Expert Systems With Applications 49 (2016) 48–59

Contents lists available at ScienceDirect

Expert Systems With Applications

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

Business health characterization: A hybrid regression and support vector machine analysis

Rudrajeet Pal[a,*], Karel Kupka[b], Arun P. Aneja[c], Jiri Militky[d]

[a] Department of Business Administration and Textile Management, University of Borås, 50190 Borås, Sweden
[b] TriloByte Statistical Software Ltd., Hradistska 300 Stare Hradiste, Pardubice 533 52, Czech Republic
[c] College of Engineering & Technology, East Carolina State University, Science and Technology Complex Suite 100, Greenville, NC 27858-4353, United States
[d] Faculty of Textile Engineering, Technical University of Liberec, Studentská 1402/2, 461 17 Liberec 1, Czech Republic

ARTICLE INFO

Keywords:
Business health
Credit rating
Predictive classification model
Support vector machine
Bankruptcy prediction
Variable selection

ABSTRACT

Business health prediction is critical and challenging in today's volatile environment, thus demand going beyond classical business failure studies underpinned by rigidities, like paired sampling, a-priori predictors, rigid binary categorization, amongst others.

In response, our paper proposes an investor-facing dynamic model for characterizing business health by using a mixed set of techniques, combining both classical and "expert system" methods. Data for constructing the model was obtained from 198 multinational manufacturing and service firms spread over 26 industrial sectors, through Wharton database.

The novel 4-stage methodology developed combines a powerful stagewise regression for dynamic predictor selection, a linear regression for modelling expert ratings of firms' stock value, an SVM model developed from unmatched sample of firms, and finally an SVM-probability model for continuous classification of business health. This hybrid methodology reports comparably higher classification and prediction accuracies (over 0.96 and ~90%, respectively) and predictor extraction rate (~96%). It can also objectively identify and constitute new unsought variables to explain and predict behaviour of business subjects.

Among other results, such a volatile model build upon a stable methodology can influence business practitioners in a number of ways to monitor and improve financial health. Future research can concentrate on adding a time-variable to the financial model along with more sector-specificity.

1. Introduction

Business insolvency, liquidation or eventual failures and bankruptcies have exacerbated amidst increased overall turbulence in the business environment in recent years. Business health in such highly competitive environment is dependent upon the financial solvency, relative ability to generate cash, capital market access, and financial capacity to deal with unplanned cash problems (Kumar & Ravi, 2007). Scholarly discussions have so far rigidly dichotomized business health in terms of a juridical definition of corporate failures (Charitou, Neophytou, and Charalambous (2004), among others) and financial distress (like several years of negative operating income, low interest coverage, or negative EBIT etc. (Platt & Platt, 2002, 2004), among others) (cf. Balcaen and Ooghe (2006), pp. 72 for a list of the studies). Since the seminal contributions by Beaver (1966) and Altman (1968), numerous business failure prediction models have been devised using various statistical methods such as regression analysis, linear discriminant analysis, multi-variate analysis, etc. which have become a major research domain within corporate finance (Balcaen & Ooghe, 2006; Huang, Tsai, Yen, & Cheng, 2008). They suggest that firms those are incapable of accruing certain benchmarked financial performance deteriorates in their financial situation and becomes unhealthy, and eventually suffers from business failure. However, another interesting way of characterizing business health for the investment community is through evaluation of stock value or investment, particularly when studies have shown that expert analysts' performance on buy and sell recommendations for firms are better than that of actual market returns. However, current stock research is still under-developed due to several reasons, in particular because of the lack of explicit knowledge in appraising the effectiveness of the research to investors. Further, there is also a high degree of undocumented and less formalized tacit knowledge associated with stock evaluation till date (Song, Adams, & Rhee, 2007).

* Corresponding author. Tel.: +46 33 435 45 30.
E-mail addresses: rudrajeet.pal@hb.se (R. Pal), kupka@trilobyte.cz (K. Kupka), anejaap@gmail.com (A.P. Aneja), jiri.militky@tul.cz (J. Militky).

http://dx.doi.org/10.1016/j.eswa.2015.11.027
0957-4174/

Apart from the above-mentioned ambiguity in the methodological standpoint while characterizing business health, classical statistical failure prediction models most widespread in extant studies have several practical limitations owing to their restrictive assumptions[1]. Significant among those assumptions is the paired sampling of failed and non-failed classes of firms based upon an arbitrary dichotomized definition of failure with rigid and static discrimination between them (Balcaen & Ooghe, 2006; Kamath and He, 2006; Shuai & Li, 2005).

In comparison, recently developed intelligent techniques or "expert systems" like artificial intelligence and machine learning applied to business health prediction studies have been seen to outperform in their prediction and performance criteria thus overcoming some of the inherent problems of the classic statistical modelling techniques (Ahn, Cho, & Kim, 2000; Huang et al., 2008; Min & Lee, 2005; Shin, Lee, and Kim, 2005). However these "expert system" studies still follow by and large the classical paradigm approach. For example, recently utilized support vector machine (SVM) technique, even though has outperformed other intelligent techniques (neural network) and traditional methods (MDA and regression analysis) in business failure prediction (Min & Lee, 2005)[2], it has still followed the classical paradigm of failure prediction to characterize business health based upon the method of pairing of failed and non-failed firms. This results in classifying firms' health into binary variables, e.g. 'good/healthy' or 'bankrupt/non-bankrupt' on the basis of the most popular financial ratios (for example, Cielen, Peeters, and Vanhoof (2004), Shin et al. (2005), among others). Even though some recent papers have modelled business failure for unmatched samples by using SVM technique (e.g. Ding, Song, and Zen (2008), Hua, Wang, Xu, Zhang, and Liang (2007) etc.), the use of continuous variables for business health classification is still underdeveloped.

Another important aspect of these bankruptcy prediction studies is the choice of the most appropriate financial features for modelling. Several qualitative and quantitative feature selection procedures have been utilized, starting from Beaver's (1966) initial adoption of financial indicators based on qualitative criteria to advanced statistical methods like stepwise regression or *t*-test in recent studies. However theorists and practitioners have always struggled to establish an effective feature set due to numerous reasons, e.g. judgment and popularity-based qualitative techniques, or due to 'brute empiricism' of advanced statistical procedures (Balcaen & Ooghe, 2006; Keasey & Watson, 1991). This always calls for research on adoption of latest and advanced techniques for feature selection.

In this context, our paper proposes a method for modelling a recent expert rating (developed by StarMine) adopted widely in stock research by using a hybrid of traditional and "expert system" methods, viz. stagewise and linear regressions, linear SVM classifier technique and SVM-probability model. Our paper addresses the above-mentioned ambiguities in three ways, by: (i) developing a continuous SVM-probability model for business health prediction by using "expert system", instead of the paired binary technique for rigid categorization, (ii) using latest stagewise regression method, instead of traditional feature selection methods in failure predictions, and (iii) creating an investor-facing model, instead of having an arbitrary definition of business failure most often used to characterize business health.

The paper is organized as follows: Section 2 provides a review of the scholarly discussion on business failure and bankruptcy prediction, followed by a brief state-of-art on feature selection methods and business failure prediction studies using SVM. In Section 3, the design of the experiment conducted is presented along with the related methodological background adopted in the paper for developing a model for business health characterization. In Section 4, the results of the empirical study is presented and discussed. Finally, from the findings of this research, we can derive some conclusions and implications for applications.

2. Literature review—business health, failure and related research

2.1. Business health research

In highly competitive environments favouring failures, businesses search for their ability to maintain a healthy state over time, despite being subjected to negative and/or destructive market events (Pearce II & Robbins, 1993; Sundström & Hollnagel, 2006). Hence it becomes a need to seek for an effective way to identify and handle business failures by achieving profitability and financial solvency (Wu, 2010). From the general system's theory perspective, few studies on business health have focussed on achieving certain business goals, viz. shareholder value, profitability, and customer equity (Sundström and Hollnagel (2006) among others). However, these have been rather qualitative interpretations. Probably because of this lack of proper characterization scholarly discussions on business financial health (of firms) have mainly fallen back on focussing towards business failure predictions or similar (financial distress or bankruptcies), or towards firms' creditworthiness.

Business failures are defined as "situations in which a firm cannot pay lenders, preferred stockholders, suppliers, etc. or a bill is overdrawn, or the firm is bankrupt according to law resulting in huge economic losses to investors and others, together with substantial social and economic costs" (Ahn et al., 2000; Dimitras, Zanakis, & Zopounidis, 1996). Business failure prediction has been a major research domain within corporate finance over 35 years resulting in the development of various modelling techniques falling into three categories: classical statistical methods, artificially intelligent expert system models and theoretical models (Aziz & Dar, 2006; Balcaen & Ooghe, 2006). In this context, the next section provides a brief conceptualization of business failure and bankruptcy studies and adopted prediction methods.

Another common research stream to reflect on business health is credit worthiness or credit risks of firms judged mainly by the banks or credit rating agencies. At the level of the banks, actions include measurement of credit risks using various internal rating based (IRB) approaches in order to judge the capital level of the businesses. Various leading rating agencies like Moody's and Standard and Poor's also uses popular qualitative rating grades (e.g. AAA and Aaa, respectively) to symbolize the economic risks (Oelerich & Poddig, 2006). Such credit ratings can be used to map or predict future rating grades, hence probability of default using various quantitative statistical models by distinguishing between solvent and insolvent firms (cf. Carey and Hrycay (2001), Jankowitsch, Pichler, and Schwaiger (2007) and others). However, Standard and Poor's guide to credit ratings suggest that these ratings are not really an indicator of investment merit (buy/hold/sell or measure of asset value) (Standard & Poor's, 2014). Even though these credit ratings have been very popular among banks and financial institutions to judge the creditability of firms to pay back debts and loans, these have been less attractive and available to investment community and market participants for usage.

Even though in this paper, we aim at developing an effective decision support system for individual investors by connecting business health studies to stock research, scholarly

[1] Such as linearity, normality, non-stationarity and data instability, feature selection and even arbitrary definition of failure (Balcaen and Ooghe, 2006).

[2] SVM has demonstrated several attractive features like high generalization performance, and superior structural risk minimization (SRM) principle making it immensely popular (Min and Lee, 2005; Shin et al., 2005).