Expert Systems with Applications 40 (2013) 385–393

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

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

Clustering and visualization of bankruptcy trajectory using self-organizing map

Ning Chen [a,*], Bernardete Ribeiro [b], Armando Vieira [a], An Chen [a,c]

[a] GECAD, Instituto Superior de Engenharia do Porto, Instituto Politecnico do Porto, Rua Dr. Antonio Bernardino de Almeida, 431, Porto, 4200-072, Portugal
[b] CISUC, Department of Informatics Engineering, University of Coimbra, Rua Silvio Lima-Polo II, Coimbra, 3030-790, Portugal
[c] Institute of Policy and Management, Chinese Academy of Sciences, No. 15, Zhongguancun Beiyitiao, Beijing, 100190, PR China

ARTICLE INFO

Keywords:
Bankruptcy risk
Trajectory pattern
Self-organizing map
Visual clustering

ABSTRACT

Bankruptcy trajectory reflects the dynamic changes of financial situation of companies, and hence make possible to keep track of the evolution of companies and recognize the important trajectory patterns. This study aims at a compact visualization of the complex temporal behaviors in financial statements. We use self-organizing map (SOM) to analyze and visualize the financial situation of companies over several years through a two-step clustering process. Initially, the bankruptcy risk is characterized by a feature self-organizing map (FSOM), and therefore the temporal sequence is converted to the trajectory vector projected on the map. Afterwards, the trajectory self-organizing map (TSOM) clusters the trajectory vectors to a number of trajectory patterns. The proposed approach is applied to a large database of French companies spanning over four years. The experimental results demonstrate the promising functionality of SOM for bankruptcy trajectory clustering and visualization. From the viewpoint of decision support, the method might give experts insight into the patterns of bankrupt and healthy company development.

1. Introduction

Bankruptcy prediction has long been a well studied topic in the literature, and remains more active in the face of the recent severe challenges of worldwide financial crisis. The increase of financial failure accelerates the economic deterioration and yields a lot of social problems. It becomes critically important to explore the potential bankruptcy behaviors and understand the implicit patterns from the perspective of early warning and decision support. To date, a large amount of research has been carried out using different methods, such as univariate and multivariate analysis, neural network, support vector machine, and rough set. Most of the prior studies are focused on financial distress prediction problem based on a static snapshot of financial situation. However, bankruptcy trajectory which characterizes the dynamic changes of financial situation receives little attention. To our knowledge, only a few works attempted to analyze the temporal sequence of financial statements (Jardin & Severin, 2011; Kiviluoto & Bergius, 1998; Schreck, Bernard, Tekusova, & Kohlhammer, 2009).

Self-organizing map (SOM) is a non-parametric neural network with the desirable combination of data abstraction and spatialization, and widely used for visual clustering in a wide range of applications. In this paper, we study the changes of financial situation of companies to examine the trajectory patterns through a two-step clustering process by extending our earlier work (Chen, Ribeiro, & Vieira, 2011). A self-organizing map clustering approach is proposed to analyze (and visualize) the effect of temporal evolution of some financial indictors in order to assess and establish eventual scenarios of bankruptcy. Initially, a feature self-organizing map (FSOM) is constructed to characterize the bankruptcy risk of companies. Afterwards, the instantaneous observations of temporal sequence are successively projected on the map and the positions are concatenated to a trajectory vector. The trajectory patterns are then learned by a trajectory self-organizing map (TSOM) and shown through appropriate visual representation. The proposed approach is applied to a large data set of French companies containing financial ratios in four consecutive years (2003–2006) and the final state in the following year (2007). The experimental results demonstrate the promising functionality of SOM for bankruptcy trajectory clustering and visualization. Taking the perspective of decision support, the described method might give experts insight into patterns of bankrupt or healthy company development.

The remainder of this paper is organized as follows. Section 2 reviews the related studies with the emphasis on bankruptcy prediction and trajectory mining. Section 3 presents the data set under exploration and the methodology of a SOM-based trajectory analysis approach. Further details on each phase of the approach are schematically illustrated. In Section 4, the experimental results are reported including the model parametrization, trajectory pattern analysis and component plan visualization. Lastly, the contributions and future remarks are discussed in Section 5.

* Corresponding author. Tel.: +351 228340500; fax: +351 228321159.
E-mail addresses: ningchen74@gmail.com (N. Chen), bribeiro@dei.uc.pt (B. Ribeiro), asv@isep.ipp.pt (A. Vieira), acn@isep.ipp.pt (A. Chen).

0957-4174/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.eswa.2012.07.047

2. Related work

Corporate bankruptcy prediction is a well-researched area in finance to predict the probability of business failure, given the historic data that describe the situation of a company over a given period. There has been a raising interest in seeking for more accurate prediction models able to better understand the financial data and prevent the sudden distress of companies. So far a large number of methods have been proposed following the research direction of statistical or intelligent approaches (Bellovary, Giacomino, & Akers, 2007). Due to the criticism on traditional statistical models, many recent efforts have been devoted into state-of-the-art intelligent approaches, which offer theories about how financial crises could be predicted. Various prediction models have been proposed using a wide range of intelligent methods including neural networks (NNs), fuzzy set theory (FS), decision tree (DT), case-based reasoning (CBR), support vector machines (SVMs), and soft computing (Ravi Kumar & Ravi, 2007). Neural networks have been the subject of many research activities in bankruptcy prediction yielding reasonably accurate models (Charalambous, Charitou, & Kaourou, 2000). Form the many studies existing in the literature, neural networks are generally superior to other methods (Atiya, 2001). A multi-layer perceptron (MLP) obtains desirable outcome on Taiwan and United States markets (Huang, Chen, Hsu, Chen, & Wu, 2004), and Iranian companies (Rafiei, Manzari, & Bostanian, 2011). In Chen, Vieira, Ribeiro, Duarte, and Neves (2011), a stable credit rating model based on Learning Vector Quantization (LVQ) neural network is successfully applied to corporate failure prediction and credit risk analysis of French companies. Likewise, SVM has shown its applicability to this problem compared with other core machine learning techniques (Chen, 2011; Tsai, 2008; Yang, You, & Ji, 2011). Some authors have mentioned that a well designed hybrid or ensemble is effective to boost the predictive capability of single predictor (Verikas, Kalsyte, Bacauskiene, & Gelzinis, 2010). Take for instance, the hybrid credit scoring model of ANN and CBR (Chuang & Hung, 2011) and the ensemble that aggregates DT, MLP and SVM (Hung & Chen, 2009).

Regardless of different techniques, most of the previous work share the common goal at an accurate prediction of the final state of companies, in other words, whether a company is bankrupt or not after a given period. In this viewpoint, they study the bankruptcy problem as a static result more than a dynamic process. Whereas from the perspective of decision support, a more important issue than accurate prediction might be the deep analysis and well understanding of the dynamic changes of financial situation of companies. Some questions have been raised: Which factors contribute mostly to the difference between the bankrupt companies and healthy ones? How to characterize the evolution of companies and identify (and visualize) the important trend? Bankruptcy trajectory analysis is a subject attempting to find solutions to these issues.

2.1. Self-organizing map and applications in bankruptcy prediction

Self-organizing map (SOM) is an unsupervised neural network proposed by Kohonen (1982) for visual cluster analysis. The neurons of the map are located on a regular grid embedded in a low (usually 2 or 3) dimensional space, and associated with the cluster prototypes by the connected weights. In the course of learning process, the neurons compete with each other through the best-matching principle in such way that the input is projected to the nearest neuron given a defined distance metric. The winner neuron and its neighbors on the map are then adjusted towards the input in proportion with the neighborhood distance, consequently the neighboring neurons likely represent the similar patterns of the input data space. Due to the data clustering and spatialization through the topology preserving projection, SOM is widely used in the context of visual clustering applications. Despite the unsupervised nature, the applicability of SOM is extended to classification tasks by means of a variety of ways, such as neuron labeling method, semi-supervised learning (Heikkonen, Koikkalainen, & Oja, 1993), and supervised learning vector quantization (LVQ) (Kohonen, 2001).

In the field of bankruptcy prediction, SOM is one of the frequently used models to analyze the high-dimensional financial data and understand the unwanted bankruptcy phenomenon. A wide range of research groups concentrate on the bankruptcy prediction problem, usually solved as a classification task to separate the companies into distress and healthy category (binary class) or a number of predefined credit rates (multi-class). The prediction models are constructed from the training data in terms of some criteria, such as the overall misclassification error, the Neyman–Pearson criterion (Kiviluoto, 1998), and the total misclassification cost (Chen, Vieira, Duarte, Ribeiro, & Neves, 2009). The capability of SOM and its supervised variants has been demonstrated in comparison with statistical and other intelligent methods. Recent examples are as follows. SOM is used to determine the credit class through a visual exploration (Merkevicius, Garsva, & Simutis, 2004). An enhanced version of LVQ can boost the prediction performance of MLP (Neves & Vieira, 2006). It is demonstrated that LVQ outperforms other neural networks, support vector machines and multivariate statistical methods on predicting the financial failure of Turkish banks (Boyacioglu, Kara, & Baykan, 2009). In cooperation with independent component analysis for dimensionality reduction, LVQ is able to recognize the distressed French companies (Chen & Vieira, 2009).

2.2. Trajectory mining and bankruptcy trajectory analysis

Trajectory data has attracted considerable attention in many applications where the object movements are routinely collected as time-dependent observation sequences, such as traffic monitoring, visual surveillance, robotic navigation, and stock prediction. Trajectory mining attempts to explore the implicit patterns from trajectory data. To achieve this, SOM has been introduced as a valuable tool in various applications. In a robotic navigational environment, SOM clusters the motion trajectory of moving objects and predicts the next instant position (Rajpurohit & Manohara, 2009). A SOM clustering model provides a powerful tool to visualize the dynamic behaviors of industrial processes for human supervision and fault detection (Fuertesa et al., 2010). In Schreck et al. (2009), a SOM-based framework is designed for general-purpose enabling users to visually monitor and interactively control the clustering process of trajectory data.

In the field of bankruptcy analysis, we may map the financial situation of a company to a 2-dimensional space, and observe the evolution of the financial situation as a trajectory in that space. These trajectories reflect the dynamic changes rather than a static snapshot of financial situation, and hence make possible to detect the time evolution of companies and recognize the trajectory patterns. SOM is regarded as an appealing information visualization technique for bankruptcy trajectory analysis due to its superiority on transforming the information into graphical representation so as to facilitate the pattern recognition and visual reasoning by decision makers. In Kiviluoto and Bergius (1998), a hierarchical SOM model is employed to explore the year-to-year trajectory of enterprises and view the time evolution of the situation. A recent study (Jardin & Severin, 2011) shows the advantages of using trajectory for medium-period forecast compared with the traditional single-period and multi-period forecast. They used SOM to design the trajectory of corporate collapse, and then forecasted the