PERGAMON

Expert Systems with Applications 18 (2000) 65–74

Expert Systems with Applications

www.elsevier.com/locate/eswa

The integrated methodology of rough set theory and artificial neural network for business failure prediction

B.S. Ahn[a,*], S.S. Cho[b], C.Y. Kim[c]

[a]Department of Business Administration and Accounting, Suwon University, San 2-2, Wau, Bongdam, Hwasung, Kyonggi 445-743, South Korea
[b]Graduate School of Management, KAIST, 207-43, Cheongryangri, Dongdaemoon, Seoul 130-012, South Korea
[c]Syntech Inc., 1556-9, Seocho, Seoul 137-070, South Korea

Abstract

This paper proposes a hybrid intelligent system that predicts the failure of firms based on the past financial performance data, combining rough set approach and neural network. We can get reduced information table, which implies that the number of evaluation criteria such as financial ratios and qualitative variables is reduced with no information loss through rough set approach. And then, this reduced information is used to develop classification rules and train neural network to infer appropriate parameters. The rules developed by rough set analysis show the best prediction accuracy if a case does match any of the rules. The rationale of our hybrid system is using rules developed by rough sets for an object that matches any of the rules and neural network for one that does not match any of them. The effectiveness of our methodology was verified by experiments comparing traditional discriminant analysis and neural network approach with our hybrid approach. For the experiment, the financial data of 2400 Korean firms during the period 1994–1997 were selected, and for the validation, *k*-fold validation was used.

Keywords: Business failure prediction; Rough sets; Neural network

1. Introduction

Evaluation of the business failure has been, for a long time, a major preoccupation of researchers and practitioners. Business failure is a general term and, according to a widespread definition, is the situation that a firm cannot pay lenders, preferred stock shareholders, suppliers, etc., or a bill is overdrawn, or the firm is bankrupt according to the law. All these situations result in a discontinuity of the firm's operations. The number of failing firms is an important indicator for the health of the economy and it can be considered as an index of the development and robustness of the economy. Clearly, failure affects a firm's entire existence and it has high cost to the firm, the collaborators (firms and organizations), the society and finally the country's economy (Warner, 1977).

The development and use of models, able to predict failure in advance, can be very important for the firms in two different ways. First, as "early warning systems", such models can be very useful for those (i.e. managers, authorities, etc.) who have to prevent failure. Second, such models can be useful in aiding decision-makers of financial institutions in charge of evaluation and selection of the firms.

At the beginning, statistical methods such as univariate statistical methods, multiple discriminant analysis, linear probability models, and logit and probit analysis have been mainly used for business classification problems (Altman, 1968; Altman, Haldeman & Narayanan, 1977; Collins & Green, 1972). Later, the development and application of artificial intelligence led some researchers to employ inductive learning and neural networks in business domain (Chung & Tam, 1992; Fletcher & Goss, 1993; Odom & Sharda, 1990; Raghupathi, Schkade & Raju, 1991; Salchenberger, Cinar & Lash, 1992; Tam & Kiang, 1992). Many other methods such as multiple criteria decision analysis (MCDA) and rough set approach have been successfully applied to real world classification problems (Siegel, de Korvin & Omer, 1993; Slowinski & Zopounidis, 1995). Rough set theory, introduced by Pawlak (1982) and Pawlak, Grzymala-Busse, Slowinski and Ziarko (1995) is a mathematical tool to deal with vagueness and uncertainty of information and proved to be an effective tool for the analysis of financial information system comprised of a set of objects (firms) described by a set of multi-valued financial ratios and qualitative variables.

The composite models of rough set components and neural network, which take advantage of each method's

* Corresponding author. Tel.: +82-331-220-2616; fax: +82-331-220-2485.

E-mail address: bsahn@mail.suwon.ac.kr (B.S. Ahn).

0957-4174/00/$ - see front matter
PII: S0957-4174(99)00053-6

generic characteristics, were constructed to predict a sample of bank holding patterns (Hashemi, Le Blanc, Rucks & Rajaratnam, 1998; Jelonek, Krawiec & Slowinski, 1995). They used rough sets as a preprocessor for neural network. Using sorting rules developed by rough sets may lead to a burdensome situation where a new case does not match any of the sorting or classification rules. On the other hand, the neural network approach classifies every object by its weighting mechanism, although in a black box fashion. Jelonek et al. (1995) employed one-dimensional (1D) reduction (attribute reduction) and Hashemi et al. (1998) employed 2D reduction (attribute and object reduction). After reduction of information system they trained neural network with the reduced information system. And then, they applied the neural network to prediction.

In this paper, we propose a hybrid intelligent system combining rough set approach and neural network. Rough set approach, by which redundant attributes in multi-attribute information system can be removed without any information loss, is utilized as a preprocessor to improve business prediction capability by neural network. The streamlined information system by rough set approach is fed into neural network for training. At the prediction step, we apply the rules developed by rough sets first, and then we apply the neural network to the objects that does not match any of the rules. The effectiveness of our hybrid approach was verified with experiments that compared traditional discriminant analysis and neural network approach with the hybrid approach.

This paper is organized as follows. Section 2 describes classification techniques used in previous researches concerned with our paper: rough set theory and neural network, respectively. In Section 3, proposed data preprocessing algorithm by rough set and hybrid models is described. In Section 4, we analyze and compare the results of each model, and conclusions in Section 5.

2. Rough sets and neural network

2.1. Rough sets

Pawlak (1982) first introduced rough set theory. The philosophy of the method is based on the assumption that with every object some information (data, knowledge) can be associated. Objects characterized by the same information are indiscernible in view of the available information. The indiscernibility relation generated in this way is the mathematical basis for the rough set theory. Slowinski and Zopounidis (1995) employed rough set approach in business failure prediction. They used 12 financial ratios and compared rough set approach with statistical approaches. Slowinski, Zopounidis and Dimitras (1997) applied the rough set approach in a real problem considered by a Greek bank which finances industrial and commercial firms in Greece presenting a great activity. The bank was interested in investing its funds in the best and dynamic firms.

2.1.1. Information system

By an information system we understand the 4-tuple $S = \langle U, Q, V, \rho \rangle$, where U is a finite set of objects, Q is a finite set of attributes, $V = \cup_{q \in Q} V_q$ and V_q is a domain of the attribute q, and $\rho : U \times Q \to V$ is a total function such that $\rho(x, q) \in V_q$ for every $q \in Q, x \in U$, called an information function. Let $S = \langle U, Q, V, \rho \rangle$ be an information system and let $P \subseteq Q$ and $x, y \in U$. We say that x and y are indiscernible by the set of attributes P in S iff $\rho(x, q) = \rho(y, q)$ for every $q \in P$. Thus every $P \subseteq Q$ generates a binary relation on U which will be called an indiscernibility relation, denoted by $IND(P)$. Obviously, $IND(P)$ is an equivalence relation for any P. Equivalence classes of $IND(P)$ are called P-elementary sets in S. The family of all equivalence classes of relation $IND(P)$ on U is denoted by $U|IND(P)$ or, in short, $U|P$.

$Des_p(X)$ denotes a description of P-elementary set $X \in U|P$ in terms of values of attributes from P, i.e.

$$Des_p(X) = \{(q, v) : \rho(x, q) = v, \qquad \forall x \in X, \ \ \forall q \in P\}$$

2.1.2. Approximation of sets

Let $P \subseteq Q$ and $Y \subseteq U$. The P-lower approximation of Y, denoted by PY, and the P-upper approximation of Y, denoted by $\bar{P}Y$, are defined as:

$$PY = \cup\{X \in U|P : X \subseteq Y\}$$

$$\bar{P}Y = \cup\{X \in U|P : X \cap Y \neq \emptyset\}$$

The P-boundary (doubtful region) of set Y is defined as

$$Bn_p(Y) = \bar{P}Y - PY$$

Set PY is the set of all objects from U which can be certainly classified as elements of Y, employing the set of attributes P. Set $\bar{P}Y$ is the set of objects from U which can be possibly classified as elements of Y, using the set of attributes P. The set $Bn_p(Y)$ is the set of objects which cannot be certainly classified to Y using the set of attributes P only.

With every set $Y \subseteq U$, we can associate an accuracy of approximation of set Y and P in S, or in short, accuracy of Y, defined as:

$$a_p(Y) = \frac{\mathrm{card}(PY)}{\mathrm{card}(\bar{P}Y)}$$

2.1.3. Approximation of a partition of U

Let S be an information system, $P \subseteq Q$, and let $\Psi = \{Y_1, Y_2, \ldots, Y_n\}$ be a partition of U. The origin of this partition is independent on attributes from P; it can follow from solving a sorting problem by an expert. Subsets Y_i, $i = 1, \ldots, n$, are categories of partition Ψ. By P-lower and P-upper approximation of Ψ in S, we mean sets $P\Psi = \{PY_1, PY_2, \ldots, PY_n\}$ and $\bar{P}\Psi = \{\bar{P}Y_1, \bar{P}Y_2, \ldots, \bar{P}Y_n\}$,