CMS 1: 293–310 (2004)

DOI: 10.1007/s10287-004-0017-6

A hybrid genetic model for the prediction of corporate failure

Anthony Brabazon, Peter B. Keenan

University College Dublin, Ireland (e-mail: anthony.brabazon@ucd.ie; peter.keenan@ucd.ie)

Abstract. This study examines the potential of a neural network (NN) model, whose inputs and structure are automatically selected by means of a genetic algorithm (GA), for the prediction of corporate failure using information drawn from financial statements. The results of this model are compared with those of a linear discriminant analysis (LDA) model. Data from a matched sample of 178 publicly quoted, failed and non-failed, US firms, drawn from the period 1991 to 2000 is used to train and test the models. The best evolved neural network correctly classified 86.7 (76.6)% of the firms in the training set, one (three) year(s) prior to failure, and 80.7 (66.0)% in the out-of-sample validation set. The LDA model correctly categorised 81.7 (75.0)% and 76.0 (64.7)% respectively. The results provide support for a hypothesis that corporate failure can be anticipated, and that a hybrid GA/NN model can outperform an LDA model in this domain.

Keywords: Corporate failure, Genetic algorithm, Neural network

MSC codes: 62M45, 68W10, 90B50, 91C20

1 Introduction

Classification is a commonly encountered decision scenario in business. Examples include decisions as to whether or not to invest in a firm, whether to extend trade credit to a new customer, or whether to extend a bank loan. In each of these scenarios, the possibility of financial loss exists if a firm is incorrectly classified as being financially healthy. Corporate bankruptcy can impose significant private costs on many parties including shareholders, providers of debt finance, employees, suppliers, customers and auditors. All of these stakeholders have an interest in being able to identify whether a company is on a trajectory which is tending towards failure. Early identification of such a trajectory could facilitate successful

intervention, to avert potential disaster. Nonetheless, it must also be recognised that corporate failure is a natural component of the market economy, facilitating the recycling of financial, human and physical resources into more productive organisations (Schumpeter, 1934; Easterbrook, 1990).
Corporate failure can arise for many reasons. It may result from a single catastrophic event or it may be the terminal point of a process of decline. Under the second perspective, corporate failure is a process which is rooted in management defects, resulting in poor decisions, leading to financial deterioration and finally corporate collapse (Altman, 1993; Hambrick and D'Aveni, 1988). Most attempts to predict corporate failure implicitly assume that management decisions critically impact on firm performance (Argenti, 1976). Although management decisions are not directly observable by external parties, their consequent affect on the financial health of the firm can be observed through their impact on the firm's financial ratios. Previous studies have utilised a wide variety of explanatory variables in the construction of corporate distress models, including data drawn from the financial statements of firms, data from financial markets, general macro-economic indicators, and non-financial, firm-specific information. In this study, we limit our attention to information drawn from financial statements.

1.1 Motivation for study

There are a number of reasons to suppose *a priori* that the use of an evolutionary algorithm combined with a neural network (NN), can prove fruitful in the prediction of corporate failure. The field is characterised by the lack of a strong theoretical framework and has a multitude of plausible, potentially interacting, explanatory variables. The first problem facing the modeller is the selection of a 'good' subset of these variables, and the second problem is the selection of an appropriate model form. In applications of NNs this is not a trivial task, as many choices are open to the modeller, including the nature of the connection structure (number of layers of nodes, number of nodes in each layer, determining which nodes should be connected to each other), the form of activation function at each node, and the choice of learning algorithm and associated parameters. The selection of quality explanatory variables and model form represents a high-dimensional combinatorial problem, giving rise to potential for an evolutionary methodology which can automate this process with little intervention from the modeller (Mitchell, 1996). Such automated methodologies have clear potential for extension to many data-mining applications.
To date only a relatively limited number of studies have applied evolutionary methodologies, including genetic algorithms (GA), genetic programming (GP), and grammatical evolution (GE), to the domain of corporate failure prediction (Varretto, 1998; Kumar et al., 1997; Back et al., 1996; Brabazon and O'Neill, 2003). This paper builds on these initial studies and employs a hybrid GA/NN methodology, which combines the global search potential of a GA with the non-linear modelling capabilities of a NN. Two prior studies have utilised a similar hybrid model (Back et

al., 1996; Wallrafen et al., 1996). This study is distinguished from the former which utilised the GA to evolve only the inputs for the NN, and from the latter which did not attempt to predict corporate failure for more than a single time-period. This study also employs a recent dataset, based on publically available information on US companies.

1.2 Structure of paper

This contribution is organised as follows. Section 2 discusses prior literature in the corporate failure domain and outlines the definition of corporate failure employed in this study. Section 3 introduces the GA and discusses how it can be combined with a NN methodology. Section 4 describes both the data utilised, and the model development process adopted. Section 5 provides the results of the constructed GA/NN models and compares these with results from a linear discriminant analysis (LDA) model. Conclusions and a discussion of the limitations of the contribution are provided in Section 6.

2 Background

Research into the prediction of corporate failure using financial data, has a long history (Fitzpatrick, 1932; Smith and Winakor, 1935; Horrigan, 1965). Early statistical studies such as Beaver (1966), adopted a univariate methodology, identifying which accounting ratios had greatest classification accuracy in separating failing and non-failing firms. Although this approach did demonstrate classification power, it suffers from the shortcoming that a single weak financial ratio may be offset (or exacerbated) by the strength (or weakness) of other financial ratios. This issue was addressed in Altman (1968) by developing a multivariate LDA model and this was found to improve classification accuracy. Altman's (1968) discriminant function had the following form:

$$Z = .012X_1 + .014X_2 + .033X_3 + .006X_4 + .999X_5$$

where:

X_1 = working capital to total assets

X_2 = retained earnings to total assets

X_3 = earnings before interest and taxes to total assets

X_4 = market value of equity to book value of total debt

X_5 = sales to total assets

A later study by Altman et al. (1977), using a larger dataset, selected the following set of explanatory variables (the study did not disclose the coefficients):

X_1 = return on assets (EBIT / Total Assets)

X_2 = stability of earnings

X_3 = debt service (EBIT / Total Interest)
X_4 = cumulative profitability (Retained Earnings / Total Assets)
X_5 = liquidity (Current Assets / Current Liabilities)
X_6 = capitalisation (Equity / Total Capital)
X_7 = firm size (Total Assets)

Since the pioneering work of Beaver (1966) and Altman (1968), a vast array of methodologies have been applied for the purposes of corporate failure prediction. In the 1970s and 1980s, attention was focussed on logit and probit regression models (Gentry et al., 1985; Zmijewski, 1984; Ohlson, 1980). In more recent times, as the field of biologically-inspired computing has flourished, the methodologies applied to the domain of corporate failure prediction have expanded to include artificial neural networks (Shah and Murtaza, 2000; Serrano-Cinca, 1996; Wilson et al., 1995; Tam, 1991), genetic algorithms (Varretto, 1998; Kumar et al., 1997), and grammatical evolution (Brabazon et al., 2002). Other methodologies applied to this problem include support vector machines (Fan and Palaniswami, 2000), rough sets (Zopounidis et al., 1999), and multicriteria decision analysis models (Zopounidis and Dimitras, 1998). Review studies covering much of the above literature can be found in Dimitras et al. (1996), and Morris (1997). Zhang et al. (1999) provide a good review of prior applications of artificial neural networks to the domain of corporate failure.

2.1 Definition of corporate failure

No unique definition of corporate failure exists (Altman, 1993). Possible definitions range from failure to earn an economic rate of return on invested capital, to legal bankruptcy, followed by liquidation of the firm's assets. Typically, financial failure occurs when a firm incurs liabilities which cannot be repaid from liquid financial resources. However, this may represent the end of a period of financial decline, characterised by a series of losses and reducing liquidity. Any attempt to uniquely define corporate failure is likely to prove problematic. While few publicly quoted companies fail in any given year (Morris, 1997) suggests that the rate is below 2% in the UK, and Zmijewski (1984) reports that this rate is less than 0.75% in the US), poorer performers are liable to acquisition by more successful firms. Thus, two firms may show a similar financial trajectory towards failure, but one firm may be acquired and 'turned-around' whilst the other may fail.
The definition of corporate failure adopted in this study is the court filing of a firm under Chapter 7 or Chapter 11 of the US Bankruptcy code. The selection of this definition provides an objective benchmark, as the occurrence (and timing) of either of these events can be determined through examination of regulatory filings. Chapter 7 of the US Bankruptcy code covers corporate liquidations and Chapter 11 covers corporate reorganisations, which usually follow a period of financial distress. Under Chapter 11, management is required to file a reorganisation plan in

bankruptcy court and seek approval for this plan. On filing the bankruptcy petition, the firm becomes a *debtor in possession*. Management continues to run the day-to-day business operations, but a bankruptcy court must approve all significant business decisions. In most cases, Chapter 11 reorganisations involve significant financial losses for both the shareholders (Russel et al., 1999) and the creditors (Ferris et al., 1996) of the distressed firm. Moulton and Thomas (1993), in a study of the outcomes of Chapter 11 filings, found that there were relatively few successful reorganisations. Out of a sample of 73 firms entering Chapter 11 between 1980 and 1986 that were examined in the study, only 44 were successfully reorganised and only 15 of these firms emerged from Chapter 11 with more than 50% of their pre-bankruptcy assets.

2.2 Explanatory variables utilised in prior literature

A comprehensive survey of the financial ratios employed in 47 journal articles on corporate failure is provided by Dimitras et al. (1996). If attention is restricted to ratios drawn from the financial statements of companies, five groupings are usually given prominence in the literature namely, liquidity, debt, profitability, activity, and size (Altman, 2000). Liquidity refers to the availability of cash resources to meet short-term cash requirements. Debt measures focus on the relative mix of funding provided by shareholders and lenders. Profitability considers the rate of return generated by a firm in relation to its size, as measured by sales revenue and/or asset base. Activity measures consider the operational efficiency of the firm in collecting cash, managing stocks and controlling its production or service process. Firm size provides information on both the sales revenue and asset scale of the firm and acts as a proxy metric on firm history (Levinthal, 1991). A range of individual financial ratios can represent the groupings of potential explanatory variables, each with slightly differing information content. The groupings are interconnected, as weak (or strong) financial performance in one area will impact on another. For example, a firm with a high level of debt may have lower profitability due to high interest costs. Whatever modelling methodology is applied in order to predict corporate distress, the initial problem is to select a quality set of model inputs from a wide array of possible financial ratios, and then to combine these ratios using suitable weightings in order to construct a high quality classifier.

3 The genetic algorithm

This section provides an introduction to the GA and a discussion of how it can be utilised in a NN modelling process.

3.1 Overview of the GA

Although the development of the GA dates from the 1960s, Holland (1975) first brought them to the attention of a wide audience. GAs have been applied in a variety of business settings, including the domain of finance (Bauer, 1994; Deboeck, 1994; Varetto, 1998), and several branches of management science such as inventory management (Sarker and Newton, 2002) and vehicle routing (Baker and Ayechew, 2003).

The GA can be considered as a mathematical optimisation algorithm with global search potential. The methodology is inspired by a biological metaphor, and applies a pseudo-Darwinian process to 'evolve' good solutions to combinatorial optimisation problems. The GA adopts a populational unit of analysis, wherein each member of the population corresponds to a potential solution to the problem of interest. Evolution in this population is simulated over time by means of a pseudo-natural selection process using differential-fitness selection, and pseudo-genetic (search) operators to induce variation in the population between successive generations of solutions.

A GA typically commences by randomly creating a population of possible solutions to a given problem. These solutions may be as diverse as a set of rules, a series of coefficient values, or a NN structure. Although many variants of GAs exist (Goldberg, 1989; Mitchell, 1996), a typical approach is to encode each potential solution as a binary string (0,1,0,1, ...). The quality of each solution is determined by reference to a problem-specific fitness function. This function maps the binary string to a real number representing the quality or fitness of that proposed solution. In this paper, the fitness of a binary string is determined by a three-step procedure:

i. Decode binary string into a NN structure
ii. Train the NN
iii. Determine the predictive accuracy of the resulting NN

The predictive accuracy corresponds to the fitness of the binary string. Once the initial population of solutions has been formed and evaluated, a reproductive process is applied in which better quality solutions have a higher chance of being selected for propagation of their genes (bits) into the next generation of candidate solutions. Over a series of generations, the better adapted solutions, in terms of the given fitness function, tend to flourish and the poorer solutions tend to disappear. Intuitively, the GA conducts a high-quality search of the parameter space corresponding to the chosen problem representation, in an attempt to ascertain the optimal values of parameters for that representation of the problem. The reproductive stage provides the engine for this search process, as it biases the search process towards high-quality existing solutions and uses information from these to navigate the search space.

The reproductive process is generally governed by two genetic operators, crossover and mutation (Mitchell, 1996). The crossover operator takes two members of the

population (chromosomes) and swaps component parts of each in order to create potential members of the next generation of the population. As an example, suppose two parent chromosomes are represented by the following binary strings, (0001 1101) and (1100 1100). If a single point crossover is applied after the fourth bit, the resulting child chromosome could be (0001 1100) or (1100 1101). The mutation operator causes small random changes in one or more of the genes of a child solution. In a binary representation of a chromosome, a 0 may mutate to a 1 or a 1 to a 0. Successive generations of candidate solutions are evaluated, selected for reproduction based on their differential fitness, and then subjected to a reproductive process until pre-defined stopping criteria are satisfied. In general, evolutionary algorithms including the canonical GA, can be characterised as:

$$x[t+1] = v(s(x[t])) \tag{1}$$

where $x[t]$ is the population of solutions at iteration t , $v(.)$ is the random variation operator (crossover and mutation), and $s(.)$ is the selection operator. Therefore the canonical GA can be described as a stochastic algorithm that turns one population of candidate solutions into another, using selection, crossover and mutation. Selection exploits information in the current population, concentrating interest on 'high-fitness' solutions. Crossover and mutation perturb these solutions in an attempt to uncover better solutions, and these operators can be considered as general heuristics for exploration. The GA can be formulated as a finite-dimension Markov chain, wherein each state corresponds to a configuration of the population of bit-strings. Depending on the form of genetic operators implemented in the algorithm, the transition probabilities between states will vary. In the canonical GA, the inclusion of a mutation operator implies that there are no absorbing states in the Markov process, and that all states can potentially be visited.

The computational power of the GA results from their explicit and implicit parallel processing capabilities. The explicit parallelisation stems from their maintenance of a population of potential solutions, rather than a single solution. The implicit parallel processing capabilities arise due to the *Schema Theorem* (Holland, 1975). This demonstrates that under general conditions, in the presence of differential selection, crossover and mutation, almost any compact cluster of components (bits) that provides above-average fitness will grow exponentially in the population between one generation and the next. The parallel nature of a GA search process makes it less vulnerable to local optima than traditional, local search, optimisation methods.

3.2 Combining GA and NN methodologies

Artificial neural networks have been applied to many problems in both finance and general business domains and a comprehensive bibliography of the literature of neural network applications to business can be found in Wong et al. (2000). NNs are inductive, data-driven modelling tools, whose inspiration is loosely drawn from

the workings of biological neurons. Particular strengths of NNs include parallel computation, universal approximator properties (Cybenko, 1989), robustness to noise, and adaptive learning. They consist of processing units (nodes or neurons) which are linked via interconnections (arcs or nerve fibres). A NN can be described as a generalised, semi-parametric, non-linear, regression model (Brown et al., 1998) and can map non-linear data structures without requiring an explicit, *a priori* specification of the relationship between model inputs and outputs. Trigueiros and Taffler (1996) note that neural networks are most likely to dominate traditional statistical modelling when 'strong non-linearities, and most importantly, interactions between independent variables, are present' (p. 353).

A wide variety of NN architectures and training algorithms exist (Gurney, 1997). This study utilises a fully-connected, feedforward multi-layer perceptron (MLP), trained using the backpropagation algorithm. An MLP generally consists of a collection of non-linear processing elements linked in a node-arc structure. This linkage of individual processing elements gives rise to emergent capabilities for the network, permitting complex input-output mappings. Nodes may serve as holding points for model inputs (input nodes), as holding points for the model's output (output node), or act as a processing unit, receiving signals from nodes, and in turn producing an output which is transmitted to other nodes. This signal can be modified (strengthened or weakened) when in transit along an interconnection (arc). In constructing an MLP the objective is to determine the appropriate node-arc structure and the appropriate arc weights which act to modify signals in transit between nodes. The MLP is constructed ('trained'), using pre-existing input/output data vectors. The connection weights, similar in concept to regression co-efficients, are determined by means of an iterative, local search, gradient descent algorithm. The MLP predicts the value (or classification) for a given input data vector. The error between the network's predicted outputs and desired (correct) outputs is propagated backwards through the network to adapt the values of connection weights. The general form of the three-layer MLP is as follows:

$$z_t = L\left(a_0 + \sum_{j=1}^{x} w_j L\left(\sum_{i=0}^{y} b_i w_{ij}\right)\right) \tag{2}$$

where b_i represents $input_i$ (b_0 is a bias node), w_{ij} represents the weight between input node$_i$ and hidden node$_j$, a_0 is a bias node attached to the output layer, w_j represents the weight between hidden node$_j$ and the output node, $z_{t)}$ represents the output produced by the network for input data vector (t), and L represents a non-linear squashing function. The size of the hidden layer is not known *a priori* and is determined heuristically by the modeller.

Despite the apparent dissimilarities between GA and MLP methodologies, the methods can complement each other. A practical problem in utilising MLPs is the selection of model inputs and model form, which can be very time-consuming tasks. The application of an evolutionary algorithm provides scope to automate this

step. MLPs develop a non-linear model linking pre-determined inputs and outputs whereas GAs are generally used to determine high-quality parameters for a given model structure. Hence, a combined GA/NN methodology contains both hypothesis generation and hypothesis optimisation components.
There are several ways that GAs can be combined with an MLP methodology. The first possibility is to use the GA to uncover good quality model inputs from a possibly large set of potential inputs. A second use of GAs is to evolve the structure of a network. Thirdly, a GA could be used to evolve the learning algorithm and relevant parameters for that algorithm. Therefore a GA could be used to select any or all of the following:

i. Model inputs
ii. Number of hidden-layers in the MLP
iii. Number of nodes in each hidden layer
iv. Nature of transfer functions at each node
v. Connection weights between each node

Evolving the connection weights removes restrictions on error functions, such as the requirement under standard backpropagation that the error function has a continuous derivative. It is also possible to use the GA to develop NNs which are not fully-connected, by allowing the GA to 'turn off' or prune arcs from the NN (Brabazon, 2002). In this study, we evolve the selection of model inputs, the number of nodes in the (single) hidden layer, and the nature of the activation functions at each node. Following the selection of these parameters, the resulting MLP is trained using the backpropagation algorithm. This GA/NN combination results in 'dual-level' learning, whereby the choice of inputs and the architecture of the MLP is encoded in a genetic structure which alters over the generations, representing *phylogenetic* learning, and the discovery of 'good' weights for this structure representing *epigenetic* or lifetime learning (Sipper and Sanchez, 1997).

4 Problem domain & experimental approach

This section describes both the data and the model development process adopted in this study.

4.1 Sample definition and model data

A total of 178 firms were selected judgementally (89 failed, 89 non-failed), from the Compustat Database. Firms from the financial sector were excluded on grounds of lack of comparability of their financial ratios with other firms in the sample. The criteria for selection of the failed firms were:

i. Inclusion in the Compustat database in the period 1991-2000

ii. Existence of required data for a period of three years prior to entry into Chapter 7 or Chapter 11
iii. Sales revenues must exceed $1M

The first criterion limits the study to publicly quoted, US corporations. The second criterion injects an element of bias into the sample in that companies without a three year financial history prior to entering Chapter 7 or Chapter 11 are omitted. Twenty-two potential explanatory variables are collected for each firm for the three years prior to entry into Chapter 7 or Chapter 11 of each failing firm. The date of entry into Chapter 7 or Chapter 11 was determined by examining regulatory filings. For every failing firm, a matched non-failing firm is selected. Failed and non-failed firms are matched both by industry sector and size (sales revenue three years prior to failure). The set of 89 matched firms is randomly divided into model building (64 pairs) and out-of-sample (25 pairs) datasets. The dependant variable is binary (0,1), representing either a non-failed or a failed firm.

The choice of explanatory variables is hindered by the lack of a clear theoretical framework that explains corporate failure (Argenti 1976; Trigueiros and Taffler 1996; Wilson et al., 1995; Dimitras et al., 1996). Prior to the selection of the financial ratios for inclusion in this study, a total of ten previous studies were examined (Beaver, 1966; Altman, 1968; Altman et al., 1977; Dambolena and Khoury, 1980; Ohlson, 1980; Serrano-Cina, 1996; Kahya and Theodossiou, 1996; Back et al., 1996; Sung et al., Lee, 1999; Moody's, 2000). These studies employed a total of fifty-eight distinct ratios. A subset of twenty-two of the most commonly used financial ratios in these studies was selected. The selected ratios were:

i. EBIT / Sales
ii. EBITDA / Sales
iii. EBIT / Total Assets
iv. Gross Profit / Sales
v. Net Income / Total Assets
vi. Net Income / Sales
vii. Return on Assets
viii. Return on Equity
ix. Return on Investment
x. Cash / Sales
xi. Sales / Total Assets
xii. Inventory / Cost of Goods Sold
xiii. Inventory / Working Capital
xiv. Fixed Assets / Total Assets
xv. Retained Earnings / Total Assets
xvi. Cash from Operators / Sales
xvii. Cash from Operations / Total Liabilities
xviii. Working Capital / Total Assets
xix. Quick Assets / Total Assets
xx. Total Liabilities / Total Assets
xxi. Leverage
xxii. EBIT / Interest

4.2 GA/NN model construction

In genetics, a strong distinction is drawn between the genotype and the phenotype. The former contains genetic information, whereas the later is the physical and behavioural manifestation of this information in an environment. In biological settings, the genotype contains an implicit form of memory, the results of past evolutionary learning. In this study, the phenotype corresponds to the developed neural network and its classification accuracy, and a binary genotype encodes the choice of inputs and the network structure. Initially, a population of 50 binary strings, each corresponding to a distinct NN structure and set of explanatory variables, was randomly generated. Each NN structure and set of inputs consists therefore of a hypothesis linking financial variables to corporate (non) failure. Hence, the evolutionary process maintains and evaluates a population of these hypotheses in an attempt to uncover 'ever-better' hypotheses. This population was allowed to evolve for 80 generations. Each network was trained three times using the backpropagation algorithm with different initial randomisations of the weight vectors, reducing the 'noisy fitness evaluation' problem which can emerge when a network architecture is evolved (Yao, 1999). The model building dataset was subdivided to provide both a training and (in-sample) test dataset. The networks were constructed using the training data and network fitness was assessed during the evolutionary process based on their performance on the test dataset, where networks with higher classification accuracy were considered more 'fit'. When constructing each NN structure, the genetic algorithm was able to choose between the following parameters:

i. Potential explanatory variables
ii. Form of activation function (linear or logistic)
iii. Number of hidden layer nodes

In developing the models, a limit of four hidden layer nodes and six input variables was initially imposed. The object in placing these constraints on the network structures being searched by the evolutionary process, was to conserve degrees of freedom and reduce the danger of model over-fit, with resulting poor classification accuracy out-of-sample. The out-of-sample, validation data was not used in the model development process. The evolutionary process at the end of each generation was as follows. A roulette selection process explicitly favouring more fit members of the current population, was initially applied to select members for the mating pool. 'Parents' were randomly selected from this pool, and single-point crossover was applied (crossover point is randomly selected), with probability of 0.5 to produce 'child' solutions. A mutation (exchange) operator was next applied. With a probability of 0.25, two randomly chosen bits of each child solution are exchanged. Finally, the current generation of solutions was replaced by the child solutions.

Table 1. LDA vs 'Best' individual GA/NN model after 80 generations, for one to three years prior to failure

Results	LDA		GA/NN	
	Train	*Out-of-sample*	*Train*	*Out-of-sample*
T-1	81.73%	76.00%	86.72%	80.67%
T-2	78.93%	69.33%	81.77%	72.00%
T-3	75.00%	64.67%	76.56%	66.00%

Table 2. LDA vs Average of 10 'Best' GA/NN models after 80 generations, for one to three years prior to failure

Results	LDA		GA/NN	
	Train	*Out-of-sample*	*Train*	*Out-of-sample*
T-1	81.73%	76.00%	85.76%	79.83%
T-2	78.93%	69.33%	80.36%	71.53%
T-3	75.00%	64.67%	76.28%	66.67%

5 Results and discussion

Accuracy of the developed models was assessed based on the overall classification accuracy, initially on the test and subsequently on the out-of-sample datasets. Summarised classification accuracies for an LDA model (averaged over three randomisations of the training/test data) constructed using forward stepwise selection, and the best GA/NN model in each time period (averaged over the same three randomisations of the training/test data), are provided in Table 1. Both LDA and GA/NN perform reasonably well at T-1 (one year prior to failure), correctly classifying 76.0% and 80.67% respectively, of all firms in the out-of-sample datasets. In years T-2 and T-3 (two and three years prior to failure), the classification accuracy of both sets of models falls off gradually. In each time period, the GA/NN models outperform their LDA counterparts on both the training and out-of-sample datasets. Calculation of Press's Q statistic (Hair et al., 1998) for each of the GA/NN and LDA models rejects a null hypothesis, at the 5% level, that the out-of-sample classification accuracies are not significantly better than chance.

Table 2 provides the classification accuracies for both the LDA models and the average of the best ten GA/NN models after 80 generations, in each time period, for the three separate randomisations of the dataset.

Additional metrics were collected on the positive accuracy (correct prediction of non-failure) and negative accuracy (correct prediction of failure) for each of the models. Table 3 provides these for the out-of-sample datasets for the LDA, the best individual, and the average of ten best individual GA/NN models. All reported results are averaged across the three different randomisations of the dataset. Generally, the out-of-sample classification accuracies for the GA/NN models are reason-

Table 3. Positive and negative classification accuracy (out-of-sample) of LDA models vs best individual (average of 10 best) GA/NN model(s)

Results	LDA		GA/NN	
	Positive	*Negative*	*Positive*	*Negative*
T-1	78.67%	73.33%	78.67% (79.00%)	82.67% (80.67%)
T-2	81.33%	57.33%	69.33% (71.07%)	74.66% (72.00%)
T-3	62.67%	66.67%	66.67% (69.73%)	65.33% (63.60%)

Table 4. Explanatory variables of 'Best' GA/NN models

T-1	T-2	T-3
Cash/Sales	Return on Assets	EBIT Margin
Sales/Total Assets	Net Income/Total Assets	Return on Investment
Quick Assets/Total Assets	Sales/Total Assets	Inventory/Working Capital
EBIT/Interest	Cash from Ops/Sales	Cash from Ops/Total Liabil
Total Liabil/Total Assets	Cash from Ops/Total Liabil	
	Working Capital/Total Assets	

ably symmetric between correct prediction of non-failure and correct prediction of failure.

Although the best GA/NN models (defined as producing the 'best' classification performance on the training dataset) for each time period, over each of the three runs, did not display identical form, there were similarities in their structure, and in the financial ratios selected. The evolved MLPs also fall within suggested heuristic guidelines (Trigueiros and Taffler, 1996; Bigus, 1996) regarding both the total number of weights in the model and the size of the hidden layer.

Table 4 provides details of the ratios utilised by the best single NN model evolved in each time period. In T-3 measures of return on both investment and sales, cash generation relative to the size of liabilities, and the size of inventory relative to working capital are included in the model. In T-2, emphasis is placed on return on asset measures as well as metrics on the level of cash generation relative to sales and the debt level of the firm. The chosen financial ratios for the model in T-1 primarily concentrate on the level of cash, and other liquid assets relative to the size of the firm, a measure of how well interest payments are covered by earnings, and a metric of the size of the total indebtedness of the firm relative to its asset base. It is interesting to note that cash, liquidity, and measures of indebtedness, rather than profit measures are emphasised in T-1. This is in accordance with a proposition that failing firms are running into liquidity crises in T-1, which prove fatal because of their high level of debt.

Examination of the internal structure of the 'best' evolved NN models showed that the T-1 model used three nodes (one linear and two logistic activation functions) in its hidden layer and a logistic activation function at its output node. The T-2 model had four nodes in its hidden layer (one logistic and three linear activation

Table 5. Average populational fitness (in-sample) for T-1 to T-3

Gen	T-1 *Average (Std Dev)*	T-2 *Average (Std Dev)*	T-3 *Average (Std Dev)*
2	71.96% (7.05)	72.78% (3.84)	66.74% (3.49)
10	77.47% (3.41)	77.14% (2.00)	72.51% (1.67)
20	84.37% (1.12)	77.71% (1.22)	72.19% (1.51)
30	84.93% (2.11)	76.24% (4.60)	72.54% (1.42)
40	85.59% (3.01)	77.90% (1.20)	72.81% (1.76)
50	85.49% (3.58)	77.44% (1.25)	72.23% (1.51)
60	86.29% (1.61)	77.39% (1.55)	72.84% (1.48)
70	86.16% (1.32)	77.29% (1.56)	72.32% (1.41)
80	86.30% (1.02)	77.40% (1.53)	72.22% (1.49)

functions) and a logistic activation function at its output node. The T-3 model had four nodes in its hidden layer (one logistic and three linear activation functions) and a logistic activation function at its output node. The models did not tend to use the most complex internal structure available, in particular we note that all the models included linear activation functions in their hidden layer.

In selecting the number of generations (80), we were guided by the results of preliminary experiments, which suggested that average fitness (classification accuracy on the training data) in the population typically converged to a plateau after about 50 generations. To demonstrate the convergence of classification accuracy, a tabulation of average populational fitness every 10 generations for one of the randomisations of the dataset is provided in Table 5. In each of the three time periods (T-1 to T-3), the convergence of average population in-sample fitness is noticeable by generation 50. In addition to examining the trajectory of average populational fitness, the time-to-discover the best 'individual' (based on in-sample fitness) in each run was recorded. The best individuals for T-1, T-2 and T-3 were discovered in generation 46, 6 and 27 respectively.

In evolving the NN structures, the number of hidden nodes and ratio inputs which could be included was limited to four and six respectively, in order to conserve degrees of freedom in the final models. A natural question is whether these limits had an adverse affect on the classification accuracies obtained, particularly in the case of the best model evolved for T-2 which employed six inputs. In order to examine the affect of imposing the constraint on the number of inputs, a series of additional experiments were conducted on the T-1 data, whereby the number of permissible inputs was varied from 4 to 12. Although the classification accuracy in-sample improved slightly as the number of inputs increased (to a maximum of 88.54% for the best model when 12 inputs were allowed, averaged over all three randomisations of the dataset), out-of-sample accuracy for the best model did not show any notable improvement (or disimprovement) beyond the six input case. This result is consistent with prior research which indicates that the available information content in a set of financial ratios can generally be captured in a relatively small subset of these ratios. Although a multitude of ratios can be calculated from a given

set of financial statements, many of them will have similar information content. For example, it would be expected that profitability ratios, whether measured against turnover or asset base, and whether profit was defined before or after tax, would have overlapping information content. Hence, a relatively concise set of financial ratios can convey much of the information content available in a set of financial statements.

A series of experiments were also undertaken using T-1 data, in which the allowed number of nodes in the hidden layer was increased to eight. This produced a marginal improvement for in-sample accuracy to 86.98% (averaged over all three dataset randomisations), but failed to produce any improvement in out-of-sample classification accuracies. A further series of experiments, using T-1 data, were undertaken to investigate whether the results from the evolved NN models were impacted by the limitations on the population size (50 individuals). In these experiments, the population size was increased to 100 individuals. The results obtained from the larger population size demonstrated no improvement in the classification accuracy in-sample or out-of-sample.

6 Conclusions

A series of classification models were constructed using a hybrid GA/NN methodology which combined the global search potential of an evolutionary algorithm with the non-linear modelling capabilities of an MLP. The predictive quality of these models was compared with that of a benchmark LDA model. The results of the study are consistent with the hypothesis that impending bankruptcy can be forecast with some degree of accuracy using ratio information drawn from financial statements. The results also indicate that a combination of an evolutionary algorithm and an MLP can develop models that are useful for this task, and that these models can outperform LDA models. It is notable that the developed GA/NN models did not tend to employ a complex internal structure suggesting that the financial ratios which form the pool of potential explanatory variables in this study, do not embed highly complex, non-linear interactions which are useful for the purposes of failure prediction. It seems plausible that lagged financial information provides an incomplete explanation of the occurrence (or not) of corporate bankruptcy, and because of this, complex models developed from this information set will not tend to substantially out-perform simpler linear models. It remains an open question as to how well models, constructed using a hybrid GA/NN methodology, perform on richer information sets which include both financial and non-financial information concerning firms and which also incorporate information drawn from the wider environment.

In assessing the performance of the models developed in this paper, a number of caveats must be borne in mind. The premise underlying these models (and all empirical work on corporate failure prediction) is that corporate failure is a process commencing with poor management decisions, and that the trajectory of this process

can be tracked using accounting ratios. This approach has inherent limitations. It will not forecast corporate failure which results from a sudden environmental event. It is also likely that the explanatory variables utilised are noisy. Commentators (Argenti, 1976; Smith, 1992) have noted that managers may attempt to utilise creative accounting practices to manage earnings and/or disguise signs of distress. Another limitation is that the underlying relationships between accounting ratios and corporate failure may not be stationary (Altman, 2000; Kahya and Theodossiou, 1996). Accounting standards and the economic environment faced by firms varies over time. Additionally, the firms sampled in this study are relatively large and are publicly quoted. Thus, the findings of this study may not extend to small businesses.

Despite these limitations, the high economic and social costs of corporate failure imply that there is value in building models that can indicate declines in the financial health of corporations. Given the lack of a clear theory underlying corporate failure, empirical modelling usually adopts a combinatorial approach, a task for which evolutionary algorithms are well suited.

The potential for application of evolutionary algorithms to automate the time-consuming process of creating NNs is noted. The fitness-evaluation step is generally the most computationally intensive component of the GA, particularly when this step is iterative (such as training a NN). This has discriminated against the widespread use of GA/NN hybrids, however as the cost of computer processing power declines, we expect to see greater application of these hybrids to classification problems.

Acknowledgements. The authors wish to thank Thomas Randles for his assistance. We also wish to thank the anonymous referees for their constructive and insightful suggestions.

References

Altman E (1968) Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance* **23**(1): 589–609

Altman E, Haldeman R, Narayanan P (1977) ZETA Analysis: A new model to identify bankruptcy risk of corporations. *Journal of Banking and Finance* **1**(1): 29–54

Altman E (1993) *Corporate Financial Distress and Bankruptcy*. New York: Wiley

Altman E (2000) Predicting Financial Distress of Companies: Revisiting the Z-score and Zeta models. http://www.stern.nyu.edu/ ealtman/Zscores.pdf, October 2001

Argenti J (1976) *Corporate Collapse: The Causes and Symptoms*. London: McGraw-Hill

Back B, Laitinen T, Sere K, van Wezel M (1996) Chosing bankruptcy predictors using discriminant analysis, logit analysis and genetic algorithms. *Technical Report no. 40, Turku Centre for Computer Science*, Turku School of Economics and Business Administration

Baker BM, Ayechew MA (2003) A genetic algorithm for the vehicle routing problem. *Computers & Operations Research* **30**(5): 787–800

Bauer R (1994) *Genetic Algorithms and Investment Strategies*. New York: Wiley

Beaver W (1966) Financial ratios as predictors of failure. *Journal of Accounting Research – Supplement: Empirical Research in Accounting*: 71–111

Bigus JP (1996) *Data Mining with Neural Networks*. New York: McGraw-Hill

Brabazon A (2002) Neural network design using an evolutionary algorithm. *Irish Accounting Review* **9**(1): 1–18

Brabazon A, O'Neill M, Matthews R, Ryan C (2002) Grammatical evolution and corporate failure prediction. In: Spector et al. (eds) *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO 2002)*, July 9–13, pp 1011–1019. New York: Morgan Kaufmann

Brabazon A, O'Neill M (2003) Anticipating bankruptcy reorganisation from raw financial data using grammatical evolution. In: Raidl G, Meyer JA, Middendorf M, Cagnoni S, Cardalda JJR, Corne DW, Gottlieb J, Guillot A, Hart E, Johnson CG, Marchiori E (eds) *Applications of Evolutionary Computing*, pp 368–378. Proceedings of EvoIASP 2003 (Lecture Notes in Computer Science 2611) Berlin Heidelberg New York: Springer

Brown S, Goetzmann W, Kumar A (1998) The dow theory: William Peter Hamilton's track record reconsidered. *Journal of Finance* **53**(4): 1311–1333

Cybenko G (1989). Approximation by superpositions of a sigmoidal function. *Math. Control Signal Systems* **2**(4): 303–314

Dambolena I, Khoury S (1980) Ratio stability and corporate failure. *Journal of Finance* **35**(4): 1017–1026

Deboeck G (1994) *Trading on the Edge: Neural, Genetic, and Fuzzy Systems for Chaotic Financial Markets*. New York: Wiley

Dimitras A, Zanakis S, Zopounidis C (1996) A survey of business failures with an emphasis on prediction methods and industrial applications. *European Journal of Operational Research* **90**(3): 487–513

Easterbrook F (1990) Is corporate bankruptcy efficient? *Journal of Financial Economics* **27**(2): 411–417

Fan A, Palaniswami M (2000) A new approach to corporate loan default prediction from financial statements. In: *Proceedings of Computational Finance / Forecasting Financial Markets Conference (CF/FFM-2000)*, May 2000, London

Ferris S, Jayaraman N, Makhija A (1996) The impact of chapter 11 filings on the risk and return of security holders, 1979–1989. *Advances in Financial Economics* **2**: 93–118

Fitzpatrick P (1932) *A Comparison of the Ratios of Successful Industrial Enterprises with Those of Failed Companies*. Washington: The Accountants' Publishing Company

Gentry J, Newbold P, Whitford D (1985) Classifying bankrupt firms with funds flow components. *Journal of Accounting Research* **23**(1): 146–160

Goldberg D (1989) *Genetic Algorithms in Search, Optimization and Machine Learning*. Boston, USA: Addison Wesley Longman

Gurney K (1997) *An introduction to neural networks*. London: University College London Press

Hambrick D, D'Aveni R (1988) Large corporate failures as downward spirals. *Administrative Science Quarterly* **33**(1): 1–23

Hair J, Anderson R, Tatham R, Black W (1998) *Multivariate Data Analysis*. Upper Saddle River, NJ: Prentice Hall

Holland J (1975) *Adaptation in Natural and Artificial Systems*. Ann Arbor, USA: University of Michigan Press

Horrigan J (1965) Some empirical bases of financial ratio analysis. *The Accounting Review* July: 558–568

Kahya E, Theodossiou P (1996) Predicting corporate financial distress: A time-series CUSUM methodology. *Review of Quantitative Finance and Accounting* **13**(4): 323–345

Kumar N, Krovi R, Rajagopalan B (1997) Financial decision support with hybrid genetic and neural based modelling tools. *European Journal of Operational Research* **103**(2): 339–349

Levinthal D (1991) Random walks and organisational mortality. *Administrative Science Quarterly* **36**(3): 397–420

Mitchell M (1996) *An Introduction to Genetic Algorithms*. Cambridge, MA, USA: MIT Press

Moody's (2000) RiskCalc For Private Companies: Moody's Default Model. http://riskcalc.moodysrms.com/us/research/crm/56402.pdf, July 2001

Morris R (1997) *Early Warning Indicators of Corporate Failure: A critical review of previous research and further empirical evidence*. London: Ashgate Publishing Limited

Moulton W, Thomas H (1993) Bankruptcy as a deliberate strategy: Theoretical considerations and empirical evidence. *Strategic Management Journal* **14**: 125–135

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

Russel P, Branch B, Torbey V (1999) Market valuation of bankrupt firms: is there an anomaly? *Quarterly Journal of Business and Economics* **38**(2): 55–74

Sarker R, Newton C (2002) A genetic algorithm for solving economic lot size scheduling problem. *Computers & Industrial Engineering* **42**(2–4): 189–198

Schumpeter J (1934) *The Theory of Economic Development*. Cambridge, MA: Harvard Business Press
Serrano-Cinca C (1996) Self organizing neural networks for financial diagnosis. *Decision Support Systems* **17**(3): 227–238
Shah J, Murtaza M (2000) A neural network based clustering procedure for bankruptcy prediction. *American Business Review* **18**(2): 80–86
Sipper M, Sanchez E (1997) A phylogenetic, ontogenetic, and epigenetic view of bio-inspired hardware systems. *IEEE Transactions on Evolutionary Computation* **1**(1): 83–97
Smith R, Winakor A (1935) Changes in the financial structure of unsuccessful corporations. *University of Illinois, Bureau of Business Research, Bulletin No. 51*
Smith T (1992) *Accounting for Growth*. London: Century Business
Sung T, Chang N, Lee G (1999) Dynamics of modelling in data mining: Interpretative approach to bankruptcy prediction. *Journal of Management Information Systems* **16**(1): 63–85
Tam K (1991) Neural network models and the prediction of bank bankruptcy. *Omega* **19**(5): 429–445
Trigueiros D, Taffler RJ (1996) Neural networks and empirical research in accounting. *Accounting and Business Research* **26**(4): 347–355
Varetto F (1998) Genetic algorithms in the analysis of insolvency risk. *Journal of Banking and Finance* **22**(10): 1421–1439
Wallrafen J, Protzel P, Popp H (1996) Genetically optimized neural network classifiers for bankruptcy prediction. *Proceedings of the 29th Annual Hawaii International Conference on System Sciences*, pp 419–425. IEEE Press
Wilson N, Chong K, Peel M (1995) Neural network simulation and the prediction of corporate outcomes: Some empirical findings. *International Journal of the Economics of Business* **2**(1): 31–50
Wong B, Lai V, Lam J (2000) A bibliography of neural network business applications research: 1994–1998. *Computers and Operations Research* **27**(11–12): 1045–1076
Yao X (1999) Evolving artificial neural networks. *Proceedings of the IEEE* **87**(9): 1423–1447
Zhang G, Hu M, Patuwo B, Indro D (1999) Artificial neural networks in bankruptcy prediction: General framework and cross-validation analysis. *European Journal of Operational Research* **116**: 16–32
Zmijewski M (1984) Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research-Supplement* **22**: 59–82
Zopounidis C, Dimitras A (1998) *Multicriteria Decision Aid Methods for the Prediction of Business Failure*. Dordrecht: Kluwer Academic Publishers
Zopounidis C, Slowinski R, Doumpos M, Dimitras A, Susmaga R (1999) Business failure prediction using rough sets: A comparision with multivariate analysis techniques. *Fuzzy Economic Review* **4**(1): 3–33