Available online at www.sciencedirect.com

SCIENCE DIRECT®

European Journal of Operational Research 169 (2006) 677–697

EUROPEAN JOURNAL OF OPERATIONAL RESEARCH

www.elsevier.com/locate/ejor

Interfaces with Other Disciplines

Bankruptcy theory development and classification via genetic programming ☆

Terje Lensberg [a], Aasmund Eilifsen [a], Thomas E. McKee [a,b,*]

[a] *Norwegian School of Economics and Business Administration, Helleveien 30, Bergen-Sandviken N-5035, Norway*
[b] *East Tennessee State University, P.O. Box 70710 ETSU, Johnson City, TN 37614-0710, USA*

Received 12 November 2002; accepted 22 June 2004
Available online 25 August 2004

Abstract

Bankruptcy is a highly significant worldwide problem with high social costs. Traditional bankruptcy risk models have been criticized for falling short with respect to bankruptcy theory building due to either modeling assumptions or model complexity.

Genetic programming minimizes the amount of a priori structure that is associated with traditional functional forms and statistical selection procedures, but still produces easily understandable and implementable models. Genetic programming was used to analyze 28 potential bankruptcy variables found to be significant in multiple prior research studies, including 10 fraud risk factors. Data was taken from a sample of 422 bankrupt and non-bankrupt Norwegian companies for the period 1993–1998. Six variables were determined to be significant.

A genetic programming model was developed for the six variables from an expanded sample of 1136 bankrupt and non-bankrupt Norwegian companies. The model was 81% accurate on a validation sample, slightly better than prior genetic programming research on US public companies, and statistically significantly better than the 77% accuracy of a traditional logit model developed using the same variables and data. The most significant variable in the final model was the prior auditor opinion, thus validating the information value of the auditor's report. The model provides insight into the complex interaction of bankruptcy related factors, especially the effect of company size. The results suggest that accounting information, including the auditor's evaluation of it, is more important for larger than smaller firms. It also suggests that for small firms the most important information is liquidity and non-accounting information.

☆ Data Availability: The data may be obtained from Compact Disclosure, Inc., a commercial data vendor.

* Corresponding author. Present address: College of Charleston, 66 George Street, Charleston, SC 29424-0001, USA. Tel.: +1 843 953 7835; fax: +1 843 953 5697.

E-mail addresses: terje.lensberg@nhh.no (T. Lensberg), aasmund.eilifsen@nhh.no (A. Eilifsen), Thomas.Mckee@nhh.no, temckee@etsu.edu (T.E. McKee).

0377-2217/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.ejor.2004.06.013

The genetic programming model relationships developed in this study also support prior bankruptcy research, including the finding that company size decreases bankruptcy risk when profits are positive. It also confirms that very high profit levels are associated with increased bankruptcy risk even for large companies an association that may be reflecting the potential for management to be "Cooking the Books".
© 2004 Elsevier B.V. All rights reserved.

Keywords: Genetic programming; Genetic algorithms; Going concern; Bankruptcy; Fraud risk

1. Introduction

Corporate bankruptcy affects the economy of every country and is monitored by policy makers seeking to promote economic growth. For example, US business bankruptcy filings declined from 53,931 in 1997, to 44,196 in 1998, and to 37,564 in 1999 (American Bankruptcy Institute, 2001), a period during which growth in the US economy, as measured by gross national product, increased from 3.4% to 4.1% (Bureau of Economic Analysis, 2001).

At the level of the individual firm, the capital markets react to data about going concern prospects for firms. For example, both Beaver (1968) and Altman (1969) showed a negative stock market price reaction as a firm approached failure. Accordingly, the International Accounting Standards (IAS) require, "When preparing financial statements, management should make an assessment of an enterprise's ability to continue as a going concern." (International Accounting Standards Board, 2002, IAS No. 1, paragraph 23). If there are significant doubts about a company's bankruptcy/non-bankruptcy status then the IAS requires information about those uncertainties to be disclosed. Disclosure responsibilities also extend to a company's auditors. International Standard on Auditing (ISA) section 570 (International Auditing and Assurance Standards Board of International Federation of Accountants, 2003) requires the auditor to modify the auditor's report by adding an emphasis of matter paragraph if there is significant doubt about the entity's going concern status and adequate disclosure is made in the financial statements. If adequate disclosure is not made in the financial statements, the auditor should express a qualified or adverse opinion. Thus, auditors are required to signal stakeholders about going concern problems. Although going concern problems can result in outcomes other than bankruptcy, this outcome is probably the one of most concern to stakeholders.

Bankruptcy prediction has been a major topic in accounting and finance for at least a century. Early research focused primarily on univariate models such as individual ratios while later research turned to multivariate models. Recent research has turned to modeling techniques like recursive partitioning, fuzzy logic, and rough sets. Despite a long research history, there is no bankruptcy prediction model based primarily on bankruptcy theory which is generally accepted. Additionally, as noted by Dimitras et al. "A unifying theory of business failure has not been developed ..." (1996, p. 487). To improve this situation either greater bankruptcy prediction research convergence or more theory building is needed.

Most prior bankruptcy research utilized techniques producing descriptions for classifying objects into classes based on the objects' properties. One of the acknowledged difficulties with this type of inductive inference is its' open-endedness which stems from the fact that there is no natural limit to the level of detail which may be used to describe reality. For example, while bankruptcy models produced by techniques such as neural networks, recursive partitioning, and rough sets theory can produce reasonable classification accuracy, they provide extremely detailed models which are difficult to generalize or develop theory from. Further, multivariate techniques are inadequate according to Michalski because, "The widely used traditional mathematical and statistical data analysis techniques, such as regression analysis, numerical taxonomy, or factor analysis are not sufficiently powerful ... for the task of ... detecting interesting conceptual patterns or in revealing structure in a collection of observations.'' (Michalski, 1983, pp. 112–113). For example, Chen and Shimerda investigated the issue of bankruptcy conceptual patterns by analysis of 26 prior studies involving 65 accounting ratios and other financial items. They found that although seven factors could describe the ratios, ''... the question of which ratio should represent a factor has yet to be resolved.'' They concluded that since each ratio contains common as well as unique information, ... selection of the best representative ratio for a factor is not independent of the ratios selected for other factors.'' (Chen and Shimerda, 1981, p. 59). Thus, there is a demand for theory building, including more insight into the pattern and interaction between bankruptcy classification factors. Genetic programming (Koza, 1992) is a recently developed technique that permits a researcher to find a solution to a problem without having to prespecify the type of model. This means the solution can be any model describable by mathematics. The aim is to let the data speak for themselves as far as possible, by minimizing the amount of a priori structure imposed by functional forms and statistical selection procedures. Koza comments that ''Genetic programming is fundamentally different from other approaches to artificial intelligence, machine learning, adaptive systems, automated logic, expert systems, and neural networks in terms of (i) its representation (namely programs), (ii) the role of knowledge (none), (iii) the role of logic (none), and (iv) its mechanism (gleaned from nature) for getting to a solution within a space of possible solutions.'' (Banzhaf et al., 1998, p. viii). However, it should be noted that there is both logic and knowledge in the programming language which is used to implement the solution mechanism.

Recent research using data from US companies showed that genetic programming is extremely powerful and could be used to generate a simple, yet feature rich model providing new insights on bankruptcy prediction and, therefore, on bankruptcy theory development (McKee and Lensberg, 2002). The current research attempts to improve our understanding of bankruptcy through research convergence and improved insight into the pattern of bankruptcy classification factors. The current study extends this prior research by developing a bankruptcy prediction model using genetic programming for a very broad sample of Norwegian companies and then compares features of the Norwegian and US models.

This research does differ significantly from prior research by (1) the use of non-US data, (2) the analysis of a broad set of variables that were significant in multiple prior studies, (3) the inclusion of fraud prediction factors, (4) the use of primarily private companies, and (5) the use of a longer prediction interval. We believe this research provides important insights for both bankruptcy theory development and bankruptcy prediction.

2. Literature review

A wide range of international research has been conducted on bankruptcy prediction. In this section, we review key literature in the following areas:

- The going-concern concept
- US bankruptcy research directions
- International bankruptcy research
- Norwegian bankruptcy research

2.1. The going-concern concept

Economic distress is the inability of a firm to obtain sufficient resources for continued normal operations. Factors such as high costs, low customer demand, and poor financial management can cause economic distress. Bankruptcy is one possible outcome from company economic distress. Other outcomes are dissolution, liquidation, merger, restructuring or continued operation. As noted by Cormier et al. (1995, p. 203), ''... each of the mentioned alternatives constitutes an end to the firm as it was known before the advent of financial difficulties''. The end of the firm is defined as economic ''discontinuity'' (McKee and Lensberg, 2002). Development of a model to predict each of these very different outcomes is a very difficult problem since the identification and timing of these events are frequently not easily determinable.

We know that sometimes a firm can become distressed and continue to operate in that condition for many years. On the other hand, some firms enter bankruptcy immediately after a highly distressing event, such as a major fraud. A number of factors influence these outcomes. From a normative perspective, one would expect a firm to enter bankruptcy when significant stakeholders, such as stockholders, lenders, or management, believe their interests will be better served by bankruptcy. This means that some bankruptcies are voluntary on the part of stakeholders and some are involuntary.

Event Graphs [1] can be used to represent event scheduling relationships. A basic event graph contains only two elements: the event node and the scheduling edge. The event node represents an event that can take place. The scheduling edge represents what can happen after some units of time have passed. Fig. 1 illustrates the only two possible options for the scheduling edge. Either a simple time delay occurs and the event stays the same [shown by a line curving from Event A back to Event A, the same event] or a condition occurs which causes another event [shown by a line from Event A to Event B].

As shown in Fig. 2, we can model the previously discussed firm continuity/discontinuity concepts using basic event graph relationships. It should be noted that this graph does not necessarily show all possible event conditions. Also, the graph does not depict the scheduling edges for events other than bankruptcy which constitute firm discontinuity status. The means that possible future events for a liquidated firm, merged firm, and restructured firm are not shown. Both merger and restructuring could result in a future non-distressed firm state but, under continuity theory, this non-distressed firm would be defined as a different firm than the non-distressed firm shown in the event graph since a return to the original firm state would not be possible after economic discontinuity.

[1] A.H. Buss, "Modeling With Event Graphs," ACM, **Proceedings of the 1996 Winter Simulation Conference,** ed. J.M. Charnes, D.J. Morrice, D.T. Brunner, and J.J. Swain, pp. 153–160.

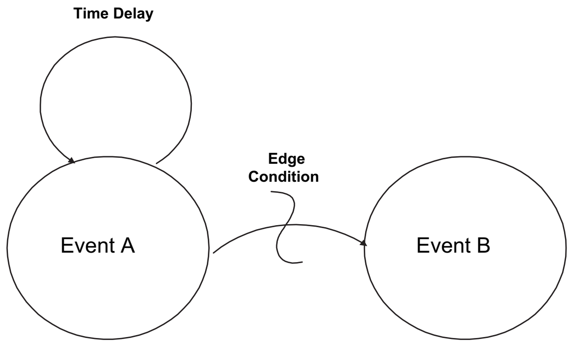
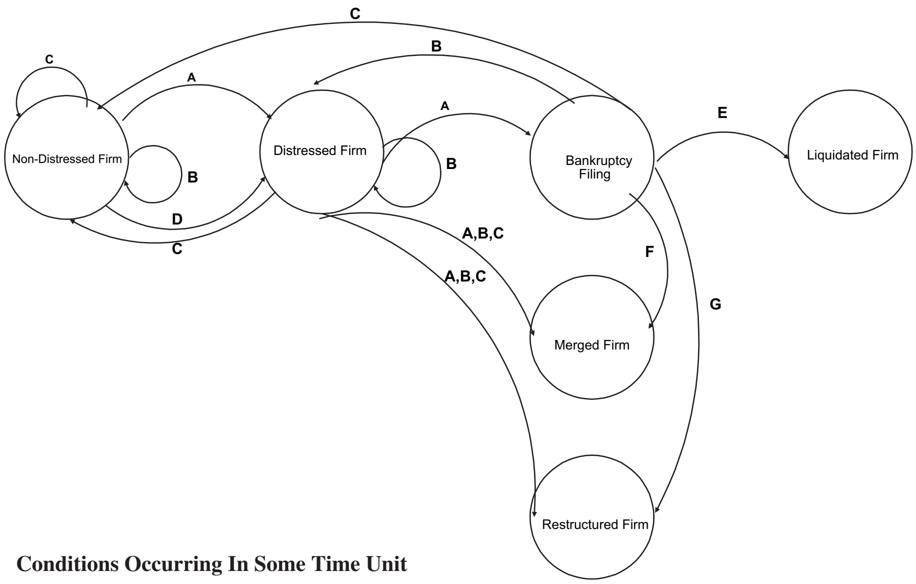


Fig. 1. Fundamental event graph.

2.2. Research directions in US bankruptcy research

One area which US bankruptcy research has consistently focused on is the development of predictive models that could be used one or more years prior to actual bankruptcy. Notable studies among the work done in this area were done by Altman (1968), Ohlson (1980), and Zmijewski (1984). They developed robust multivariate predictive models for which the overall accuracy depended on Type I and Type II error cutoff choices. Subsequent research dealt with a variety of issues surrounding these types of models, such as use on stressed versus non-stressed firms (Hopwood et al., 1994), or effect of business cycle (Richardson et al., 1998). This predictive model research subsequently included a number of other techniques such as recursive partitioning, survival analysis, mathematical programming, neural networks, and rough sets (McKee, 2000). Recent research on auditors' going concern decisions by Lenard, Alam, and Booth reports comparative results for a fuzzy clustering model and a hybrid model that consisted of a discriminant model combined with an expert system. The hybrid model, with a 98% accuracy on a holdout sample of 49 companies, statistically significantly outperformed the fuzzy clustering model (Lenard et al., 2000). A very recent US study by McKee and Lensberg (2002) utilized another hybrid model building approach involving genetic programming and rough sets to construct a bankruptcy prediction model from data on 291 US companies from the period 1991 to 1997. The genetic programming model



Conditions Occurring In Some Time Unit

A = financial condition deterioration
B = financial condition unchanged
C = financial condition improvement
D = impairment of future financial condition [e.g. threat of legal action]
E = economic benefit to creditors
F = economic benefit to creditors and shareholders
G = economic benefit to creditors, shareholders, and management

Fig. 2. Event graph of possible firm states.

developed from the rough sets developed variables was 80% accurate on a validation sample.

Since many bankrupt firms are audited, an additional line of research has concerned the information value of an auditor's report with respect to predicting future bankruptcy. A number of studies have shown that auditors only signal, via a modified audit report, in about half of the cases where companies ultimately go bankrupt (Altman and McGough, 1974; Menon and Schwartz, 1986; Chen and Church, 1992; Johnson and Khurana, 1995). Other studies have shown that only about one-third of the firms receiving a modified opinion actually go bankrupt during the year after receiving the going concern modified audit report (Altman, 1982; Little and McAlum, 1991). Recent research in this area has shown that a qualified opinion is positively associated with a firm subsequently successfully emerging from bankruptcy, presumably by providing an early warning about possible jeopardy which permits stakeholders to take protective actions (Bryan et al., 2000).

Possible auditor bias based on firm size has also been a research direction. McKeown et al. (1991) found that large firms received going concern opinions at a lower rate than small firms. Recently, Nelson et al. (2002) found that auditors are less likely to require earnings adjustments for large firms than for small ones. This latter study helps to confirm that there is a size effect for auditor actions.

Another area of bankruptcy research has concerned stressed firms undertaking actions to present themselves more favorably such as by choosing income-increasing accounting methods (Zmijewski and Hagarman, 1981; Sweeney, 1994) or by switching auditors (Schwartz and Menon, 1985; Knapp and Elikai, 1990; Schwartz and So, 1995).

2.3. Limited review of international bankruptcy research

Dimitras et al. (1996) did an extensive literature review to find journal articles concerning business failure which were written in English or French during the time period 1932–1994. They found 158 published journal articles. They analyzed the subset, 47 of these articles, which presented failure prediction models. This subset of 47 articles represented studies in 11 countries other than the USA. They found the primary model building methods to be discriminant and logit analysis. An analysis of variables employed in these studies found "... there are no financial characteristics common to all predictive studies." (Dimitras et al., 1996, p. 492).

Vermeulen et al. (1998) recently published a paper describing the possible use a multi-factor model failure prediction model employing exogenous risk factors rather than internal financial ratios. They did not empirically test the model so its' real world application remains undetermined.

Greco et al. (1998) published a study in which they applied a dominance relation approach to rough sets evaluation of bankruptcy risk for a sample of 39 Greek companies. The found the dominance relation approach resulted in only four reducts as opposed to 26 when the more classical rough sets indiscernibility approach was used. A reduct is a reduced set of variables which can predict the item of interest so this means the dominance relation approach potentially comes closer to finding a single model. They did not use a hold-out sample so the predictive accuracy of their model is difficult to evaluate.

Laitinen and Kankaanpää (1999) performed a study using Finish data in which they compared the six most popular failure prediction methods to see if there were systematic differences in prediction accuracy. The methods compared were:

- Linear discriminant analysis
- Logit analysis
- Recursive partitioning
- Survival analysis
- Neural networks and
- Human information processing

A statistically significant difference in predictive accuracy was found only between logistic analysis and survival analysis for one year prior to failure, leading to the conclusion "... no superior method has been found" (Laitinen and Kankaanpää, 1999, p. 67).

2.4. Norwegian bankruptcy research

Norway was not included as one of the 11 countries in the Dimitras et al. (1996) survey previously discussed. This may be because few prior bankruptcy prediction studies have been executed in Norway. Most studies have been student theses with limited objectives, typically replication of some of the early financial ratio bankruptcy prediction studies on industry sector segment data. The most extensive of these studies reports use of discriminant and logit analysis and results similar to Altman et al. (1977), i.e., 97% and 83% accurate on the validation sample in year one and five prior to bankruptcy (Olsen, 1991). On a broader data set and using discriminant analysis only (Eklund, 1990), reports an average of 80% accuracy the four years prior to bankruptcy. Eilifsen and Gjesdal (2001) have found that the fraction of bankruptcy companies receiving auditor's going concern modifications in Norway is higher than figures reported in US studies, 62% compared with US figures of around 50% (McKee, 1995).

2.5. Research on variable selection

There have been a large number of research studies which were concerned with variable selection factors. For example, Jones (1987) surveys the substantial volume of bankruptcy prediction studies which have been published since Beaver's (1968) and Altman's (1968) seminal works. One finding by Jones (1987) was that prior research indicates that start-up (new) companies have a higher failure rate than established companies. He also determined that newly established companies may exhibit some unique characteristics.

Most of the early bankruptcy studies concentrate on investigating the predictive power of various financial ratios. Courtis (1978) made an early attempt to determine which of the financial ratios had proven to be useful in predictive studies. He found 79 financial ratios in a wide variety of studies and classified them into three categories, profitability ratios, solvency ratios, and managerial performance ratios. Later studies proposed that macroeconomic variables be included (Foster, 1986) as well as a wide variety of qualitative variables such as quality of management etc. (Zopounidis, 1987). Other studies have investigated the predictive power of the content of auditor's report (e.g., McKeown et al., 1991; Hopwood et al., 1994; Nogler, 1995).

In 1994 Hopwood et al. suggested that standard statistical bankruptcy models do not apply simultaneously to both stressed and non-stressed firms. They argued that including non-stressed firms was not actually reflective of the auditors' decision problem.

Many failing firms also experience management fraud (illegal earnings management) because of the economic pressure that stressed firms are subject to. Recently, several models have been proposed to better predict such fraud (e.g., Benish, 1997; Green and Choi, 1997).

3. Research design

Since genetic programming is quite different from classical statistics, the research approach requires some explanation. This section discusses the following main elements of the research approach:

- General explanation of genetic programming
- Variable selection process for this study
- Data selection for this study
- Variable reduction procedures
- Final model development

3.1. General explanation of genetic programming

"Genetic programming is a technique for programming computers by means of natural selection" (Koza, 1992). It is a variant of the genetic algorithm developed by Holland (1975) and others, which is based on the concept of adaptive survival in natural organisms. Simply stated, the organisms that are best suited to a particular environment are more likely to survive and reproduce. By surviving to reproduce, they will transfer their survival traits to the succeeding generation. If the environment has constraints that limit survivability of the population, such as ability to perform a certain task, then the fitness standards should increase for each successive generation.

Classical genetic algorithms are implemented using bit strings to represent behavior conditional on values of the variables that are relevant to the problem. Each bit string is a possible structure (individual behavior rule) that might survive to the next generation. These bit strings are analogous to chromosomes in nature, and evolve through time via genetic recombination and selection pressure. Genetic programming differs by using a subset of some programming language such as LISP (Koza, 1992), or assembly language (Nordin, 1997) to represent the individual behavior rules. [2] In this case, an individual behavior rule is a string of elementary instructions in the given language, which is then subject to genetic recombination and selection as in the classical genetic algorithm. One advantage with this is that evolved rules consist of arithmetic and logical expressions, which can easily be analyzed to detect structural properties of the surviving behavior rules.

Genetic algorithms use some type of fitness measure to evaluate the performance of each individual in a population. Fitness is typically determined from the objective function for some optimization problem, e.g., the number of correctly classified cases in a set of bankrupt and non-bankrupt firms.

In order to find a computer program which best solves a given task, the algorithm starts by randomly generating a large population of candidate programs. It then continues for a large number of iterations by replacing low performing programs with genetic recombinations of high

[2] If genetic programming manipulation of either LISP or assembly language results in nonsense programs, these programs would receive low fitness scores and be dropped from the pool of candidate programs.

performing ones. Each computer iterations is therefore a type of "tournament" where the developed computer programs compete for survival. A program with a low fitness measure is deleted and does not survive for the next generation [computer iteration]. The genetic operators used to perform recombinations of programs [create new programs] are crossover and mutation. These are used by all GP-algorithms and will be explained below. However, GP-algorithms differ in the way they mimic natural selection. In this paper, we will use a steady state algorithm with tournament selection, which works as follows:

1. *Initialization:* Randomly generate a population of N bankruptcy classification programs, B_n, $n = 1, \ldots, N$.
2. *Loop:*
   - 2.1. *Tournament:* Randomly select four programs from the population, evaluate them on the set S of training cases, and rank them according to fitness.
   - 2.2. *Reproduction:* Replace the least fit two programs by copies of the best fit 2.
   - 2.3. *Mutation:* For each of the two copied programs, do the following with mutation probability p_m: Randomly select a single instruction in the program, and replace it with a randomly selected instruction. This operation has the effect of increasing the genetic diversity of the population.
   - 2.4. *Crossover:* With crossover probability p_c, recombine the genetic material of the two copied (and possibly mutated) programs by randomly selecting a subarray of instructions from each program, and exchanging it with the subarray from the other program. The resulting two new individuals take the place of the least fit two, and are different from their parents with probability $1 - (1 - p_m)(1 - p_c)$.
3. *Go to 2.*

The population of programs evolves through this training process as long as it continues to produce significantly better programs. When the learning process slows down, the process must be stopped in order to avoid overtraining. This phenomenon occurs when the population has ceased to learn about the underlying structure of the problem and begins to memorize the data and the associated target response. Further training will then only result in a loss of the ability of the programs to generalize beyond the training data set.

GP-algorithms also differ with respect to the type of building blocks they use to construct individual programs. Typically, one uses a subset of elementary instructions from some existing programming language, such as LISP, Java or machine code. In the present paper, we use a machine code implementation of GP due to Nordin (1997). Each program is then a linear string of machine instructions which operates on variables and constants stored in memory, using the CPU floating point registers to store and manipulate temporary variables. The return value of a given program is taken to be the contents of register 0 after all its machine instructions have been processed by the CPU.

Fig. 3 illustrates some aspects of the machine code GP algorithm by means of an example where the maximum program length is 6 instructions. In practice, the maximum program length will be much larger, e.g. 100–500 instructions. The left part of the figure depicts two programs, A and B, with 5 and 6 machine instructions, respectively. X1 and X2 are variables, whose values are stored in memory, and R0 and R1 refer to floating point registers 0 and 1 of the Intel Pentium chip, which has a total of 8 such registers. The GP-algorithm clears these registers by loading them with the value 0.0 before it passes a program to the CPU for execution.

Consider program A. Execution begins with instruction 1, which loads the value of variable X1 into register 0, and next copies it into register 1 in instruction 2. In instruction 3, the contents of register 0, which is still X1, is multiplied into register 1 to yield $X1^2$, and in instruction 4, the contents of register 0 is divided by 3 to yield X1/3. This is also the return value from program A, as indicated by instruction 5.

In order to mutate program A, one selects one of the first four instructions at random, leaving the return instruction in slot 5 unchanged, and re-

		Before crossover			*After crossover*		
Program:		A		B	A'		B'
Machine instructions	*1*	R0 = X1		R0 = X2	R0 = X1		R0 = X2
	2	R1 = R0		R0 = R0 2	R1 = R0		R0 = R0 - 2
	3	R1 = R1*R0		R0 = R0*R0	R1 = R1*R0		R0 = R0*R0
	4	R0 = R0/3	⇔	R0 = R0+1	R0 = R0+1		R0 = R0/3
	5	Return R0	⇔	R0 = R1/R0	R0 = R1/R0		Return R0
	6		⇔	Return R0	Return R0		
Return value		X1/3		0	$X1^2/(X1+1)$		$(X2-2)^2/3$

Fig. 3. Program structure and crossover operation in machine code GP.

places the selected instruction with a randomly chosen machine instruction. To cross two programs, A and B say, one chooses a common subarray in both of them and exchanges the instructions in that subarray between them. This is illustrated in Fig. 3, where the subarray consists of instruction slots 4–6. The result is two new programs A′ and B′, as depicted in the rightmost part of the figure.

In our context, each program is a non-linear bankruptcy prediction rule $B: X \to R$, where $x \in X$ is a vector of explanatory variables, and $B(x)$ is a real number which is interpreted by the modeler as a bankruptcy probability. The corresponding classification rule $C: [0, 1] \to \{0, 1\}$ is defined as follows: Choose a cutoff value c, and for each $b = B(x)$, let $C(b) = 1$ (bankruptcy prediction) if $b > c$ and $C(b) = 0$ (non-bankruptcy prediction) otherwise. Note that the cutoff value cannot be used to change the distribution of Type I and Type II errors of the evolved classification rules, because the GP-algorithm will compensate for any change in c by corresponding changes in the non-linear bankruptcy prediction rules B. Therefore, we may arbitrarily set the cutoff value to 1/2.

Given a set $S = \{(y_1, x_1), (y_2, x_2), \ldots, (y_n, x_n)\}$ of accounting data (x_i) for bankrupt $(y_i = 1)$, and non-bankrupt $(y_i = 0)$ firms, the *fitness* $F(B)$ of rule B is defined by

$$F(B) = \sum_{i \in Z} d(y_i, B(x_i)),$$

where d is a suitable distance function. Since we are primarily interested in bankruptcy *classification*, it would be natural to use the distance function $|y_i - C(b_i)|$, i.e. a penalty per misclassified case. However, since this is not sensitive to marginal improvements in deviations from the target, we can improve upon the search process performed by the GP-algorithm by combining it with a term that introduces such a sensitivity. We will therefore use the distance function

$$d(y_i, b_i) = (y_i - b_i)^2 + (\pi/2) * |y_i - C(b_i)|,$$

which yields a fitness function given by the sum of squared deviations from the targets plus a penalty per misclassified case of π. In order to find a suitable value for π, we did a few exploratory runs with the GP-algorithm, varying π from 0 to 1 in steps of 1/4, and found that $\pi = 1/2$ gave the best results in terms of hit rates on the set of training cases.

In order to inspect and analyse the programs generated by the GP-algorithm, one needs to translate them into some humanly readable language, such as C or arithmetic expressions. The implementation of machine coded GP that we use here has a built-in facility for generating this kind of output in a form which can be read into various software packages for doing symbolic math manipulation, simplification and analysis (McKee and Lensberg, 2002).

Banzhaf et al. (1998) provide a more extended discussion of genetic programming for those readers who want more information.

3.2. Variable selection process for this study

As previously discussed in the literature review, bankruptcy research been conducted from a number of different perspectives. We reviewed each of these perspectives to see what lessons could be learned regarding potential predictive variables which might contribute to a more unified theory of bankruptcy.

Since financial ratios had been widely used in prior research, we judgmentally selected 15 basic ratios which were strong bankruptcy indicators in multiple prior bankruptcy prediction models studies with a view toward including ratios from each major analysis perspective, i.e., liquidity, profitability, solvency, etc.

Based on research concerning the predictive power of the auditor's report, we decided to include the prior audit opinion as a potential explanatory variable. We classify the auditor's opinion either unmodified, unqualified with explanatory paragraph, or some other type of modified reporting.

Since the professional literature indicates fraud is a factor which might influence bankruptcy, we decided to include 10 possible fraud indicators as potential predictive variables for the bankruptcy status.

As it had been suggested that stressed and non-stressed firms represent different firms' states, we included a financial stress dummy variable as a potential predictive variable.

Since new or start-up companies have a higher failure rate and may exhibit unique characteristics compared to older, established companies, we used a dummy variable in this study to reflect start-up status. This allowed us to control for the higher failure rate experienced by new companies.

To summarize, after extensive consideration of prior literature, we ultimately selected a total of 28 financial and non-financial variables with potential power to predict bankruptcy as summarized below:

- financial ratios [15 variables]
- fraud indicators [10 variables]
- prior auditor's opinion [1 variable]
- presence of financial stress [1 variable] and
- survival status of the company (start-up year) [1 variable]

These variables are described in more detail in Table 1.

Many research approaches, either formally or informally, link explanatory variables with decision outcomes. For example, rough sets theory incorporating dominance relations utilizes information about the ordering properties of explanatory variables. Likewise, isotonic separation, a linear programming technique, also incorporates predetermined ordering properties for explanatory variables. Indeed, it can be argued that even forms of regression analysis use an implied hypothesized relationship since solution variable form possibilities may be limited by the technique selected.

A lack of hypothesized relationship between explanatory variables is a comparative advantage of genetic programming. In contrast to past research approaches, we do not hypothesize individual relationships between explanatory variables and bankruptcy. Since genetic programming is capable of detecting complex interrelationships between multiple explanatory variables and bankruptcy status, establishing one-to-one relationships between individual variables and bankruptcy status is an unnecessarily limiting approach for theory development. McKee and Lensberg (2002) found that the change in bankruptcy status as one variable changes (e.g., profitability), may depend on the level of several other variables (e.g., liquidity and company size).

3.3. Data selection

The Norwegian Register of Bankruptcies maintains a record of every business entity entering bankruptcy. This register indicated that during the first half of 1998, 1953 entities entered bankruptcy in Norway. Using these entities as our population, we searched a financial accounting database developed by Dun and Bradstreet, containing approximately 100,000 companies, for bankrupt companies (both private and public) for usable accounting records for a prediction period ending in 1996. Almost all of this population consisted of private companies since there were

Table 1
Potential predictive variables

	#	Variable names	Definition of variables
AUDIT	1	Audit opinion	Unmodified opinion (coded 1), unqualified opinion with explanatory paragraph (coded 2) or other modified opinion (coded 3)
FINANCIAL RATIOS	2	Firm size	$\log_{10}$ total assets$_t$[a]
	3	Profitability I	Net income$_t$/Total assets$_t$
	4	Profitability II	(Operating income + financial income)$_t$/Total assets$_t$
	5	Working capital management	Current assets$_t$/Current liabilities$_t$
	6	Long-term investment I	(Net investment-cash flow tangible fixed assets − net income)$_t$/Total assets$_t$
	7	Long-term investment II	(Net investment-cash flow fixed assets − net income)$_t$/Total assets$_t$
	8	Financial management	(Dividends − net income)$_t$/Total assets$_t$
	9	Quality of earnings	(Operating cash flow − net income)$_t$/Total assets$_t$
	10	Capital turnover	Revenues$_t$/Total assets$_t$
	11	Cash position I	Cash$_t$/Current liabilities$_t$
	12	Cash position II	(Cash + short-term investment)$_t$/Current liabilities$_t$
	13	Earned capitalization	Retained earnings$_t$/Total assets$_t$
	14	Leverage	Liabilities$_t$/Total assets$_t$
	15	Age I	$\log_{10}(1900 + 1 + t -$ year founded)
	16	Age II	Share capital$_t$/Total assets$_t$
FRAUD INDICATORS	17	Profitability I index	(Net income/Total assets)$_t$/(Net income/Total assets)$_{t-1}$
	18	Long-term investment II index	[(Net investment-cash flow − net income)$_t$/Total assets$_t$]/ [(Net investment-cash flow − net income)$_{t-1}$/Total assets$_{t-1}$]
	19	Leverage index	(Liabilities/total assets)$_t$/(Liabilities/Total assets)$_{t-1}$
	20	Receivables index	(Receivables/sales)$_t$/(Receivables/sales)$_{t-1}$
	21	Gross margin index	(Operating income − sales)$_t$/(Operating income − sales)$_{t-1}$
	22	Positive accrual dummy I	Coded 1 if total accruals were positive in year $t-1$ and in year t, 0 otherwise [Accrual calculation adopted from Benish (1997)]
	23	Positive accrual dummy II	1 if total accruals were positive in year $t-1$ and in year t, 0 otherwise [Accrual calculation adopted from Healy (1985)]
	24	Declining cash flow dummy	Coded 1 if cash sales in year t were lower than cash sales in year $t-1$, 0 otherwise
	25	Interest paying ability	(Operating income + interest expenses)$_t$/Interest expenses$_t$
	26	Interest paying ability index	[(Operating income + interest expenses)$_t$/Interest expenses$_t$]/ [(operating income + interest expenses) $_{t-1}$/ Interest expenses$_{t-1}$]
START-UP	27	Start-up dummy	Coded 1 if (t + 1900 − year founded) less than 2, 0 otherwise
STRESS	28	Financial stress dummy	Classified as stressed and coded 1 (0 otherwise) if a company meet at least one of these four criteria: 1. Negative working capital in current year 2. Operating loss in any of the 3 years prior to bankruptcy 3. Net income loss in any of the 3 years before bankruptcy 4. Negative retained earnings in any of the 3 years before bankruptcy

[a] t is the predictive year.

only 235 public traded companies in Norway by the end of 1998. In order to calculate all our 28 potential predictors we needed accounting data for the four years 1993–1996. Only 211 of the 1953 bankrupt companies in the Dun and Bradstreet database had data available for all 28 variables. Accordingly, our initial sample size for developing a model was restricted to 211 bankrupt companies.

The 211 bankrupt companies were then matched with non-bankrupt firms using 5-digit industry codes to yield a total sample of 422 bankrupt and non-bankrupt companies. Table 2 shows the means of the 28 variables for the year 1996 when partitioned into bankrupt and non-bankrupt sub-samples.

As Table 2 reveals, there is quite a wide range of means for some variables while others differ only marginally. For example, the means for the variables such as Firm size, Age I, and the Start-up are relatively close, while the means for the variables such as Cash position, Age II, and Interest paying ability differ substantially. The range for these variables mirrors the real life problem that auditors and lenders have in actual bankruptcy prediction, that is, they must predict bankruptcy or non-bankruptcy status for companies that differ on some, but not on other attributes. Companies that eventually go bankrupt, and those that do not, frequently exhibit very similar characteristics.

3.4. Variable analysis process

In order to identify the significant variables from the overall set of 28 variables, we performed an initial variable analysis on the sample of 422

Table 2
Comparison of sub-sample characteristics

#	Variable names	Means non-bankrupt companies (211)	Means of bankrupt companies (211)
1	Audit opinion	0.39	1.02
2	Firm size	3.26	2.98
3	Profitability I	−0.06	−0.37
4	Profitability II	0.06	−0.32
5	Working capital management	2.00	0.82
6	Long-term investment I	0.00	0.37
7	Long-term investment II	0.01	0.35
8	Financial management	0.08	0.37
9	Quality of earnings	0.17	0.09
10	Capital turnover	3.22	4.29
11	Cash position I	0.68	0.10
12	Cash position II	0.71	0.10
13	Earned capitalization	−0.08	−0.37
14	Leverage	0.94	5.43
15	Age I	2.39	2.22
16	Age II	−0.15	−5.09
17	Profitability I index	−0.07	−0.16
18	Long-term investment II index	−0.03	0.11
19	Leverage index	0.00	3.64
20	Receivables index	0.00	−0.01
21	Gross margin index	0.25	2.77
22	Positive accrual dummy I	0.17	0.10
23	Positive accrual dummy II	0.08	0.12
24	Declining cash flow dummy	0.42	0.51
25	Interest paying ability	16.82	−3.32
26	Interest paying ability index	1.14	1.83
27	Start-up dummy	0.00	0.00
28	Financial stress dummy	0.74	0.95

bankrupt and non-bankrupt companies. Our objective was to identify the significant variables so we could subsequently expand our sample size prior to performing the actual model development. We believed the model developed would be more robust if we were able to utilize more than 211 bankrupt companies in the model development. We also wanted to make sure we would be able to utilize a significant holdout sample to subsequently validate any model developed.

Our initial variable analysis was done by performing 70 computer runs with the GP-algorithm, each computer run consisting of 2,020,000 tournaments. [3] For each computer run we reported variable results after every 40,000 tournaments which resulted in 50 reports. At the time of each report, we checked each of the 28 variables to determine whether or not it had any effect on the classification ability of the best program at that point in the run. We assigned a value of 1 if a variable had a non-zero effect on the classification ability of the best program and zero otherwise.

We analyzed the variables for reports $30, \ldots, 50$, and obtained summary statistics d_i across runs and periods for each variable i. We did not deem it necessary to analyze results for runs for reports 0–29 since genetic programming models are randomly started, develop slowly, and take a while to stabilize on a meaningful model.

Thus, we did 70 computer runs, reported results 21 times in the latter phase of each run, and thereby obtained 1470 observations for each variable ($70 \times 21 = 1470$). Table 3 provides data about the strength of the individual variables. The value reported in the table tells the fraction of periods in which a specific variable had a non-zero effect on classification ability.

As shown in Table 3, the leading five variables which were significant predictors in at least 50% of the tournaments were:

- audit opinion [variable 1]
- cash position II [variable 12]
- age II [variable 16]
- firm size [variable 2] and
- interest paying ability [variable 25]

Table 3
Summary of variable strengths across tournament rounds

#	Variable names	Mean	Standard deviation
1	Audit opinion	1	0
12	Cash position II	0.972789	0.162753
16	Age II	0.886395	0.317439
2	Firm size	0.832653	0.373412
25	Interest paying ability	0.646259	0.478293
9	Quality of earnings	0.470068	0.499273
5	Working capital management	0.463265	0.498818
19	Leverage index	0.458503	0.498445
6	Long-term investment I	0.37415	0.484067
7	Long-term investment I	0.329252	0.470101
10	Capital turnover	0.319728	0.46653
21	Gross margin index	0.27415	0.446237
11	Cash position I	0.212925	0.409515
14	Leverage	0.187075	0.390104
23	Positive accrual dummy II	0.17483	0.379951
3	Profitability I	0.168708	0.374621
20	Receivable index	0.167347	0.373412
26	Interest paying ability index	0.163946	0.370352
28	Financial stress dummy	0.131293	0.337835
15	Age I	0.131293	0.337835
4	Profitability II	0.12517	0.331024
22	Positive accrual dummy I	0.12449	0.330252
13	Earned capitalization	0.120408	0.325549
27	Start-up dummy	0.106803	0.308967
17	Profitability I index	0.106122	0.308099
8	Financial management	0.1	0.300102
18	Long-term investment II index	0.088435	0.284024
24	Declining cash flow dummy	0.081633	0.273897

The number of observations is 1470 for all variables.

There are two interesting factors about these variables that are interesting to note. First, the prior audit opinion was so high that it appeared in all models created in all tournaments. This is consistent with prior research by Hopwood et al. (1994) and Nogler (1995). Second, despite considerable discussion in the literature suggesting they might be significant, only one of the fraud indicators (interest paying ability) was significant, and the financial stress variable did not turn out to be significant.

It is not surprising that neither the Age I or Start-up Dummy variables were included since for a firm to be included in the data set employed

[3] The 70 computer runs took a total of approximately 18 hours on a 1.8 GHz AMD processor.

on the 28 variables it must have had accounting data for the four-year period 1993–1996. Therefore, the Age I variable and the Start-up dummy were not likely to possess classification ability since all 422 firms analyzed were at least 4 years old. After individually analyzing these two variables in the larger sample, we judgmentally decided to include the Age I variable with the other five variables in the final model in order to improve prediction capability for companies that had been in existence only a short time.

3.5. Model development

The variable analysis process reduced the number of variables from 28 to 6. An analysis of the overall data set indicated that it contained 568 bankrupt companies with data for those six variables. This is a significant increase from the 211 bankrupt companies which had data for all 28 variables. Thus, the reduction from 28 to 6 variables in the variable analysis process enabled us to increase the number of bankrupt companies utilized in model development and validation from 211 to 568.

The 568 bankrupt companies were then matched with 568 non-bankrupt companies by 5-digit industry code to yield a paired sample of 1136 observations. The 568 pairs were then reordered randomly. Table 4 shows the means of these variables based on non-bankrupt and bankrupt sub-samples.

In contrast to Table 3, all means in Table 4 differ substantially. Since space does not permit a detailed analysis of each of the six variables only the two more unusual ones are analyzed in more detail. Fig. 4, does however, provide additional insight into variable number 1, Audit Opinion. Confirming prior research, it shows that non-bankrupt companies have substantially more opinions in category 1, unmodified, and substantially less opinions in categories 2 and 3, unmodified with explanatory paragraph and modified. Fig. 5 provides additional insight in variable number 15, Age I. As suggested by prior literature, it shows that younger companies have a substantially higher bankruptcy rate.

Table 4
Comparison of sub-sample characteristics

#	Variable names	Means non-bankrupt companies (568)	Means of bankrupt companies (568)
1	Audit opinion	0.37	1.06
2	Firm size	3.21	2.86
12	Cash position II	0.96	0.14
15	Age I	2.09	1.64
16	Age II	−0.63	−15.18
25	Interest paying ability	13.70	−7.39

We judgmentally determined that, of the 1136 total firms, the first 900 paired firms would be used for training, and the remaining 236 for model validation. Based on a review of the literature, we believed that a validation sample of 236 firms would be adequate.

The annual accounts of the prediction year 1996 are required to be filed with the register no later than August 1, 1997. Thus, the prediction interval ranges from 12 months to 18 months prior to bankruptcy.

We did one computer run with the genetic programming algorithm for the six variables which consisted of 2,020,000 tournaments. [4] In order to avoid over training, we introduced some random noise in the training data. [5]

The genetic program converged on a stable training hit rate of some 81.5% prior to the end of the 2,020,000 tournaments. We simplified the

[4] This computer run was approximately 30 minutes on a 1.8 GHz AMD processor.

[5] For each variable, we selected the set of training observations between the 5th and the 95th percentile and calculated the standard deviation s_i for each variable i of that subsample, along with the minimum vmin_i and maximum vmax_i for each variable i over the whole training set. For each period, we scrambled each variable v_i by replacing it with w_i, defined as follows:

$$w_i = \mathrm{Max}(\mathrm{vmin}_i, \mathrm{Min}(\mathrm{vmax}_i, v_i + x_i)),$$

where x_i is uniformly distributed on $[-as_i, +as_i]$, where a is a constant. We did some preliminary tests with different a's and found that $a = 1/3$ eliminated the overtraining problem while still yielding good classification results.

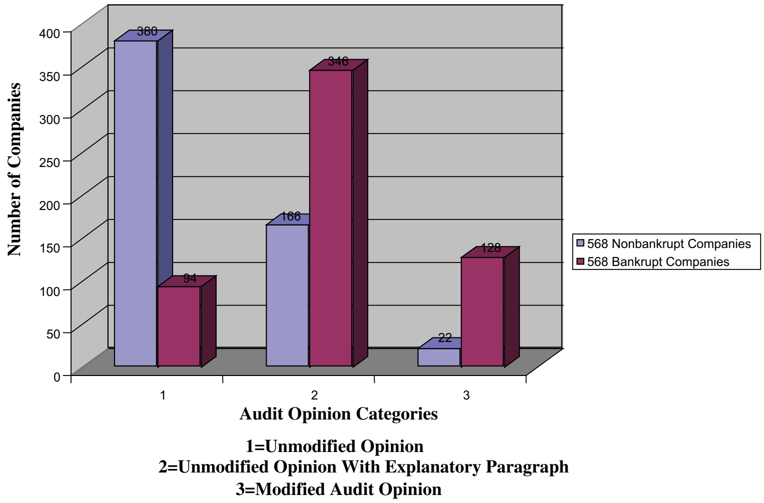


Fig. 4. Comparison of audit opinion categories by bankruptcy status.

programs by introducing a penalty for complexity. [6]

In order to check whether or not the population had appropriately converged to a single model after the initial 2,020,000 tournaments, we continued model development for an additional 10,100,000 tournaments. [7] We then selected the best model after every 20,000 tournaments in order to compare their classifications. This resulted in the development of 500 models in all. Results indicated that these 500 models agree on the classification on 99.1% of the 568,000 [500×1136] classifications. This indicates that the model developed after 2,020,000 tournaments was reasonably stable and represented an appropriate configuration for the variables employed.

[6] This was done by letting M be the maximum length of a program, measured in 16 bit words, and for each program i, let L_i be the number of words occupied by non-dummy instructions. (Dummy instructions occupy one word, and are included in the instruction set of our GP-algorithm because they seem to work well in nature, cf "junk DNA".) The penalty for complexity assigned to program i is then $b * L_i/M$. We then ran the algorithm for another 12,120,000 tournaments. The complexity measure L_i/M was reduced from slightly below 1 to approximately 1/3 within the first 2,020,000 tournaments without significantly affecting the training fitness of the best programs.

[7] This computer run took approximately 3 hours on a 1.8 GHz AMD processor.

4. Results

The 500 models generated in the final model development had hit rates of approximately 82% on the training sample and 81% on the validation sample. We then randomly selected one of the models for further analysis. It had a hit rate of 81.7% on the training sample and 80.9% on the validation sample. Table 5 presents the summary statistics for the 500 models.

Further analysis was conducted in order to determine if the model selected was appropriately representative of the 500 models. Our goal was to determine if the differences in the models were insignificant or significant. We did a classification test using each of the 500 models for all 1136 companies in the overall data set [500×1136 companies $= 568{,}000$ classifications]. We then computed the percentage of correct classifications by each individual model for the classification [bankrupt or non-bankrupt] made by the majority of the 500 models. Results are listed in Table 6.

We found the following agreement percentages between the 500 classification models:

90% or more of the 500 models agreed on 97.5% of all cases
95% or more of the 500 models agreed on 96.0% of all cases

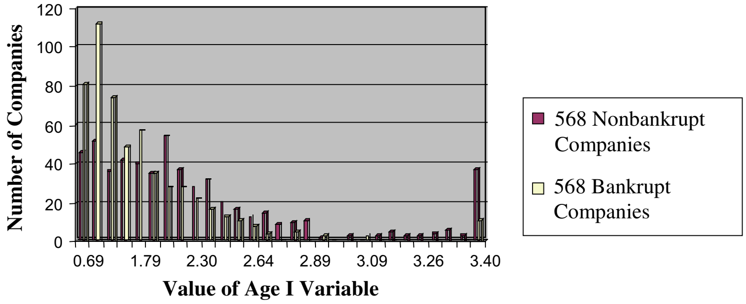


Fig. 5. Comparison of Age I variable by bankruptcy status.

Table 5
Forecasting accuracy statistics for 500 models

	Training sample [900 companies]	Validation sample [236 companies]	Total sample [1136 companies]
Mean	0.818	0.805	0.815
Variance	0.149	0.158	0.151
Standard deviation	0.386	0.397	0.388

Table 6
Comparison of individual model classification success against classifications made by majority of models

	Majority classification model (%)	Single selected model (%)
236 validation cases	80.5	80.9
900 training cases	81.8	81.7
All 1136 cases	81.5	81.5

99% or more of the 500 models agreed on 90.0% of all cases

This relationship between the percentage of the 500 models making a classification and the percentage of cases correctly classified is graphically displayed in Fig. 6. As shown in Fig. 6, there is only a small area in the upper right quadrant where the 500 models had differences.

These results further support the conclusion that the genetic programming algorithm had converged to an appropriately representative and stable classification model after 2,020,000 tournaments. Additionally, they show that the single selected classification model represents the "core" knowledge contained in the population of 500 models and is, therefore, appropriately representative of the population of 500 models.

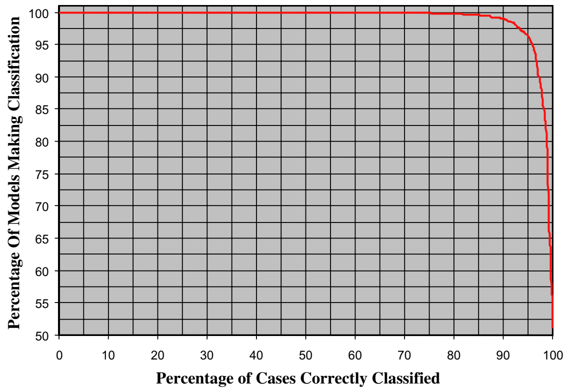


Fig. 6. Agreement between 500 classification models.

4.1. Benchmarking GP model against logit models

We wanted to benchmark the GP model against a more standard methodology to determine its' relative prediction effectiveness. Accordingly, for comparison purposes, we developed two logit models using the same six variables and the same data set. The first logit model used the audit variable as is (with values 0, 1 and 2). Since we believed the bankruptcy probability increased in a nonlinear way for the audit variable, we used a second logit model which represented all possible non-linear functions of the audit variable through the use of two dummy variables. [8] The logit models were 77% and 76% accurate in predicting bankruptcy on the training and development samples respectively.

We performed a McNemar test to determine if the GP model was significantly different from the first logit model. This test examines the predictions where the two models differ, which were 105 out of 1136 observations. The z for this test was -5.953 which means the models were significantly different at the .0000 level. We confirmed this result by also performing a Wilcoxon test which compared all 1136 observations.

4.2. Model structure and comparison

The model selected for detailed analysis had the following structure:

$$(V_1, V_2, V_3, V_4, V_5, V_6) = \frac{X^2}{X^2 + Y} \in [0, 1],$$

where V_1 is the audit opinion (variable 1), V_2 is the firm size (variable 2), V_3 is the cash position II (variable 12), V_4 is the age I (variable 15), V_5 is the age II [share capital/total assets] (variable 16) and V_6 is the Interest paying ability (variable 25),

$$X = 2 + (V_1 * V_2) - V_5 * (3 * V_2 - 5) + 2 * f(V_1, V_4, V_5, V_6) * (V_2 - 2),$$

$$Y = (V_2)^2 * (1 + V_3)^2 * V_4,$$

$$f(V_1, V_4, V_5, V_6) = x^2/(x^2 + V_6 * x + y),$$

where $x = 4 + V_5 * (2 * V_1 * V_6 - 1)$ and $y = (V_4 * V_6^2)/2$.

Note that the preferred genetic model produces values in the interval [0, 1], which suggests that these values can be easily logically interpreted as a measure of bankruptcy risk.

The selected genetic programming model reveals some interesting new features. Although prior research by Hopwood et al. (1989, 1994) and Flagg et al. (1991) revealed that modified audit opinions were significant in distinguishing failed firms, these studies did not indicate significant firm size effects for this variable. A later study by McKeown et al. (1991) found an inverse relationship between client size and the going-concern qualification. The current research generally supports these prior findings and found that an unfavorable audit report has a more negative bankruptcy status impact for a large company than a small one.

Also, the interaction of the x and y terms which include interest paying ability, suggests that interest paying ability has a more positive bankruptcy status impact for large firms than small ones. Taking these two features in the aggregate, this might be interpreted to indicate that the model is suggesting that accounting information (including the auditor's evaluation of it) is more important for larger firms than smaller ones. It also suggests that for small firms the most important information is liquidity and non-accounting information. One might hypothesize that accounting data is perceived as more reliable for large firms than for small ones.

It is interesting to note that only one of the variables in the model, variable 6-Interest Paying Ability, has been considered to be a fraud prediction variable in other research studies. Thus, the presence of fraud risk factors is not necessarily indicative of bankruptcy.

The model validates some findings from a hybrid approach genetic programming-rough sets model developed from US data (McKee and Lensberg, 2002). These findings were (1) liquidity improves non-bankruptcy status regardless of the value of other variables like profitability and size, and unprofitable companies can maintain high levels of liquidity to offset low profitability, and (2) bankruptcy risk decreases with increased size except when profits are negative. This result validation is noteworthy considering (1) the different methodologies employed in the two studies, (2) the substantial asset size differences in the two studies, \$327,000,000 mean for US study versus \$1,276,000 mean for current study, (3) data from different countries were used in the two studies, and (4) the US study only included public

[8] A third dummy variable was not needed since that would make the set of dummy variables collinear with the constant term.

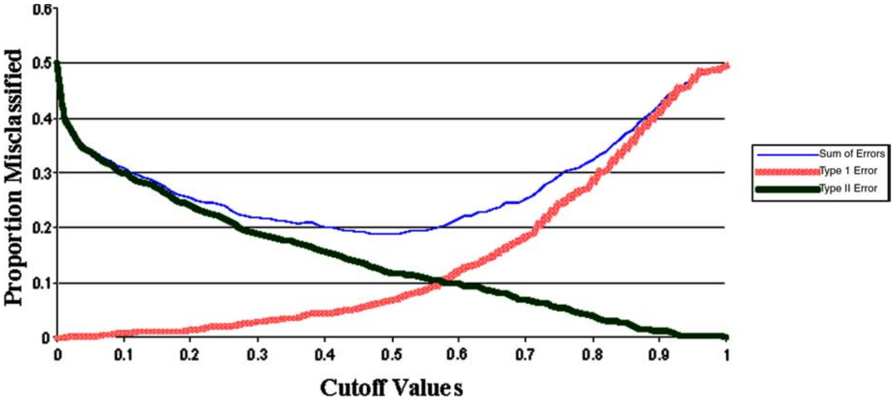


Fig. 7. Error analysis with different cutoffs.

companies while the current research included primarily private companies.

4.3. Error analysis

With a bankruptcy prediction model derived with a more traditional procedure such as logit, the transfer function is part of the model. Therefore, it is possible to analyze Type I and Type II errors with a varying model form. Due to the way the genetic model is developed, the structure of the model and the transfer function are determined simultaneously, thus the model produces bankruptcy probabilities between 0 and 1.

Fig. 7 describes how the predictive accuracy of the genetic programming model would change if the cutoff scores were varied. The flatness of the curve summing the Type I and II errors over the cutoff interval from 0.4 to 0.6, indicates that the model accuracy was relatively insensitive to the cutoff point. Decision makers wanting to further minimize Type II errors could choose a cutoff point higher than 0.5.

4.4. Research limitations

The need for relevant data required us to perform our initial GP analyses for the 28 potential predictive variables on only 211 of the 1953 companies going bankrupt during the first half of calendar year 1998. It is possible that different variables might have been revealed to be significant if data were available for more of these companies.

We limited the GP mathematical operations in order to develop a final model that was less mathematically complex. If the GP algorithm were permitted to perform more complex operations it is possible that the final model may have been of a more complex mathematical form with a higher accuracy percentage.

Genetic programming operations continually produce new programs rather than an "optimum" program. It is possible that had we selected a different program from the 500 final programs, the alternate selection could have been less complex or have had higher classification accuracy although our analysis of the 500 programs indicates they made similar classifications for the most part.

5. Summary and conclusions

Bankruptcy is a highly significant worldwide problem that affects the economic well being of all countries. The high social costs incurred by various stakeholders associated with bankrupt firms have spurred searches for better theoretical understanding and prediction capability.

This research used genetic programming to further improve our understanding of both bankruptcy theory and prediction by extending genetic programming to a different national environment. This research differed significantly from prior US research by (1) the use of non-US data, (2) the analysis of a broad set of variables that were significant in multiple prior studies, (3) the inclusion of fraud prediction factors, (4) the use of primarily private companies, and (5) the use of a longer prediction interval.

A genetic programming model consisting of six variables was developed from a large set of 28 variables which had been significant in prior research. The model was 82% and 81% accurate on the 900 firm training and 236 firm validation samples, respectively. Two logit models developed using the same six variables were only 77% and 76% accurate on the 900 firm training and 236 firm validation samples, respectively.

The in-depth analysis of the interaction of the variables for in genetic programming model reveals several very interesting results. The model shares some features previously observed in a hybrid approach genetic programming- rough sets model developed from US data. That is that (1) the liquidity improves non-bankruptcy status regardless of the value of other variables like profitability and size, and that unprofitable companies can maintain high levels of liquidity to offset low profitability, and (2) the bankruptcy risk decreases with increased size except when profits are negative.

The selected genetic programming model also reveals some interesting new features. One of these is that an unfavorable audit report has a more negative bankruptcy status impact for a big firm than a small one. The model also suggests that interest paying ability has a more positive bankruptcy status impact for large firms than small ones. This might be interpreted to indicate that the model is suggesting that accounting information (including the auditor's evaluation of it) is more important for larger firms than smaller ones. It also suggests that for small firms the most important information is liquidity and non-accounting information. One might hypothesize that, in Norway, accounting data is perceived as more reliable for large firms than for small ones.

In summary, genetic programming produced what can be deemed a highly accurate bankruptcy prediction model considering both the nature of the primarily non-public companies in the sample and the prediction interval up to 18 months. The model employed six variables derived from multiple prior bankruptcy and fraud related studies. The genetic programming model was more accurate than a traditional logit model using the same variables. The model provides insight into the complex interaction of bankruptcy related factors and suggests avenues for future research.

References

Altman, E.I., 1968. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. Journal of Finance 23 (4), 589–609.

Altman, E.I., 1969. Corporate bankruptcy, potential stockholder returns and share valuation. Journal of Finance (December), 887–900.

Altman, E.I., 1982. Accounting implications of failure prediction models. Journal of Accounting, Auditing and Finance 6 (1), 4–19.

Altman, E.I., McGough, T., 1974. Evaluation of a company as a going concern. Journal of Accountancy 138 (6), 51–57.

Altman, E.I., Haldeman, R.C., Narayanan, P., 1977. Zeta analysis: A new model to identify bankruptcy risk of corporations. Journal of Banking and Finance 1, 29–54.

American Bankruptcy Institute, 2/20/2001. Available from: <http://www.abiworld.org/stats/1994quarterbus.html>.

Banzhaf, W., Nordin, P., Keller, R.E., Francone, F.D., 1998. Genetic Programming—An Introduction. Morgan Kaufmann Publishers, San Francisco, CA.

Beaver, W.H., 1968. Market prices, financial ratios and the prediction of failure. Journal of Accounting Research (Autumn), 179–192.

Benish, M.D., 1997. Detecting GAAP violations: Implications for assessing earnings management among firms with extreme financial performance. Journal of Accounting and Public Policy, 271–309.

Bryan, D., Tiras, S.L., Wheatley, C.M., 2000. The relation of audit opinion and auditor change with bankruptcy emergence. Paper presented at 2001 AAA Audit Section Midyear Meeting.

Bureau of Economic Analysis—US Department of Commerce, 2/21/2001. Available from: <http://www.bea.doc.gov/bea/dn1.htm>.

Chen, K., Church, B., 1992. Default on debt obligations and the issuance of going-concern opinions. Auditing: A Journal of Theory and Practice 11 (2), 30–39.

Chen, K.H., Shimerda, T.A., 1981. An empirical analysis of useful financial ratios. Financial Management (Spring), 51–60.

Cormier, D., Magan, M., Morard, B., 1995. The auditor's consideration of the going concern assumption: A diagnostic model. Journal of Accounting Auditing and Finance 10 (2), 201–222.

Courtis, J.K., 1978. Modeling a financial ratio categoric framework. Journal of Business, Finance, and Accounting 5 (4), 371–386.

Dimitras, A.I., Zanakis, S.H., Zopounidis, C., 1996. A survey of business failures with an emphasis on prediction methods and industrial applications. European Journal of Operational Research 90, 487–513.

Eilifsen, Aa., Gjesdal, F., 2001. Short-term prediction of bankruptcy based on non-financial indicators. Manuscript, Norwegian School of Economics and Business Administration.

Eklund, T., 1990. Konkursindikator—et nyttig verktøy for analyse av distriktsbedrifter? BETA 1, 13–25.

Flagg, J.C., Giroux, G.A., Wiggins Jr., C.E., 1991. Predicting corporate bankruptcy using failing firms. Review of Financial Economic 1 (1), 67–78.

Foster, G., 1986. Financial Statement Analysis. Prentice-Hall, NJ.

Greco, S., Matarazzo, B., Slowinski, R., 1998. A new rough set approach to evaluation of bankruptcy risk. In: Zopounidis, C. (Ed.), Operational Tools in the Management of Financial Risks. Kluwer Academic Publishers, The Netherlands.

Green, B.P., Choi, J.H., 1997. Assessing the risk of management fraud through neural network technology. Auditing: A Journal of Practice & Theory 16 (1), 14–28.

Holland, J.H., 1975. Adaptation in Natural and Artificial Systems. University of Michigan Press.

Hopwood, W., McKeown, J.C., Mutchler, J.F., 1989. A test of the incremental explanatory power of opinions qualified for consistency and uncertainty. The Accounting Review 66 (1), 28–48.

Hopwood, W., McKeown, J.C., Mutchler, J.F., 1994. A reexamination of auditor versus model accuracy within the context of the going concern opinion decision. Contemporary Accounting Research 10 (2), 409–431.

International Accounting Standards Board, 2002. International Accounting Standards No. 1-Presentation of Financial Statements.

International Auditing and Assurance Standards Board of International Federation of Accountants, 2003. International Standards on Auditing.

Johnson, V., Khurana, I., 1995. Auditor reporting for bankrupt companies: Evidence on the impact of SAS No. 59. Research in Accounting Regulation 9, 3–32.

Jones, F.L., 1987. Current techniques in bankruptcy prediction. Journal of Accounting Literature 6, 131–164.

Knapp, M., Elikai, F., 1990. Auditor changes and information suppression. Research in Accounting Regulation 4, 3–20.

Koza, J.R., 1992. Genetic programming: On the programming of computers by means of natural selection. Massachusetts Institute of Technology.

Laitinen, T., Kankaanpää, N., 1999. Comparative analysis of failure prediction methods: The Finnish case. The European Accounting Review 8 (1), 67–92.

Lenard, M.J., Alam, P., Booth, D., 2000. An analysis of fuzzy clustering and a hybrid model for the auditor's going concern assessment. Decision Sciences 31 (4), 861–884.

Little, P., McAlum, H., 1991. Subsequent events for companies receiving going concern audit opinions. Southern Business Review 17 (2), 22–32.

McKee, T.E., 1995. Predicting bankruptcy via induction. Journal of Information Technology 10, 26–36.

McKee, T.E., Lensberg, T., 2002. Genetic programming and rough sets: A hybrid approach to bankruptcy classification. European Journal of Operational Research 138, 436–451.

McKee, T.E., 2000. Developing a bankruptcy prediction model via rough sets theory. International Journal of Intelligent Systems in Accounting, Finance, & Management 9, 159–173.

McKeown, J.C., Mutchler, J.F., Hopwood, W., 1991. Towards an explanation of auditor failure to modify the audit opinions of bankrupt companies. Auditing: A Journal of Practice & Theory 10 (supplement), 1–13.

Menon, K., Schwartz, K., 1986. The auditor's report for companies facing bankruptcy. The Journal of Commercial Bank Lending (January), 42–52.

Michalski, R.S., 1983. A theory and methodology of inductive learning. Artificial Intelligence 20, 111–161.

Nelson, M., Elliott, J., Tarpley, R., 2002. Evidence from auditors about managers' and auditors earnings-management decisions. Presented at Accounting Review Quality of Earnings Conference.

Nogler, G., 1995. The resolution of auditor going concern opinions. Auditing: A Journal of Practice & Theory 14 (2), 54–73.

Nordin, P., 1997. Evolutionary Program Induction of Binary Machine Code and its Application. Krehl Verlag, Munster, Germany.

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

Olsen, J.P., 1991. En empirisk analyse av modeller for konkursprediksjon. Thesis, Norwegian School of Economics and Business Administration.

Richardson, F., Kane, G., Lobingier, P., 1998. The impact of recession on the prediction of corporate failure. Journal of Business, Finance and Accounting 25, 167–186.

Schwartz, K., Menon, K., 1985. Auditor switches by failing firms. The Accounting Review 60 (2), 248–261.

Schwartz, K., So, B., 1995. An analysis of form 8-K disclosures of auditor changes by firms approaching bankruptcy. Auditing: A Journal of Practice & Theory 14 (1), 125–136.

Sweeney, A., 1994. Debt-covenant violations and managers' accounting responses. Journal of Accounting and Economics 17, 281–308.

Vermeulen, E.M., Spronk, J., Van Der Wijst, N., 1998. The application of the multi-factor model in the analysis of corporate failure. In: Zopounidis, C. (Ed.), Operational Tools in the Management of Financial Risks. Kluwer Academic Publishers, The Netherlands.

Zopounidis, C., 1987. A multicriteria decision making methodology for the evaluation of the risk of failure and an application. Foundations of Control Engineering 12 (1), 45–67.

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

Zmijewski, M.E., Hagarman, R.L., 1981. An income strategy approach to the positive theory of accounting standards setting/choice. Journal of Accounting and Economics 3, 129–149.