarding non-financial firms, the primary variables used for predicting bankruptcies are those based on the prediction models of Altman et al. (1977) and Ohlson (1980). The strong performance of Ohlson's model that researchers have demonstrated over time supports its effectiveness (Begley, Ming, & Watts, 1996; Grice & Dugan, 2003; Hillegeist, Keating, Cram, & Lundstedt, 2004; Karamzadeh, 2013).

Due to the large number of bank failures in the recent crisis in the U.S., numerous studies since have aimed to forecast the insolvency of financial institutions. Jordan, Rice, Sanchez, Walker, and Wort (2010) examined 225 failed banks and 225 solvent ones for the period 2007–2010. They performed discriminant analysis and achieved an out-of-sample forecasting accuracy of 78.10%. Mayes and Stremmel (2014) examined a large number of 16,188 U.S. banks from 1992 to 2012 with the aim of forecasting bank failures during the period 2008–2012. They employed a logit model and achieved an out-of-sample forecasting accuracy of 83%. Papadimitriou, Gogas, Plakandaras, and Mourmouris (2013) used a sample of 300 U.S. banks for forecasting U.S bank failures with SVM, and obtained an out-of-sample forecasting accuracy of 76.40% using only six input variables that refer to banks' efficiency, leverage and market appreciation in terms of goodwill and other intangibles. Iturriaga and Sanz (2015) compared NN and SVM for forecasting U.S. bank failures over the period 2002–2012. Their dataset consisted of 386 failed U.S. banks and 386 solvent ones. They found the NN to outperform the SVM for the short-term (one year) horizon, but the SVM to outperform the NN for the medium and long-term forecasting horizons (two and three years before failure). Their optimum one-year horizon model achieved a forecasting accuracy of 94.23% with NN. The two- and three-year horizon models produced forecasting accuracies of 86.54% and 82.69% via SVM. According to their study, the most important variables are provisions, risk concentration on the construction industry and equity support to loans. Cleary and Hebb (2016) examined 323 banks that failed in the U.S. over the period 2002–2011 and an equal sample of non-failed ones. They used discriminant analysis and variables related to bank capital, loan quality and profitability to forecast bank failures in out-of-sample data, and achieved a forecasting accuracy of 89.50%.

This paper presents an SVM-based methodology for forecasting the bankruptcy of U.S. financial institutions over the period 2007–2013 using financial data taken from the banks' publicly-available financial statements. The proposed approach includes a two-step feature selection process that is used to find the most relevant variables for the identification of soon-to-fail banks. These variables are then fed into an SVM model that has been optimized through a training and testing procedure. This study introduces three innovations. First, in contrast to the relevant literature, which uses one predetermined cut-off level in the SVM decision function, here we identify the optimum cut-off level out of several alternatives. Our results show that the optimum cut-off level is different from the standard one used in the literature. Second, even though the proposed model deals with a binary classification problem ("solvent" or "failed"), the corresponding sensitivity analysis offers a quantitative tool for measuring the confidence of our forecast. This can also be extended to an alternative stress-testing tool: for each explanatory variable (or a combination of them), we can measure the change that would be necessary in order to reclassify the bank from "solvent" to "failed" or vice versa. Thus, this procedure provides us with a sensitivity analysis of the resulting classification. Finally, the third innovation has to do with the sample size and is two-fold: (a) we attempt to construct a forecasting model for all U.S. bank failures for the period 2007–2013, and (b) we employ a realistic ratio of 1:10 of failed to solvent banks in the out-of-sample data, instead of the 1:1 that has been used in most previous studies.

[1] For an extensive review of forecasting banks' bankruptcy via statistical and intelligent techniques, see Kumar and Ravi (2007), Demyanyk and Hasan (2010) and Chen, Ribeiro, and Chen (2016).

[2] CAMELS is a rating system that was introduced in 1979 by U.S. regulators for assessing the financial condition of banks by assigning ratings from 1 (strong) to 5 (weak).