Applied Soft Computing 44 (2016) 144–152

Contents lists available at ScienceDirect

Applied Soft Computing

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

Financial distress prediction using the hybrid associative memory with translation

L. Cleofas-Sánchez [a], V. García [b], A.I. Marqués [c], J.S. Sánchez [a,*]

[a] Institute of New Imaging Technologies, Department of Computer Languages and Systems, Universitat Jaume I, 12071 Castelló de la Plana, Spain
[b] División Multidisciplinaria de Ciudad Universitaria, Universidad Autónoma de Ciudad Juárez, 32310 Ciudad Juárez, Chihuahua, Mexico
[c] Department of Business Administration and Marketing, Universitat Jaume I, 12071 Castelló de la Plana, Spain

ARTICLE INFO

Article history:
Received 22 July 2015
Received in revised form 4 April 2016
Accepted 6 April 2016
Available online 12 April 2016

Keywords:
Associative memory
Neural network
Financial distress
Bankruptcy
Credit risk

ABSTRACT

This paper presents an alternative technique for financial distress prediction systems. The method is based on a type of neural network, which is called hybrid associative memory with translation. While many different neural network architectures have successfully been used to predict credit risk and corporate failure, the power of associative memories for financial decision-making has not been explored in any depth as yet. The performance of the hybrid associative memory with translation is compared to four traditional neural networks, a support vector machine and a logistic regression model in terms of their prediction capabilities. The experimental results over nine real-life data sets show that the associative memory here proposed constitutes an appropriate solution for bankruptcy and credit risk prediction, performing significantly better than the rest of models under class imbalance and data overlapping conditions in terms of the true positive rate and the geometric mean of true positive and true negative rates.

1. Introduction

A large number of techniques have been developed to help decision-makers and analysts in predicting financial distress. Traditionally, decisions on credit risk of a corporate borrower were exclusively based upon subjective judgments made by human experts, using past experiences and some guiding principles [69]. However, two major problems with this approach are the difficulty to make consistent estimates and the fact that it tends to be reactive rather than predictive. The world financial crisis has led to increasing attention of banks and financial institutions on this question because of its significant impact on the decisions made [14], resulting in the development of numerous techniques to face the important challenge of credit risk and bankruptcy prediction from financial ratios using mathematical models. From the pioneer work by Altman [7], based on multivariate discriminant analysis, a variety of statistical and operations research methods have subsequently been applied to credit risk and bankruptcy prediction, including linear and logistic regression, multivariate adaptive regression splines, survival analysis, linear and quadratic programming, and multiple criteria programming. Most of these techniques typically rely on the assumptions of linear separability, multivariate normality and independence of the predictive variables, but they are very often violated in real-life problems [25,34,55].

Popular computational intelligence tools such as decision trees, neural networks, support vector machines, fuzzy systems, rough sets, artificial immune systems, and evolutionary algorithms are techniques that can deal with non-linearity. Besides, these methods are highly capable of extracting meaningful information from imprecise data and detecting trends that are too complex to be discovered by either humans or conventional systems. Despite various studies have concluded that no technique is clearly superior to other competing algorithms because it depends on the characteristics of the problem analyzed [13,15,16], different neural network architectures have shown good performance in comparison to other methods for a range of financial applications [10,19,48,53,78]. However, when the number of examples is relatively small, several works have demonstrated that the accuracy and generalization performance of a support vector machine (SVM) is usually better than that of statistical and other soft computing techniques [23,24,65,67]. While typical neural networks used in this context are the multi-layer perceptron (MLP), the radial basis function (RBF) and the probabilistic or Bayesian network (BN), other neural models such as the associative memories have not been explored as yet.

The ability of human brain to make associations from partial information has historically attracted great interest among

* Corresponding author. Tel.: +34 964 728350.
E-mail address: sanchez@uji.es (J.S. Sánchez).

http://dx.doi.org/10.1016/j.asoc.2016.04.005
1568-4946/

researchers, leading to a variety of theoretical neural networks that act as associative memories. An associative memory [39] is an early type of artificial neural network that relates an input vector **x** with an output vector **y**. The functionality of associative memories is reached in two phases: learning and recall. The learning process consists of building a connection matrix **W** with a value for each association ($\mathbf{x}^k$, $\mathbf{y}^k$). In the recall phase, an output vector **y**, which corresponds to the most similar to the input vector **x**, is obtained from the associative memory. These models are powerful computational tools due to their conceptual and implementational simplicity, their strong mathematical foundation, and their capability of storing huge amounts of data that allow to properly recover the most similar patterns to an input vector with low computational efforts [77].

Representative examples of associative memories are lernmatrix [66], the linear associator [8,38], the Moore–Penrose generalized inverse associative memory [40], the Hopfield network [28], the bidirectional associative memory [41], the fuzzy associative memory [42], the morphological associative memory [58], and the alpha–beta associative memory [2]. Some of these models have been used to solve very different problems. Sabourin and Mitiche [59] developed a Kohonen associative memory with selective multiresolution for OCR. A fuzzy associative memory was introduced to determine rock types from well-log signatures [17]. The bidirectional associative memory networks were used to find the relations between various cancers and elemental contents in serum samples with the aim of diagnosing cancer [81]. A hybrid classifier based on self-organizing maps and associative memories was designed for speaker recognition [31]. Zhang et al. [79] proposed a modular face recognition scheme by combining the wavelet subband representations and kernel associative memories. An associative memory based on the restricted Coulomb energy was also applied to human face recognition [49]. Namba and Zhang [50] devised an associative memory to recognize Braille images. A novel system for medical diagnosis based on associative memories was proposed by Aldape-Pérez et al. [5]. Itkar and Kulkarni [32] developed an efficient algorithm for mining frequent patterns using an auto-associative memory.

Apart from the associative memories just mentioned, Santiago-Montero [63] introduced the hybrid associative classifier and its extension, the hybrid associative classifier with translation (HACT). Both these associative memories are based on the learning phase of the linear associator and the recall phase of the Steinbuch's lernmatrix. This paper applies the HACT neural network to decision making problems for financial distress prediction and presents an empirical comparison with other popular prediction methods. To the best of our knowledge, this model has not been used for classification purposes, and even less in the context of finance and management. The aim of this paper therefore is four-fold:

1. To explore the capability of the HACT model in the prediction of bankruptcy and credit risk.
2. To analyze the behavior of this neural network under the presence of imbalance in class distribution, which constitutes a data complexity often neglected in financial applications.
3. To investigate how the class overlapping affects the performance of the associative memory.
4. To compare the performance of HACT with that of other prediction techniques.

From now on, the paper is organized as follows. Section 2 provides a review of works related to neural networks used for corporate bankruptcy and credit risk prediction. Section 3 introduces the fundamental concepts of the associative memories and describes the bases of the HACT model. The experimental set-up and databases are given in Section 4, while the results are discussed in Section 5. Finally, Section 6 presents the concluding remarks and outlines some directions for future research.

2. A review of neural networks applied to financial distress prediction

From the beginning of the 1990s, the development of artificial neural network technologies for bankruptcy and credit risk prediction problems has been the subject of considerable attention and research efforts. The first reference to using neural networks can be found in the paper by Odom and Sharda [51], showing that a three-layer feed-forward perceptron is more accurate and robust than multi-variate discriminant analysis. After this seminal work, many other studies have proposed the use of neural networks in credit scoring, bankruptcy or business failure prediction. For instance, Tam and Kiang [68] compared neural network models to linear discriminant analysis, logistic regression, nearest neighbors and decision tree for evaluating bank status. Salchenberger et al. [61] reported that the neural networks produced fewer or equal number of total errors, type-I errors and type-II errors compared to the logit model. Lacher et al. [43] investigated the use of the Cascade-Correlation neural network architecture and compared its performance with that of the multivariate discriminant analysis approach. Chang et al. [18] applied the theory and numerical algorithms of the BN to risk scoring and compared the results with traditional methods for computing scores and posterior predictions of performance variables.

Desai et al. [22] concluded that the MLP and the modular neural network can be especially useful to correctly predict the bad loans, but logistic regression models are comparable to the neural networks when the performance is measured by the percentage of good and bad loans correctly classified. West [74] analyzed the credit scoring accuracy of the MLP, the RBF network and several statistical techniques, suggesting that the MLP may not be the most accurate neural network model. An auto-associative memory trained with only data of non-bankrupt firms was developed by Baek and Cho [11]. Baesens et al. [12] used Markov Chain Monte Carlo search to learn unrestricted Bayesian network classifiers for credit scoring, which gave a very good performance in terms of accuracy and area under the ROC curve. Also, Leong [46] showed that the BN performs well against logistic regression and MLP particularly with class imbalance, higher dimensions and a rejection sample, and it can be scaled efficiently when implemented onto a large data set.

The power of probabilistic and MLP neural networks was compared to that of discriminant analysis, probit analysis and logistic regression to evaluate credit risk in Egyptian banks [1]. Khashman [36] explored various back-propagation learning schemes to train three models (each with a different number of hidden neurons) of a three-layer supervised neural network. Angelini et al. [9] developed two neural network systems with a four-layer feed-forward topology, proving their applicability to credit risk prediction. An algorithm based on the threshold accepting meta-heuristic to train the principal component neural network architecture was investigated by Ravi and Pramodh [56], who inferred that their proposal outperformed other classifiers.

Lee and Chen [44] developed a credit scoring system using a hybrid modeling procedure with artificial neural networks whose input nodes were the variables obtained by multivariate adaptive regression splines. Hsieh [29] designed a credit scoring model that employed the SOM and K-means clustering algorithms to obtain the best inputs to a feed-forward MLP. Similarly, Lee et al. [45] explored the performance of credit scoring by integrating the linear discriminant analysis approach into a three-layer back-propagation neural network, revealing that the proposed hybrid approach converges much faster than the conventional neural network model