Developing Financial Distress Prediction Models Using Cutting Edge Recursive Partitioning Techniques: A Study of Australian Mining Performance

Review of Integrative Business & Economics Research

Nikita Shah
School of Business, Bond University, Gold Coast, Queensland, Australia
Nikita.Shah@student.bond.edu.au

ABSTRACT

The purpose of this paper is to analyze financial ratios of Australian Mining companies in order to specify and quantify the variables which are effective indicators and predictors of corporate distress. Using financial ratios, the paper explores the quantifiable characteristics of potential bankrupts using cutting edge Recursive Partitioning techniques like Discriminant Analysis, Decision Tree Method, Artificial Neural Network and Hybrid Method, and constructs financial distress prediction models. Australian mining industry is considered for the experiment data set and a sample of 351 healthy firms and 44 distressed firms are studied over a 12 month period from 2012 to 2013 as our experimental targets.

The recursive partitioning, Decision and Hybrid Intelligence methods are found to have higher classification power and obtain higher accuracy than the other methods. It proves that this model for prediction of corporate financial crisis is a good solution and can also help investors to make the correct investment decisions.

Keywords - bankruptcy prediction; financial distress; insolvency; decision trees; survival analysis; artificial neural network

1. INTRODUCTION

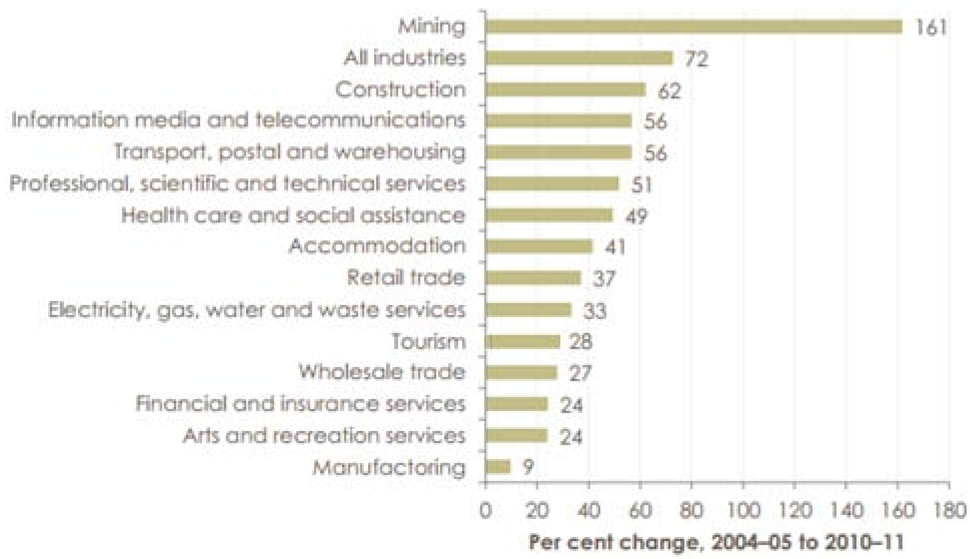
Australia is in the midst of a mining boom. This industry supports the Australia, provides thousands of jobs and is considered the darling of the economy. The mining boom helped to cushion Australia from the global recession in 2008-09, with output and income in many other developed countries still at or below previous peaks (Shann, 2012).

The rise in mining sector employment levels recorded over the 2007-2012 period is the second largest of any industry, equivalent to 130,900 workers. Unsurprisingly, the mining industry during this period recorded the strongest growth nationally in terms of percentage, increasing by 94.3 percent to reach 269,700 workers, an historic high according to Australia Bureau of Statistics (ABS) data. To date, the mining boom has delivered strong growth in revenues and investment dollars to those regions and industries aligned with mining, such as construction.

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

Australia is experiencing a period of unprecedented investment activity in mining projects, and due to the complex nature of these projects, policy makers require assistance to understand its size and impact on the Australian economy. Moreover, the recent surge in capital expenditure ($40,000M in 2013) due to the mining industry has also affected other industries, particularly those in downstream production process such as the manufacturing industry; as a result of which, there is a growing need for the Australian community to monitor these investments that affect other projects, get a complete picture of its current investment activity as well as the *likelihood* of planned investment activity being realized in the future. (Australian Bureau of Statistics, 2010) (Australian Bureau of Statistics, 2012)

Figure 2: Growth in Gross Value Added, by industry (including tourism), 2004–05 to 2010–11



Sources: ABS Cat. No. 5206.0, Australian National Accounts
ABS Cat. No. 5249.0, Australian National Accounts, Tourism Satellite Accounts 2010–11

Using financial ratios of current companies and other multivariate techniques, this paper attempts to construct models to predict the success and failure of upcoming companies in the Australian mining industry.

Benefits from accurate Business Failure Prediction (BFP) are many. Some of them include (Kumar & Chaturvedi, 2010)-

- Banks, investment banks, credit unions and other financial investment institutions that invest in these industries can be careful to lend to business that can potentially fail. Particularly for the Mining Industry, that require large amounts of capital to establish a world-class and internationally competitive mine. More so because, as cited earlier, investments and expectations from mining industry has also affected other industries in Australia and long term returns from this industry requires banks to monitor and analyze the potential returns from their projects.
- Businesses could establish long term relationships with other businesses such as suppliers and service providers including legal, accounting, architectural, engineering, technical and research services, that will not fail in future and thus create longevity and viability of their business relationships.

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

- BFP models could also improve investor's confidence in investment, lending and development of possible relationships in these growing industry.
- Companies themselves can understand their financial position and act upon the probable outcomes.

Therefore, a model of predicting corporate distress would serve to reduce such losses by providing a pre-warning to stakeholders of firms. Such a model could provide an early warning signal of probable distress which could help both management and investors to take preventive actions and shorten the length of time whereby losses are incurred (Jaikengkit, 2004). Hence, an accurate prediction of firms’ financial distress has become an important issue in finance (Cybinski, 2001).

1.1 Research Objectives

This paper focuses on identifying financial distress of growing mining enterprises in Australia within the theory framework of financial distress prediction. This study includes the following research objectives:

- To identify the nature of causes of financial distress using financial ratios.
- To create a model that uses financial and non-financial factors to predict financial distress of the companies.

1.2 Data Source

All the data of growth enterprises used in the present study was derived from the official portal of Morningstar on 20th March 2014. Morningstar provides data on approximately 433,000 investment offerings, including stocks, mutual funds, and similar vehicles, along with real-time global market data on nearly 10 million equities, indexes, futures, options, commodities, and precious metals, in addition to foreign exchange and Treasury markets. Morningstar also offers investment management services through its registered investment advisor subsidiaries and has approximately $166 billion in assets under advisement and management as of June 30, 2013. The data from Morningstar has been widely used in prior researches.

2. LITERATURE REVIEW

The literature on corporate financial distress is extensive. The generally recognized pioneers in this area are Beaver (1966) and Altman (1968). Since then, there has been great work of research and advancements in this field by many other expert researchers, including Muller, Steyn-Bruwer, & Hamman (2009) and Kumar & Gepp (2010)

2.1 Multivariate Discriminant Analysis

Beaver (1966), an important paper in accounting research which employs statistical analysis to a similar matched sample, cites the paper. Beaver’s (1966) study was one of the first studies using financial ratios, which were based on the data from financial statements, to predict failure. It was designed to provide an empirical verification of the predictive ability of financial statements.

After Beaver’s study (1966), instead of ratio analysis, using rigorous statistical techniques to assess the performance of the enterprises had gradually become more popular. Altman (1968) attempted to create a link between the rigorous statistical techniques and the traditional ratio analysis.

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

In order to solve the problems of Beaver's study, Altman used a Multivariate Discriminant Analysis, MDA, as the statistical technique. This is mainly used to predicting or classifying issues with qualitative dependent variables. In this model, the "Z-score" indicator provided a forecast of whether the company would enter into financial distress within a two-year period.

Altman's pioneering work (1968) used multivariate discriminant analysis with a set of five financial ratios for distinguishing failed firms from non-failed firms. The multivariate discriminant analysis is based on the development of a linear equation. This equation provides an overall score used to predict whether the subject lies in either of the groups which should be no less than two (Muller, Steyn-Bruwer, & Hamman, 2009). In addition, this resulting equation firstly combines all the variables (ratios) and weighs the variables in such a way as to maximize its ability to discriminate between different groups (Muller, Steyn-Bruwer, & Hamman, 2009). The score was calculated based on the following general discriminant function:

$Z = a1x1 + a2x2 +.. + anxn + c$

Where, Z was the score,

x1 were the independent variables,

a1 and c were the estimated parameters.

Therefore, the discriminant function of this equation could transform individual variables' values with their corresponding coefficients to a single discriminant score (Altman, 1968). In terms of predicting financial distress, enterprises are classified as 'distressed' or 'non-distressed' based on whether the overall score of the discriminant function is less than or greater than the predetermined cut-off value. Altman's MDA used 2 cut off scores (1.8 and 2.7) to classify businesses into three categories as shown in the table below:

Z-score lookup	**Prediction**
Z > 2.7	Success
Z < 1.8	Failure
1.8 ≤ Z ≥ 2.7	Inconclusive

Although the probability of failure and success is not explicit output of this model, a relative measure of probability can be obtained by calculating the difference between Z-Scores and cut-off values. For instance, a business with a score of 1.0 is more likely to fail than a business with a Z-Score of 1.7.

Altman's DA model outperformed Beaver's univariate model for one year prediction intervals; however, Altman's model was not as accurate for longer predictions. It is also interesting to note that Martin (1977) concluded that using DA with more than two classification groups (MDA) was more accurate than only using two classification groups (Linear Discriminate Analysis).

2.3 Logistic analysis

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

In 1980, Ohlson carried out a research into the probabilistic prediction of financial distress using logistic analysis. The application of logistic analysis requires four steps: (1) calculate a series of financial ratios; (2) multiply each ratio with its corresponding coefficient; (3) sum the result of each coefficient to form a new variable y and (4) calculate the probability of financial distress for a company as $1/(1+e^{-y})$. Here the independent variables with a negative coefficient increase the probability of financial distress due to the fact that they reduce e^{-y} toward zero, with the result that the financial distress (probability function) approaches 100 per cent or 1. Likewise, the independent variables with a positive coefficient decrease the probability of financial distress (Ohlson, 1980).

Ohlson (1980) is the seminal study for applying LA to BFP. He produces three separate LA models to predict failure for one, two and three years in advance. Fourteen ratios were used as predictors, consisting of standard accounting ratios, dummy variables based comparisons of balance sheet figures, and a variable representing the change in net income over the last year. The empirical results for his model were disappointing, but he showed that LA is mode statistically valid and easier to interpret than DA.

This was supported by Collins and Green (1982) who compared forecasting results by using a logistic model, a discriminate analysis and a linear probability model, respectively. Their results show that the logistic model performs better. Hall (1994) set up a logistic model with non-financial variables and the model could distinguish distressed firms from non-distressed firms with as high as 95 per cent of accuracy. In addition, the subsequent studies on LA have shown that it is usually slightly empirically superior to DA in both classification and prediction accuracy (Laitinen & Kankaanpaa, 1999).

2.4 Artificial Neuron Network

Odom and Sharda (1990) used the same financial ratios employed by Altman (1968) and applied ANN to a sample of 65 failed and 64 non-failed firms.

In their study, three layer feed forward networks are used and the results are compared to those of MDA. Using different ratios of bankrupt firms to non-bankrupt firms in training sample, they test the effects of different level on the predictive capability of neural networks and DA. Their model correctly identified all failed and non-failed firms in the training sample, compared to 86.8% accuracy by MDA. Regarding the performance with holdout samples, ANN had an accuracy rate of 77% or higher, whereas MDA could hit the target only between 59% and 70%. ANNs were found to be more accurate and robust in both training and test results.

Following Odom and Shrada (1990), a number of studies further investigated the use of ANN in BFPs. For instance Salchenberger et al (1992) presented an ANN approach to bankruptcy of savings and loan institutions. ANNs were found to be as good as or better than Logit models across three different lead times of 6, 12 and 18 months. To test the sensitivity of networks to different cutoff values in classification decisions, they compare the results of threshold of 0.5 and 0.2. The information is useful when one expects Type I or Type II error. There are other vast majority of ANNs studies with positive results that used forward feeding back-propagation NNs, such as Coleman et al. (1991), Coats and Fant (1992), Tam and Kiang (1992), and Fletcher & Goss (1993).

Most researchers in BFP using NNs focus on the relative performance of NNs over other statistical techniques. While empirical studies show that ANNs produce better results for many classification or prediction problems, they are not always uniformly superior. Bell et al (1990) report disappointing findings in predicting commercial banks failures. Boritz and Kennedy (1995) have found that ANNs

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

perform reasonably well in predicting failures but their performance is not any systematic way superior to conventional statistical techniques such as LA and DA.

Previous studies have often included an equally matched sample of firms to ensure the robustness of the models and the ability of these models to discriminate firms that thrive from those that fail. This study will attempt to construct prediction models employing some multivariate techniques specifically for the mining industry in Australia and compare the prediction accuracies of these models by testing them on upcoming projects.

2.5 Decision Trees

Frydman, Altman, & Kao (1985) first introduced recursive partitioning decision rule for nonparametric classification. As suggested by Pompe and Feelders (1997), 'the basic idea of recursive partitioning is to fit a tree to the training sample by successively splitting it into increasingly homogeneous subsets until the leaf nodes contain only cases from a single class or some other reasonable stopping criterion applies'.

These have no distribution assumptions to violate and thus there is no need to consider transforming variables. The application of DT requires the right choice of algorithm that influences the accuracy of the final DT. The advantage of a DT is that there are no assumptions, except the common assumptions in the BFP such as - the successful firms are discrete, non overlapping and distinctly identifiable. However, certain disadvantages of the model was that it requires prior probabilities of successful and failed businesses as inputs (Gepp, Kumar, & Bhattacharya, 2010)

Kumar, Gepp, & Bhattachariya (2010) provided empirical evidence to support the claim that less complex and more parsimonious models are better predictors than more complex models. There was further evidence to suggest that the DT techniques are superior classifiers and predictors of business failures.

2.6 Hybrids

Combinations of statistical techniques are frequently accompanied by artificial intelligence systems for better model performance in practice. For example, McKee & Lensberg, (2002), present a hybrid financial analysis model combining genetic programming and rough sets. The authors use a sample of 291 US firms referring to the period from 1991 to 1997 and they select 11 variables to describe the cases. They conclude that the hybrid model reaches an accuracy of 80% on the validation set, while the simple rough set performs considerably lower on the same data (67%). Ahn, Cho & Kim, (2000) work on the combination of rough sets and neural networks for business failure prediction. They also use Korean data referring to the period between 1994 and 1997 and they compare their results to different standard neural network techniques. Accuracy exceeds 80% in some cases.

3. METHODOLOGIES

Australian mining industry is considered for the experiment data set and a sample of 351 healthy firms and 44 distressed firms are studied over a 12 month period from 2012 to 2013 as our experimental targets. The variable 'status' is a dichotomous variable coded 1 if the company was a success and 0 if financially distressed.

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

The software package IBM SPSS was used for experiments. All experiment results in this report are from the implementation of IBM SPSS 20.0. The methods used are as below:

3.1 Direct Logistic Regression

Regression methods have become an integral component of any data analysis concerned with describing the relationship between a response variable and one or more explanatory variables. Over the last decade, logistic regression model has become the standard method of analysis of this situation in the field of forensic analysis. (Hosmer & Lemeshow, 2000). Martin (1977) used the logit model for bank failure prediction. Subsequently, Ohlson (1980) also used the logit model to predict business failure with a relatively unbiased sampling procedure. Zmijewski (1983) examined the "choice-base" sample bias and "sample selection" bias typically faced by financial distress researchers.

Logistic regression, which includes binary and multinomial logistic regressions, is used for prediction of the categorical outcomes. The multinomial logistic regressions are used for two and more categories of outcomes respectively. In most cases, the dependent variable for logistic regression only assumes two discrete values and the binary logistic regression is used (Anderson, Sweeney, & Williams, 2008). For example, the dependent variable can be code as $y = 1$ if an event occurs, and $y = 0$ if this event does not occur. In terms of independent variables, they can be either continuous or categorical or a mixture of both in one model (Pallant, 2007, p.166). Using logistic regression, researchers can estimate the probability of occurrence of the event.

Logistic regression does not make any assumptions of normality, linearity, and homogeneity of variance for the independent variables. Because it does not impose these requirements, it is preferred to discriminant analysis when the data does not satisfy these assumptions.

According to Pallant (2007), there are three assumptions underpinning the use of logistic regression. The first assumption concerns the number of cases in the sample and the number of independent variables included in the model. If there is a small sample with a large number of independent variables, the research might have problems with the analysis. It becomes a real problem when the categorical independent variables have limited cases in each category.

The second assumption refers to checking for inter-correlations among independent variables or namely, multi-co linearity. Ideally, independent variables have to be strongly related to dependent variables but not strongly related to each other. Therefore, the highly inter-correlating variable has to be removed (Pallant, 2007, p.167).

Finally, the third assumption is about checking for the presence of outliers, because outliers can influence the results of logistic regression. If there are some cases that are not well explained by the model, checking the outlying cases would become a particularly important step.

As for the sample size, the minimum number of cases per independent variable is 10, using a guideline provided by Hosmer & Lemeshow (2000). For our case study, we use a ratio of 20:1. The method we use is simultaneous logistic regression, in which all independents are included at the same time.

The response variable is classified into 0 and 1 dummy variables, depending on the experimental unit, for instance in this case, failure or success of the company.

LR Model for the binary dependant variable can be represented as:

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

$$\text{Prob (event)} = \frac{e^z}{1+e^z}$$

where, the link function indicates the cumulative standard logistic probability distribution function. This report examines how the financial distress of mining companies is affected by various financial ratios. Z is the linear combination:

$$z = B_0+B_1X_1 + B_2X_2 + ... + B_kX_k$$

Where,

$X_{1,}$ X_2 .... X_k are continuous categorical independent variables, in this case the values of financial ratios

B_0 is the Coefficient of constant

B_1 to B_k are the Coefficients of independent variables in model

K stands for the number of independent variables

Once each business has an associated probability of failure, n cut - off (or critical) values can be established to separate the business into n + 1 groups. However, it is usually one cut - off that separates businesses into failure and success groups. The cut-offs can be changed to cater the different misclassification costs: as Type I error is more serious, the chosen critical value is < 0.5. The probabilities closer to 0.5 are more sensitive to changes in the independent variables. (Kumar & Chaturvedi, 2010)

In this study, if the Prob(event) > threshold (i.e, 0.5), the case is classified as successful firm, vice versa, when Prob(event) < threshold, the case is classified as a bankrupt firm.

We performed a Logistic analysis selecting 'Forced Entry Method'. In this method, all variables are entered in a single step. All independent variables were continuous and no categorical variables were used. The probabilities and group memberships are requested using standardized residuals to analyze the classifications within the range 0 and 1, failure and success respectively.

3.2 Linear Discriminate Analysis

The discriminant analysis is a multivariate technique that allows to differentiate between two or more groups of objects with respect to several variables simultaneously. DA is used to classify an observation into one of several a priori groupings dependent upon the observation's individual characteristics (Malhotra, 2007). This technique also helps to determine the most parsimonious way to separate groups and discard variables which are little related to group distinctions.

In recent years, this technique has become increasingly popular in the practical business world. DA helps us examine whether there is significant differences that exist among the groups, in terms of predictor variables and if so, which predictor variables contribute to most of the intergroup differences. DA has been developed in recent years as an alternative to Logistic Regression.

Under usual assumptions of regression analysis such as:

(1) The assumption of full rank,

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

(2) equality of variance-co-variance matrices, and

(3) normal distribution, the MDA model is a linear combination of the discriminatory variables of the following form:

$D = B_0+B_1X_1 + B_2X_2 + ... + B_kX_k$

Where,

D is the Discriminant score

$X_{1,}\ X_2 .... X_k$ are the values of the independent variables,

B_0 is the Coefficient of constant,

B_1 to B_k are the Discriminant Coefficients estimated from the data,

K stands for the number of independent variables.

One way to assess the contribution of variables in the discriminant function is by examining the magnitude and sign of the standardized coefficients. Another way is to examine the correlation between the values of the function and values of the variables. The higher the correlation, the higher the contribution of the variable to the discriminant function.

In DA, the three most commonly used algorithms for variable selection available in SPSS Package are forward entry, stepwise selection and backward elimination.

We performed a discriminant analysis selecting 'Enter independents together'. In this method all variables in a block are entered in a single step. All independent variables were continuous and no categorical variables were used. The descriptive Univariate Anova's Box's M and unstandardized function coefficients are requested to analyze the classifications within the range 0 and 1, failure and success respectively.

3.3 Principal Component Analysis (PCA) and Factor Analysis (FA)

Principal Component Analysis (PCA) and Factor Analysis (FA) are two set of techniques similar in many ways and are often used interchangeably by researchers. Both are designed for data reduction and summarization and not to discriminate one group from another. They can be used to reduce a large number of related independent variables to a smaller more manageable clump of related variables or groups, prior to using them in another analysis such as multiple regression or MANOVA. (Pallant, 2007)

PCA involves attempting to express a system with p components in a linear combination of k principal components, where $k<p$ and the k components are representative of the system. That is, the goal is to explain the variance-covariance matrix with linear combination of variables that is less than the number of original variables. The analysis often extra relationships between components that lead to new interpretations. (Kumar & Chaturvedi, 2010)

There are three assumptions underlying the application of factor analysis.

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

- Sample Size - Although there is no agreement in the literature concerning how large the sample should be, the general recommendation for the data size is: the larger, the better. Generally, the overall sample size of no less than 100 is acceptable and a minimum of five cases for each of the variables is required for factor analysis (Coak, 2005). In our study, a sample of 351 healthy firms and 44 distressed firms are used over a 12 month period from 2012 to 2013 as our experimental targets.
- Factor analysis is sensitive to outlying cases or outliers. These cases should be either removed from the data set or recoded to a less extreme value. There were no outliers found in our data set.
- Factorability of the correlation matrix - the correlation matrix of all variables should have at least some correlations, with r being no less than 0.3. Moreover, the Kaiser-Meyer-Olkin value ranges from 0 to 1 and should be no less than 0.5 (Child, 2006). Bartlett's test of Sphericity should have a p value less than 0.05 (Pallant, 2007)

If the variables are standardized, the factor model is represented as:

$X_i = A_{i1}F_1 + A_{i2}F_2 + A_{i3}F_3 + \ldots\ldots + A_{im}F_m + V_iU_i$

Where,

X_i is the ith standardized variable

A_{i1j} is standardized multiple regression coefficient of variable i on common factor j

F is common factor

V_i is the standardized regression coefficient variable of i on unique factor i

U_i is the unique factor for variable i

m is the number of common factor.

The unique factors are correlated with each other and with the common factors. The common factors themselves can be expressed as linear combinations of the observed variables.

$F_{i\,=}\,W_{i1} X_1 + W_{i2} X_2 + W_{i3} X_3 + \ldots. + W_{ik} X_k$

Where,

F_i is the estimated observed factor

W_{i1} is the weight or factor score coefficient

K is the number of variables.

Three techniques can be used in the SPSS package to decide what factors to retain -

- Kaiser's Decision - Eigenvalue represents the amount of total variance explained by that factor. If Eigenvalue > 1.0, retain for further investigation.
- Scree Test - We plot each of the Eigenvalue and find the point at which the shape of the curve changes and becomes horizontal. We retain the variables until the curve as its cumulative contribution is the most.
- Parallel Analysis - Compare the eignevalue with those of a set of randomly collected data.

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

We use all three techniques in our experiment for interpretations.

The next step is to determine the rotation of the variables. Rotation presents a pattern of loadings in a manner that is easier to interpret. This can result in either uncorrelated or correlated solutions.

Orthogonal (uncorrelated) solution is easy to interpret and usually uses Varimax Technique that minimizes the number of variables that have high loadings on each factor. We use the varimax technique to generate our results. Whereas the Oblique (correlated) solution is difficult to interpret and usually uses Direct Oblimin that provides that degree of correlation between the factors.

3.4 Artificial Neuron Network

Neural networks are the preferred tool for many predictive data mining applications because of their power, flexibility, and ease of use. Neural computing is a computer system that consists of a network of interconnected units called artificial neurons (AN). AN are organized in layers inside the network. The first layer is the input layer, and the last is the output layer. Hidden layers exist between the input and output layers, and there can be several hidden layers for complex applications.

The Topology of BPF involves connected layers of neurons:

- Input layer of financial ratios
- Hidden layer of inter connected neurons
- Outer layer of just one Boolean fail or success neuron

As per Kumar & Chaturvedi (2010) Using Algorithms in these layers, the computing system recognizes patterns and can tell us what data is important to the patterns and how to use the patterns to interpret new data, including often how to predict future patterns. Neural computing is more adaptive to the real world situation because it is not subject to distribution constraints. Prior to its use, ANN is trained like a human brain by processing data with known results. A common ANN with back-propagation is then rewarded for a correct classification by increasing the weight of the neuron connections that led to the correct classification. Similarly, the weights of the neuron classifications that lead to an incorrect classification are reduced. Eventually, a suitable weighted interconnection of neurons is established and can be used for making predictions. This is known as supervised learning, but there are also different training techniques such as unsupervised or graded learning.

The advantage of ANN is that they do not have the same restrictive assumptions as traditional statistical methods, such as normality, linearity and independence among input variables. Due to their flexibility, ANN can also deal with outliers, missing data multicolinearity better than traditional techniques. The major criticism is that ANN is a black box approach. Although they output a continuous score that can be compared with cut-off values to generate failure/success predictions, the internal logic is hidden from the users. (Kumar, Gepp, & Bhattachariya, 2010)

Odom & Sharda (1990) used the same financial ratios employed by Altman (1968) and applied ANN to their studies. A three-layer neural network was created with five hidden nodes. Their model correctly identified all failed and non-failed firms in the training sample, compared to 86.8% accuracy by MDA. There were other vast majority of studies with positive results such as Coats & Fant (1992) and Coleman, Graettinger & Lawrence (1991)

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

SPSS Package includes two methods for NN algorithms, Multilayer perceptron and radical basis function. In this study, all independent features are fed as 'Standardized' covariates to perform the multilayer perceptron.

3.5 Decision Trees

Decision Tree (DT) Techniques generate a set of tree based classification rules use to construct a DT also called as a classification tree.

Different algorithms can be used for building decision trees, such as classification and regression trees (CART), chi squared automatic interaction detection (CHAID), Quest, C4.5, C5.0, or entropy reduction algorithm (Ravi Kumar & Ravi, 2007). Decision trees have been popularly used for classification problems, because their rules are easy to understand and communicate (Cho, Hong & Ha, 2010). However, they may not be as robust to cyclical changes as classic LDA (Bardos & Rasson, 2001).

DTs assign data to predefined classification groups: in the case of BFP, a DT usually assigns each business to a successful or failing group. In general, DTs are binary trees, which consists of root nodes, non leaf nodes and leaf nodes connected by ranches, whereby each non leaf node ahs two branches leading to two distinct nodes. (Kumar, Gepp, & Bhattachariya, 2010)

In this study, each root node represent classification groups - fail or success and non leaf nodes each contain a splitting or decision rule. Thus the tree is built by recursive process of splitting the data when moving from higher to lower level of the tree.

The splitting rule comprises of an expression (usually the financial ratio) that is evaluated for each case (business) and compared to cut-off value. For instance, the splitting rule may classify a business into Left Sub-tree if current ratio is < 2.5 or Right Sub-tree if current ratio is > 2.5.

Splitting rules are usually univariate but the same variable can be used in zero, one or many splitting rule. (Kumar & Chaturvedi, 2010)

Similar to supervised learning with ANNs, DTs build algorithms and are used to manage the creation of DTs. There are two main tasks that a DT building algorithm performs:

- Choosing the best splitting rule at each non-leaf ndoe that sidcriminates between successful and failing firms
- Managing complexity of the DT (number of nodes), which includes the decision of when to stop the process and use current DT as the best DT.

The major advantage of DT is that they are non-parametric. This means that DTs make no assumptions of underlying data and consequently there is no violation of distribution assumptions and there is no need to consider transforming variables. DTs can also handle missing values and qualitative data, as well as easily represented in a user friendly graphical format. (Kumar, Gepp, & Bhattachariya, 2010)

3.6 Hybrids

Hybrid Models combine several individual techniques to maximize their advantages while minimizing the combined model’s weaknesses. The general idea is that the gains achieved by precision and certainty, as in more conventional methods (i.e. DA, Logit, ANN, etc.), are not justified by their costs (Ravi Kumar & Ravi, 2007). This technique has recently become very popular among researchers and practitioners and is

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

seen as one of the latest trend in corporate prediction modeling (Demyanyk & Hasan, 2010). There are many different possibilities of combinations and associations. Combinations of techniques are not exclusively reserved to solely artificial intelligent techniques, which are often found complementary (Ravi Kumar & Ravi, 2007). Statistical techniques, operations research, as well as other techniques found useful in predicting bankruptcies can be combined to develop the ultimate model.

As some of the researchers mentioned earlier, the ANN produces better results but not relatively higher than the conventional statistical methods. The idea of hybrid system is to analyze the advantages and weak points of the classifiers. We use two hybrid methods for our experiments -

Hybrid I is a combination of ANN and the results obtained from the LR. The probability (0,1) from the Logistic Regression Model is used as second order independent variable combining with the original features. The features are then fed to Neuron classifiers.

The process can be described using following steps.

Step I: Use all independent variables for logistic regression classification to extract and save the probabilities of group membership.

Step 2: Feed all independent variables including the saved probabilities of group membership from LR into 'Standardized' Neural Network for further classifications.

Similarly, Hybrid II uses a combination of ANN and the results obtained from the DA. The probability from the Discriminate Function Analysis is used as an independent variable.

4. RESULTS AND INTERPRETATIONS

Tests were performed on 20 independent variable - Asset Turnover, Current Ratio, Depreciation/PPE, Financial Leverage, Gross Debt/CF, Gross Gearing, Invested Capital Turnover, LT Asset Turnover, Net Gearing, Net Interest Cover, PER, PPE Turnover, Price/Book Value, Price/Gross Cash Flow, Quick Ratio, Receivables/Op. Rev., ROA, ROE, ROIC and Working Cap Turnover to develop model for BFP.

4.1 Logistic Regression Model

Direct Logistic Regression was performed to assess the impact of a number of factors on the likely hood that the companies would be a success. The model contained Twenty independent variables of which only six were found to be significant in predicting financial distress of the mining companies. The Variables are - Depreciation/PPE, Gross Debt/CF, PER, Price/Gross Cash Flow, ROA, ROIC.

The full model containing all predictors was statistically significant, x^2 (20, N = 395) = 74.18, p<.001, indicating that the model was able to distinguish between those companies successful and those that failed. As presented in Appendix B, Table 1 indicates that the model was explained between 17.1% (Cox and Snell R Square) and 34.1% (Nagelkerke R Squared) of the variance in the bankruptcy status and correctly classifies 91.9% of the cases.

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

As shown in table 2 of Appendix B, two of these six significant independent variables made a unique statistically significant contribution to the model - Price Gross Cash Flow and ROA, where the significance was < .05.

The strongest predictor of a success of a company was Price Gross Cash Flow, recording an odds ratio of 1.068 with a positive B value. This indicated that for companies that made 1% higher Price Gross Cash Flow, the odds for them becoming successful was 1.066 times higher than companies that did not, all factors being equal.

The Odds Ratio of .971 for Price Earnings Ratio (PER) was less that 1 with a -ve B, indicating that for every additional 1% of PER, companies were .971 times less likely to become bankrupt, controlling for other factors in the model.

The derived estimated equation model for distress prediction is:

Z = 4.182 + .052 (Depreciation/PPE) - 0.024 (Gross Debt/CF) -0.029 (PER) + 0.066 (Price/Gross Cash Flow) + 0.053 (ROA) - 0.019 (ROIC)

As per Table 2, the model also indicated that the true positives were 348 of the 351 cases reported. The positive predictive value is 99.1%, indicating that of the companies predicted to be successful our model accurately picked 99.1% of them.

The true negatives were 15 of 44 companies. The negative predicted value is 34.1%, indicating that of the companies predicted to be a failure our model accurately picked 43.1% of them. The overall predictability accuracy of the logistic regression model was 91.9%. In our study (Table 2), none of the independent variables had a standard error (S.E) larger than 2.0.

Using the above model, we can predict the financial distress of the upcoming companies in mining sector. Using the publicly available data such as Depreciation/PPE, Gross Debt/CF, PER, Price/GCF, ROA and ROIC, from the company financial reports, we can determine the probability (z) of the company. If this probability is higher than the threshold i.e 0.5, we can be 95% confident that the firm will be successful. With a Z<0.5, we can be 95% confident that the firm is financially distressed. With results as low as 0.01 - 0.03, it is safe to assume that the company will be bankrupt.

4.2 Discriminant Analysis

The purpose of this study is to discriminate successful and bankrupt companies based on the 20 independent variables . As per the Tests of Equality of Group Means (Table 4 of Appendix C) using stepwise regression, nine of twenty independent variables made unique statistically significant contribution to the model - Asset Turnover, Current Ratio, Invested Capital Turnover, LT Asset Turnover, Net Gearing, PPE Turnover, Price/Book Value, Price/Gross Cash Flow and Quick Ratio.

Amongst which, Quick Ratio and Net Gearing were most important independent variables to discriminate the functions.

In table 5 of Appendix C, Eigenvalue.170 accounts for 100% of the explained variance. The Canonical Correlation associated with this function is .381 and therefore the square of this correlation (.381^2 = 0.13), indicated the 14.51% of the variance in dependent variable - successful or bankrupt, is explained by

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

this model. The Wilks' Lambda.855 > .05, (in Table 6) shows that the functions are statistically significant.

Standardized Canonical Discriminant Function Coefficients (Table 7) indicated that Quick Ratio and Asset Turnover Ratio made a fairly strong contribution to classifying companies as successful or bankrupt.

Using Fisher's Linear Discriminant Function, two equations were derived:

Successful (1)

Y = -1.096 + 0.109(Quick Ratio) - 0.019(Price/Gross Cash Flow) + 0.190(Price/Book Value) + 1.577(Asset Turnover)

Bankrupt (0)

Y = -3.099 + 0.278(Quick Ratio) - 0.056(Price/Gross Cash Flow) + 0.355(Price/Book Value) + 4.068Asset Turnover)

As per Table 9 of Appendix C, the Linear Discriminant Function reported that 86.89%, i.e 305 of 351 cases were correctly classified as successful firms. As a further check of efficiency of the Linear Discriminant Function, cross validated error rate was calculated and the same results were reported as that in the original. In total, 82.8% of original grouped cases correctly classified.

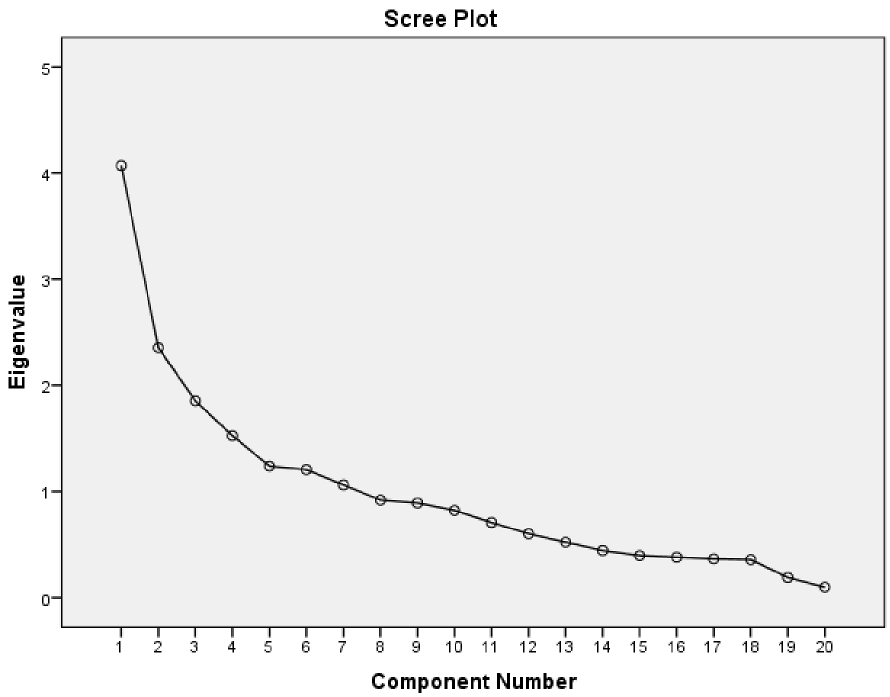
4.3 Principal Component Analysis (PCA) and Factor Analysis (FA)

The present study used factor extraction to determine the smallest number of financial factors that could best represent the interrelations among a group of financial ratios. The most commonly used extraction technique (principal components) was then used to extract the underlying financial factors.

The 10 independent financial ratios were subject to Principal Component Analysis. Prior performing PCA, the suitability of data for factor analysis was assessed. Inspection of the correlation matrix revealed the presence of many coefficients of .3 and above. As presented in Appendix D, Table 10, the KMO value was .701, exceeding the recommended value of .6 and Bartlett's Test of Sphericity reached statistical significance, supporting the factorability of correlation matrix.

As per (Table 11), PCA revealed the presence of seven components with eigenvalues exceeding 1, explaining 15.96%, 12.03%, 9.93%, 9.17%, 6.78%, 6.45% and 6.19% of the variance respectively. An inspection of the Scree plot revealed a clear break after the second component. These seven components explained a total of 66.54% of variance.

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)



Based on the Rotated Component Matrix, we can see the pattern of factor loadings.

Factor 1 comprised of ratios LT Asset Turnover, Invested Capital Turnover, Asset Turnover, PPE Turnover. These factor can be termed as **Turnover Ratios** as they measure the efficient use of assets and how well they produce revenue during the corresponding period.

Factor 2 comprised of ROA, ROE and ROIC and hence can be termed as **Profitability Index** as they measure the profitability of the company based on the returns they earn on these acquired assets.

Factor 3 comprised of Quick Ratio, Current Ratio and Price/Gross Cash Flow. These can be termed **as Liquidity Measures**, as they measure the company's ability to pay immediate cash by selling out its current assets.

Factor 4 comprised of Quick Gross Gearing (D/E), Financial Leverage, Net Gearing and Price/Book Value. These can be termed as **Solvency Ratios**, as gearing or leverage measures the level of debt or borrowings of the company.

Factor 5 comprised of PER and Receivables / Op Revenue, hence can be termed as **Earnings Ratio**. They measure the fair value or price of the assets.

Factor 6 comprised of Working Cap Turnover and Gross Debt/CF, hence can be termed as **Expense Ratios**. It indicates the value of sales that the company has generated per dollar invested in capital.

Factor 7 comprised of Depreciation/PPE and Net Interest cover. These ratios assess if the company has the ability to pay out fixed interests on the assets or debts, and hence can be termed as **Interest on asset Ratios**.

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

These factors can be used as operational representatives for future decision making process.

Table 12 - Rotated Component Matrix[a]

	Component						
	1	2	3	4	5	6	7
LT Asset Turnover	.925						
Invested Capital Turnover	.879						
Asset Turnover	.846						
PPE Turnover	.722						
ROA		.855					
ROE		.832					
ROIC		.810					
Quick Ratio			.852				
Current Ratio			.790				
Price/Gross Cash Flow			-.458		.436		.428
Gross Gearing (D/E)				.801			
Financial Leverage				.755			
Net Gearing			-.394	.509			
Price/Book Value				.508		-.303	
PER					.755	-.318	
Receivables/Op. Rev.					.671		
Working Cap Turnover						.753	
Gross Debt/CF		.335				.583	.397
Net Interest Cover							.613
Depreciation/PPE							.597

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 11 iterations.

4.4 Artificial Neuron Network

ANN was performed on the data set using the Multilayer Perceptron. Using the Standardized rescaling of covariates, the case processing summary showed that 286 cases were assigned to the training sample and 109 to the testing sample. The most important variables in predicting the financial distress of the firms as per the normalized importance were Gross Gearing and Price Gross Cash Flow. Current Ratio was considered the least important of all independent variables.

As shown in Table 14 of Appendix E, of the cases used to create the model, 251 of the 253 successful companies are classified correctly. 17 of the 33 bankrupt companies are classified correctly. Overall, 93.7% of the training cases are classified correctly, corresponding to the 6.3% incorrect shown in the

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

model summary table. This is considered a good model as it classifies more than 90% of the cases correctly.

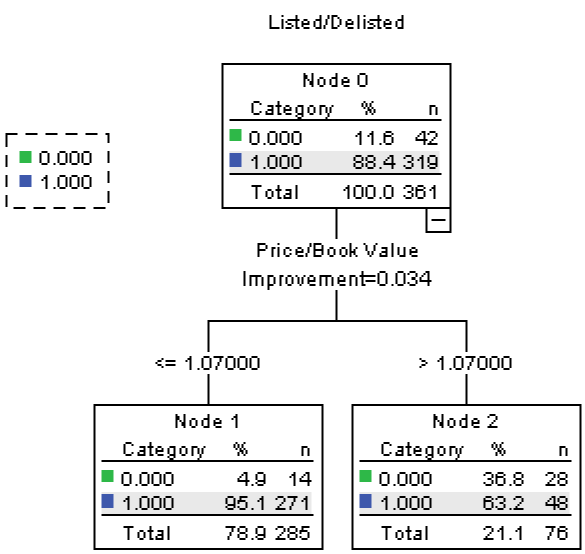
Classifications based upon the cases used to create the model tend to be too “optimistic” in the sense that their classification rate is inflated. The testing sample helps to validate the model; here 88.9% of these cases were correctly classified by the model. This suggests that, overall, this model is in fact correct about four out of five times.

4.5 Decision Tree

Decision Tree was build using the CART and CHAID as its growing method on this data set. The ratio for training to testing was taken as 9:1.

Twenty independent variables were specified for CART, but only one was included in the final model. None of the variables except Price/Book Value made any significant contribution and hence the rest were excluded from the model.

As shown in Table 15 of Appendix F, of the cases used to create the model, all 319 cases of successful companies are classified correctly but none of the bankrupt companies are correctly classified. Overall, 88.4% of the training cases are classified correctly. The results for testing was the same as training where all successful cases were correctly classified and overall 84.1% of the testing cases are classified correctly.

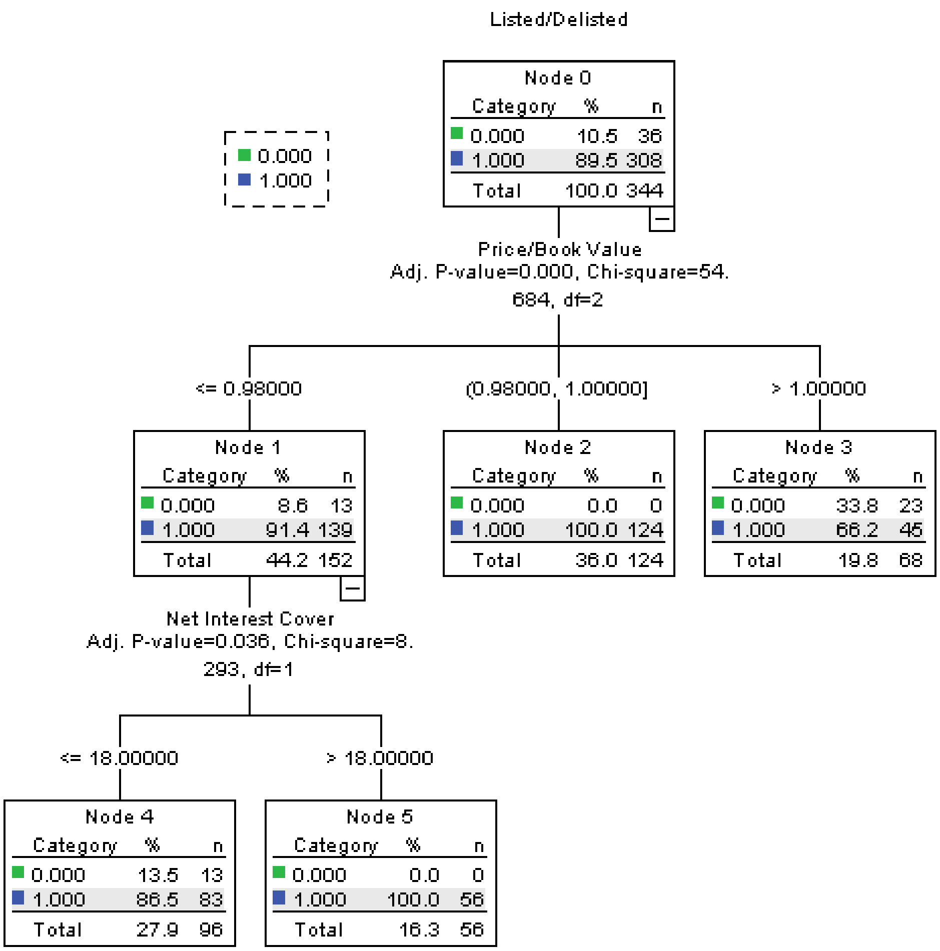


The tree diagram is a graphic representation of the tree model. As show in Figure 1 of Appendix F, this tree diagram shows that, using the CART method, Price/Book Value is the best and only predictor of Financial Distress. Of the all the companies, 95.6% have been successful companies. Since there are no child nodes below it, this is considered a terminal node.

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

Twenty independent variables were specified for CHAID, but only two were included in the final model. None of the variables except Price/Book Value and Net Interest Cover made any significant contribution and hence the rest were excluded from the model. The ratio for training to testing was taken as 9:1.

As shown in Table 16 of Appendix F, of the cases used to create the model, all 308 cases of successful companies are classified correctly but none of the bankrupt companies are correctly classified. Overall, 89.5% of the training cases are classified correctly. The results for testing was the same as training where all successful cases were correctly classified and overall 84.3% of the testing cases are classified correctly.



As show in Figure 2 of Appendix F, the tree diagram represents that:

- Using the CHAID method, Price/Book Value is the best predictor of credit rating.

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

- For the low Price/Book category (below .98) , the next best predictor is Net Interest Cover of the firm. For over 86% of those firms, those with Net Interest cover of 18% or lower are financially strong, while all of those over Net Interest cover of 18% are financially strong.
- For the medium and high Price/Book category (> .98) , Net Interest Cover of the firm is the only significant predictor of financial distress of the firm. Since there are no child nodes below it, this is considered a terminal node.

4.6 Hybrid Model I - ANN and LR

A hybrid method usually integrates two or more technologies. The purpose of integrating technologies is to strengthen the best features of each. This study uses the previously analyzed data using Logistic Regression and Discriminate Analysis and compares the multilayer perception with the classification tool.

ANN was performed on the same data set and the case processing summary (Table 17 in Appendix G) shows that 267 cases were assigned to the training sample and 128 to the testing sample. The most important variables in predicting the financial distress of the firms as per the normalized importance were PER and Price Gross Cash Flow. ROE was considered the least important of all independent variables.

As observed in Table 18 of Appendix F, of the cases used to create the model, all 240 successful companies are classified correctly. 11 of the 27 bankrupt companies are classified correctly. Overall, 94.0% of the training cases are classified correctly, corresponding to the 6% incorrect shown in the model summary table. This is considered a good model as it classifies more than 90% of the cases correctly.

4.7 Hybrid Model II - ANN and DA

ANN was performed on the same data set and the case processing summary (Table 19 in Appendix H) shows that 267 cases were assigned to the training sample and 128 to the testing sample. The most important variables in predicting the financial distress of the firms as per the normalized importance were Net Gearing and Quick Ratio. Invested Capital Turnover was considered the least important of all independent variables.

As observed in Table 20 of Appendix H, of the cases used to create the model, all 234 successful companies are classified correctly. 4 of the 28 bankrupt companies are classified correctly. Overall, 91.0% of the training cases are classified correctly, corresponding to the 9% incorrect shown in the model summary table.

5. DISCUSSIONS AND CONCLUSIONS

In the existing literature, financial ratios or factors are the most frequently used predictors in the models that forecast corporate financial distress using variables for firms from various sectors and/or from firms around the globe. The present study's findings are restricted to the corporate financial ratios of the Australian Mining Industry. Furthermore, the various techniques with underlying assumptions were analyzed and tested to compare results of the existing researches. In this study, the cutting edge recursive methods delivered better results than the traditional univariate and multivariate analytical models.

5.1 Comparison of Classification Rates

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

This study applied multivariate and recursive partitioning techniques to the financial data of various Australian mining companies for constructing prediction models and analysis. All six classifiers - logistic regression (LR), discriminant analysis (DA), neural networks (NN) and decision trees (C5.0), Hybrid I and Hybrid II are investigated separately for predicting corporate failure.

From the classification result, it is concluded that Hybrid I is the best model as it gives a prediction accuracy of 94.0%. This was close to results from ANN, a 93.70% prediction accuracy. As the results from H-I are as good as results from LR, ANN, and H-II, it could be concluded that every model has its own parameters and the classification results depend on the parameters. The more experiments are conducted the greater the chance to improve the result with optimal parameter set.

As per the comparison of all techniques seen in Table 21 of Appendix K, the total prediction accuracy of LR, DA, NN, CART, CHAID, Hybrid I and II in Training are 91.90%, 82.80%, 93.70%, 88.4%, 89.5%, 94.0% and 91.00% respectively and in Testing are 91.90%, 82.80%, 91.7% 94.1%, 84.3%, 89.10 and 89.80%.

Table 21 - Classification Table

Model	Mass Classification Rate (%)	
	Training	Testing
LR	91.90%	91.90%
DA	82.80%	82.80%
ANN	93.70%	91.70%
DT - CART	88.4%	94.1%
DT - CHAID	89.5%	84.3%
Hybrid I - ANN and LR	94.00%	89.10%
Hybrid II - ANN and DA	91.00%	89.80%

Using the Multilayer Perceptron procedure, the report has constructed a network for predicting the probability that a given firm will default for bankruptcy. The model results are comparable to those obtained using Logistic Regression or Discriminant Analysis, so we can be reasonably confident that the data do not contain relationships that cannot be captured by those models; thus, we can use them to further explore the nature of the relationship between the dependent and independent variables.

Factor analysis was used to summarize the independent variables into seven different categories - Turnover Ratios, Profitability Index, Liquidity Measures, Solvency Ratios, Earnings Ratio, Expense Ratios and Interest on asset Ratios.

5.2 Comparison of Significant Independent Variables

It was attempted to examine which Independent variable was most significant in determining the financial distress of the firm. It is hard to conclude which features are more important than the others, especially when these are correlated. Table 22 of Appendix L, below lists the features which are commonly selected or considered as the significant variables in the experiments using different models.

Few variables were considered significant in all or most of the models such as PER and Price/Gross Cash Flow.

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

	Table 22 - Comparison of Significant Independent Variables						
Independent Variables	LR	DA	ANN	DT - CART	DT - CHAID	Hybrid I	Hybrid II
Asset Turnover		×					
Current Ratio		×					×
Depreciation/PP&E	×						
Financial Leverage							
Gross Debt/CF	×						
Gross Gearing (D/E)			×				
Invested Capital Turnover		×					
LT Asset Turnover		×					
Net Gearing		×					
Net Interest Cover					×		
PER	×					×	×
PPE Turnover		×					
Price/Book Value		×		×	×		
Price/Gross Cash Flow	×	×	×				×
Quick Ratio		×					
Receivables/Op. Rev.							
ROA	×		×				
ROE							
ROIC	×						
Working Cap Turnover							

5.3 Future Work

Economists have been constantly working on the accuracy of BFPs and have evolved models from statistical techniques to recursive partitioning techniques. The proposed Hybrid system has shown the potential for improving classifications for predicting bankruptcy of firms in the growing mining industry. The current results from Hybrid I is recommendable, but needs to be tested on more data. The results from this hybrid is not satisfying, although it is better than produced by other models. Different ways of combination of intelligent techniques can be developed and tested in the future.

The technique presented in this study can be used for business failure predictions in the Australian mining sector. Also these models can be modified and used in other research areas such as in accountancy, business, economics etc.

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

Appendix

Appendix A - Ratios selected for the study and their definitions

Ratios	Definition
Asset Turnover	Operating revenue / total assets. Asset Turnover measures the efficient use of assets and how well they produce revenue during the corresponding period. It is calculated by dividing net sales by total assets.
Current Ratio	Current assets / current liabilities. Current assets divided by current liabilities. This ratio is a useful measure of the short term debt-paying ability of the company. The higher the ratio, the more liquid the company is. Whilst a ratio of 2 or more was traditionally considered desirable many companies have reduced this in recent years as operating cycles have shortened. It is more relevant to understand the ratio in the context of the sector average and the trend over the last few years
Depreciation/PP&E	Depreciation / gross property, plant & equipment. Depreciation will be correlated with capital expenditure, but will lag as capital expenditure is gradually expensed in the profit and loss statement. High levels of current capital expenditure will generally be followed up with higher depreciation in future periods, which will lower earnings
Financial Leverage	Total assets / shareholders equity. Financial leverage is the degree to which a company uses fixed-income securities such as debt and preferred equity. The more debt financing a company uses, the higher its financial leverage. A high degree of financial leverage means high interest payments, which negatively affect the company's bottom-line earnings per share.
Gross Debt/CF	(Short term debt + long term debt) / gross cash flow. Gross cash flow is defined as NOPLAT + depreciation.
Gross Gearing (D/E)	(Short term debt + long term debt) / Shareholders equity. Gearing ratio refers to the fundamental analysis ratio of a company's level of long-term debt compared to its equity capital.
Invested Capital Turnover	Operating revenue / operating invested capital before goodwill. This ratio shows the value of sales revenue that the company has generated per unit of capital invested in the business. Companies that are highly capital intensive such as airlines and steel producers will tend to have lower Capital Turnover than services companies or distributors that require lesser capital expenditures. In the cases where operating invested capital before goodwill is less than zero, we have set the value of Capital Turnover to null.
LT Asset Turnover	
Net Gearing	(Short term debt + long term debt - cash) / Shareholders equity.
Net Interest Cover	Earnings before interest and tax / interest expense. A ratio used to determine how easily a company can pay interest on outstanding debt.

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

PER	"Price/Earnings Ratio" = {(market value of share)/ (Earnings per Share)}. A relatively high PE ratio can be an indication that the market expects earnings growth to be relatively high
PPE Turnover	Operating revenue / (property, plant & equipment - accumulated depreciation). PPE turnover allows the analyst to determine how productively the company has been utilizing their PPE to generate sales.
Price/Book Value	The ratio of the current price per share divided by book value per share. The book value measures the value of the shareholders ownership in the company, as measured by the last full year balance sheet. The price to book ratio is usually greater than one as the market value will usually exceed the balance sheet value attributed to the assets of the company. This is because assets are generally recorded at their original cost, less any accumulated depreciation. The market, on the other hand, is concerned with the cash generating ability of the company's assets rather than its historical cost. If an asset can generate returns in excess of its cost of capital, then a premium will be paid for the asset. This premium is the price to book ratio.
Price/Gross Cash Flow	Closing share price on the last day of the company's financial year / gross cash flow. Gross cash flow is defined as NOPLAT + depreciation.
Quick Ratio	(Current assets - current inventory) / Current liabilities. Also known as the "acid test", the quick ratio is similar to the current ratio but excludes the value of inventory or stocks in the current asset calculation. The reasoning for this is that inventories are not always immediately realizable as a source of cash. Inventory can also be subject to valuation problems. The formula is current assets less inventory divided by current liabilities. As with the current ratio it, it is important to understand the ratio in the context of the sector average and the trend over the last few years.
Receivables/Op. Rev.	Debtors / Operating revenue. The receivables turnover ratio is used to calculate how well a company is managing their receivables. The lower the amount of uncollected monies from its operations, the higher this ratio will be. In contrast, if a company has more of its revenues awaiting receipt, the lower the ratio will be.
ROA	Earnings before interest / (total assets less outside equity interests). ROA is a key measure of a company's profitability, equal to a fiscal year's earnings divided by its total assets. Return on assets essentially shows how much profit a company is making on the assets used in its business.
ROE	NPAT before abnormals / (shareholders equity - outside equity interests). ROE is an evaluation of profit earned in relation to equity resources invested (the viewpoint of equity holders). It is calculated by dividing net profit before abnormal by shareholders equity. In the cases where shareholders equity is less than zero, we have set the value of ROE to null.
ROIC	NOPLAT / operating invested capital before goodwill. Return on invested capital is a key measure of how effectively a company uses the money and invested in its operations. In the cases where invested capital before goodwill is less than zero, we have set the value of ROIC to null.
Working Cap	Operating revenue / operating working capital.

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

Turnover	Indicates the value of sales that the company has generated per dollar invested in working capital. An increase in this ratio over time indicates the company is becoming more efficient in generating revenues from its resources. A company whose costs are mainly fixed costs as opposed to variable will be able to increase revenues at a faster rate than expenses and this will result in a higher working capital turnover ratio. Operating working capital is defined as (Current assets - cash) - (current liabilities - short term debt).

Appendix B - Results and Interpretations for Logistic Regression Model

Table 1:

Table 1 - Model Summary			
Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	201.857[a]	.171	.341
a. Estimation terminated at iteration number 8 because parameter estimates changed by less than .001.			

Table 2:

Table 2 - Variables in the Equation									
		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B)	
								Lower	Upper
Step 1[a]	AssetTurnover	.325	.940	.120	1	.729	1.385	.219	8.737
	CurrentRatio	.127	.208	.371	1	.542	1.135	.755	1.705
	DepreciationPP&E	.052	.025	4.098	1	.043	1.053	1.002	1.107
	FinancialLeverage	-.009	.012	.645	1	.422	.991	.968	1.014
	GrossDebtCF	-.024	.008	10.496	1	.001	.976	.962	.990
	GrossGearingDE	.016	.023	.498	1	.480	1.016	.972	1.063
	InvestedCapitalTurnover	-.242	.199	1.484	1	.223	.785	.532	1.159
	LTAssetTurnover	-.278	.517	.288	1	.591	.758	.275	2.087
	NetGearing	.013	.013	1.021	1	.312	1.013	.988	1.040
	NetInterestCover	.004	.010	.189	1	.664	1.004	.985	1.024
	PER	-.029	.012	5.426	1	.020	.971	.948	.995
	PPETurnover	-.119	.084	2.013	1	.156	.888	.753	1.047
	PriceBookValue	-.079	.057	1.956	1	.162	.924	.827	1.032
	PriceGrossCashFlow	.066	.022	9.142	1	.002	1.068	1.024	1.115
	QuickRatio	-.223	.209	1.131	1	.288	.800	.531	1.206
	ReceivablesOp.Rev	.027	.033	.672	1	.412	1.027	.963	1.096

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

	ROA	.053	.016	11.057	1	.001	1.054	1.022	1.087
	ROE	-.008	.011	.493	1	.483	.992	.972	1.014
	ROIC	-.019	.008	6.154	1	.013	.981	.967	.996
	WorkingCapTurnover	.023	.021	1.274	1	.259	1.024	.983	1.066
	Constant	4.182	1.405	8.856	1	.003	65.492		
a. Variable(s) entered on step 1: AssetTurnover, CurrentRatio, DepreciationPP&E, FinancialLeverage, GrossDebtCF, GrossGearingDE, InvestedCapitalTurnover, LTAssetTurnover, NetGearing, NetInterestCover, PER, PPETurnover, PriceBookValue, PriceGrossCashFlow, QuickRatio, ReceivablesOp.Rev, ROA, ROE, ROIC, WorkingCapTurnover.									

Table 3:

Table 3 - Classification Table					
	Observed		Predicted		
			Listed/Delisted		Percentage Correct
			0	1	
Step 1	Listed/Delisted	0	15	29	34.1
		1	3	348	99.1
	Overall Percentage				91.9
a. The cut value is .500					

Appendix C - Results and Interpretations for Discriminant Analysis Model

Table 4:

Table 4 - Tests of Equality of Group Means					
	Wilks' Lambda	F	df1	df2	Sig.
Asset Turnover	.976	9.738	1	393	.002
Current Ratio	.982	7.100	1	393	.008
Depreciation/PP&E	1.000	.180	1	393	.672
Financial Leverage	.999	.245	1	393	.621
Gross Debt/CF	.996	1.540	1	393	.215
Gross Gearing (D/E)	.995	2.131	1	393	.145
Invested Capital Turnover	.989	4.330	1	393	.038
LT Asset Turnover	.980	7.879	1	393	.005
Net Gearing	.967	13.346	1	393	.000
Net Interest Cover	.996	1.603	1	393	.206
PER	1.000	.086	1	393	.769
PPE Turnover	.985	5.934	1	393	.015
Price/Book Value	.984	6.198	1	393	.013

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

Price/Gross Cash Flow	.975	10.207	1	393	.002
Quick Ratio	.919	34.549	1	393	.000
Receivables/Op. Rev.	.999	.361	1	393	.548
ROA	.994	2.551	1	393	.111
ROE	.998	.723	1	393	.396
ROIC	.993	2.760	1	393	.097
Working Cap Turnover	.999	.254	1	393	.614

Table 5:

Table 5 - Eigenvalues				
Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	.170[a]	100.0	100.0	.381
a. First 1 canonical discriminant functions were used in the analysis.				

Table 6:

Table 6 - Wilks' Lambda				
Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	.855	61.395	4	.000

Table 7:

Table 7 - Standardized Canonical Discriminant Function Coefficients	
	Function
	1
Asset Turnover	.631
Price/Book Value	.284
Price/Gross Cash Flow	-.318
Quick Ratio	.763

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

Table 8:

Table 8 - Classification Function Coefficients		
	Listed/Delisted	
	0	1
Asset Turnover	4.068	1.577
Price/Book Value	.355	.190
Price/Gross Cash Flow	-.056	-.019
Quick Ratio	.278	.109
(Constant)	-3.099	-1.096
Fisher's linear discriminant functions		

Table 9:

Table 9 - Classification Results[a,c]					
		Listed/Delisted	Predicted Group Membership		Total
			0	1	
Original	Count	0	22	22	44
		1	46	305	351
	%	0	50.0	50.0	100.0
		1	13.1	86.9	100.0
Cross-validated[b]	Count	0	22	22	44
		1	46	305	351
	%	0	50.0	50.0	100.0
		1	13.1	86.9	100.0
a. 82.8% of original grouped cases correctly classified.					
b. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.					
c. 82.8% of cross-validated grouped cases correctly classified.					

Appendix D - Results and Interpretations for Principal Component Analysis and Model

Table 10:

Table 10 - KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.701
Bartlett's Test of Sphericity	Approx. Chi-Square	2655.483
	df	190
	Sig.	.000

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

Table 11:

Table 11 - Total Variance Explained									
Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4.070	20.352	20.352	4.070	20.352	20.352	3.193	15.967	15.967
2	2.354	11.771	32.123	2.354	11.771	32.123	2.406	12.031	27.998
3	1.853	9.267	41.390	1.853	9.267	41.390	1.986	9.931	37.929
4	1.525	7.626	49.016	1.525	7.626	49.016	1.836	9.178	47.107
5	1.238	6.192	55.209	1.238	6.192	55.209	1.357	6.785	53.891
6	1.205	6.027	61.236	1.205	6.027	61.236	1.291	6.455	60.346
7	1.061	5.306	66.542	1.061	5.306	66.542	1.239	6.196	66.542
8	.919	4.593	71.134						
9	.892	4.458	75.593						
10	.822	4.108	79.700						
11	.706	3.531	83.231						
12	.603	3.014	86.245						
13	.522	2.609	88.854						
14	.443	2.215	91.069						
15	.396	1.981	93.049						
16	.381	1.903	94.952						
17	.365	1.824	96.776						
18	.357	1.787	98.562						
19	.189	.947	99.509						
20	.098	.491	100.000						
Extraction Method: Principal Component Analysis.									

Table 12:

Table 12 - Rotated Component Matrix[a]							
	Component						
	1	2	3	4	5	6	7
LT Asset Turnover	.925						

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

Invested Capital Turnover	.879						
Asset Turnover	.846						
PPE Turnover	.722						
ROA		.855					
ROE		.832					
ROIC		.810					
Quick Ratio			.852				
Current Ratio			.790				
Price/Gross Cash Flow			-.458		.436		.428
Gross Gearing (D/E)				.801			
Financial Leverage				.755			
Net Gearing			-.394	.509			
Price/Book Value				.508		-.303	
PER					.755	-.318	
Receivables/Op. Rev.					.671		
Working Cap Turnover						.753	
Gross Debt/CF		.335				.583	.397
Net Interest Cover							.613
Depreciation/PPE							.597
Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.							
a. Rotation converged in 11 iterations.							

Appendix E - Results and Interpretations for Artificial Neural Network Model

Table 13:

Table 13 - Model Summary		
Training	Cross Entropy Error	48.969
	Percent Incorrect Predictions	6.3%
	Stopping Rule Used	1 consecutive step(s) with no decrease in error[a]
	Training Time	0:00:00.22
Testing	Cross Entropy Error	22.600

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

	Percent Incorrect Predictions	8.3%
Dependent Variable: Listed/Delisted		
a. Error computations are based on the testing sample.		

Table 14:

Table 14 - Classification				
Sample	Observed	Predicted		
		0	1	Percent Correct
Training	0	17	16	51.5%
	1	2	251	99.2%
	Overall Percent	6.6%	93.4%	93.7%
Testing	0	4	7	36.4%
	1	2	96	98.0%
	Overall Percent	5.5%	94.5%	91.7%
Dependent Variable: Listed/Delisted				

Appendix F - Results and Interpretations for Decision Tree

Table 15:

Table 15 - Classification				
Sample	Observed	Predicted		
		0	1	Percent Correct
Training	0	0	42	0.0%
	1	0	319	100.0%
	Overall Percentage	0.0%	100.0%	88.4%
Test	0	0	2	0.0%
	1	0	32	100.0%
	Overall Percentage	0.0%	100.0%	94.1%
Growing Method: CRT Dependent Variable: Listed/Delisted				

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

Figure 1:

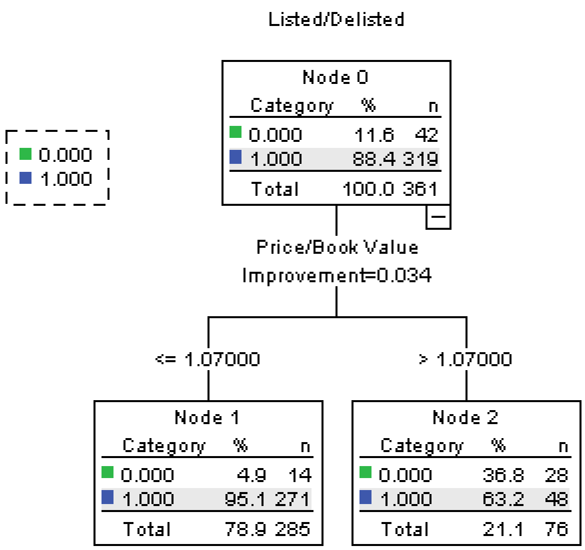
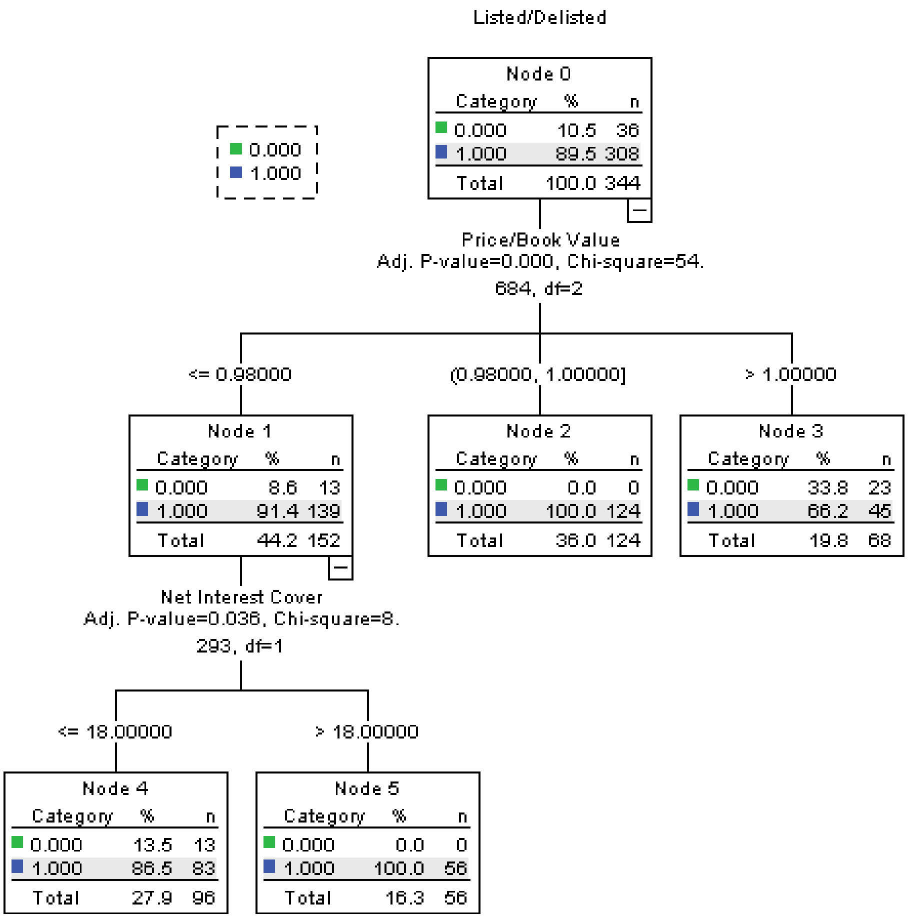


Table 16:

Table 16 - Classification				
Sample	Observed	Predicted		
		0	1	Percent Correct
Training	0	0	36	0.0%
	1	0	308	100.0%
	Overall Percentage	0.0%	100.0%	89.5%
Test	0	0	8	0.0%
	1	0	43	100.0%
	Overall Percentage	0.0%	100.0%	84.3%
Growing Method: CHAID Dependent Variable: Listed/Delisted				

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

Figure 2:



Appendix G - Results and Interpretations for Hybrid Model I - ANN and LR

Table 17:

Table 17 - Model Summary		
Training	Cross Entropy Error	61.764
	Percent Incorrect Predictions	6.0%
	Stopping Rule Used	1 consecutive step(s) with no decrease in error[a]
	Training Time	0:00:00.34
Testing	Cross Entropy Error	34.284

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

	Percent Incorrect Predictions	10.9%
Dependent Variable: Listed/Delisted		
a. Error computations are based on the testing sample.		

Table 18:

Table 18 - Classification				
Sample	Observed	Predicted		
		0	1	Percent Correct
Training	0	11	16	40.7%
	1	0	240	100.0%
	Overall Percent	4.1%	95.9%	94.0%
Testing	0	5	12	29.4%
	1	2	109	98.2%
	Overall Percent	5.5%	94.5%	89.1%
Dependent Variable: Listed/Delisted				

Appendix H - Results and Interpretations for Hybrid Model II - ANN and DA

Table 19:

Table 19 - Model Summary		
Training	Cross Entropy Error	71.859
	Percent Incorrect Predictions	9.0%
	Stopping Rule Used	1 consecutive step(s) with no decrease in error[a]
	Training Time	0:00:00.22
Testing	Cross Entropy Error	39.274
	Percent Incorrect Predictions	10.2%
Dependent Variable: Listed/Delisted		
a. Error computations are based on the testing sample.		

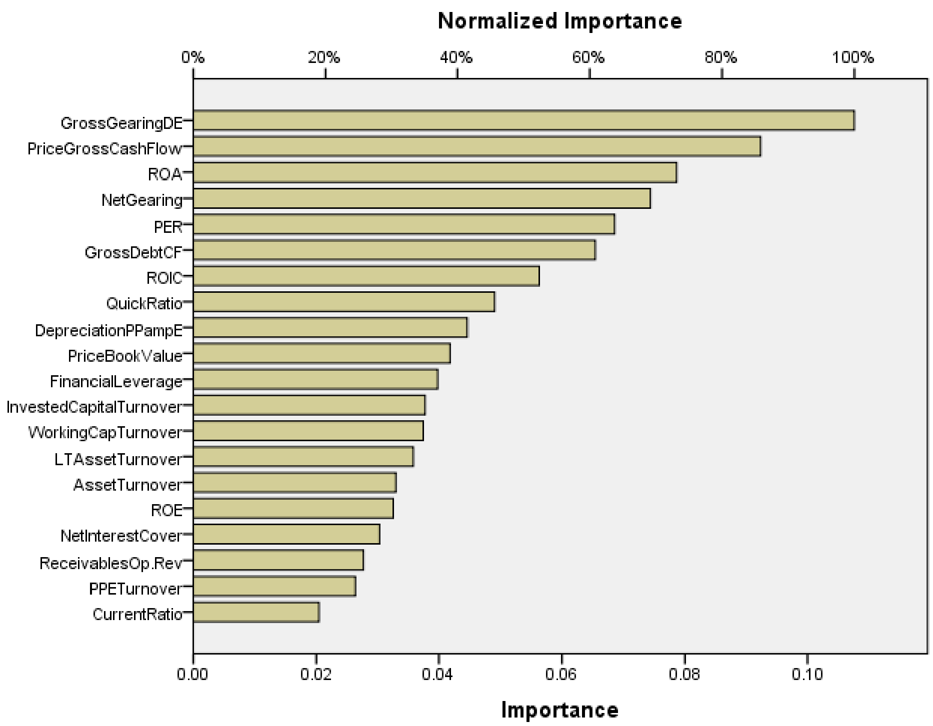
Table 20:

Table 20 - Classification				
Sample	Observed	Predicted		
		0	1	Percent Correct
Training	0	4	24	14.3%
	1	0	239	100.0%
	Overall Percent	1.5%	98.5%	91.0%

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

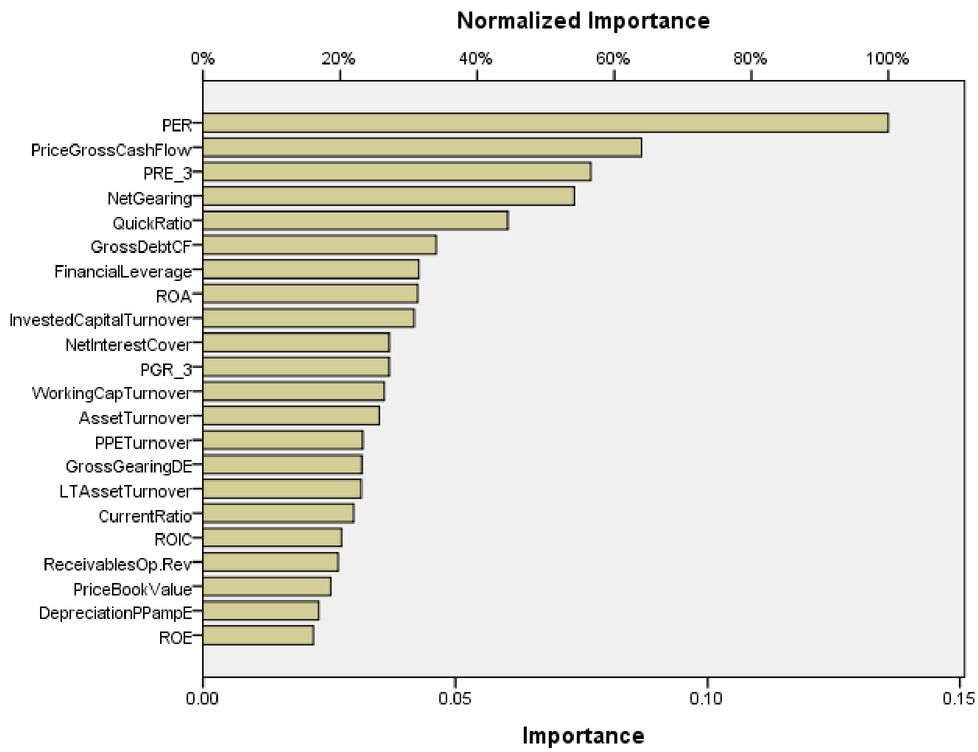
Testing	0	3	13	18.8%
	1	0	112	100.0%
	Overall Percent	2.3%	97.7%	89.8%
Dependent Variable: Listed/Delisted				

Appendix I - Importance of each variable from ANN



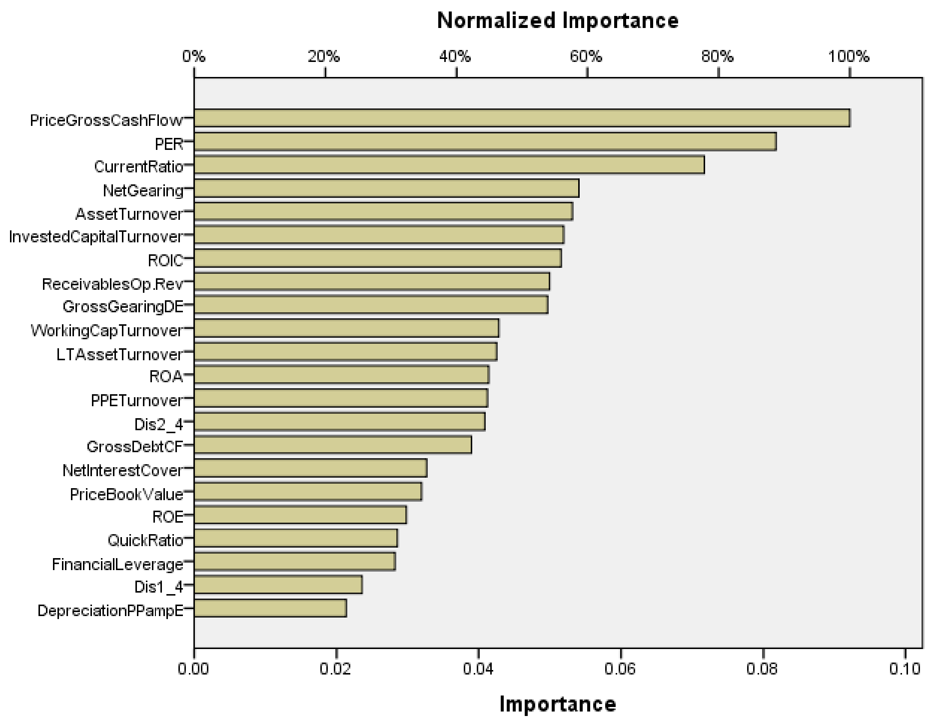
ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

Appendix J - Importance of each variable from ANN and LR



ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

Appendix K - Importance of each variable from ANN and DA



Appendix L - Comparison of Classification Rates

Table 21:

Table 21 - Classification Table

Model	Mass Classification Rate (%)	
	Training	Testing
LR	91.90%	91.90%
DA	82.80%	82.80%
ANN	93.70%	91.70%
DT - CART	88.4%	94.1%
DT - CHAID	89.5%	84.3%
Hybrid I - ANN and LR	94.00%	89.10%

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

Hybrid II - ANN and DA	91.00%	89.80%

Appendix M - Comparison of Significant Independent Variables

Table 22:

	Table 22 - Comparison of Significant Independent Variables						
Independent Variables	LR	DA	ANN	DT - CART	DT - CHAID	Hybrid I	Hybrid II
Asset Turnover		×					
Current Ratio		×					×
Depreciation/PP&E	×						
Financial Leverage							
Gross Debt/CF	×						
Gross Gearing (D/E)			×				
Invested Capital Turnover		×					
LT Asset Turnover		×					
Net Gearing		×					
Net Interest Cover					×		
PER	×					×	×
PPE Turnover		×					
Price/Book Value		×		×	×		
Price/Gross Cash Flow	×	×	×				×
Quick Ratio		×					
Receivables/Op. Rev.							
ROA	×		×				
ROE							
ROIC	×						
Working Cap Turnover							

Acknowledgements

I would like to express my sincere thanks to Prof. Kuldeep Kumar, from Bond University, my Primary Supervisor, for his expertise, guidance and assistance during the course of my candidacy. His enthusiasm and motivation is ever inspiring and his invaluable ideas greatly contributed to the successful completion of this paper.

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

I am also grateful to my colleague Adrian Gepp for his time and patience, and sharing his expertise.

I would like to thank my parents, Mr. Sunil Kumar Shah and Mrs. Sunita Shah, for their eternal patience, unconditional support and always believing in me. Many other people who have been a part of this journey with me and each have carried me along my path in some way and express my love for everyone of them.

REFERENCES

[1] Altman, E. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *Journal of Finance,23(4)* , 589 - 609.

[2] Anderson, D. R., Sweeney, D. J., & Williams, T. A. (2008). *Statistics for business and Economics.* Mason: South-Western College.

[3] Australian Bureau of Statistics. (2010). *Mining Capital Expenditure in Australia.* Canberra: Australian Bureau of Statistics.

[4] Australian Bureau of Statistics. (2012). *Mining Investments in ABS Publications.* Canberra: Australian Bureau of Statistics.

[5] Bardos, M., & Rasson, J. (2001). *Analyse discriminante: Application au risque et scoring financier.* Paris: Dunod.

[6] Beaver, W. (1966). Financial ratios as predictors of bankruptcy. *Journal of Accounting Research, 4(3)* , 71 - 111.

[7] Bell, T., Ribar, G., & Verchio, J. (1990). Neural nets vs. Logistic regression: A comparison of each model's ability to predict commercial bank failures. *Proceedings of the Deloitte Touche/University of Kankas Symposium on Auditing Problems*, (pp. 29 - 53). Lawrence.

[8] Boritz, J., & Kennedy, D. (1995). Predicting Corporate Failure Using a Neural Network Approach. *Intelligent Sysyems in Accounting, Finance and Management 4* , 95 - 111.

[9] Child, D. (2006). *The essentials of factor analysis.* London: Continuum International Publishing.

[10] Cho, S., Hong, H., & Ha, B. (2010). A hybrid approach based on the combination of variable selection using decision trees and case-based reasoning using the mahalanobis distance: For bankruptcy prediction. *Expert Systems with Applications* , 3482-3488.

[11] Coak, S. J. (2005). *SPSS Version 12.0 for Windows: Analysis Without Anguish.* Brisbane: John Wiley & Sons Australia.

[12] Coats, P. K., & Fant, L. F. (1992). A neural Network Approach to forecasting Financial Distress. *Journal of Business Forecasting, 10(4)* , 9 - 12.

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

[13] Coleman, K. G., Graettinger, T., & Lawrence, W. (1991). Neural Networks for Bankruptcy Prediction: the power to solve financial problems. *Artificial Intelligence Review* , 48 - 50.

[14] Collins, R. A., & Green, R. D. (1982). Statistical method for bankruptcy forecast. *Journal of Economics and Business, 34(4)* , 348-354.

[15] Cybinski, P. (2001). Description, explanation, prediction – the evolution of bankruptcy studies? *Managerial Finance, 27(4)* , 29 - 44.

[16] Demyanyk, Y., & Hasan, I. (2010). Financial Crises and Bank Failures: A Review of Prediction Methods. *Social Science Research Network Working Paper Series* , 315-324.

[17] FitzPatrick, P. J. (1932). A Comparison of the Ratios of Successful Industrial Enterprises With Those of Failed Companies. *Journal of Accounting Research* , 3 - 21.

[18] Fletcher, D., & Gross, E. (1993). Forecasting with Neural Networks: an application using Bankruptcy data. *Journal ofInformation and Management, 24(3)* , 159 - 167.

[19] Frydman, H., Altman, E. I., & Kao, D. L. (1985). Introducing Recursive Partitioning for Financial Classification: The Case of Financial Distress. *Journal of Finance, 40(1)* , 269 - 291.

[20] Gepp, A., Kumar, K., & Bhattacharya, S. (2010). Business Failure Prediction using Decision Trees. *Journal of Forecasting, 29(6)* , 536 - 555.

[21] Hall, G. (1994). Factors distinguishing survivors from failures among small firms in the U.K construction sector. *Journal of Management Studies, 31* , 737-760.

[22] Hansen, J. V. (1991). *Artificial neural networks: Foundations and application to a decision problem.* New York: Elsevier Ltd.

[23] Hosmer, D. W., & Lemeshow, S. (2000). *Applied Logistic Regression.* New York: Wiley-Interscience Publication.

[24] Jaikengkit, A. (2004). *Corporate governance and financial distress : an empirical analysis - the case of Thai financial institutions.* USA: Case Western Reserve University.

[25] Kumar, K., & Chaturvedi, A. (2010). *Some Recent Developments in Statistical Theory and Applications.* Florida: BrownWalker Press.

[26] Kumar, K., Gepp, A., & Bhattachariya, S. (2010). Business Failure Predictions Using Decision Trees. *Journal of Forecasting, 29(6)* , 536 - 555.

[27] Laitinen, T., & Kankaanpaa, M. (1999). Comparative ANalysis of failure prediction methods: the Finnish case. *The European Accounting Review* , 67 - 92.

[28] Malhotra, N. (2007). *Essentials of Marketing Research: An Applied Orientation.* Frenchs Forest, NSW: Pearson Education Australia.

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)

[29] Martin, D. (1977). Early warning of bank failure: a logit regression approach. *Journal of Banking and Finance* , 249-276.

[30] Muller, G., Steyn-Bruwer, B., & Hamman, W. (2009). Predicting financial distress of companies listed on the JSE - a comparison of techniques. *South African Journal of Business Management, 40(1)* , 21 - 32.

[31] Odom, M., & Sharda, R. (1990). A neural network model for bankruptcy prediction. *Proceedings of the EEE international conference on neural networks* , 1163 - 1168.

[32] Ohlson, J. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research (Spring)* , 109-131.

[33] Pallant, J. (2007). *SPSS survival manual: a step by step guide to data analysis using SPSS for Windows.* Maidenhead: Open University Press.

[34] Pompe, P., & Feelders, A. (1997). Using Machine Learning, Neural Networks and Statistics to Predict Corporate Bankruptcy: A Comparative Study. *Microcomputers in Civil Engineering* , 3 - 19.

[35] Ravi Kumar, P., & Ravi, V. (2007). Bankruptcy prediction in banks and firms via statistical and intelligent techniques – A review. *European Journal of Operational Research* , 1-28.

[36] Salchenberger, K. M., Cinar, E. M., & Lash, N. A. (1992). Neural Networks: A new tool for predicting thrift failures. *Decision Sciences* , 185 - 205.

[37] Shann, E. (2012, March). Maximizing Growth in a Mining Boom. *Minerals Council of Australia* .

[38] SPSS. (2012). *Neural Network Applications in Real Estate Analysis.* IBM Corporation 1989, 2012.

[39] The Australian Institute. (2013). *Just how important is the mining sector to Australia?* Canberra: The Australia Institute.

[40] Y., T. K., & Kiang, M. (1992). Managerial Applications of Neural Networks: The case of bank failure predictions. *Management Science, 38(7)* , 926 - 947.

[41] Zmijewski, M. (1983). Methodological issues related to the estimation of finance. *Journal of Accounting Research, Supplement on Current* , 59 - 82.

ISSN: 2304-1013 (Online); 2304-1269 (CDROM)