PERGAMON

Expert Systems with Applications 25 (2003) 637–646

Expert Systems with Applications

www.elsevier.com/locate/eswa

The discovery of experts' decision rules from qualitative bankruptcy data using genetic algorithms

Myoung-Jong Kim*, Ingoo Han

Graduate School of Management, Korea Advanced Institute of Science and Technology, 207-43 Cheongryangri-Dong, Dongdaemoon-Gu, Seoul 130-012, South Korea

Abstract

Numerous studies on bankruptcy prediction have widely applied data mining techniques to finding out the useful knowledge automatically from financial databases, while few studies have proposed qualitative data mining approaches capable of eliciting and representing experts' problem-solving knowledge from experts' qualitative decisions. In an actual risk assessment process, the discovery of bankruptcy prediction knowledge from experts is still regarded as an important task because experts' predictions depend on their subjectivity. This paper proposes a genetic algorithm-based data mining method for discovering bankruptcy decision rules from experts' qualitative decisions. The results of the experiment show that the genetic algorithm generates the rules which have the higher accuracy and larger coverage than inductive learning methods and neural networks. They also indicate that considerable agreement is achieved between the GA method and experts' problem-solving knowledge. This means that the proposed method is a suitable tool for eliciting and representing experts' decision rules and thus it provides effective decision supports for solving bankruptcy prediction problems.

Keywords: Genetic algorithms; Decision rules; Bankruptcy prediction

1. Introduction

Rapid progress in digital data acquisition and storage technology has led to a fast-growing tremendous amount of data stored in databases, data warehouses, or other kinds of data repositories including the World Wide Web. Information collection has become easier, but the effort required to retrieve relevant knowledge from databases has become significantly greater. As a consequence, there has been a growing interest in data mining which is capable of facilitating the discovery of interesting and useful knowledge from a huge amount of data.

The discovery of knowledge in business data is an important task capable of providing significant competitive advantage for a business organization by exploiting the potential of large databases. Data mining has been applied to various business domains such as marketing, finance, banking, manufacturing and telecommunications (Brachman, Khabaza, Kloesgen, Piatesky-Shpiro, & Simoudis, 1996). Classification is one of the important issues in many business applications. The typical examples of business classification problems include credit approval, securities trading, product selection, risk estimation, personnel selection, and corporate bankruptcy.

Corporate bankruptcy triggers economic losses for management, stockholders, employees, customers and others, together with great social and economic costs to the nation. Thus, the accurate prediction of bankruptcy has been a critical issue in finance. The applications of data mining to bankruptcy prediction have used three major approaches. A popular data mining approach is to develop quantitative models for bankruptcy prediction. Since the study of Altman (1968) on bankruptcy prediction, numerous studies have tried to further develop appropriate quantitative models by applying data mining techniques including discriminant analysis (Altman, Marco, & Varet, 1994), logit (Ohlson, 1980), probit (Zmijewski, 1984), and neural networks (Fletcher & Goss, 1993; Odom & Sharda, 1990; Tam & Kiang, 1992). The core of this approach is learning classification functions consisting of a set of weights among financial variables.

Another quantitative approach is to extract bankruptcy prediction rules automatically from a huge amount of financial database. The data mining techniques, such as inductive learning methods, neural networks, and genetic

* Corresponding author. Tel.: +82-2-958-3613; fax: +82-2-958-3604.
E-mail address: mjkim@kgsm.kaist.ac.kr (M.-J. Kim).

0957-4174/03/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/S0957-4174(03)00102-7

algorithms (GAs), have been successful in obtaining useful bankruptcy prediction rules (Messier & Hansen, 1988; Shaw & Gentry, 1990; Shin & Lee, 2002).

The third data mining approach is to construct qualitative models called subjective models based on experts' problem-solving knowledge. Experts' knowledge plays an important role in an actual risk assessment process. Experts do not wholly depend on the results provided by quantitative approaches to determine the creditworthiness of a particular borrower. Instead they work with their subjective knowledge framework to induce appropriate conclusions from the integration of quantitative and qualitative information that can be used in estimating the default risk of the borrower. They classify various loan applications into categories such as approval, pending, and disapproval using their subjective knowledge framework. Therefore, the risk assessment process heavily relies on the subjective judgment of experts. Interactive techniques such as interviews can be applied to investigating experts' knowledge framework associated with bankruptcy prediction. However, the knowledge acquisition process and verification process are difficult and time-consuming.

Several studies on bankruptcy prediction use quantitative approach with the nonfinancial variables such as the number of employees and the years from establishment. But, few studies on bankruptcy prediction have reported the systematic approaches that can discover experts' subjective knowledge framework. This may be a result of the difficulties in collection of qualitative information and resolving the inconsistencies among knowledge of experts who differ in logics and the way of thinking.

This paper proposes a GA-based data mining method capable of extracting decision rules from experts' qualitative bankruptcy decisions. This study is the first application of GAs for the purpose of discovering experts' problem-solving knowledge associated with bankruptcy prediction. Two data mining techniques, neural networks and inductive learning methods, are used to compare the performance with the GA method. The results of the experiment show that the GA method has significantly better performance than neural networks and inductive learning methods in terms of predictive accuracy and coverage. They also indicate that a reasonable level of agreement is achieved between the GA and experts' knowledge. This means that the proposed GA method can be a suitable tool for eliciting and representing the problem-solving knowledge of experts.

The remainder of this paper is organized as follows. Section 2 presents a review and comparison of the three kinds of data mining techniques used in this paper. Section 3 explains the details of the GA for extracting experts' subjective knowledge. In Section 4, we discuss the design process and the results of an empirical experiment. Concluding remarks and further research issues are described in Section 5.

2. Data mining techniques for extracting experts' decision rules

In an actual default risk estimation process, experts use their subjective knowledge to deal with qualitative information in the following analysis steps: preliminary investigation, financial statement analysis, industry level analysis, company level analysis, financing strategy evaluation, and default risk estimation (Caouette, Altman, & Narayanan, 1998). Qualitative information that can be used for the default risk estimation process consists of numerous risk components. These components can be categorized into six risk factors established and used by one of the largest Korean commercial bank. They include industry risk (IR), management risk (MR), financial flexibility (FF), credibility (CR), competitiveness (CO), and operating risk (OP). IR is measured by the stability and the growth of the industry, the degree of competition within the industry, and the overall conditions of the industry. MR is concerned with the efficiency and stability of management and organization structure. It is measured by the ability of management, the stability of top management, the stability of organization structure, management performance, and the feasibilities of business plans. FF means the firm's financing ability from direct and indirect financial market and other sources such as affiliates and the third parties. CR is concerned with the reputation of a company associated with credit history, reliability of information provided by the company, and the relationship with financial institutions. CO means the degree of competitive advantage determined by market position and the capacity of core technology. OP is the volatility and stability of procurement, the efficiency of production, the stability of sales, and the efficiency of collection policy of accounts receivable. The details of qualitative risk factors are listed in Table 1.

Experts evaluate the qualitative risk factors through the risk estimation process and assign appropriate levels such as positive, average and negative to these factors using their subjective knowledge. Finally, they classify various cases into categories such as accept, hesitate and reject.

It is expected that the qualitative data mining approach using experts' problem-solving knowledge can provide the more understandable models and results for bankruptcy prediction. Data mining techniques used in quantitative rule extraction approach can also be applied to discover decision rules from the qualitative predictions of experts. They include inductive learning methods, neural networks, and GAs. Inductive learning methods are typical rule extraction techniques which operate a successive partitioning of cases until all subsets belong to a single class (Quinlan, 1986, 1993). Several studies have used inductive learning methods for predicting corporate failure (Messier & Hansen, 1988; Shaw & Gentry, 1990). The criticism is that one-step-ahead node splitting without backtracking may generate a suboptimal tree in solving certain types of problems such as multiplexer problems (Weiss & Kulikowski, 1991).

Table 1
The details of qualitative risk factors

Risk factor	Variables	Risk components
Industry risk	IR	Government policies and International agreements Cyclicality Degree of competition The price and stability of market supply The size and growth of market demand The sensitivity to changes in macroeconomic factors Domestic and international competitive power Product Life Cycle IR
Management risk	MR	Ability and competence of management Stability of management The relationship between management/owner Human resources management Growth process/business performance Short and long term business planning, achievement and feasibility
Financial Flexibility	FF	Direct financing Indirect financing Other financing (Affiliates, Owner, Third parties)
Credibility	CR	Credit history The reliability of information The relationship with financial institutes
Competitiveness	CO	Market position The level of core capacities Differentiated strategy
Operating Risk	OP	The stability and diversity of procurement The stability of transaction The efficiency of production The prospects for demand for product and service Sales diversification Sales price and settlement condition Collection of A/R Effectiveness of sale network

Another criticism is that even the best tree may not be able to represent the best set of rules (Quinlan, 1988; Tsumoto, 1998).

Neural networks are dynamic models developed to mimic the biological neural systems in performing learning control and pattern recognition. Since the mid-1980s, neural networks have been the most widely used techniques in developing quantitative bankruptcy prediction models (Altman et al., 1994; Dutta, Shekhar, & Wong, 1994; Wilson & Sharda, 1994). Recently, neural networks have also been employed to extract rules for solving crisp and fuzzy classification problems (Giles, Lee, Lawrence, & Tsoi, 1997; Hayashi & Imura, 1990; Lin & Lee, 1991). The results show that neural networks are suitable in building a logic system with a relatively small number of numerical variables. However, neural networks lack analytical guidance in determining the network configuration. They may also be trapped into local optima in the learning process. These problems put some limitations on the quality of rules generated from the neural networks.

A more recent technique applied to classification problems is GAs which are heuristic search techniques based on the theory of natural selection and evolution (Holland, 1975). Most data mining-related GAs are used in the task of rule extraction in propositional and first-order logic (Anglano, Giordana, Lo Bello, & Saitta, 1997; Augier, Venturini, & Kodratoff, 1995; Giordana, Saitta, & Zini, 1994; Noda, Freitas, & Lopes, 1999). GA-based methods are also used for choosing appropriate sets of fuzzy if–then rules for classification problems (Ishibuchi, Kozaki, Yamamoto, & Tanaka, 1994; Peña & Sipper, 1999). Hybrid classification learning systems involve a combination of GA and neural networks (Yao & Liu, 1997), GA and linear discriminating models (Fogel, Wasson, Boughton, Porto, & Angeline, 1998). Shin and Lee (2002) applied a GA in bankruptcy prediction to extract rules from financial ratios.

The limitation of GAs is that a given problem cannot easily be encoded by GAs due to the fixed-length genomes. Another limitation is that they offer no guarantee of optimality. Nonetheless, numerous studies report that GAs are suitable for obtaining relevant knowledge from a huge number of databases due to the following distinctive features from conventional search algorithms such as inductive learning and neural networks (Goldberg, 1989; DeJong, 1990; Yuan & Zhuang, 1996). First, it may be very hard to apply conventional search algorithms to the classification problem when the number of possible different combinations of parameters is high. GAs are suitable in dealing with such problems because they can effectively explore large solution spaces without performing exhaustive searches.

Second, GAs consider not a single point but various points in the search space simultaneously reducing the chance of converging to local optima into which the conventional search algorithms may be trapped. This increases the chance of discovering the better or optimal rules through the learning process.

Third, GAs can generate relevant knowledge for the objectives of classification systems. Data mining can often be applied to solve multi-objective optimization problems. The conventional search algorithms have difficulties in solving such a problem while GAs can find out the optimal or near optimal knowledge by defining the composite fitness function associated with multi objectives. The discovery of classification rules in database is essentially to generate the rules that satisfy the conditions such as accuracy, generality, and compactness (Yuan & Zhuang, 1996). GAs can be a suitable tool for obtaining the rules which have such multi-requirements.

3. The discovery of experts' decision rules using a GA

This section describes the genetic evolution process for extracting decision rules from experts' qualitative decisions. The genetic coding, the fitness function and genetic operators for effective GA learning are explained.

3.1. Initialization of the population

Traditional GAs use the initial population consisting of chromosomes randomly distributed by the system while the initial population of the proposed GA method consists of the rules directly converted from experts' decision cases. It can provide a better starting point for reproduction. We use a binary string to represent a rule or a case. A rule can be coded as one chromosome which consists of several segments. Each segment is corresponding to either an attribute in the condition part of the rule or a class in the conclusion part of the rule. Each segment consists of a string of genes that take a binary value of 0 or 1. Each gene corresponds to one discrete linguistic term of the attribute or class.

Assume that we have a decision case made by an expert u as shown in Fig. 1. Each black circle in the risk factor table means the level of each risk factor assigned by an expert, while the black circle in the class table means actual class for the same case.

We use the following binary string for representing the case u by converting each black circle into binary value 1 and each white circle into binary value 0:

$u = [(100)(010)(100)(010)(001)(001); (10)]$.

where parentheses are to separate segments, and semicolon is to separate the IF part and the THEN part of the rule. This binary string can be interpreted as the following rule: "IF IR is negative and MR is average and FF is negative and CR is average and CO is positive and OP is positive THEN Nonbankrupt".

This genetic coding schema can represent cases with 'OR' relations. As the initial population used in this study evolves from old generation to new generation, new children might be created. The new children, such as 110 and 111, represent the 'OR' relation for an attribute. A binary string 110 represents the 'OR' relation between two linguistic terms. It can be interpreted as the following condition for an attribute: negative or average. The all-one string 111 represents the 'OR' relation among all linguistic terms in an attribute.

Risk Factor	Negative	Average	Positive	Genetic Coding
Industry Risk	●	□	□	(100)
Management Risk	□	●	□	(010)
Financial Flexibility	●	□	□	(100)
Credibility	□	●	□	(010)
Competitiveness	□	□	●	(001)
Operating Risk	□	□	●	(001)

Class	Nonbankrupt	Bankrupt	;
Actual Class	●	□	(10)

Fig. 1. An example of rule conversion from a case.

This string includes all possible cases for an attribute, i.e. unconditional such that the corresponding attribute is not involved in the condition part of the rule. But the all-zero string such as 000 is not allowed in our coding schema because each attribute must take at least one term.

3.2. Fitness evaluation of rules

The role of the fitness function is to encode the performance of the rule numerically. In this study, the objective of the GA method is to find the accurate and general rules among all the rules in the population. Thus, the GA method uses the composite fitness function consisting of accuracy and coverage. To measure the accuracy and coverage of the rule, we use the following definitions adapted from the pattern matching definitions of Yuan and Zhuang (1996).

Definition 1. The condition match of a rule r and a case u is defined by

$$mA(r,u) = \begin{cases} 1\text{—matched}, & \text{if } r(A_{kj}) \geq u(A_{kj}) \text{ for all } k \text{ and } j \\ 0\text{—mismatched}, & \text{otherwise}, \end{cases} \quad (1)$$

where A_{kj} means the jth linguistic term of the kth attribute. In this study, j represents one of three levels of Positive, Average, and Negative, while k corresponds to one of six risk factors in Table 1. This definition means that the rule r can be applicable to the case u and $mA(r,u) = 1$ if all linguistic values in the IF part of the rule r are equal to or greater than those of the case u. Otherwise, the conditions between the rule and the case are mismatched and $mA(r,u) = 0$.

Definition 2. The conclusion match of a rule r and a case u is defined by

$$mC(r,u) = \begin{cases} 1\text{—matched}, & \text{if } r(C_i) = u(C_i) \text{ for all } i, \\ 0\text{—mismatched}, & \text{otherwise}. \end{cases} \quad (2)$$

where i means the ith class. The conclusions between the rule and the case are matched and $mC(r,u) = 1$ if the rule and the case have the same class. Otherwise, $mC(r,u) = 0$.

Definition 3. The rule match between a rule r and a case u is defined by

$$mR(r,u) = \begin{cases} 1\text{—matched}, & \text{if } mA(r,u) = mC(r,u) = 1, \\ 0\text{—mismatched}, & \text{otherwise}. \end{cases} \quad (3)$$

This definition means that the case is accurately classified by the rule and $mR(r,u) = 1$ if the IF part of the rule r is applicable to the case u and has the same conclusion as the case. Otherwise, $mR(r,u) = 0$.

For example, assume that we obtain a rule $r1$ and a case $u1$ as follows:

$$r1 = [(011)(100)(111)(001)(011)(111); (10)]$$

and

$$u1 = [(001)(100)(001)(001)(001)(001); (10)].$$

For the rule $r1$ and the case $u1$, the condition match is $mA(r1, u1) = 1$ because the rule $r1$ with 'OR' operation covers the case $u1$, that is, all linguistic values of the rule $r1$ are equal to or greater than those of the case $u1$. The conclusion match is $mC(r1, u1) = 1$ because the rule $r1$ and the case $u1$ have the same conclusion, and thus the rule match is $mR(r1, u1) = 1$ because condition and conclusion matches are the value of 1.

Definition 4. The coverage means how well the condition part of the rule is universally applicable to all cases. Therefore, the coverage is a proportion of the number of cases that the rule r can be applied to all cases used in learning and thus it can be defined as

$$\mathrm{COV}(r) = \frac{\sum(mA(r,u))}{n} \tag{4}$$

where n is the number of all cases. The larger the coverage is, the more general the rule is.

Definition 5. The predictive accuracy of a rule r meaning the quality of the rule is defined by

$$\mathrm{PA}(r) = \frac{\sum mR(r,u)}{\sum mA(r,u)} \tag{5}$$

$\mathrm{PA}(r)$ is the proportion of the number of accurately classified cases to the cases to which the rule r can be applicable.

Definition 6. The fitness function of the rule r is defined by

$$\text{Objective}: \ \text{Maximize}(\mathrm{PA}(r) + \mathrm{Cov}(r))$$

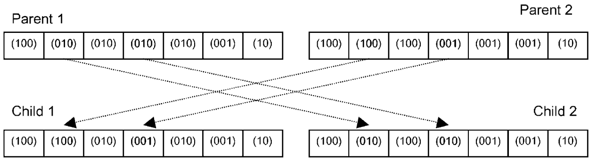
The objective of the fitness function is defined as composite measure of accuracy and coverage. This composite measurement provides an effective selection environment which balances the accuracy and generality of the rules.

3.3. Genetic operators

Selection is a process for choosing the rules with high fitness value as parents for reproduction. The mating selection of the proposed GA method is restricted within the same species, that is, the parents for reproduction are selected from the rules with the same class because the genetic operation between two rules with different classes tends to generate low-performance offsprings (Holland, 1975).

Crossover is a GA process for exchanging the information between two parent chromosomes to generate two child chromosomes while mutation is to reach the optimal point through an occasional alternation. These processes are performed at the bit position in the traditional GAs while those of the proposed GA method are performed segment by segment rather than bit by bit because each segment of a rule has a special meaning. This reduces the possibility of generating useless rules. Figs. 2 and 3 show examples of crossover and mutation. The two parent chromosomes can generate two child chromosomes through crossover on their second and fourth segments as shown in Fig. 2. The mutation on the second segment of the chromosome may generate one of the six possible new chromosomes as shown in Fig. 3.

Replacement will be performed between the rules with the same conclusion part because the replacement between two rules with different classes tends to generate low-quality rules. After replacement, new population is evaluated based on the fitness function. This process continues iteratively until reaching a predefined stopping condition. Fig. 4 shows



Semantic meanings of the GA strings are as follow:

Parent 1: IF IR is negative and MR is average and FF is average and CR is average and CO is average and OP is positive THEN Nonbankrupt

Parent 2: IF IR is negative and MR is negative and FF is negative and CR is positive and CO is positive and OP is positive THEN Nonbankrupt

Child 1: IF IR is negative and MR is negative and FF is average and CR is positive and CO is average and OP is positive THEN Nonbankrupt

Child 2: IF IR is negative and MR is average and FF is negative and CR is average and CO is positive and OP is positive THEN Nonbankrupt

Fig. 2. An example of crossover.

Parent

(110)	(100)	(011)	(010)	(011)	(001)	(10)

Children

(110)	(001)	(011)	(010)	(011)	(001)	(10)
(110)	(010)	(011)	(010)	(011)	(001)	(10)
(110)	(011)	(011)	(010)	(011)	(001)	(10)
(110)	(101)	(011)	(010)	(011)	(001)	(10)
(110)	(110)	(011)	(010)	(011)	(001)	(10)
(110)	(111)	(011)	(010)	(011)	(001)	(10)

Fig. 3. An example of mutation.

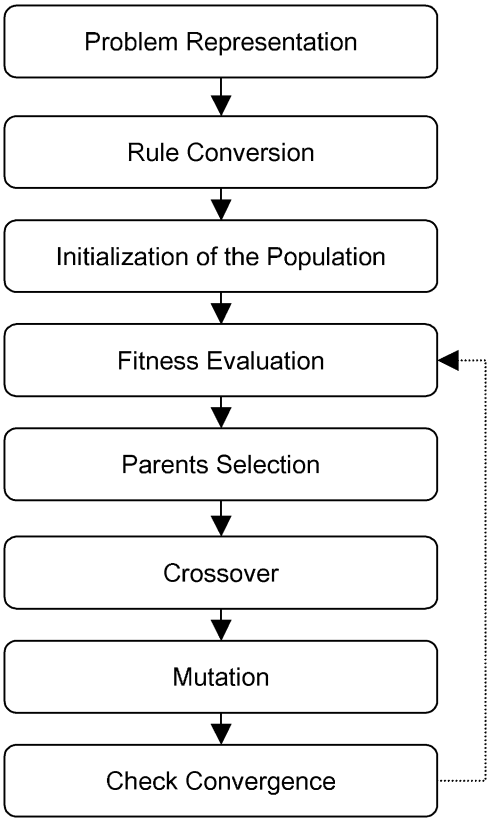


Fig. 4. The basic steps of the genetic evolution process.

the basic steps of the genetic evolution process for the proposed GA method.

4. Experiments and results

4.1. Data and experimental design

The samples collected during the period of 2001–2002 from one of the largest commercial banks in Korea consist of 772 manufacturing and service companies. Half are classified into the bankrupt group which is in receivership or liquidation while the other half are in the nonbankrupt group.

The experts who rated these companies are experienced loan officers working for the commercial bank. These officers have 9 years of rating experience on average. They evaluate six qualitative risk factors and assign an appropriate level to each of six qualitative risk factors as shown in Fig. 1. The data of qualitative risk factors are actual experts' decisions collected from loan management database of the bank.

The data set for the GA is split into two subsets which are a training set and a validation set of 70 and 30% of all data, respectively. The crossover rate ranges from 0.5 to 0.7 and the mutation rate ranges from 0.06 to 0.12. The set of individuals is evolved to 100 generations.

For inductive learning methods, we assign the value to three levels of each risk factor as follows: negative (1), average (2), and positive (3). The CHAID algorithm (Kass, 1980), which is capable of facilitating multiway split based on chi-square test, is applied to generate the induction rules from the same training data set for GA.

The two stages of the learning procedure are applied to generate rules from neural networks. In the first stage, 3-layered backpropagation neural networks learn classification functions from the training cases. The data set for learning backpropagation neural networks is split into three subsets: a training set, test set, and validation set of 40, 30 and 30% of all data. In the second stage, the CHAID algorithm is used to generate the rules from trained neural networks. The software package used for training inductive learning and neural networks is SAS Enterprise Miner 4.0.

4.2. Results

The genetic evolution process finally extracts eleven bankruptcy prediction rules, seven of which are

Table 2
The rules generated from genetic evolutionary process

No.	Rule							Training data (540 cases)			Validation data (232 cases)		
	IR	MR	FF	CR	CO	OP	Class	Coverage	Accuracy	Fitness	Coverage	Accuracy	Fitness
1	(111)	(111)	(011)	(011)	(011)	(111)	(10)	0.238	1.000	1.238	0.235	1.000	1.235
2	(111)	(111)	(100)	(100)	(100)	(100)	(01)	0.150	1.000	1.150	0.157	1.000	1.157
3	(111)	(111)	(001)	(111)	(001)	(111)	(10)	0.127	1.000	1.127	0.151	1.000	1.151
4	(011)	(111)	(111)	(011)	(001)	(111)	(10)	0.189	1.000	1.189	0.175	0.966	1.140
5	(011)	(011)	(011)	(111)	(011)	(011)	(10)	0.181	1.000	1.181	0.163	0.963	1.126
6	(111)	(011)	(111)	(011)	(011)	(111)	(10)	0.155	1.000	1.155	0.187	0.968	1.154
7	(111)	(110)	(100)	(100)	(100)	(111)	(01)	0.218	0.988	1.206	0.211	0.943	1.154
8	(001)	(011)	(111)	(111)	(001)	(111)	(10)	0.155	0.967	1.122	0.108	0.944	1.053
9	(011)	(111)	(111)	(111)	(001)	(011)	(10)	0.158	0.967	1.125	0.187	0.935	1.122
10	(111)	(100)	(100)	(100)	(110)	(110)	(01)	0.298	0.870	1.167	0.247	0.854	1.101
11	(100)	(100)	(100)	(111)	(100)	(111)	(01)	0.220	0.800	1.020	0.217	0.694	0.911
	Average							0.190	0.963	1.153	0.185	0.933	1.119

Average $= \dfrac{\sum M_{R_i}}{\text{the number of rules}}$, where M is accuracy, coverage, or fitness of rule R_i.

Table 3
The descriptions of the rules generated from the GA

Rule	Description
Rule1	IF FF is average or positive and CR is average or positive and CO is average or positive THEN Nonbankrupt
Rule2	IF FF is negative and CR is negative and CO is negative and OP is negative THEN Bankrupt
Rule3	IF FF is positive and CO is positive THEN Nonbankrupt
Rule4	IF IR is average or positive and CR is average or positive and CO is positive THEN Nonbankrupt
Rule5	IF IR is average or positive and MR is average or positive and FF is average or positive and CO is average or positive and OP is average or positive THEN Nonbankrupt
Rule6	IF MR is average or positive and CR is average or positive and CO is average or positive THEN Nonbankrupt
Rule7	IF MR is negative or average and FF is negative and CR is negative CO is negative THEN Bankrupt
Rule8	IF IR is positive and MR is average or positive and CO is positive THEN Nonbankrupt
Rule9	IF IR is average or positive and CO is positive and OP is average or positive THEN Nonbankrupt
Rule10	IF MR is negative and FF is negative and CR is negative and CO is negative or average and OP is negative and average THEN Bankrupt
Rule11	IF IR is negative and MR is negative and FF is negative and CO is negative THEN Bankrupt

nonbankrupt rules and the others are bankrupt rules. The rules and the corresponding descriptions are illustrated in Tables 2 and 3. Their performances are stable for both the training and validation data set in terms of coverage and accuracy as shown in Table 2.

The rules generated from inductive learning methods and neural networks are illustrated in Tables 4 and 5, respectively. The rules generated from inductive learning methods consist of 16 rules, 10 of which are nonbankrupt rules while the others are bankrupt rules as shown in Table 4. The numbers of rules generated from neural networks are 12 as listed in Table 5. Half are nonbankrupt rules while the other half are nonbankrupt rules.

When the rules are applied to cases, two or more rules can be applied to classify the same case. We use the following application steps to deal with such a situation. First, the case is predicted as the class of the corresponding rule if only a single rule is applied to the case. Second, the case is predicted as the class of the rule with the highest accuracy if two or more rules are applied to the case at the same time. Finally, the case is predicted as the class of the rule with the largest coverage if two or more rules with the same accuracy are applied to classify the case at the same time.

Table 6 compares the performances of three rule extraction methods over the validation data set. Table 6 shows that the GA method can generate the rules with higher accuracy and larger coverage than other techniques while the rule structure is more compact. Overall classification accuracy means the accuracy level when the rules are applied to cases according to application steps. The GA method also shows better predictive accuracy than neural networks and inductive learning methods.

We use the Wilcoxon matched-pairs signed-ranks test to examine whether the predictive accuracy of the rules

Table 4
The descriptions of the rules generated from inductive learning methods

Rule	Description
Rule1	IF FF is positive and CO is positive THEN Nonbankrupt
Rule2	IF FF is positive and CO is average and CR is average or positive THEN Nonbankrupt
Rule3	IF FF is positive and CO is average and CR is negative THEN Bankrupt
Rule4	IF FF is positive and CO is negative and MR is average or positive THEN Nonbankrupt
Rule5	IF FF is positive and CO is negative and MR is negative THEN Bankrupt
Rule6	IF FF is average and MR is positive and CO is average or positive THEN Nonbankrupt
Rule7	IF FF is average and MR is positive and CO is negative THEN Bankrupt
Rule8	IF FF is average and MR is average and OP is average or positive THEN Nonbankrupt
Rule9	IF FF is average and MR is average and OP is negative THEN Bankrupt
Rule10	IF FF is average and MR is negative THEN Bankrupt
Rule11	IF FF is negative and OP is positive THEN Nonbankrupt
Rule12	IF FF is negative and OP is average and IR is average or positive THEN Nonbankrupt
Rule13	IF FF is negative and OP is average and IR is negative THEN Bankrupt
Rule14	IF FF is negative and OP is negative and CR is average or positive THEN Nonbankrupt
Rule15	IF FF is negative and OP is negative and CR is negative and MR is average or positive THEN Nonbankrupt
Rule1	6IF FF is negative and OP is negative and CR is negative and MR is negative THEN Nonbankrupt

Table 5
The descriptions of the rules generated from neural networks

Rule	Description
Rule1	IF FF is positive and CO is average or positive THEN Nonbankrupt
Rule2	IF FF is positive and CO is negative and MR is average or positive THEN Nonbankrupt
Rule3	IF FF is positive and CO is negative and MR is negative THEN Bankrupt
Rule4	IF FF is average and MR is positive THEN Nonbankrupt
Rule5	IF FF is average and MR is average and OP is average or positive THEN Nonbankrupt
Rule6	IF FF is average and MR is average and OP is negative THEN Bankrupt
Rule7	IF FF is average and MR is negative THEN Bankrupt
Rule8	IF FF is negative and OP is positive THEN Nonbankrupt
Rule9	IF FF is negative and OP is average and IR is average or positive THEN Nonbankrupt
Rule10	IF FF is negative and OP is average and IR is negative THEN Bankrupt
Rule11	IF FF is negative and OP is negative THEN Nonbankrupt
Rule12	IF FF is negative and OP is negative and CR is negative and MR is negative THEN Nonbankrupt

Table 6
The performances of data mining techniques (232 cases)

Data mining techniques	The number of rules extracted	Average of coverage (%)	Average of accuracy (%)	Overall accuracy (%)
GA	11	18.5	93.3	94.0
Inductive learning	16	15.3	87.7	89.7
Neural networks	12	15.6	88.4	90.3

Table 7
The results of Wilcoxon's matched-pairs signed-ranks test

	GA	Inductive learning	Neural networks
GA	–		
Inductive learning	3.563*	–	
Neural networks	3.312*	1.346	–

*Significant at 1%.

generated from three data mining techniques is significantly different. Table 7 shows the results of the test. The results show that the rules of the GA method are significantly better than those of other data mining techniques while predictive accuracies between the rules generated from inductive learning method and neural networks are not significantly different.

It is important to measure the agreement between the classification made by expert and the classifier because such a measure indicates the degree to which the subjectivity of experts is incorporated in the model. We adopt Cohen's *kappa* (1960) as the measure of agreement. Cohen's *kappa* measures the agreement between two rater (e.g. an expert and a data mining technique) classifying the same set of cases. Cohen's *kappa* defines the measure of agreement as the ratio of the percentage of agreement minus the chance agreement to the largest possible nonchance agreement. This measure, thus, takes into account the classifications that could match merely by chance. The chance agreement actually depends upon the percentage of matches in each class, and it reduces as the number of classes increases. Using the above definition, a *kappa* value of 1 indicates a perfect agreement and a *kappa* value of 0 indicates that agreement is no better than chance.

Table 8 shows the combinations between classifications of bankrupt (BK) and nonbankrupt (NBK) provided by experts and data mining techniques. The value in each cell indicates the number of cases in each combination, and the value in the parenthesis indicates the percentage of the number of cases in each combination to all cases, respectively. For example, the number of cases classified as nonbankrupt by experts and the GA method is 111 and the percentage of this combination is 47.8% (111/232). From Table 8, we can compute the agreement between experts and the GA method by chance as 50.0% ($50.9 \times 50.4 + 49.1 \times 49.6$). Therefore, the value of *kappa* is $(94.0 - 50.0)/(100 - 50.0) = 0.8799$. This measure of agreement is reasonably high considering that it represents the agreement over and above the chance agreement. The GA method generates a more reasonable level of agreement between experts' problem-solving knowledge than inductive learning methods and neural networks. This indicates that the GA method is effective in the discovery of experts' subjective knowledge.

5. Concluding remarks

Data mining has been widely applied to discovering quantitative bankruptcy knowledge from financial databases. However, few studies have reported the potential of data mining that can investigate the qualitative problem-solving knowledge from experts' decisions. This paper demonstrated the GA-based data mining approach to discover decision rules from experts' decision process. This study is the first work on GAs for the purpose of discovering experts' qualitative knowledge on bankruptcy. The fitness function of the GA is the composite measure to discover decision rules that satisfy two different conditions: accuracy and coverage. This composite fitness function can provide an efficient environment for reproduction. Effective learning strategies are implemented in genetic operators including selection, crossover, mutation, and replacement to generate useful rules.

Two data mining techniques, neural networks and inductive learning methods, are applied to compare their performance with that of the GA method. The results of the experiments show that the performance of the GA method is significantly better than neural networks and inductive

Table 8
The results of *kappa* test between predictions of experts and data mining techniques

Experts	GA			Inductive learning			Neural networks		
	NBK	BK	Total	NBK	BK	Total	NBK	BK	Total
NBK	111 (47.8)	7 (3.0)	118 (50.9)	103 (44.4)	18 (7.8)	121 (52.2)	105 (45.3)	15 (6.5)	120 (51.7)
BK	6 (2.6)	108 (46.6)	114 (49.1)	16 (6.9)	95 (40.9)	111 (47.8)	13 (5.6)	99 (42.7)	112 (48.3)
Total	117 (50.4)	115 (49.6)	232 (100.0)	119 (51.3)	113 (48.7)	232 (100.0)	118 (50.9)	114 (49.1)	232 (100.0)
Chance agreement	50.0			50.1			50.0		
Kappa	0.8799			0.7929			0.8054		

NBK: nonbankrupt, BK: bankruptcy.

learning methods in terms of predictive accuracy and coverage. They also show that considerable agreement is achieved between the GA method and experts' problem-solving knowledge.

This study provides useful implications for bankruptcy prediction. First, model formulation for bankruptcy prediction is a difficult task requiring experts' subjective knowledge because of the complexity of the problem. This study provides effective supports in incorporating experts' subjective knowledge into the models, and thus it facilitates efficient development of bankruptcy prediction models.

Second, the financial databases used in quantitative approaches are concerned with financial stability and trends over the past years. Bankruptcy predictions using past information has a critical limitation on predicting default risk that might occur in near future. In contrast, the information used in qualitative approach reflects experts' prospects on companies' status in the near future. Two approaches are complementary, and thus the combination of two approaches can generate an improved performance. Several studies propose a variety of techniques to combine quantitative and qualitative models (Kim, Kim, & Lee, 2002). This qualitative study can be helpful for developing hybrid models.

However, there remain several limitations and issues for further research. First, the current rules have a redundant and overlapping structure. The reason for this structure is that they are generated from the data set with a small number of input features. This structure can be considerably refined by introducing additional input features. This refinement can be helpful for generating more efficient learning as well as more effective decision support.

Second, the GA method uses several learning strategies to improve the efficiency. Unfortunately, we did not keep track of their impacts on learning efficiency. Although it seems that these strategies significantly reduce learning time required to obtain the rules by preventing the generation of useless rules, more advanced research is needed to further improve the algorithm. For example, further development can be obtained by applying the niching method which makes the population eventually converge around a single point in the solution space (Mahfoud & Mani, 1995).

References

Altman, E. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance*, *23*(3), 589–609.

Altman, E. I., Marco, G., & Varet, F. (1994). Corporate distress diagnosis: comparisons using linear discriminant analysis and neural networks. *Journal of Banking and Finance*, *18*, 505–529.

Anglano, C., Giordana, A., Lo Bello, G., & Saitta, L. (1997). *A network genetic algorithm for concept learning. Proceedings of the ICGA'97*, San Francisco: Morgan Kaufmann, pp. 434–441.

Augier, S., Venturini, G., & Kodratoff, Y. (1995). *Learning first order logic rules with a genetic algorithm. Proceedings of the First International Conference on Knowledge Discovery and Data Mining*, Menlo Park, CA: AAAI Press, pp. 21–26.

Brachman, R. J., Khabaza, T., Kloesgen, W., Piatesky-Shpiro, G., & Simoudis, E. (1996). Mining business databases. *Communication of the ACM*, *39*(11), 42–48.

Caouette, J. B., Altman, E. I., & Narayanan, P. (1998). *Managing credit risk: The next great financial challenge*. New York: Wiley & Sons Inc.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*, 37–46.

DeJong, K. A. (1990). Genetic algorithm-based learning. *Machine Learning*, *3*, 611–638.

Dutta, S., Shekhar, S., & Wong, W. Y. (1994). Decision support in non-conservative domains: generalization with neural networks. *Decision Support Systems*, *11*(5), 527–544.

Fletcher, D., & Goss, E. (1993). Forecasting with neural networks: an application using bankruptcy data. *Information and Management*, *24*(3), 159–167.

Fogel, D. B., Wasson, E. C., Boughton, E., Porto, M., V, W., & Angeline, P. J. (1998). Linear and neural models for classifying breast masses. *IEEE Transactions on Medical Imaging*, *17*(3), 485–488.

Giles, C., Lee, C., Lawrence, S., Tsoi, A. C (1997). Rule inference for financial prediction using recurrent neural networks, *Proceedings of IEEE/IAFE Conference on Computational Intelligence for Financial Engineering* (pp. 253–259). Piscataway, NJ

Giordana, A., Saitta, L., & Zini, F. (1994). *Learning disjunctive concepts by means of genetic algorithms. Proceedings of the 11th International Conference on Machine Learning*, pp. 96–104.

Goldberg, D. E. (1989). *Genetic algorithm in search, optimization, and machine Learning*. Reading, MA: Addison-Wesley.

Hayashi, Y., & Imura, A. (1990). *Fuzzy neural expert system with automated extraction of fuzzy if–then rules from a trained neural network. The First International Symposium on Uncertainty Modeling and Analysis*, pp. 489–494.

Holland, J. H. (1975). *Adaptation in natural and artificial systems*. Ann Arbor: The University of Michigan Press.

Ishibuchi, H., Kozaki, K., Yamamoto, N., & Tanaka, H. (1994). Construction of fuzzy classification systems with rectangular fuzzy rules using genetic algorithms. *Fuzzy Sets and Systems*, *65*, 237–253.

Kass, G. (1980). An exploratory technique for investigating large quantities of categorical data. *Applied Statistics*, *29*(2), 119–127.

Kim, E. J., Kim, W. J., & Lee, Y. B. (2002). Combination of multiple classifiers for the customer's purchase behavior prediction. *Decision Support Systems*, *34*, 167–175.

Lin, C. T., & Lee, C. S. G. (1991). Neural network-based fuzzy logic control and decision system. *IEEE Transactions on Computer*, *12*, 1320–1336.

Mahfoud, S., & Mani, G. (1995). Genetic algorithm for predicting individual stock performance. *Proceedings of the Third International Conference on Artificial Intelligence Applications on Wall Street*, 174–181.

Messier, W. F., & Hansen, J. (1988). Inducing rules for expert system development: an example using default and bankruptcy data. *Management Science*, *34*(12), 1403–1415.

Noda, E., Freitas, A. A., & Lopes, H. S. (1999). *Discovering interesting prediction rules with a genetic algorithm. Proceedings of the Congress on Evolutionary Computation*, Piscataway, NJ: IEEE Press.

Odom, M., & Sharda, R. (1990). *A neural networks model for bankruptcy prediction. Proceedings of the IEEE International Conference on Neural Network*, pp. 163–168.

Ohlson, J. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, *18*(1), 109–131.

Peña, C. A., & Sipper, M. (1999). *Designing breast cancer diagnosis system via a hybrid fuzzy-genetic methodology* (*Vol. 1*). *Proceedings of IEEE International Fuzzy Systems Conference*, pp. 135–139.

Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, *1*, 81–106.

Quinlan, J. R. (1988). *An empirical comparison of genetic and decision tree classifiers*. *Proceedings of the Fifth International Conference on Machine Learning*, pp. 135–141.

Quinlan, J. R. (1993). *C4.5-Programs for machine learning*. Palo Alto: Morgan Kaufmann.

Shaw, M., & Gentry, J. (1990). Inductive learning for risk classification. *IEEE Expert*, 47–53.

Shin, K. S., & Lee, Y. J. (2002). A genetic algorithm applications in bankruptcy prediction modeling. *Expert Systems with Applications*, *23*(3), 321–328.

Tam, K., & Kiang, M. (1992). Managerial applications of neural networks: The case of bank failure prediction. *Management Science*, *38*(7), 926–947.

Tsumoto, S. (1998). Extraction of experts' decision rules from clinical databases using rough set model. *Intelligent Data Analysis*, *2*, 215–227.

Weiss, S. M., & Kulikowski, C. A. (1991). *Computer systems that learn: Classification and prediction methods from statistics, neural nets, machine learning, and expert systems*. San mateo, CA: Morgan Kaufmann.

Wilson, R. L., & Sharda, R. (1994). Bankruptcy prediction using neural networks. *Decision Support Systems*, *11*(5), 545–557.

Yao, X., & Liu, Y. (1997). A new evolutionary system for evolving artificial neural networks. *IEEE Transactions on Neural Networks*, *8*(3), 694–713.

Yuan, Y., & Zhuang, H. (1996). A genetic algorithm for generating fuzzy classification rules. *Fuzzy Sets and Systems*, *84*, 1–19.

Zmijewski, M. E. (1984). Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research*, *22*(1), 59–82.