Chapter 1 Introduction

1.1 Background

Corporate failure always brings huge economic losses to investors and others, together with a substantial social and economical cost to the nation. Failure can be defined in many ways, depending on the specific interest of the firms under examination. According to a general definition (Dimitras et al., 1996), failure is the situation that a firm cannot pay lenders, preferred stock shareholders, suppliers, etc., or a bill is overdrawn, or the firm is bankrupt according to law. All these situations result in a discontinuity of the firm's operations.

Numerous researchers have studied failure prediction over the last decades. As a result, various theories have evolved in an effort to explain or distinguish between firms that have failed. Altman (1968) used multivariate discriminant analysis to differentiate between failed and non-failed US firms. Following the study of Altman, a large number of methods such as logit analysis, probit analysis and linear programming have been applied to model this problem.

These conventional statistical methods, however, have some restrictive assumptions such as the linearity, normality and independence among predictor or input variables. Considering that the violation of these assumptions for independent variables frequently occurs with financial data, these methods can have limitations to obtain the effectiveness and validity. The development of a cause-effect relationship between attributes that may cause bankruptcy and the actual occurrence of bankruptcy is difficult for several reasons. One reason is that current theory may not allow decision makers to clearly identify all relevant attributes, thus some could be omitted or misspecified. A second reason is that some of the attributes may be quantitative while others may be qualitative, thus creating measurement problems. A third reason is that the attributes may occur in one or more time periods prior to bankruptcy, thereby introducing the problem of temporal measurement. A fourth reason is that there is not a one-to-one mapping between attributes and results. Thus, multiple attributes may measure the same construct and confound attempts to appropriately describe the construct. When dealing with real companies, these difficulties mean that analysis of the cause-effect relationship is usually inconsistent in case classifications. That is, a nonbankrupt company may have the same attributes as a bankrupt company (Mckee, 2003).

Recently, a number of new techniques emerging to assist the failure prediction, such as expert systems, neural networks, rough set theory and genetic programming. A key advantage of these contemporary methods over their traditional counterparts is that they do not require pre-specification of a function form, nor the adoption of restriction assumptions concerning the distributions of model variables and errors. However, most classification systems lack the ability to systematically conduct sensitivity analysis. Sensitivity analysis in bankruptcy prediction problem has several decision support advantages. This information, if available, provides substantial decision support advantages for investors as well as managers. Recently, Troutt et al. (1996) and Seiford and Zhu (1998) showed that the DEA model can be used for classification. Pendharkar (2002) showed how to use the sensitivity analysis procedure in DEA model to solve the inverse classification on bankruptcy prediction. Cielen et al. (2004) compared the bankruptcy classification performance of a linear programming model, a DEA model and a rule induction (C5.0) model. Paradi et al. (2004) use a layered worst practice DEA technique, where the sequential layers of poor performance are found with decreasing risk rating. This layering technique enables incorporation of risk attitudes and risk-based pricing. A limitation of these DEA studies is that the difficulty to treat qualitative data, therefore all these studies use quantitative financial data.

On the other hand, rough set theory has been applied to a wide variety of financial decision analysis problems. A limitation of rough set is that the continuous data used to derive the rough set rules, have been discretised with the aid of a selected expert.

Rough set analysis produces better results when the attribute domains for continuous variables are finite sets of low cardinality. Therefore it is necessary to recode continuous variables, such as financial ratios, into qualitative terms such as 'low, medium, high'. Using sorting rules developed by rough sets may lead to a burdensome situation where a new case does not match any of the sorting or classification rules.

In this research we propose a hybrid system combining rough set approach and DEA. Our system has two agents: one is a visual display agent that helps users monitor the risk by placing the distressed firms on layered frontiers based on how efficient they are at being bad. At the prediction step, the mining agent apply the rules developed by rough set, and help users make decision about the risk analysis. The effectiveness of our hybrid approach was verified with experiments that compared to the worst DEA model.

Figure 1 illustrates that a firm might go through various stages of financial distress. Normally, an inability to compete successfully in the market place and/or an inability to manage liquid assets precedes financial distress and leads to inadequate liquid assets. The inadequate income may continue for many years but, either because it is not particularly severe or due to the availability of additional financing, may not lead to an inadequate liquid asset position. However, if management does not eventually adopt successful operating strategies, at some point the inadequate income can lead to an inadequate liquid asset position. The firm subsequently may experience one or more of the following conditions: bankruptcy, liquidation, reorganization, debt or other restructuring, or merger.

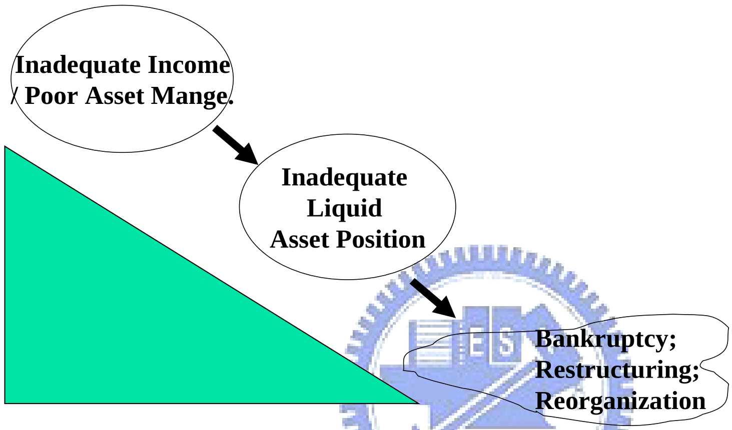

Fig.1.1. Temporal model of financial failure and possible outcomes [Mckee 2003]

The end to a firm is defined as economic 'discontinuity'. Investors, creditors and other interested parties would like to predict is when this economic discontinuity will occur. As a practical matter discontinuity is difficult to predict since it may involve so many alternative conditions and definitions. To simplify the problem, many researchers have chosen to focus on predicting bankruptcy or non-bankruptcy, which is just one possible outcome which may occur when discontinuity exists. This research uses the border definition of predicting financial crisis (includes all kinds of discontinuity) for failure prediction.

1.2 Thesis Objectives

The objectives of this research are as follows:

- to introduce a hybrid Rough set and DEA approach as a modeling tool that addresses many of the shortcomings of previous business failure prediction;
- to propose an efficiency model to find attribute disjunction decision rules by the integer linear programming under rough set concept;
- to test the feasibility of using Rough Set DEA as a tool to assess corporate failure risk by developing a rough set rule that captures the qualitative characteristics of failure analysts and measure credit risk by determining the worst practice DEA ranking of a company;
- to introduce and test worst practice DEA theory within the failure risk evaluation problem; it is a layering technique that does not require the specification of a cut off point. Unlike many other techniques, this method is not subject to sample data specificity issues;
- to introduce a new use for DEA, moving away from its traditional role as a tool

for efficiency measurement. It uses the peer groups and the layered techniques to find companies that are the most similar to each other.

1.3 Thesis Outline

The chapters of the thesis come together in the following manner:

- Chapter 2 provides a comprehensive overview of the literature on business failure prediction and associated approaches used. The shortcomings of these methods are outlined.
- Chapter 3 proposes an efficiency model to find decision rules by the integer linear programming under rough set concept. It provides a description of the rough set basics and its application on economic and financial forecasting.
- Chapter 4 provides a description of the construction of layering DEA for failure prediction. This chapter provides a brief description of basic DEA as well as its limitation, and defines the respective terminology and mathematical formulation of the layering worst practice DEA production model.
- Chapter 5 presents the hybrid rough set/DEA classification model. It discusses the data and the selected companies used in this research. The data

acquisition process, limitations and related assumptions are provides.

- Chapter 6 presents our results and discusses the classification accuracy of the proposed model.
- Chapter 7 provides conclusions and suggestions for future research.
- Appendix A – Contains the financial information of the public companies used in the DEA analysis.
- Appendix B – Contains the risk level prediction obtained through the hybrid model.
- Appendix C – Contains the distress related regulations in Taiwan.
- Appendix D – Contains the LINGO code for rule induction

Chapter 2 Methods Review for Business Failure Prediction

This section provides a thorough review of the literature in the area of business failure prediction and credit risk analysis. This section reviews the relevant literature and outlines the traditional analytical approaches used for this purpose. The section also highlights the shortcomings associated with some of the methodologies.

2.1 Overview of Business Failure Prediction Models and Classification Tools

2.1.1 Statistical Techniques

Over the last thirty years, many statistical techniques have been used for business failure prediction.

2.1.1.1 Discriminant Analysis

Altman's multivariate study in 1968 built on the findings of Beaver (1966) by combining several measures into a predictive model with the use of a statistical technique called multiple discriminant analysis (MDA). The discriminant function is the following:

Altman's Z-Score

Z=1.2X_1 + 1.4X_2 + 3.3X_3 + 0.6X_4 + 1.0X_5

X_1= Working Capital / Total Assets

X_2 = Retained Earnings / Total Assets

X_3 = Earnings Before Interest and Taxes / Total Assets

X_4 = Market Value of Equity / Book Value of Total Liabilities

X_5 = Sales / Total Assets

Using this equation, a Z score can be calculated for any company.

Z > 2.67 the model classifies the company as healthy.

Z < 1.81 the model classifies the company as becoming distressed.

According to Altman, bankruptcy could be explained quite completely by using a combination of five (selected from an original list of 22) financial ratios. Altman utilized a paired sample design, which incorporated 33 pairs of manufacturing companies. The pairing criteria were predicated upon size and industrial classification. The classification of Altman's model based on the value obtained for the Z score has a predictive power of 96% for prediction 1 year prior to bankruptcy.

2.1.1.2 Logit Analysis

Ohlson (1980) used a logit of the maximum likelihood method to build and analyze a model, which sampled 105 failed companies and 2058 non-failed companies during 1970 to 1976. He set up 3 models from 9 explanatory variables to predict corporate failure. From this it was possible to identify four basic factors as being statistically significant in affecting probability of failure (within one year). These are: (1) the size of the company; (2) measures of the financial structure; (3) measures of performance; and (4) measures of current liquidity.

Keasey and Watson (1987) employed logit to build a prediction model. They sampled 73 failed companies and 73 non-failed companies from 1970 to 1983, using 28 financial variables and 18 non-financial variables in their study. For the logit functions presented below, the dependent variable is failure / non-failure and the sets of independent variables are financial variables.

These conventional statistical methods, however, have some restrictive assumptions such as the linearity, normality and independence among predictor or input variables. The most common assumption, the one that is required for discriminant analysis, is the assumption of multivariate normality. However many financial ratios are not normally distributed from the fact that they are bounded on one side. Considering that the violation of these assumptions for independent variables frequently occurs with financial data, the methods can have limitations to obtain the effectiveness and validity.

2.1.2 Artificial Intelligence and Neural Networks

Recently, numerous studies have demonstrated that artificial intelligence such as inductive learning and neural networks can be an alternative method for classification problems to which traditional statistical method have long been applied. In neural networks simple nonlinear processing elements are interconnected in a large network in which an input signal is propagated towards one or more designated output nodes (Hertz et al., 1991). Because neural networks are capable of identifying and representing non-linear relationships in the data set, they have been studied extensively in the fields of financial problems including bankruptcy prediction (Atiya, 2001; Barniv et al., 1997; Bell, 1997; Boritz & Kennedy, 1995; Charalambous et al., 2000; Etheridge & Sriram, 1997; Fletcher & Goss, 1993; Grice & Dugan, 2001; Lee et al., 1996; Leshno & Spector, 1996; Odom & Sharda, 1990; Salchenberger et al., 1992; Tam & Kiang, 1992; Wilson & Sharda, 1994; Zhang et al., 1999)

Neural networks fundamentally differ from statistical models. Parametric statistical models require the developer to specify the nature of the functional relationship such as linear or logistic between the dependent and independent variables. Once an assumption is made about the functional form, optimization techniques are used to determine a set of parameters that minimizes the measure of error. In contrast, neural networks with at least one hidden layer use data to develop an internal representation of the relationship between variables so that a priori assumptions about underlying parameter distributions are not required. As a consequence, better results might be expected with neural networks when the relationship between the variables does not fit the assumed model (Salchenberger et al., 1992).

The first attempt to use neural networks for bankruptcy prediction is done by Odom and Sharda (1990). They took a research sample of 65 bankrupt firms between 1975 and 1982, and 64 non-bankrupt firms, overall 129 firms. Among these, 74 firms (38 bankrupt and 36 non-bankrupt firms) were used to form the training set, while the remaining 55 firms (27 bankrupt and 28 non-bankrupt firms) were used to make holdout sample. As a result, neural networks correctly classified 81.81% of the hold out sample. Zhang et al. (1999) also compared a neural network models' performance with a logit model, and employed a five-fold cross-validation procedure, on a sample of manufacturing firms. The robustness and performance of the neural network model improved significantly from small sets to large sets. Overall, neural networks are comparable to their statistical counterparts. For real-world problems with high nonlinearity neural networks usually perform better at prediction and classification accuracy. Furthermore, neural network models are more robust, more easily adaptive to a changing environment, and less sensitive to changes in sample size, number of variables, and data distribution.

Neural networks are useful when a large amount of data has to be modeled and a physical model is not known well enough to use statistical methods. However, predicted outputs of the model are limited to the scope of the training set used. As they are not able to discover new relationships in the data, neural nets are not true data mining (Kittler and Wang, 1999).

In recent research works (Yasdi, 1995; Hashemi et al. 1998; Ahn et al. 2000), the rough set theory combined with neural network has been used in economic and financial prediction. In these hybrid models, the rough set theory took the role of preprocessor for the neural network by reducing the decision table. This is very useful for neural network in that reduction of attributes prevents overfitting and saves training time. Furthermore, removing conflicting objects and training neural network with consistent cases can improve the performance as well as reduce the training time. Ahn et al. (2000) applied this hybrid model to predict the business failure for over 1200 healthy firms and 1200 failed firms in Korea. The results showed this hybrid model outperformed discriminant analysis model and neural network model.

A major disadvantage of neural network is their lack of transparency. The internal structure of the network is hidden and may not be easy to duplicate. This leads to a lack of accountability because the system’s intermediate steps cannot be checked.

2.1.3 The Rough Set Theory

The concept of rough sets theory (RST) is based on the assumption that every object of the universe is associated with some information. Objects characterized by the same information are indiscernible (similar) in view of their available information. The rough sets theory provides a technique of reasoning from imprecise data, discovering relationships in data and generating decision rules. Szladow and Mills (1993) presented a comparative study of rough set model against multivariable discriminant analysis for prediction of corporate bankruptcy from five financial ratios, namely, working capital, retained earnings, earnings before interest and taxes, market value of equities and sales to total assets volumes.

The application of the RST in business failure prediction was investigated by Slowinski et al. (2000) and Dimitras et al. (1999). In their works, the RST was tested for its prediction ability and was compared with three other methods, namely, C4.5

inductive algorithm, discriminant analysis and logit analysis. Dimitras et al. (1999) employed indiscernibility relationships for a sample of 80 Greek firms even though variables were preference ordered. The comparison of predictive accuracy with the discriminant analysis also showed that the rough sets approach was a strong alternative. More recently, Beynon and Peel (2001) employed the variable precision rough sets model to predict between failed and non-failed UK companies. The results are compared to those generated by the classical logit and multivariate discriminant analysis, together with non-parametric decision tree methods. The above comparison results showed that the prediction model based on the RST has more advantages over classical statistical models, such as discriminant analysis, logit analysis and probit analysis.

In summary, rough set has the following advantages (Dimitras et al., 1999; Greco et al., 1998):

∗ It is based on the original data only and does not need any external information, unlike probability in statistics or grade of membership in fuzzy set theory;

∗ It discovers important facts hidden in data and expresses them in the natural language of decision rules;

∗ The set of decision rules derived by the Rough Set Theory gives a generalized description of the knowledge contained in the financial information tables, eliminating any redundancy typical of the original data;

∗ The decision rules obtained from the Rough Set Theory are based on facts, because each decision rule is supported by a set of real examples;

∗ The results of the Rough Set Theory are easy to understand, while the results from other methods (credit scoring, utility function and outranking relation) require an interpretation of the technical parameters, with which the user may not be familiar.

2.1.4 Data Envelopment Analysis

DEA is a nonparametric programming method. The methodology compares the ability of a DMU (decision making unit) to convert its inputs to outputs with "similar" DMUs. In this case, similar means those DMUs with measurable inputs and outputs that operate in the same environment.

The majority of the research in the area of failure prediction using DEA has been in the banking industry. Most work in this area involved the measurement of bank efficiency with some emphasis on forecasting bank failure (Barr 1993). Barr et al. (1993) analyzed 930 banks over five years, which validated this approach and showed that the DEA scores for the surviving institutions are significantly higher than the scores for the failing banks. Sima (1997) was the first study which used DEA for predicting bankruptcy of publicly traded companies using non-ratio inputs and outputs. The results outperform the results of the popular Z score model. Pendharkar (2002) showed how to use the sensitivity analysis procedure in DEA model to solve the inverse classification on bankruptcy prediction. Cielen et al. (2004) compared the bankruptcy classification performance of a linear programming model, a DEA model and a rule induction (C5.0) model.

Paradi et al. (2004) use a layered worst practice DEA technique, where the sequential layers of poor performance are found with decreasing risk rating. This layering technique enables incorporation of risk attitudes and risk-based pricing. A limitation of these DEA studies is that the difficulty to treat qualitative data, therefore most these studies uses quantitative financial data.

The approach adopted in this research, DEA, has several advantages (Charnes et al 1996) :

- It gives a single measure of performance, which can take into account all dimensions of corporate activity, by simultaneously handling multiple inputs and outputs without making judgments on their relative importance;

- It does not require an a priori specification of a functional form for the input-output relation;
- Focuses on individual observations in contrast to population averages;
- Focuses on revealed best-practice frontiers rather than on central-tendency properties of frontiers;
- It ensures that the firms being examined are compared to those that have similar environment (its peers), as indicated by the financial data;
- Value free and does not require specification or knowledge of a priori weights or prices for the inputs and outputs.

In most studies the DEA approach has been used as a tool for evaluating accomplishments in the past. The results highlight the status of the operational performance and are helpful for planning future activities for improving the performance. In this paper we use DEA as a prediction and risk analysis tool for the public companies in Taiwan. The results are regarded as forward-looking information to enhance the decision quality.

The same characteristics that make DEA a powerful tool can also create problems. An analyst should keep these limitations in mind when choosing whether or not to use DEA.

- Since DEA is an extreme point technique, noise data such as measurement error can cause significant problems;
- DEA is good at estimating "relative" efficiency of a DMU but it converges very slowly to "absolute" efficiency. In other words, it can tell you how well you are doing compared to your peers but not compared to a "theoretical maximum" ;
- Since DEA is a nonparametric technique, statistical hypothesis tests are difficult and are the focus of ongoing research;
- Since a standard formulation of DEA creates a separate linear program for each DMU, large problems can be computationally intensive;
- The lack of discrimination among efficient DMUs that occurs when the number of DMUs is small in comparison with the total number of variables in the analysis.

2.2 Failure Prediction Using Financial Statement Data

There is a long history of research attempting to develop bankruptcy prediction models based on financial variables and other indicators of financial distress. This research is listed below in an approximate historical development order:

* Univariate ratio models (Beaver, 1966);

* Multiple discriminant analysis (Altman, 1968);

* Multivariate conditional probability models (Ohlson, 1980);

* Linear programming (Wallin and Sundgren, 1995);

* Expert Systems (Messier and Hansen, 1998; Shaw and Gentry, 1990; Chung and Tam, 1992);

* Neural networks (Bell et al., 1997; Hansen and Messier, 1991; Tarn and Kiang, 1992; Koh and Tan, 1999);

* Rough sets theory (Slowinski and Zopounidis, 1995; McKee, 1998);

* Genetic programming (McKee and Lensberg, 2002).

In most bankruptcy identification studies, financial ratios are the key variables that are used to explain differences between failed and non-failed businesses. The first modern analysis of distress indicators was performed by Beaver in 1967. His major finding was that financial ratios have failure predictive capabilities for at least five years prior to failure.

2.3 Failure Prediction Using Market Value Information

The underlying concept of this methodology is that the liabilities of a firm are claims on the firm's assets. If the market value of the assets at time T (A_T) is below the interest and principle payments (L) due at time T, then it is reasonable to assume that the firm will default on the obligation and the value of the equity is zero. On the other hand, if the assets are worth more than the repayment amount, the company can make the payment and the value of the equity becomes $A_T - L$. Therefore, the model gives the value of the firm's equity at time T to be :

$$\text{Max}\ (A_T - L,\ 0)$$

This is a call option on the value of the company's assets with a strike price equal to the repayment value of the debt.

The main weakness of this methodology is that it does not take into account liquidity consideration. A firm might have sufficiently valuable assets, but they are not liquid enough to meet payment obligations as they become due.

2.4 Failure Prediction Using Non-Financial Data Information

A company's performance and future may be influenced by characteristics other than financial data. Zopounidis (1997) employs a set of 'strategic criteria' to assess the risk of failure of French firms. These were: quality of management, research and development level, diversification stage, market trend, market niche/position, cash out method and world market share.

Keasey and Watson (1989) utilized a number of non-financial variables, either alone or in conjunction with financial ratios, and were able to predict a small company's failure more accurately than models based solely upon financial ratios. Similar propositions were taken by Shaw and Gentry (1990), while Peel et al. (1989) proposed other qualitative variables, such as the changes in the lag in reporting accounts of a firm, the number of director resignations, and appointments and the changes in directors' shareholdings.

Previous studies have reported that failing firms are more likely to be characterized by auditor switches (up to three years before failure), largely in consequence of disputes between auditors and managers over accounting methods, together with disagreements in respect of audit opinions/ qualifications (Morris 1997). The social importance of the firm and the strength of bank relationship (Suzuki 2004 and Wright 2000) could be also critical. The decomposition measures analysis were suggested by Booth et al. (1989).

Chapter 3 Approximation of Decision Rules in Rough Set Concept

The focus of this chapter is on the development of efficient techniques for rule generation based on rough set theory. A new optimization model for feature reduction, which can select the desired reducts will be developed.

Rough set theory is a formal construction to transform data into knowledge. A set of data is generally complicated and disorganized but knowledge is not. The framework of rough set is related to data reduction to diminish the level of redundancy or noise, and to facts discovery to make the data more observable. One of the main advantages of rough set theory is that it does not need any preliminary or additional information about data, such as probability distribution in statistics, or grade of membership or the value of possibility in fuzzy set theory.

3.1 Basic rough Set Concept

The concept of Rough Sets Theory (Pawlak, 1981) is based on the assumption that every object of the universe is associated with some information. Objects characterized by the same information are indiscernible (similar) in view of their available information. Rough sets theory provides a technique of reasoning from imprecise data, discovering relationships in data and generating decision rules. In this section, some basic notation of rough set for complete information system, which is related to our research will be briefly introduced. For more detailed introduction, see Pawlak (1991).

3.1.1 Information Table

Information about objects is represented in the form of an information table. The rows of the table are labeled by objects, whereas columns are labeled by attributes and entries of the table are attribute values. An information table, in which the set of attributes is split into condition and decision attributes is called a decision table. An information table may be defined by the 4-tuple $S = < U, A, V, f >$, where U is the universe (a finite set of objects), A is a finite set of attributes (features, variables), V is the value of attribute and f is the total decision function called the information function. For any application of rough set, the first step is to transfer the original data into a decision table.

For example, consider the following diagnosis data (Table 3.1). We transform it into a rough set decision table

Table 3.1 Diagnosis data example

	Auditor_Changed	CEO_Changed	Debt	Bankruptcy
e_1	Yes	Yes	normal	no
e_2	Yes	Yes	High	yes
e_3	Yes	Yes	Very high	yes
e_4	No	Yes	normal	no
e_5	No	No	High	no
e_6	No	Yes	Very high	yes

Using the terminology of rough set theory, this data set can be considered as an information system $T=(U, A \cup d)$, where universe U, attributes A and decision feature d are:

$U = \{e_1, e_2, e_3, e_4, e_5, e_6\}$, where e_i is object, $i = 1,\ldots,6$

$A = \{F_1, F_2, F_3\}$

Where

F_1 = Auditor_Changed

F_2 = CEO_Changed

F_3 = Debt

d = Present or absence of Bankruptcy (decision)

The domains of the particular attributes are:

$V_1 = \{0, 1\}$, 0 = No, 1 = Yes

$V_2 = \{0, 1\}$, 0 = No, 1 = Yes

$V_3 = \{0, 1, 2\}$, 0 = Normal, 1 = High, 2 = Very High

Rows of a table, labeled e_1, e_2, e_3, e_4, e_5, and e_6 in Table 3.1, are called examples (objects, entities). Properties of examples are perceived through assigning values to some variables. We will distinguish between two kinds of attributes: condition attributes and decision attributes. That is, the domain of each attribute is the set of values of that attribute. The decision table *T* for this system is presented in Table 3.2.

Table 3.2 Decision table of example

	Auditor_Changed	CEO_Changed	Debt	Bankruptcy
e_1	1	1	0	0
e_2	1	1	1	1
e_3	1	1	2	1
e_4	0	1	0	0
e_5	0	0	1	0
e_6	0	1	2	1

3.1.2. Indiscernibility

Indiscernibility means similarity and it is the mathematical basis of rough sets theory. It is normally associated with a set of attributes. Let $S = (U, Q, V, \rho)$ be an information system and let $P \subseteq Q, x, y \in U$, so that x and y are indiscernible by the set of attributes P in S ,denoted by $x\tilde{P}y$, iff $r(x,q) = r(y,q)$ for every $q \in P$ (Dimitras et al., 1999). The equivalence classes of relation $\tilde{P}$ (or IND_P) are called $P-$elementary sets in S, whereas the $Q-$elementary sets are called atoms in S.

These elementary sets represent the smallest discernible groups of objects, and the construction of elementary sets is the primary step to perform the classification through rough sets (Walczak & Massart, 1999).

For example, in the set consisting of attributes Auditor_Changed and CEO_Changed from Table 3.1, objects e_l and e_2 are characterized by the same values of both attributes: for the attribute Auditor_Changed the value is yes for e_1 and e_2 and for the attribute CEO_Changed the value is yes for both e_1 and e_2. Moreover, example e_3 is indiscernible from e_1 and e_2. Examples e_4 and e_6 are also indiscernible from each other. The indiscernibility relation is an equivalence relation. Sets that are indiscernible are called elementary sets. Thus, the set of attributes Auditor_Changed and CEO_Changed defines the following elementary sets: $\{e_l, e_2, e_3\}$, $\{e_4, e_6\}$, and $\{e_5\}$.

Objects in elementary sets are those that can be clearly distinguished in terms of the available information or knowledge. However, in practice, sets of objects will probably not be determined unambiguously (by an elementary set), hence, objects will have to be described roughly through a pair of sets: i.e. a lower and an upper approximation. The lower approximation contains all objects that certainly belong to that category. The upper approximation consists of all objects that possibly belong to that category. The boundary region is the group of objects that cannot be decisively assigned as being either a member or a non-member of that category. By using the lower and upper approximation of a set, we can define the accuracy and the quality of approximation which are numbers from interval $[0,1]$ (Pawlak, 1984). A rough set is thus any subset defined through its lower and upper approximation.

3.1.3. Reduct and Core

In large data sets some attributes may be redundant, and thus can be eliminated without losing classification information. By definition, a reduct is the minimal subset still providing the same object classification as with the original set of attributes. The intersection of all reducts is called the core. The core is a collection of the most relevant attributes in the table. Let the $S=(U,Q,V,\rho)$ be an information system and let $P,R\in Q$. Then, the set of attributes P is said to be dependent on set of attributes R in S (denotation $R\to P$) iff $IND_R\subseteq IND_P$, whereas the set of attributes P,R are called independent in S iff neither $R\to P$ nor $P\to R$ hold (Pawlak, 1982).

Moreover, finding the reduction of attributes is another important thing. Let the minimal subset of attributes $R\subseteq P\subseteq Q$ such that $\eta_P(\ddot{Y})=\eta_R(\ddot{Y})$ is called $\ddot{Y}-$reduct of P and is denoted by $RED_{\ddot{Y}}(P)$. Then the intersection of all $\ddot{Y}-$reducts is called the $\ddot{Y}-$core of P. Especially, the core is a collection of the most relevant attributes in the table, and is the common part of all reducts.

In the example from Table 3.1, let the set of attributes be the set {Auditor_Changed, Debt} and its superset be the set of all three attributes, i.e., the set {Auditor_Changed, CEO_Changed, Debt}. Elementary sets of the indiscernibility relation defined by the set {Auditor_Changed, Debt} are singletons, i.e., sets $\{e_1\}$, $\{e_2\}$, $\{e_3\}$, $\{e_4\}$, $\{e_5\}$, and $\{e_6\}$, and so are elementary sets of the indiscernibility relation defined by the set of all three attributes. Thus, the attribute CEO_Changed is redundant. Table 3.3 presents a new information table based on this reduct.

Table 3.3 Information reduct table

	Auditor_Changed	Debt	Bankruptcy
e_1	1	0	0
e_2	1	1	1
e_3	1	2	1
e_4	0	0	0
e_5	0	1	0
e_6	0	2	1

3.1.4. Decision Rule

One of the most important reasons for applying rough sets is the generation of decision rules. The decision rule reflects a relationship between a set of conditions and a conclusion or a decision. Decision rules derived from a decision table can be used in predictions concerning new objects. A decision rule can be expressed as a logical statement:

IF conjunction of elementary conditions

THEN disjunction of elementary decisions

An information system can be seen as a decision table in the form of $S=(U,C\cup D,V,\rho)$, in which $C\cup D=Q$ dictating that condition attributes C and decision attributes D are two disjoint classes of attributes (Greco et al., 2002). Through analyzing the decision table, valuable decision rules can be extracted. To generate decision rules from the data in the decision table, it is required to reduce unnecessary conditions. According to Pawlak (2002), a decision rule in S is an expression $\Phi \to \Psi$, read if Φ then Ψ, where Φ and Ψ are conditions and decisions of the decision rule, respectively; most importantly, $\sigma_S(\Phi,\Psi) = \mathrm{supp}_S(\Phi,\Psi)/card(U)$ is the strength of the decision rule $\Phi \to \Psi$ in S, where the $\mathrm{supp}_S(\Phi,\Psi)$ is called the support of the rule $\Phi \to \Psi$ in S.

Decision rules induced from a decision table can be applied to classify new objects. The classification of a new object can be supported by matching its description to one of the decision rules. The matching may lead to one of the four situations (Tay and Shen 2002):

(i) The new object matches exactly one of the deterministic decision rules;

(ii) The new object matches exactly one of the non-deterministic decision rules;

(iii) The new object does not match any of the decision rules;

(iv) The new object matches more than one rule.

Each condition rule is characterized by the strength of its suggestion, which means the number of objects satisfying the rule. The modular nature of decision rules makes it easy for decision makers to insert new decision rules or to modify existing decision rules without affecting the original information.

Problems of inducing decision rules have been extensively investigated in many fields, particularly in the machine learning domain. Rough set can also be applied to different stages of rule induction and data processing. However, one aspect that distinguishes rough set from typical machine learning systems is that rough set does not correct or aggregate the inconsistency in the input data. Lower and upper approximation are applied to describe the inconsistency, and consequently, certain and approximate rules are induced.

By analogy with attributes, we can define elementary sets associated with the decision as subsets of the set of all examples with the same value of the decision. Such subsets will be called concepts. For Tables 3.2 and 3.3, the concepts are $\{e_1, e_4, e_5\}$ and $\{e_2, e_3, e_6\}$. The first concept corresponds to the set of all companies free from Bankruptcy, the second one to the set of all companies sick with Bankruptcy. The question is whether we may tell who is free from Bankruptcy and who is sick with Bankruptcy on the basis of the values of attributes in Table 3.3. To answer this question, we may observe that in terms of rough set theory, decision Bankruptcy depends on attributes Auditor_Changed and Debt, since all elementary sets of indiscernibility relation associated with {Auditor_Changed, Debt} are subsets of some concepts. As a matter of fact, one may induce the following rules from Table 3.3:

(Debt, normal) -> (Bankruptcy, no),

(Auditor_Changed, no) and (Debt, high) -> (Bankruptcy, no),

(Auditor_Changed, yes) and (Debt, high) -> (Bankruptcy, yes),

(Debt, very_high) -> (Bankruptcy, yes).

In the real world, data sets are very large. We need a systematic procedure to derive all possible rule-reducts. A decision rule generation algorithm based on the developments in Pawlak (1991) is as follows:

Step 0. Initialize object number i =1, feature number j = 1.

Step 1. Select feature j = 1~n, for all

$k \neq i, if\ a_{ij} \neq a_{kj}\ or\ a_{ij} = a_{kj} \wedge d_i = d_k$ then a_{ij} can generate r - reduct

, if all found, go to step 2;

Step 2. Set i = i +1, If all objects have been considered, go to Step 3; Otherwise, go to Step 1.

Step 3. Select two features and go to Step 1 until all (n-1)-features r-reducts have been considered.

3.2 Approximation of Decision Rule

The focus of this section is on the development of efficient techniques for rule generation based on rough set theory. A new optimization model for feature reduction, which can select the desired reducts will be developed. The goal is to provide techniques for extracting useful information from large data sets in a very short time, which is the most important challenge of data mining and rule induction research areas.

Procedures for derivation of decision rules from decision table were presented by Grzymala-Busse (1992), Skowron (1993) and Ziarko et al. (1993). More advanced rule induction methods have been studied in Bazan (1998) for comparing the dynamic and non-dynamic methods of induction rules from decision tables. Grzymala-Busse and Stefanowski (1998) carried out work in this area with the focus on induction rules from inconsistent decision table. Lin (1996) and Lin and Yao (1996) studied the rule induction from very large databases combined with database technologies.

3.2.1 Approximation of Sets and Approximation of Accuracy

In the rough set theory, the approximations of sets are introduced to deal with the vague concept. Let $P \subseteq Q$ and $Y \subseteq U$. The lower approximation of Y is denoted by $\underline{P}Y = \{x \mid IND_P(x) \subseteq Y\}$, the $P-$ upper approximation of Y is denoted by $\bar{P}Y = \{x \mid IND_P(x) \cap Y \neq \phi\}$, and the $P-$boundary of set Y is the doubtful region denoted by $BN_P(Y) = \bar{P}(Y) - \underline{P}(Y)$. If the lower and upper approximations are identical i.e. $\bar{P}Y = \underline{P}Y$, the set Y is definable; otherwise, the set Y is undefinable in S. The set $\underline{P}Y$ is the set of elements of U, which can be certainly classified as elements of Y by the set of attributes P; the set $\bar{P}Y$ is the set of elements of U, which can be possibly classified as elements of Y by the set of attributes; and the set $BN_P(Y)$ is the set of elements, which cannot be certainly classified to Y by the set of attributes P.

According to Pawlak (1982), the accuracy of the approximation $\mu_P(Y)$, the quality of classification $\eta_P(\ddot{Y})$, and the accuracy of the classification $\beta_P(\ddot{Y})$ can be measured as follows. To measure the accuracy of the approximation $\mu_P(Y)$ of the set Y by P in S, we can use the way of that $\mu_P(Y) = card(\underline{P}Y) / card(\bar{P}Y)$, in which $0 \leq \mu_P(Y) \leq 1$; the Y is definable by P in S if $\mu_p(Y) = 1$, whereas the Y is undefinable by P in S if $\mu_P(Y) < 1$. In addition, let S be an information system, a subset of attributes $P \subseteq Q$; and let $\ddot{Y}$ be the classification of U by P, the subsets $Y_i \in \{Y_1, Y_2, \ldots, Y_m\}$ are the classes of the classification $\ddot{Y}$, the $P-$lower approximation of $\ddot{Y}$ is denoted as $\underline{P}\ddot{Y}$, and the $P-$upper approximation of $\ddot{Y}$ is denoted as $\bar{P}\ddot{Y}$. Then, the accuracy of the classification $\beta_P(\ddot{Y})$ by P it can be measured with the way that $\beta_P(\ddot{Y})=\sum_{i=1}^{m}card(\underline{P}Y_i)\Big/\sum_{i=1}^{m}card(\bar{P}Y_i)$; the higher ratio of $\beta_P(\ddot{Y})$ means that the classification is less ambiguous. As to the quality of classification $\eta_p(\ddot{Y})$ by P can be measured with the way that $\eta_P(\ddot{Y})=\sum_{i=1}^{m}card(\underline{P}Y_i)\Big/card(U)$; the higher ratio of $\eta_p(\ddot{Y})$ means that the classification is better correctly.

3.2.2 Current Approximation Method

One of the central problems of rough sets theory is classification analysis. Rough sets theory can be used to do classification but the classification must be completely correct or certain. In practice, however, some level of uncertainty in the classification process may lead to a better utilization of properties of the data being analyzed. This approach operated in the context of the variable precision rough sets (VPRS) model in the literature. Ziarko (1993) constructed VPRS, which includes a probabilistic generalisation on Rough sets theory. He extended rough sets theory by introducing a probability value β. The β value represents a bound on the conditional probability of a proportion of objects in a condition class which are classified to the same decision class. This type of reduct preserves the sum of objects in β lower approximations of all decision classes. But the derived decision rules from the β-reduct may be in conflict with the ones from the original system. To overcome this kind of drawback, we introduce new concepts of rule reduction.

Unlike most of the rough set based classification and rule induction methods which induce knowledge from lower and upper approximation concepts or from tedious procedures of finding reducts and cores, the rule induction techniques proposed in our research applies rule reduct algorithm to extract knowledge directly from the minimal set of attributes.

3.2.3 Proposed Rule Induction Model

3.2.3.1 Presentation of Data and Rules

Here we use an example to illustrate the way of presenting data and rules. Consider a data set in Table 3.4 which has 5 objects, (x_1, x_2, x_3, x_4, x_5), four attributes (a_1, a_2, a_3, a_4) and one group index g. The domain values of a_1, a_2, a_3, a_4 are respectively {1, 2, 3}, {1, 2}, {1, 2, 3, 4} and {1, 2, 3}. The domain value of g is {g_1, g_2, g_3}. This study intends to deduce the classification rules for the objectives with a specific group.

Table 3.4 Decision table for proposed mode data.

U	a_1	a_2	a_3	a_4	g
x_1	2	1	3	3	1
x_2	3	2	1	1	2
x_3	2	2	3	3	2
x_4	1	1	4	2	3
x_5	3	1	2	1	3

Table 3.5 Binary value table converted from decision table

U		a_1			a_2		a_3				a_4		G_k
	a_{11}	a_{12}	a_{13}	a_{21}	a_{22}	a_{31}	a_{32}	a_{33}	a_{34}	a_{41}	a_{42}	a_{43}	
x_1	0	1	0	1	0	0	0	1	0	0	0	1	1
x_2	0	0	1	0	1	1	0	0	0	1	0	0	2
x_3	0	1	0	0	1	0	0	1	0	0	0	1	2
x_4	1	0	1	1	0	0	0	0	1	0	1	0	3
x_5	0	0	1	1	0	0	1	0	0	1	0	0	3

Firstly, we can convert this data set into a new one presented by binary values as showing in Table 3.5. An objective x_i can then be written as

$x_i \ = \ (a^i_{11}, a^i_{12}, a^i_{13}, a^i_{14}\ ; a^i_{21} \ldots ; \ldots), x_i \in G_k$

(means x_i belongs to group k) For instance, x_1 is expressed as x_1 = (0, 1, 0; 1, 0; 0, 0, 1, 0; 0, 0, 1), $x_1 \in G_1$

Denote $R_l(k)$ as the *l'th* rule of classifying the *k'th* group from others.

$R_l(k)$ can be expressed as a binary vector below:

$R_l(k) = \ (d^l_{11}, d^l_{12}, d^l_{13}, d^l_{14}\ ; d^l_{21} \ldots ; \ldots)$

where

$d^{\,l}_{jp} = 1$ if attribute a_{ij} is chosen in classifying group k from others.

$d^{\,l}_{jp} = 0$ if attribute a_{ij} is not chosen in classifying group k from others.

Such an expression is very useful in expressing rules with conjunction and disjunction rules.

For the data set contains n objects, m attributes where each attribute a_j having $g(j)$ levels. All those n objects belong to p groups. A general form for expressing an object x_i is written as

$$x_i = (a^i_{11}, a^i_{12} \ldots a^i_{1q(1)}\,; a^i_{21} \ldots a^i_{2q(2)}\,; \ \ldots; a^i_{m1} \ldots a^i_{mq(m)}), x_i \in G_k \qquad (3.1)$$

where a^i_{jp} are binary values.

A general form of expressing a rule $R_l(k)$, which is the l'th rule of classifying k'th group, is expressed as:

$$R_l(k) = (d^i_{11}, d^i_{12} \ldots d^i_{1q(1)}\,; d^i_{21} \ldots d^i_{2q(2)}\,; \ \ldots; d^i_{m1} \ldots d^i_{mq(m)}) \ , \qquad (3.2)$$

where d^i_{jp} are binary values.

For the small example in Table 3.4, we can list intuitively related classification rules in Table 3.5. Descript as follows:

(1) There may have more than one classification rule for a specific group. For instance, both $R_1(1)$ and $R_2(1)$ are used to classify the 1st group.

(2) A rule with more supporting objects is better than a rule with less supporting objects. For instance, $R_1(2)$ is better than $R_2(2)$.

Table 3.6 Decision rules table

Rules	d_{11}………… d_{43}	Meaning	Support	AR	SR	CR
$R_1(1)$	$d_{12} = 1$ & $d_{21} = 1$	If a_1=2 & a_2=1 then g = 1	x_1	1	1	1/6
$R_2(1)$	$d_{21} = 1$ & $d_{43} = 1$	If a_2=1 & a_4=3 then g = 1	x_1	1	1	1/6
$R_1(2)$	$d_{22} = 1$	If a_2 = 2 then g = 2	x_2 , x_3	1	1	1/12
$R_2(2)$	$d_{31} = 1$	If a_3 = 1 then g = 2	x_2	1	0.5	1/12
$R_1(3)$	$d_{32} = 1$	If a_3 = 2 then g = 3	x_5	1	0.5	1/12
$R_2(3)$	$d_{34} = 1$	If a_3 = 4 then g = 3	x_4	1	0.5	1/12
$R_3(3)$	$d_{32} = 1$ & $d_{34} = 1$	If a_3 = 2 or 4 then g = 3	x_5 , x_4	1	1	1/6

(3) A rule may be obtained by integrating related rules, thus to have more supporting objects. For instance, $R_3(3)$ is the integration of $R_1(3)$ and $R_2(3)$, which is compacter and is supported by more objects.

(4) The rules can be expressed in both the conjunction and the disjunction forms. For instance, $R_3(3)$ is expressed in disjunction form which $R_1(1)$ and $R_2(1)$ are in conjunction form.

Propositions

Proposition 1

Given n objects x_i (i=1,2,…n), either $x_i \in G_k$ or $x_i \notin G(k)$. There is a rule $R_l(k)$ expressed in (3.2), $R_l(k)$ can separate these n objects into groups if following conditions are satisfied.

(i) $\sum_p a^i_{jp} d^l_{jp} = 1$ for all i where $x_i \in G(k)$ and for all j where $\sum_p d^l_{jp} \geq 1$.

(ii) $\sum_p a^r_{jp} d^l_{jp} = 0$ for all r where $x_i \notin G(k)$ and at least a j where $\sum_p d^l_{jp} \geq 1$.

Proof :

Clearly $\sum_p a^i_{jp} = 1$ and $\sum_p a^r_{jp} = 1$ for all j.

Case 1: For an object $x_i \in G_k$, if a^i_{jp} satisfies condition (i) then it will not satisfy condition (ii).

Case 2: For an object $x_i \notin G(k)$, if a^r_{jp} satisfies condition (ii) then it will not satisfy condition (i).

Such a rule $R_l(k)$ therefore can let all $x_i \in G_k$ fit condition (i) and let all $x_i \notin G(k)$ fit condition (ii). The proposition is then proven.

Theorem 1 For the rule $R_l(k)$ described in Proposition 1, if there are h criteria where $\sum_p d^l_{jp} = 0$, then

(i) $\sum_{j=1}^{m} \sum_{p} a^{i}_{jp} d^{l}_{jp} = m - h$ for all i where $x_i \in G(k)$

(ii) $\sum_{j=1}^{m} \sum_{p} a^{r}_{jp} d^{l}_{jp} \le m - h - 1$ for all r where $x_r \notin G(k)$

Take $R_1(1)$ in Table 3.6 for instance, to check Theorem 1, here $d_{12} = d_{21}$ and the number of criteria with $\sum_{p} d^{l}_{jp} = 0$ is 2. It is convenient to check that

(i) $a^{1}_{12} d_{12} + a^{1}_{21} d_{21} = 2$ for x_1

(ii) $a^{2}_{12} d_{12} + a^{2}_{21} d_{21} = 0 \le 1$ for x_2

$a^{3}_{12} d_{12} + a^{3}_{21} d_{21} = 0 \le 1$ for x_3

$a^{4}_{12} d_{12} + a^{4}_{21} d_{21} = 0 \le 1$ for x_4

$a^{5}_{12} d_{12} + a^{5}_{21} d_{21} = 0 \le 1$ for x_5

Similarly, checking $R_3(3)$ where $d_{32} = d_{34}$=1 and the number of criteria with $\sum_{p} d^{l}_{jp} = 0$ is 3, to have following results :

(i) $a^{4}_{32} d_{32} + a^{4}_{34} d_{34} = 1$ for x_4 ;

$a^{5}_{32} d_{32} + a^{5}_{34} d_{34} = 1$ for x_5 ;

(ii) $a^{1}_{32} d_{32} + a^{1}_{34} d_{34} \le 0$ for x_1 ;

$a^{2}_{32} d_{32} + a^{2}_{34} d_{34} \le 0$ for x_2 ;

$a^{3}_{32} d_{32} + a^{3}_{34} d_{34} \le 0$ for x_3 ;

We then have following remark:

Remark : A $x_i \in G(k)$ is said to "support" a rule $R_l(k)$ if x_i fit condition (i) of

Theorem 1. A $x_r \notin G(k)$ is said not to violet a rule $R_l(k)$ if x_i fit condition (ii) of Theorem 1.

Here we specify binary variable u_i and v_r defined as follows:

$u_i = 1$ if $x_i \in G(k)$ support $R_l(k)$, and otherwise $u_i = 0$.

$v_r = 1$ if $x_r \notin G(k)$ does not violet $R_l(k)$, and otherwise $v_r = 0$.

Proposition 2 For n subjects x_i and a rule $R_l(k)$, there exist u_i and v_r for satisfying following inequalities.

(i) $n(u_i - 1) + m - h \le \sum_{j=1}^{m} \sum_{p} a^{i}_{jp} d^{l}_{jp} \le m - h + n(1 - u_i)$ for $x_i \in G(k)$

(ii) $n(u_i - 1) + m - h \le \sum_{j=1}^{m} \sum_{p} a^{r}_{jp} d^{l}_{jp} \le m - h - 1 + n(1 - v_r)$ for $x_r \notin G(k)$

where u_i , v_r $\in \{0.1\}$

Proof : If u_i =1 then (i) is equivalent to (i) in Theorem 1.

If v_r =1 then (ii) is equivalent to (ii) in Theorem 1.

The followings are three criteria for evaluating the goodness of a value.

(i) The rule should be supported by most objects of a specific group. That is, the coverage rate of a good rules should be high.

(ii) The rule should be accurate. That is, the rule had better not be supported by the objects of non-specific groups. In other words, the accurate rate of a good rule should be high.

(iii) The rule should be expressed in a compacter way.

Consider a data set that contains *n* objects. Denote the number of objects belonging to a specific *k* group as *n*(*k*). By referring to (3.1) and (3.2), the meanings of the accuracy rate, the coverage rate, and the compactness rate are specified as following remarks:

<u>Remark 1</u>

The accuracy rate of a rule $R_l(k)$ is given by $AR_l(k) = \frac{1}{n - n(k)}(\sum_{x_r \notin G(k)} v_r)$ That means, if none of $x_r \notin G(k)$ violets the rule, then the accuracy rate of the rule is 1.

<u>Remark 2</u>

The support rate of a rule $R_l(k)$ is given by $SR_l(k) = \frac{1}{n(k)}(\sum_{x_r \in G(k)} u_i)$ That means, if all $x_i \in G(k)$ support the rule, then its support rate of the rule is 1.

<u>Remark 3</u>

The compactness of a rule $R_l(k)$ is specified as $CR_l(k) = \frac{\sum_{j,p} d_{jp}^{l}}{\sum_{j} j(p)}$.

AR, *SR* and *CR* for example rules are also listed in Table 3.6.

3.2.3.1 Models of Deducing Rules

From the basis of above discussion, a model of deducing classification rules is formulated below:

Min $\quad CR_l(k) = \dfrac{\sum_{j,p} d^{l}_{jp}}{\sum_{j} j(p)}$

s.t. (i) $AR_l(k) = \dfrac{1}{n-n(k)}\left(\sum_{x_r \notin G(k)} v_r\right) \geq a$

(ii) $SR_l(k) = \dfrac{1}{n(k)}\left(\sum_{x_i \in G(k)} u_i\right) \geq s$

(iii) Inequalities (i) of proposition 2.

(iv) Inequalities (ii) of proposition 2.

u_i, v_r $\in \{0,1\}$.

Take table 3.4 for instance, the model of deducing the rule for G = 1 is formulated below.

$$MIN \quad CR = \frac{\sum_{j,p} d_{jp}}{3+2+4+3}$$

$$s.t. \quad AR = \frac{1}{5-1}(v_2 + v_3 + v_4 + v_5) \geq a$$

$$SR = \frac{1}{1}u_1 \geq s$$

$5(u_1\text{-}1) + 4 - h \;\leq\; d_{12} + d_{21} + d_{33} + d_{43} \;\leq\; 4 - h + 5\,(1\text{-}u_1)$

$d_{13} + d_{22} + d_{31} + d_{41} \;\leq\; 4 - h\text{-}1 + 5\,(1\text{-}v_2)$

$d_{12} + d_{22} + d_{33} + d_{43} \;\leq\; 4 - h\ \text{-}1 + 5\,(1\text{-}v_3)$

$d_{11} + d_{21} + d_{34} + d_{42} \;\leq\; 4 - h\ \text{-}1 + 5\,(1\text{-}v_4)$

$d_{13} + d_{21} + d_{32} + d_{41} \;\leq\; 4 - h\ \text{-}1 + 5\,(1\text{-}v_5)$

where u_i, v_r, d_{jp} $\in$ {0.1}, h is non-negative integer.

By specifying a = 1 and s = 1, the solution obtained is :

$d_{12} = d_{21} = 1,$ all other $d_{jp} = 0, h = 2, u_1 = v_2 = v_3 = v_4 = v_5 = 1, AR = 1$ and $SR = 1,$
$CR = \frac{2}{12} = \frac{1}{6}.$

This rule is exactly $R_1(1)$ in Table 2 .

Similarily, the model of deducing the rule for G = 3 is formulated below

$$MIN \quad CR = \frac{\sum_{j,p} d_{jp}}{12}$$

$$s.t. \qquad AR = \frac{1}{5-2}(v_1 + v_2 + v_3) \geq a$$

$$SR = \frac{1}{2}(u_4 + u_5) \geq s$$

$$5(u_4 - 1) + 4 - h \leq d_{11} + d_{21} + d_{34} + d_{42} \leq 4 - h + 5(1 - u_4)$$

$$5(u_5 - 1) + 4 - h \leq d_{13} + d_{21} + d_{32} + d_{41} \leq 4 - h + 5(1 - u_5)$$

$$d_{12} + d_{21} + d_{33} + d_{43} \leq 4 - h - 1 + 5(1 - v_1)$$

$$d_{13} + d_{22} + d_{31} + d_{41} \leq 4 - h - 1 + 5(1 - v_2)$$

$$d_{12} + d_{22} + d_{33} + d_{43} \leq 4 - h - 1 + 5(1 - v_3)$$

$u_i, v_r, d_{jp} \in \{0.1\}$, h is non-negative integer.

By specifying a = 1 and s = 1, the solution obtained is

$d_{32} = d_{34} = 1,$ all other $d_{jp} = 0, h = 3, u_4 = u_5 = v_1 = v_2 = v_3 = 1, AR = 1$ and $SR = 1,$
$CR = \frac{2}{12} = \frac{1}{6}$

3.3. Numerical Example

In this section we illustrate the proposed model and compare the proposed rule induction procedures as well as results with the rough set reduction model. We investigate the relationships among alternative types of knowledge reduction in information systems. These results provide more flexible approaches to rule reduction based on rough sets model, which are significant both in the theoretic and applied perspectives.

Before the data analysis, it is required to construct the decision table. As shown in Table 3.7, the decision table contains 14 records characterized by one decision attribute (d) and four conditional attributes: C_1, C_2, C_3 and C_4. Further, the four attributes and their values are denoted as: $V_{c_1} = \{1,2,3\}$, $V_{c_2} = \{1,2,3\}$, $V_{c_3} = \{1,2\}$ and $V_{c_4} = \{1,2\}$.

Table 3.7 Decision table for inconsistent data.

U	C_1	C_2	C_3	C_4	d
O_1	2	2	1	2	0
O_2	1	1	1	1	0
O_3	1	1	1	2	0
O_4	2	3	2	2	0
O_5	1	2	1	1	0
O_6	3	1	1	1	1
O_7	2	2	1	1	1
O_8	2	3	2	1	1
O_9	3	3	2	2	1
O_{10}	1	3	2	1	1
O_{11}	2	2	2	1	1
O_{12}	1	2	2	2	1
O_{13}	3	2	1	2	1
O_{14}	3	1	2	1	1

3.3.1 Rough Set Rule Induction Approach

Step 1: Calculating the approximation.

The first step of data analysis using rough set theory is to calculate the approximations of decision classes. As shown in Table 3.8, each decision class is well describable due to its high accuracy of 1.000 shown in the last column. This is to say that all two decision classes are characterized exactly by those data in the decision table. In addition, there are totally 14 atoms in the decision table. On the whole, the accuracy of the entire classification is 1.000, and also the quality of the entire classification is 1.000.

Table 3.8 Lower and upper approximations

Class number	Number of objects	Lower approx.	Upper approx.	Accuracy
1	5	5	5	100%
2	9	9	9	100%

Step 2: Finding the reducts of attributes and the core of attributes.

In this step, the indiscernibility relation method is used for dealing with the reduction of attributes and finding the core of attributes, due to all the condition attributes are nominal attributes (unordered qualitative attributes) with linguistic values. Employing the indiscernibility relation method, it may find all potential reducts in the information table. As a result, we obtained two reducts of attributes and two core of attributes. These two reducts are: {C_1, C_2, C_4} and {C_1, C_3, C_4}. The core of attributes is the attributes {C_1, C_4}. This means that C_1 and C_4 are the most meaningful among the four attributes.

Step 3: Creating the decision rules.

The most important step of data analysis is to generate decision rules. In order to find the minimal covering rules, the minimal covering method is employed, which attempts to find the minimal number of attribute values for a decision rule. As a result, 6 rules are created. These 6 exact rules are shown in Table 3.9, from which we can acquire several valuable implications for making decisions. In particular, we can find the most important determinant for each decisions class through using the covering ratio of Covering Index (CI).

Table 3.9 Decision rules generated by rough set

	Rules generated by rough set		accuracy	CI
Rule 1	$(C_3 = 1)$ & $(C_1 = 1)$	=> $(D = 0)$	1.00	0.60
Rule 2	$(C_1 = 2)$ & $(C_4 = 2)$	=> $(D = 0)$	1.00	0.40
Rule 3	$(C_3 = 2)$ & $(C_4 = 1)$	=> $(D = 1)$	1.00	0.44
Rule 4	$(C_1 = 3)$	=> $(D = 1)$	1.00	0.44
Rule 5	$(C_1 = 2)$ & $(C_4 = 1)$	=> $(D = 1)$	1.00	0.33
Rule 6	$(C_1 = 1)$ & $(C_2 = 2)$ & $(C_4 = 2)$	=> $(D = 1)$	1.00	0.11

3.3.2 Proposed Rule Induction Approach

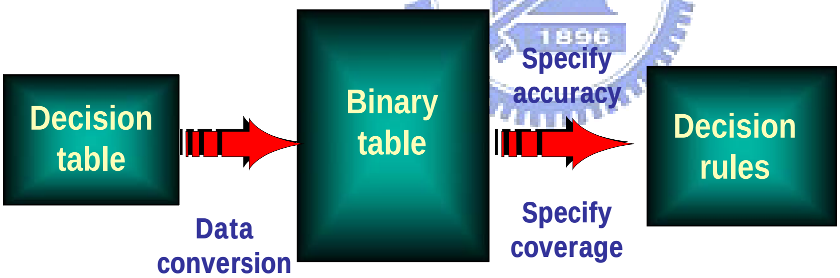

Figure 3.1.Proposed rule induction Model

Step 1: Convert the decision table to a binary value table.

First, we convert the decision table into a binary table which each attribute value is mapped to a binary value in the derived table.

Step 2: Generate the decision rules using integer programming

The most important step in data analysis is to generate decision rules. In order to find the maximal covering rules, the maximal covering method is employed, which attempts to find the maximal covering rate with the disjunction of attribute values for a decision rule.

Table 3.10 Binary value table converted from decision table II

U	C_1			C_2			C_3		C_4		d
	C_{11}	C_{12}	C_{13}	C_{21}	C_{22}	C_{23}	C_{31}	C_{32}	C_{41}	C_{42}	
O_1	0	1	0	0	1	0	1	0	0	1	0
O_2	1	0	0	1	0	0	1	0	1	0	0
O_3	1	0	0	1	0	0	1	0	0	1	0
O_4	0	1	0	0	0	1	0	1	0	1	0
O_5	1	0	0	0	1	0	1	0	1	0	0
O_6	0	0	1	1	0	0	1	0	1	0	1
O_7	0	1	0	0	1	0	1	0	1	0	1
O_8	0	1	0	0	0	1	0	1	1	0	1
O_9	0	0	1	0	0	1	0	1	0	1	1
O_{10}	1	0	0	0	0	1	0	1	1	0	1
O_{11}	0	1	0	0	1	0	0	1	1	0	1
O_{12}	1	0	0	0	1	0	0	1	1	0	1
O_{13}	0	0	1	0	1	0	1	0	0	1	1
O_{14}	0	0	1	1	0	0	0	1	1	0	1

Table 3.11 Rules generated by proposed model

	Rules generated by integer programming	accuracy	coverage
Rule 1	(C_1 = 1 or 2) & (C_3 = 1) & (C_4 = 2) => (D = 0)	1.00	0.60
Rule 2	(C_1 = 1) & (C_3 = 1) => (D = 0)	1.00	0.60
Rule 3	(C_1 = 2 or 3) => (D = 1)	0.60	0.78
Rule 4	(C_2 = 2 or 3) & (C_3 = 2) => (D = 1)	0.80	0.56
Rule 5	(C_1 = 3) => (D = 1)	1.00	0.44

Chapter 4 Construction of Layered Data Envelopment Analysis

This chapter discusses the use of Data Envelopment Analysis (DEA) to address the failure prediction problem. It gives a brief review of the techniques and introduces a new concept called layering worst practice DEA. The goal of this approach is to set up DEA models that will place the bad companies to the frontier.

The basic DEA results group the DMUs into two sets, those that are efficient and define the Pareto frontier and those that are inefficient. One problem that has been discussed frequently in the literature has been the lack of discrimination in DEA applications. The layering DEA can divide DMUs into different levels of efficient frontiers. If one removes the original efficient frontiers, then the remaining DMUs will form a new second-level efficient frontier. If one removes this new second-level efficient frontier, a third-level efficient frontier is formed, and so on, until no DMU is left.

4.1 DEA Basic Model

The DEA model was developed by Charnes, et al. (1978). The DEA model formulates the ratio between output and input of resources as mathematical programming to measure relative efficiency. The model will not be affected by the unit of input and output, has no pre-set function form, and can accommodate multiple inputs and outputs. The DEA technique defines an efficiency measure of a Decision making Unit (DMU) by its position relative to the frontier of the best DMU performance established mathematically by the ratio of weighted sum of outputs to weighted sum of inputs. The estimated frontier of best performance, also referred to as the envelopment surface, characterizes the efficiency of DMUs and identifies inefficiencies. A DEA model can be analyzed in two ways, an input orientation and an output orientation. An input orientation provides information as to how much proportional reduction of inputs is necessary while maintaining the current levels of outputs for an inefficient DMU to become DEA-efficient. On the other hand, an output orientation analysis provides information on how much augmentation to the levels of outputs of an inefficient DMU is necessary while maintaining current input levels for it to become DEA-efficient.

DEA in evaluating any number of DMUs, with any number of inputs and outputs:

* Requires the inputs and outputs for each DMU to be specified;
* Defines efficiency for each DMU by an objective function. The objective function in DEA can be ratio oriented (output/inputs), or net profit oriented (outputs-inputs);

* In calculating the efficiency of a particular DMU, weights are chosen to maximize its efficiency, thereby presenting the DMU in the best possible light.

Many DEA models and extensions can be found in literature. We discuss only the CCR and BCC models.

4.1.1 The CCR Model

The CCR Model determines the set of weights that maximizes any DMU efficiency relative to other DMUs of the sample, provided that no other DMU or convex combination of DMUs could achieve the same output vector with a smaller vector. In the input-oriented model, the objective is to produce the observed outputs using a minimum level of resources. In the output-oriented model, the objective is to produce the maximum level of outputs given an observed level of inputs.

Within an input oriented CCR model, it is assumed that there are n Decision-Making Units, with m inputs and p outputs, while the efficiency evaluation model of k^{th} DMU can be defined as in Eq. (1).

$$Max\ f_k = \frac{\sum_{r=1}^{p} u_r y_{rk}}{\sum_{i=1}^{m} v_i x_{ik}} \tag{1}$$

$$\begin{aligned} s.t.\quad & \frac{\sum_{r=1}^{p} u_r y_{rk}}{\sum_{i=1}^{m} v_i x_{ik}} \le 1, \quad k = 1,2,\ldots,n; \\ & u_r \ge \varepsilon \ge 0,\ r = 1,2,\ldots,p; \\ & v_i \ge \varepsilon \ge 0,\ i = 1,2,\ldots,m. \end{aligned}$$

Where, x_{ik} is the i^{th} input value for k^{th} DMU, y_{rk} is the r^{th} output value for the k^{th} DMU, u_r and v_i are the virtual multiplier of the output and input, respectively, and ε is a very small positive value.

It is difficult to obtain the solution from Eq. (1) because the Eq. (1) is a nonlinear programming problem. Therefore, Eq. (1) is modified as Eq. (2) by Charnes, et al., resulting in a linear programming problem where a solution can be more easily obtained.

$$Max\ \theta_k = \sum_{r=1}^{p} u_r y_{rk} \tag{2}$$

$$\begin{aligned} s.t.\quad & \sum_{r=1}^{p} u_r y_{rk} - \sum_{i=1}^{m} v_i x_{ik} \le 0, \quad k = 1,2,\ldots,n; \\ & \sum_{i=1}^{m} v_i x_{ik} = 1, \quad k = 1,2,\ldots,n; \\ & u_r \ge \varepsilon \ge 0,\ r = 1,2,\ldots,p; \\ & v_i \ge \varepsilon \ge 0,\ i = 1,2,\ldots,m. \end{aligned}$$

where, θ_k is the efficiency value for k^{th} DMU, while θ_k is a crisp number under x_{ik} and y_{rk}, the crisp number for the k^{th} DMU.

4.1.2 The BCC Model

In an input oriented model, one focuses on maximal movement toward the frontier through proportional reduction of inputs, whereas in an output orientation one focuses on maximal movement via proportional augmentation of outputs. The BCC model relaxes the CCR requirement of the original CCR ratio model, and make it possible to investigate local returns to scale.

Similar to the CCR model, the BCC input oriented model can be expressed as linear programming formulations. The objective is to produce the observed outputs with a minimum resource level. If a DMU is efficient in a CCR model it will also be efficient with the BCC model, but the converse does not necessarily hold.

$$Max \ \theta_k = \sum_{r=1}^{p} u_r y_{rk} - u_0 \qquad (2)$$

$$\begin{aligned} s.t. \quad & \sum_{r=1}^{p} u_r y_{rk} - \sum_{i=1}^{m} v_i x_{ik} - u_0 \le 0, \quad k = 1,2,\ldots, n; \\ & \sum_{i=1}^{m} v_i x_{ik} = 1, \quad k = 1,2,\ldots, n; \\ & u_r \ge \varepsilon \ge 0, r = 1,2,\ldots, p; \\ & v_i \ge \varepsilon \ge 0, i = 1,2,\ldots, m. \end{aligned}$$

The interpretation for the envelopment problem of CCR is the selection of a point in the cone that allows maximal input reduction of x. This point lies on a ray through the origin via the most north-western DMU. In Fig. 1, C is the efficient (productive) units under constant return of scale (CRS). The rest of the DMUs are being compared to either one of the efficient units or being compared to a virtual unit, which is a linear combination of two efficient units. Under the variable return to scale (VRS) of BCC model, A, B, C and E are on the efficient frontier.

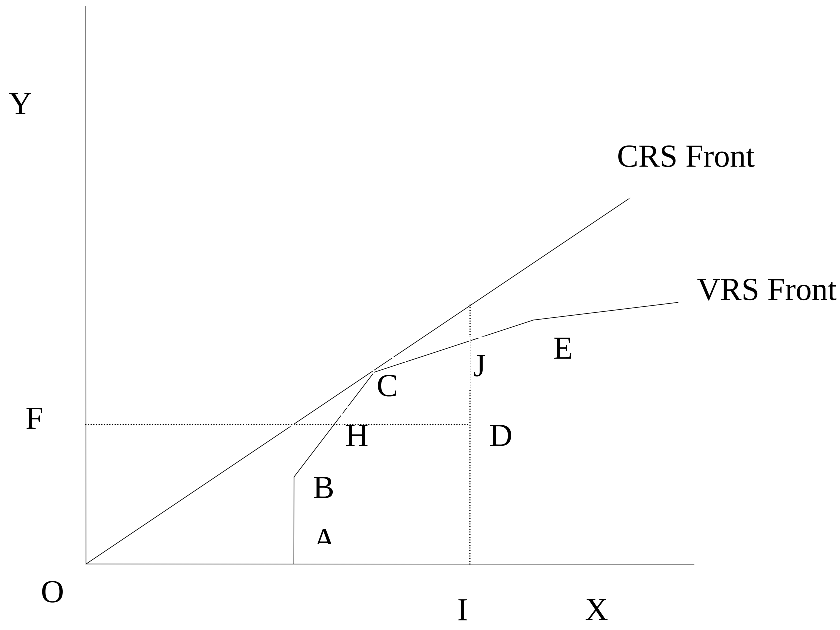

Fig. 4.1 DEA Efficiency Model

4.2 Worst Practice DEA Model

In this section, we further develop the use of DEA in this context by introducing the concept of worst practice DEA. The worst practice DEA introduced by Paradi et al (2004) uses the same model formulation of DEA, but instead of picking out the good performers, the goal is to identify the bad performer. Where normal DEA selects potentially distressed firms by measuring how inefficient they are at being good, worst practice DEA picks out distressed firms based on how efficient they are at being bad. This is achieved by selection variables reflect poor utilization of resources as output. This approach is an ideal fit for the business failure prediction problem, where it is the worst companies that need to be clearly identified.

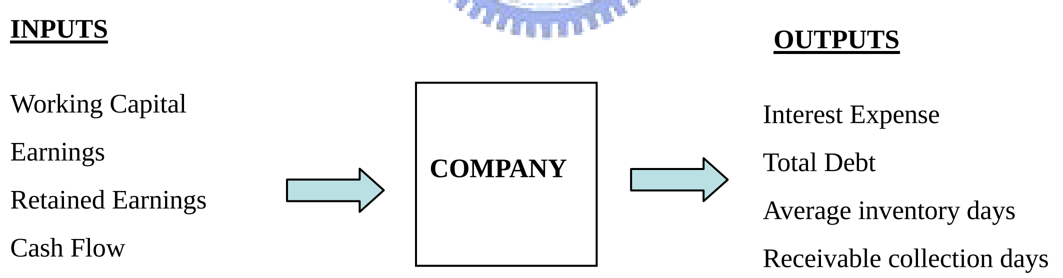

Figure 4.2 Worst Practice DEA Model Variables

The variables that go into a worst practice DEA model are chosen to make the failure companies look as efficient as possible. The figure 4.2 gives an example of variables that could be part of such a model. In a worst practice DEA model, financial ratios with a negative correlation to business failure are defined as input factors (x).

The financial ratios with a positive correlation to business failure are defined as output factors (y)

The companies that will make up the frontier will be the ones that have the lowest earnings, the lowest amount of cash flow and the lowest value, while having the highest level of leverage or interest expense.

The worst practice DEA modeling approach also offers a new possibility in identifying distressed companies using a technique that does not rely on an optimal cut-off value, which is very unique within the failure prediction literature.

4.3 Layered DEA Technique

In the DEA literature, a context-dependent DEA is developed to provide finer evaluation results by examining the efficiency of DMUs in specific performance levels based upon radial DEA efficiency scores. Barr (1993), Seiford (2003) have demonstrated the use of the layering technique in prior studies, but not in a failure prediction context. The context-dependent DEA (Tversky and Simonson, 1993) is introduced to measure the relative attractiveness of a particular DMU when compared to others. In the context-dependent DEA, the evaluation contexts are obtained by partitioning a set of DMUs into several levels of efficient frontiers. Each efficient frontier provides an evaluation context for measuring the relative attractiveness.

4.3.1 Context-dependent DEA Model

Assume that there are *n* DMUs which produce *s* outputs by using *m* inputs. We define the set of all DMUs as J^1 and the set of efficient DMUs in J^1 as E^1. Then the sequences of J^1 and $E^1$are defined interactively as $J^{l+1} = J^{1} - E^{1}$. The set of E^1 can be found as the DMUs with optimal value θ_0^l of 1 to the following linear programming problem:

$$
\begin{aligned}
&\max_{\lambda\theta} \theta_0^l = \theta \\
&st. \quad \sum_{j \in J^l} \lambda_j x_{ij} \le x_{i0}, \; i = 1,....m, \\
&\qquad \sum_{j \in J^l} \lambda_j y_{rj} \le \theta \, y_{rj}, \; i = 1,....s, \\
&\qquad \lambda_j \ge 0, \;\; j \in J^l
\end{aligned}
$$

where x_{ij} and y_{rj} are *i* th input and *r* th output of DMU *j*. When *l* = 1, model (1) becomes the original output-oriented CCR model (Charnes, Cooper and Rhodes, 1978) and E^l consists of all the efficient DMUs. The DMUs in set E^1 define the first-level efficient frontier. When *l* = 2, model (1) gives the second-level efficient frontier after the exclusion of the first-level efficient DMUs. In this manner, we identify several levels of efficient frontiers. Then E^l consists the *l* th level efficient frontier.

With the concept of content DEA, instead of identifying fixed cut-off points to classify a firm as distressed or not, as is usually done in this kind of performance evaluation, we suggest using a layering technique. Using this approach the firms on the frontier in the (worst practice) DEA analysis are removed, after which the model is run a second time resulting in a new set of frontier units, which are then removed before the model is run a third time and so on. Thereby sequential layers of "efficient" performance can be found, which have changing risk ratings.

4.3.2 Layered DEA Model

In the following we briefly present the idea of Layering DEA by introduce an example (Table 4.1). Suppose the following financial data is given for five firms (A, B, C, D and E). In Fig.4.3 the worst practice DEA is illustrated with two outputs and fixed inputs. The units A and B are on the frontier and thus the companies have the highest liabilities (Y1) and receivable days (Y2). When removing these frontier units and running the DEA model again, a second layer of frontier units, C and D are identified. The companies on the first layer are the companies with the highest risk and the companies on the second layer are assumed to have a lower risk rating on DEA.

Table 4.1. Layering DEA example

	A	B	C	D	E	Variables
X1: Profit	1	1	1	1	1	DEA input variable
Y1: Liabilities	2	8	3	5	2	DEA output variable
Y2: Receivable days	8	3	4	2	2	DEA output variable

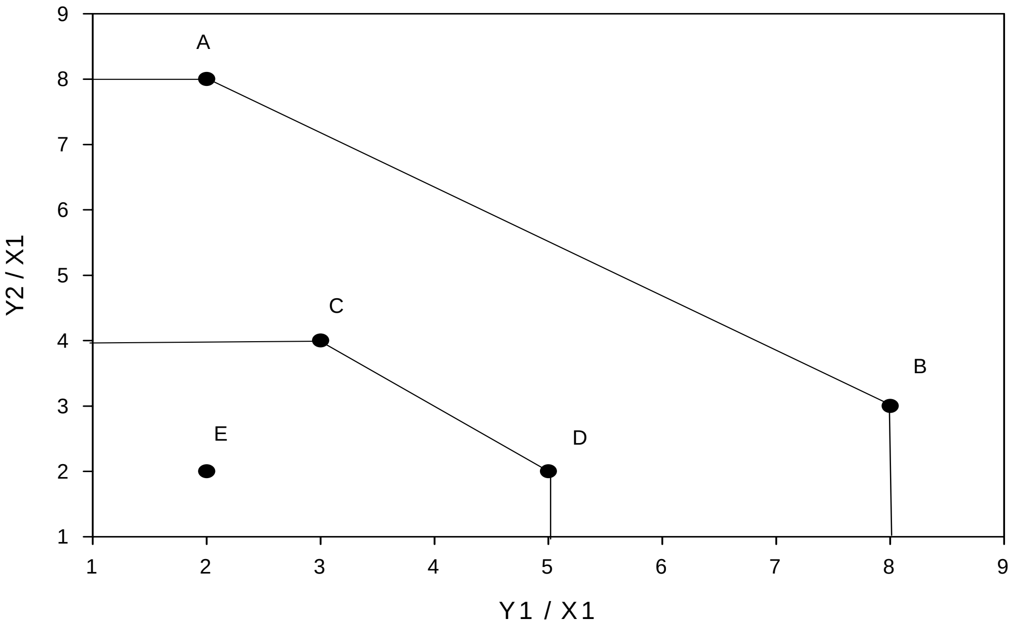

Fig. 4.3. The worst practice DEA and the layering technique.

Chapter 5 Empirical Models and Data

5.1 Hybrid System Development

In this paper we propose a hybrid system combining rough set approach and DEA. Our system has two steps: In the first step, irrelevant and redundant attributes are removed from the table by rough set approach without any classification information loss. Then the knowledge—a rule set is generated from the decision table. A risk monitor by DEA helps users monitor the risk by placing the distressed firms on layered frontiers based on how efficient they are at being bad. In the prediction phase, a new object is first predicted by the rule set, if it does not match any of the rules, it is fed into the DEA to get its risk level. The mining agent apply the rules developed by rough set, and help users make decision about the risk analysis. The effectiveness of our hybrid approach was verified with experiments that compared to the worst practice DEA model. The hybrid model can get high classification accuracy.

In the following we briefly present the idea of rough set DEA by introduce an example (Table 5.1). Suppose the following financial data is given for five firms (S_1, S_2, S_3, S_4 and S_5). In Table 5.2 a rule set is generated from the decision table based on original rough set theory. The proposed rule induction model generates another set of rules which has higher coverage as shown in Table 5.3. In Fig.5.1 the worst practice

DEA is illustrated with two outputs and fixed inputs. The units S_1 and S_2 are on the frontier and thus the companies have the highest liabilities (*Y1*) and receivable days (*Y2*). When removing these frontier units and running the DEA model again, a second layer of frontier units, S_3 and S_4 are identified. The companies on the first layer are the companies with the highest risk and the companies on the second layer are assumed to have a lower risk rating on DEA. However, the DEA model contains only quantitative data; with the decision rule generated from rough set we can identify that firm S_3 also have high risk since the firm changed auditor and financial manager recently. Information like this can be used to determine which companies at risk and the risk level.

Table 5.1. Rough Set DEA example

	S_1	S_2	S_3	S_4	S_5	Variables
X1: Profit	1	1	1	1	1	DEA input variable
Y1: Liabilities	2	8	3	5	2	DEA output variable
Y2: Receivable days	8	3	4	2	2	DEA output variable
C1: Auditor changed	2	0	1	0	0	Rough condition var.
C2: CFO changed	1	1	1	1	0	Rough condition var.
D : Failure ?	Yes	Yes	Yes	No	No	Rough decision var.
Failure Risk	High	High	High	Medium	Low	

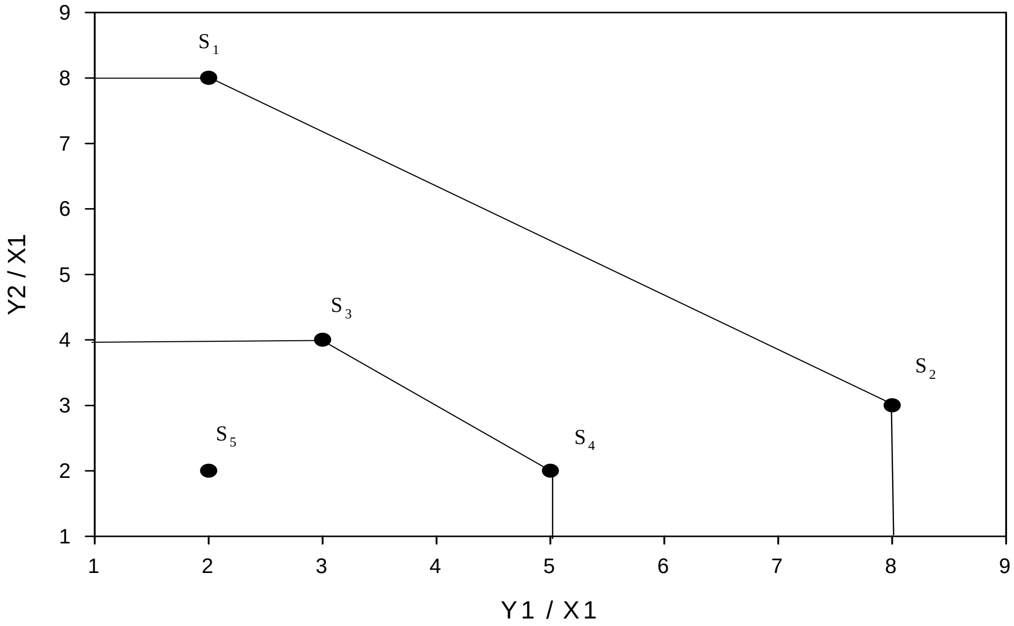

Fig. 5.1. The worst practice DEA and the layering technique.

Table 5.2. Example decision rules induced by rough set

Rule		Strength	Coverage	Accuracy
1	(C2 = 0) => (D = No);	1	50.00%	100.00%
2	(C1 = 1) => (D = Yes);	1	33.33%	100.00%
3	(C1 = 2) => (D = Yes);	1	33.33%	100.00%

Table 5.3. Example decision rules induced by proposed model

Rule		Strength	Coverage	Accuracy
1	(C1 = 0) $\wedge$ (C2 = 0) => (D = No);	1	50.00%	100.00%
2	(C1 = 1 $\vee$ 2) $\wedge$ (C2 = 1) => (D = Yes);	2	66.67%	100.00%

5.2 Structure of hybrid system

The proposed system solves failure prediction problem in a hierarchical framework, as illustrated in Fig.5-2 and Fig. 5-3.

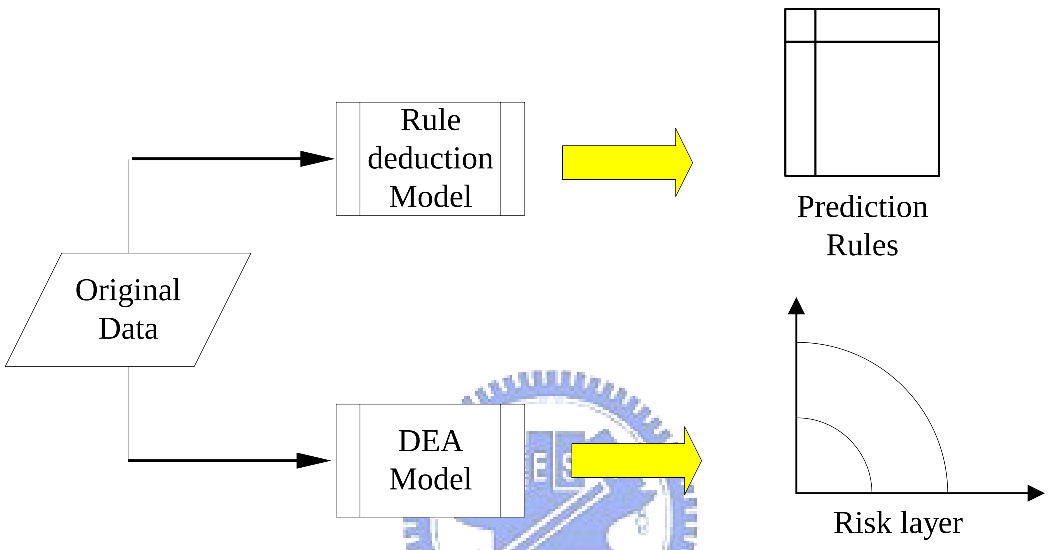

Fig. 5.2 Hybrid Process Flows: reduction, rule generation and layering DEA

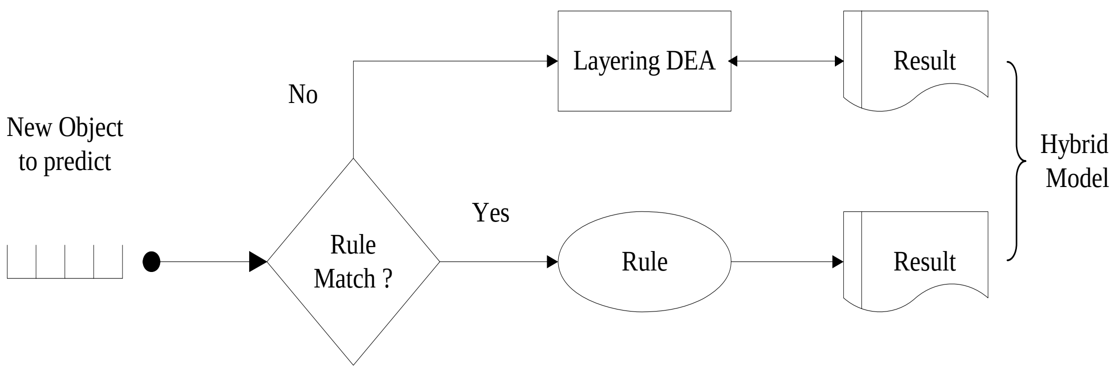

Fig. 5.3 Hybrid Prediction Model

Our hybrid approach of rough sets and DEA for failure prediction rules consists of three major phases:

- Using rough set approach, a reduct of history data decision table is obtained. Then the knowledge—a rule set is generated from the reduced decision table.
- Using DEA to display and group new companies by their quantitative financial performances and risk levels.
- Mining Agent generates rules for firms with different risks level. The decision rules generated by rough set can identify potential risky firms for each risk level as well as get the change curve with multi period comparison.

5.3 Sample Selection

The population of `failed´ companies comprise all companies that have been declared "special arrangement" by authorities when the company has operation difficulties during the years 2002 and 2003. According to Operation Rules of the Taiwan Stocks Exchange Corporation article: 49, 50 and 50-1 (Appendix C), a company in an unhealthy financial condition is recognized as a company in financial crisis. The population of `running´ companies include all other companies.

A number of firms which failed in Taiwan in the year 2002-2003 were collected. Among the publicly traded firms only a very small percentage of companies went bankrupt in a given year. This reduces to a smaller sample when the analysis focuses on a single industry. The companies selected for this analysis were the companies in the electronic sector. This restriction was introduced because the results are expected to differ across industries, and the electronic sector is the sector with the largest number of documented public failures. At the same time outliers have been removed from the subset (i.e. companies that show extreme scores for some discriminating ratios), as they can influence the results.

We drew a state-based sample instead of a pure random sample. A pure random selection would lead to a very small sample of `failed´ companies and inaccurate parameter estimation in the models. The healthy companies are not necessarily good companies, but at least healthy enough so that they did not file for bankruptcy in that time period. Some of them might be in poor financial condition, so the assumed distribution of their efficiency is normal. The resulting sample comprises 400 annual reports broken down in two subsets running: 378 and failed: 22. For each annual report 8 financial ratios and five non-financial ratios have been calculated.

Note that this control group selection is different from the matched pair sample approach commonly adopted in the majority of bankruptcy studies, which usually consists of half failed and half non-failed firms. The most obvious problem with the latter approach is that the ratio of healthy to failed companies is not 1 to 1 in the real world, but more like 100 to 1 for public companies. As a result much information is lost in the paired sample approach since it means truncating the sample of healthy companies.

5.4 Data collection

A large number of ratios have been proposed in the literature. Courtis (1978) made an attempt to identify the variables useful in predictive studies. In his survey 79 financial ratios were identified and grouped in three main categories: (a) profitability ratios; (b) managerial performance ratios and (c) solvency ratios. Table 5.4 lists some of the most important financial ratios in the from 1964 to 1994 (Dimitras et al. 1996).

Table 5.4 Financial ratios included in industrial failure models

No.	Financial Ratios	Cited out of 59 literature
1.	Working Capital / Total Asset	16
2.	Liabilities / Total Assets	15
3.	Current Ratio	12
4.	Profit before tax to paid-in capital	12
5.	Operating income to paid-in capital	11
6.	Quick Ratio	09
7.	Return on shareholders' equity	06
8	Average inventory days	04

In the analysis, we consider data for the year prior to failure, that is the 2002 data for the companies that went bankrupt during 2003 and their control group, 2003 data for the 2004 bankruptcies. The financial-related group consists of 18 financial ratios from the database of Taiwan Economic Journal. For DEA analysis, we have selected 8 ratios because they were proved to be efficient to predict bankruptcy in prior research.

Those variables selected in the one year before failure, including the liabilities/assets ratio, the current ratio, the quick ratio, the profit before tax to paid-in capital, the average accounts receivable collection days, the average inventory days, Operating income to paid-in capital, the return on shareholders' equity and the total assets turnover.

When dealing with corporate failures it is very common that some of the variables take on negative values. The translation invariance property of the variable returns to scale DEA model (Ali and Seiford, 1990) is used to eliminate negative values in all these variables. For each of the variables that take on negative values, the most negative value among the DMUs plus one was added to the value of all DMUs. For the rough set analysis, five variables (conditional attributes) were collected for potential rule generation. Table 5.5 contains the variable definitions and the corporate attributes they refer to.

By using a worst practice DEA model, financial ratios with a positive correlation to a healthy firm are defined as input factors. The financial ratios with a negative correlation are defined as output factors. The companies that make up the frontier in this analysis are those with the lowest of those good inputs while having the highest level of the bad outputs. This approach is an ideal fit for the failure risk evaluation, where it is the worst companies that need to be clearly identified. Using the layering technique, the firms on the frontier are removed, after which the model is run a second time resulting a new set of frontier units and so on. Thereby sequential layers of worst performance can be found which have changing risk rating. The rough set is used to provide a set of rules able to discriminate between healthy and failing firms in order to predict business failure.

The five variables (C9-C13) described in Table 5.5 relate to various non-financial corporate attributes which have been found to be associated with corporate failure. The first two variables (C9, C10) are binary in nature and relate to the management quality and fraud detection since several Taiwanese firms declare bankruptcy after the abnormal change of chief executive and/or financial officers.

Table 5.5. Definition of variables

Attribute Description	Variables value	DEA
Financial structure		
C1	Liabilities/ Total assets ratio (%)	output
Solvency		
C2	Current ratio (%)	input
C3	Quick ratio (%)	input
Management Quality		
C4	Average accounts receivable collection days	output
C5	Average inventory days	output
Profitability		
C6	Operating income to paid-in capital	input
C7	Return on shareholders' equity	input
C8	Profit before tax to paid-in capital	input
Non-financial Qualitative data for Rough set analysis		
C9	Code 1 if abnormal changed CEO in previous year, 0 otherwise;	
C10	Code 1 if changed financial manager in previous year, 0 otherwise;	
C11	Code 1 if auditor qualified audit report, 0 otherwise	
C12	Code 1 if changed auditor in previous year, 0 otherwise	
C13	Code 1 if changed financial statement forecast in previous year, 0 otherwise	

The two variables (C11, C12) are related to auditor characteristics. The first is the auditor opinion signaling for going-concern problems. The second is whether or not a company had changed its auditor the previous year. Previous studies have reported that failing firms are more likely to switch auditor, largely in consequence of disputes between auditors and managers over disagreements in respect of audit opinions/qualifications.

Using the layering technique enables a more flexible approach to classification which can take into managerial judgment and risk attitude. The more risk aversion the institution is, the more layers in the worst practice models should be considered. The decision rules generated by rough set can identify potential risky firms for each risk level as well as get the change curve with multi period comparison.

Chapter 6 Empirical Results and Discussions

A large number of firms which failed in Taiwan in the year 2003-2004 were collected. The companies selected for this analysis were the companies in the electronic sector as indicated by their SIC code. This restriction was introduced because the results are expected to differ across industries, and the electronic sector is the sector with the largest number of documented public failures. The non-bankrupt companies needed for comparative purposes were selected from the "Corporate Information on Public Companies Filing with the SEC". The only criterion for the healthy companies was that they did not go bankrupt before 2003.

We drew a state-based sample instead of a pure random sample. A pure random selection would lead to a very small sample of failed companies and inaccurate parameter estimation in the models. The resulting sample comprises 420 annual reports broken down in two subsets running: 400 and failed: 20. The best classification results are achieved when combining the rough set with the worst practice DEA model.

6.1 Prediction Rules Induction

In practice, for the analysis of decision table, there are some main steps mentioned by Walczak and Massart (1999), such as: (1) construction of elementary sets; (2) calculation of upper and lower approximations of the elementary sets; (3) finding the core and reducts of attributes; and (4) finding the core and reducts of attribute values. Hence, for the data analysis in rough set approach, we suggest the following three-step analytical procedure: (1) calculating the approximation; (2) finding the reducts of attributes and the core of attributes; and (3) creating the decision rules.

6.1.1 Rules deducted from qualitative data

Before the data analysis, it is required to construct the decision table. As shown in Table 6.1, the decision table contains 47 records characterized by one decision attribute (failure) and five qualitative condition attributes: changed chief executive officer last year (CEO), changed chief financial officer last year (CFO), auditor qualified opinion (Opinion), changed auditor last year (Auditor) and changed financial forecast last year (Forecast). Further, the six attributes and their values are denoted as binary values.

Table 6.1 Sample qualitative criteria decision table

Company	CEO	CFO	Opinion	Auditor	forecast	Failure
2305	0	0	0	0	0	0
2311	0	0	0	0	0	0
2312	0	0	0	0	1	0
2314	1	1	0	0	0	0
2323	0	1	0	0	0	0
2325	0	0	0	0	0	0
2327	1	0	0	0	0	0
2331	1	0	1	0	1	0
2332	0	0	0	1	1	0
2336	0	0	0	0	0	0
2337	0	0	0	0	0	0
2340	0	0	0	0	1	0
2343	1	0	0	0	0	0
2345	1	0	1	0	0	0
2349	0	0	0	0	0	0
2361	0	0	0	0	0	0
2366	0	0	0	0	1	0
2376	0	0	0	0	0	0
2393	0	0	0	0	0	0
2406	0	0	0	1	0	0
2408	0	0	0	0	1	0
2422	0	0	0	0	1	0
2426	0	0	0	0	1	0
2432	0	0	1	0	0	0
2448	0	0	0	0	0	0
2460	0	0	0	1	0	0
2479	1	1	0	0	1	0
2483	1	0	1	0	0	0
3026	0	1	0	0	0	0
4903	0	0	1	0	1	0
1602	0	0	1	1	0	1
2326	1	0	0	0	1	1
2329	1	0	1	0	1	1
2342	0	1	1	1	1	1
2359	0	0	1	0	0	1
2393	0	1	1	0	0	1
2445	1	1	1	0	1	1
2490	0	0	1	1	1	1
2494	1	0	1	0	1	1
3004	0	0	1	0	1	1
3021	0	0	1	0	1	1
5307	1	0	0	0	1	1
5325	0	1	0	1	1	1
5336	1	1	1	1	1	1
5347	1	0	0	0	1	1
6193	0	1	0	0	1	1
8012	1	1	0	0	0	1

Step 1: Calculating the approximation.

The first step of data analysis using rough set theory is to calculate the approximations of decision classes. As shown in Table 6.2, each decision class is describable by the lower and upper approximations accuracy shown in the last column.

Table 6.2 Lower and upper approximations

Class number	Number of objects	Lower approx.	Upper approx.	Accuracy
1	30	26	36	0.722
2	17	11	21	0.524

Step 2: Finding the reducts of attributes and the core of attributes.

In this step, the indiscernibility relation method is used for dealing with the reduction of attributes and finding the core of attributes, due to all the condition attributes are nominal attributes (unordered qualitative attributes) with linguistic values. Employing the indiscernibility relation method, it may find all potential reducts in the information table. As a result, we obtained four cores of attributes. The core of attributes is the the attribute {CEO, CFO, Auditor, Opinion}. This means that these attributes are the most meaningful attribute among those five attributes.

Step 3: Creating the decision rules.

The most important step of data analysis is to generate decision rules. In order to find the minimal covering rules, the minimal covering method is employed, which attempts to find the minimal number of attribute values for a decision rule. As a result, 8 rules are created. These 8 exact rules are shown in Table 6.3, from which we can acquire several valuable implications for making decisions. In particular, we can find the most important determinant for each decisions class through using the covering ratio.

Table 6.3. Qualitative decision rules deducted from rough set

	Elementary conditions					Decision		
Rule	CEO	CFO	Opinion	Auditor	Forecast	Failure	Accuracy	Coverage
1	0	0	0			No	1	0.633
2	1	0			0	No	1	0.133
3	0		0		0	No	1	0.467
4	0	1			1	Yes	1	0.176
5	1	0	0		1	Yes	1	0.176
6			1	1		Yes	1	0.235
7	0	1			1	Yes	1	0.176
8		1	1			Yes	1	0.235

Table 6.4. Qualitative decision rules deducted from proposed model

	Elementary conditions					Decision		
Rule	CEO	CFO	Opinion	Auditor	Forecast	Failure	Accuracy	Coverage
1			0			No	0.807	0.833
2					0	No	0.833	0.667
3	0	0				No	0.808	0.700
4	0			0		No	0.800	0.667
5	1				1	Yes	0.778	0.412
6			1		1	Yes	0.800	0.471

Comparing Table 6.3 and Table 6.4 we can find that the proposed model using fewer attributes and generate rules with higher coverage rates.

6.1.2 Rules deducted from both quantitative and qualitative data

As shown in Table 6.5, the decision table contains 41 records characterized by one decision attribute (failure) and 13 conditional attributes with the same definition on Table 5.5. The eight quantitative data are all converted to three ranges {high, medium, Low}. Further, the five qualitative data and their values are denoted as binary values.

Step 1: Calculate the approximation.

The first step of data analysis using rough set theory is to calculate the approximations of decision classes. As shown in Table 6.5, each decision class is well describable due to its high accuracy of 1.000 as in the last column. This is to say that all two decision classes are characterized exactly by those data in the decision table. As the whole, the accuracy of the entire classification is 1.000, and also the quality of the entire classification is 1.000.

Table 6.5. Lower and upper approximations of both qualitative and quantitative data

Class number	Number of objects	Lower approx.	Upper approx.	Accuracy
1	27	27	27	1.000
2	12	12	12	1.000

Step 2: Find the reducts of attributes and the core of attributes.

We obtained 135 reducts of attributes and four core of attributes. The core of attributes is the attributes {C_3, C_7, C_9, C_{11}}. This means that these attributes are the most meaningful attributes among those thirteen attributes.

Table 6.6. 2002 Sample criteria decision table

Company	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	**D**
2490	1	3	3	3	3	1	1	1	0	0	1	1	1	1
2326	3	3	3	3	3	1	1	1	1	0	0	0	1	1
2337	3	1	1	1	3	1	1	1	0	0	0	0	0	0
2359	3	1	1	2	1	1	1	1	0	0	0	0	0	1
2329	3	1	1	1	1	1	1	1	1	0	1	0	1	1
2406	1	1	2	3	2	1	1	1	0	0	0	1	0	0
2407	2	1	1	2	1	1	1	1	1	1	1	0	1	1
2349	2	2	3	3	2	1	1	1	0	0	0	0	0	0
2327	1	1	1	3	3	2	1	1	1	0	0	0	0	0
2342	3	1	1	1	2	1	1	1	0	1	1	1	1	1
2314	2	3	3	2	3	2	1	1	1	1	0	0	0	0
2445	3	1	1	3	2	1	1	1	1	1	1	0	1	1
2494	1	3	2	1	2	1	1	1	1	0	1	0	1	1
2432	1	1	1	2	3	2	2	2	0	0	1	0	0	0
2340	3	2	2	3	3	2	2	2	0	0	0	0	1	0
2311	2	1	1	2	1	2	2	2	0	0	0	0	0	0
2323	2	1	2	3	1	2	2	2	0	1	0	0	0	0
2325	2	2	3	1	1	2	2	2	0	0	0	0	0	0
2305	2	1	1	2	2	1	2	2	0	0	0	0	0	0
2336	2	2	2	3	1	2	2	2	0	0	0	0	0	0
2479	3	2	2	3	2	1	2	2	1	1	0	0	1	0
3004	3	3	3	3	3	3	2	2	0	0	1	0	1	1
2426	2	2	2	3	3	2	2	2	0	0	0	0	1	0
2408	2	2	1	1	3	2	2	2	0	0	0	0	1	0
2422	2	3	2	2	3	2	3	2	0	0	0	0	1	0
2343	1	2	2	1	2	2	2	2	1	0	0	0	0	0
3021	3	3	3	2	2	3	2	3	0	0	1	0	1	1
2312	1	2	3	1	1	2	3	3	0	0	0	0	1	0
2361	1	3	3	1	1	3	3	3	0	0	0	0	0	0
2483	1	3	3	1	3	3	3	3	1	0	1	0	0	0
2332	2	3	3	2	1	3	3	3	0	0	0	1	1	0
2348	3	1	1	3	1	3	3	3	0	1	1	0	1	1
2345	1	3	3	2	2	3	3	3	1	0	1	0	0	0
2366	3	2	1	1	2	3	3	3	0	0	0	0	1	0
2393	1	2	2	2	2	3	3	3	0	1	1	0	0	1
2460	2	3	2	2	2	3	3	3	0	0	0	1	0	0
2448	1	2	2	2	3	3	3	3	0	0	0	0	0	0
2376	1	3	3	1	1	3	3	3	0	0	0	0	0	0
2331	3	2	2	1	1	3	3	3	1	0	1	0	1	0

Step 3: Creating the decision rules.

10 rules are created as shown in Table 6.7.

Table 6.7. All decision rules deducted from rough set

	Elementary conditions													Decision		
Rule	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	Failure	Accuracy	Coverage
1		2										0		No	1	0.407
2	1									0			0	No	1	0.333
3	2										0			No	1	0.444
4		2								0				No	1	0.407
5	1				3									No	1	0.148
6		1											1	Yes	1	0.417
7	3		3											Yes	1	0.250
8						1					1			Yes	1	0.500
9				2					0	1				Yes	1	0.083
10	3			2										Yes	1	0.167

Table 6.8. All decision rules deducted from proposed model

	Elementary conditions													Decision		
Rule	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	Failure	Accuracy	Coverage
1	1or 2										0			No	1	0.704
2		1or 3									1		1	Yes	1	0.750
3											1		1	Yes	0.900	0.750
4	3										1			Yes	0.857	0.500
5						1					1			Yes	1	0.500
6						1							1	Yes	0.875	0.583
7													1	Yes	1	0.500
8							1				1			Yes	1	0.500
9							1						1	Yes	1	0.583
10								1			1			Yes	1	0.500
11								1					1	Yes	1	0.583

Rules generated by proposed model are listed on Table 6.8. Comparing Table 6.7 and Table 6.8 we can find that the proposed model generate rules with much higher coverage rate than the original rough set. The decision rules deducted from sample set is verified by the 2004 data as shown in Table 6.9. We can predict 10 out of 19 failure firms from the decision rule. For those firms that can not match with any rules, we will use layering DEA technique to do prediction.

Table 6.9 Verified decision table

Company	CEO	CFO	Opinion	Auditor	Forcast	Hit
2318	1	0	1	0	0	
2335	1	1	0	1	1	
2348	0	1	0	0	0	
2398	0	1	0	1	1	V
2407	1	1	1	0	1	V
2491	0	0	1	0	1	
3039	1	1	1	0	1	V
3053	1	1	1	1	0	V
3054	1	1	1	0	1	V
5344	1	1	0	0	1	
5348	1	1	0	0	0	
5385	1	1	1	0	1	V
5386	0	1	1	0	1	V
5442	1	0	0	1	0	
5497	0	1	1	0	1	V
6130	1	1	1	0	1	V
6145	0	1	0	0	1	
6181	1	0	0	1	1	
6241	0	1	1	1	1	V

6.2 Layering Risk Analysis

The companies were divided into two groups based on whether they filed for bankruptcy in 2003 or 2004, with 12 and 19 failure companies in the groups, respectively. The failed companies were matched up with 196 and 200 healthy companies in the two years, respectively, with different companies in these two groups. Note that this control group selection is different from the traditionally accepted matched pair sample approach adopted in the majority of bankruptcy studies, which usually consists of half failed and half non-failed firms. The most obvious problem with the latter approach is that the ratio of healthy to failed companies is not 1 to 1 in the real world, but more like 100~200 to 1 for public companies. As a result much information is lost in the paired sample approach since it means truncating the sample of healthy companies.

In the analysis, we are considering data for the year prior to failure, that is the 2002 data for the companies that went bankrupt during 2003 and their control group, and 2003 data for the 2004 bankruptcies. It should be noted that different models and distress indicators may be relevant at different points of time, as discussed in Paradi et al. (2001). Here, however, we choose to look only at the data for the year prior to failure. All variable and model selections are performed on the 2002 data set alone.

The 2003 data are used later to test the (out-of-sample) classification accuracies for the developed models.

Running first using worst practice DEA model gives the average efficiency scores for the bankrupt and the non-bankrupt companies, respectively, shown in Table 6.10. In Table 6.10, we observe considerable differences in the average efficiency scores between the bankrupt and the healthy companies with all the model formulations. The average efficiency scores for the failing companies are higher than the score of the healthy companies the year prior to their failure. Note that in the worst practice analysis, the units with high-efficiency scores means that the companies are efficient at being bad.

What really relevant is not the average scores, but how to distinguish between failure and non-failure companies. Usually this is done by selecting a cut-off value. The optimal cut-off value depends on the costs of the two types of misclassification. Consequently, the Type I errors (loss resulting from invest in failure firms) are much more expensive for the bank than the Type II (misclassification of healthy firms to failure firms) errors. However, Type I errors are a lot less frequent, because there are a lot fewer companies that have financial crisis than firms that do not.

Instead of this optimal cut-off value approach we suggest using the layering technique. The results of this approach for the best of the suggested worst practice are shown in Table 6.11. From Table 6.11, we see that on the first layer of the worst practice model, 38% of the failure companies are found and only 13% of the non-bankrupt companies, so looking at this layer means correctly classifying 49% of the failure companies and 87% of the healthy companies. More and more bankrupt companies are identified on each consecutive layer, until by the third layer all failure companies are found. For each layer, however, there is also an increasing misclassification of the non-bankrupt companies. Using the layering technique enables a more flexible approach to classification which can take into account subjective consideration, managerial judgment, risk attitude, etc., by the choice of the number of layers one wish to consider. The more risk averse the decision maker is, the more layers in the worst practice models should be considered, in order to eliminate more of the risky companies. This will come at the expense of excluding more healthy companies as well, as indicated by the falling non-failure classification accuracies for each layer.

Finally, what we propose here is to include both rough set and worst practice DEA models. The results from this combination are given in Table 6.11. The first row in Table 6.5 repeats the failure and non-failure classification accuracies for the layered worst practice model alone. Again, it shows how looking only at a few layers of this model means correctly classifying many of the non-failure companies but also results in a low bankruptcy classification accuracy. The more risk averse the lending institution is, the more layers of the worst practice model should be considered, resulting in a higher accuracy in identifying failure companies but at the expense of excluding more well performing companies.

The results are significantly improved by combining this model with the rough set decision rule, as shown in the remaining rows of Table 6.11. The best classification results are achieved when combining the worst practice and the rough set model and including three layers of each of those models. This gives an impressive 100% bankruptcy classification accuracy as well as 78% non-failure classification accuracy.

These results are the within sample classification accuracies, so to validate the approach we test the models using the 2003 data set which consists of 20 companies that went failure in 2004 and the corresponding 400 healthy companies. The results from this analysis are shown in Table 6.11. Note that these results are out of sample accuracies and therefore are the level of accuracy that would be expected on a new data set.

Table 6.11 shows that when testing the models on the 2003 data set we observe, that three layers deep in the worst practice model 3 means correctly identifying all failure companies, but at the cost of a 15% misclassification of the healthy companies. Again, note that the number of layers to include depends on the investors' risk attitude. The more risk averse, the more layers of the worst practice models and fewer layers of the normal models should be considered.

Table 6.10**.** Average efficiency scores for normal and worst practice DEA

Model	Inputs	Outputs	Failure	Non-Failure
Worst practice	ROE, CR, IN	TL, RE, IV	0.58	0.25
BCC	Asset, Equity	IN, Profit	0.29	0.69

Table 6.11. Failure and non-failure classification accuracies

	Failure Prediction			Non-Failure Prediction		
	Layer 1	Layer 2	Layer 3	Layer 1	Layer 2	Layer 3
Worst Practice DEA	0.38	0.54	0.82	0.87	0.95	1.00
Rough Set DEA	0.77	0.92	100	0.91	0.97	1.00
Rough Set	0.48			0.89		

6.3 Discussions

In this section, we made a comparison among these three models (hybrid model, DEA and rough set) to see advantages and limitations of them. On a real world, the data may not be clear and clean. How practical these models can deal with different issues which often occur in data analysis. The results are summarized on Table 6.12.

Table 6.12 Issues of real world

Methods / Real world issues	DEA	Rough Set	Hybrid Approach
very large data set			▲
mixed types of data		▲	●
noisy data		●	▲
incomplete instances	●	●	
use of background knowledge	●		●

● Okay ▲ possible

Chapter 7 Conclusions

This study illustrates the usefulness of the rough set DEA approach as an operational tool for the prediction of company failure. This prediction model has an advantage over models in the form of functions.

We have shown how worst practice DEA analysis, aimed at identifying the companies that are efficient at being bad, can be used to identify worst performers, as particularly relevant for failure prediction. Furthermore, we have illustrated how the use of a layering technique gives much higher classification accuracies and is less sample specific than the traditional fixed cut-off point approach. The layering approach also has the advantage of giving flexibility through the choice of layers one wishes to consider, which enables incorporation of risk attitudes.

Inclusion of non-financial characteristics in the acquisition evaluation methods has already been recommended in several other studies in order to improve the validity of the decision rules model. The rough set DEA approach adapts very easily to this need since it accepts both qualitative and quantitative attributes. In contrast to classical statistical techniques such as discriminant analysis, the strength of this hybrid approach is that it requires no underlying statistical assumptions, especially the rough set can provide rules which cover only subsets of the basic objects or data records available (Curry, 2003). It is not only free from such the unrealistic assumption of statistical hypotheses (Dimitras et al., 1999), but also it has no need of a huge data. In particular, it can directly analyze the original data with either quantitative attributes or qualitative attributes, as well as does not need additional information. Finally, by combining rough set and DEA models the classification accuracies are increased. The results from combining three layers of the worst practice DEA models and rough set are as good as 100% out-of-sample classification accuracy for the failure companies and 87% for the healthy companies. The evaluation of corporate performance via Rough set DEA discussed in this research is able to provide part of the early-warning information needed beforehand. It is a tool worthy of consideration by investors, auditors and government officials for decision and control.

References

[1] Ahn, B.S., Cho, S. and Kim, C., "The integrated methodology of rough set theory and artificial neural network for business failure prediction", Volume: 18, Issue: 2, Expert Systems with Applications, pp. 65-74, 2000.

[2] Ali, A. I. and Seiford L. M., "Translation invariance in data envelopment analysis." Journal of Operations Research, 9(99), pp. 403-405, 1990.

[3] Altman, E. I., “Financial ratios, discriminant analysis and the prediction of corporate bankruptcy”. Journal of Finance, pp. 589-609, September 1968.

[4] Altman, E. I., “Accounting implications of failure prediction models”. Journal of Accounting, Auditing, and Finance 6(1), pp. 4-19, 1982.

[5] Ambrose, J. M. and Carol A. M., "Using best ratings in life insurer insolvency prediction", Journal of Risk and Insurance, Volume: 61, pp. 317-327, 1994.

[6] Atiya, A., “Bankruptcy prediction for credit risk using neural networks: a survey and new results”, IEEE Transactions on Neural Networks, Volume: 12, Issue: 4 , 2001.

[7] Bahnson, P. R. and Bartley, J. W., “The sensitivity of failure prediction models to alternative definitions of failure”, Advances In Accounting 10, pp. 255-278, 1992.

[8] Barniv, R., Agarwa,l A. and Leach, R., “Predicting the outcome following bankruptcy filing: a three-state classification using neural networks”, Intelligent Systems in Accounting, Finance and Management, Volume: 6, (1997), pp. 177--194

[9] Barniv, R., "Accounting procedures, market data, cash-flow figures and insolvency classification: the case of the insurance industry", The

Accounting Review, Volume: 65, Issue: 3, pp. 578-604, 1990.

[10] Barr, R. S., Seiford, L. M. and Siems, T. F., "An envelopment-analysis approach to measuring the managerial efficiency of Banks." Annals of Operations Research 45, 1-19, 1993.

[11] Beaver, W., “Financial ratios as predictors of failure”, Journal of Accounting Research, 4, pp. 71-111, 1966.

[12] Bazan, J., "A comparison of dynamic and non-dynamic rough set methods for extracting laws from decision tables", Polkowski L., Skowron A., (Ed.) Rough Sets in Knowledge Discovery, (1998), pp. 321-365 Phisica-Verlag, Heidelberg.

[13] Bell, T., “Neural nets or the logit model? A comparison of each model's ability to predict commercial bank failures”, Intelligent Systems in Accounting, Finance and Management, Volume: 6, pp. 249—264, 1997.

[14] Beynon, M., Curry, B. and Morgan, P., "Classification and rule induction using rough set theory", Expert System, Volume: 17, pp. 136-148, 2000.

[15] Beynon, M. J. & Peel, M. J., “Variable precision rough set theory and data discretisation: an application to corporate failure prediction.”, Omega: International Journal of Management Science, 29(6), pp.561-576, 2001.

[16] Booth, D., Alam, P., Ahkam, S. N. and Osyk B., “A robust multivariate procedure for the identification of problem savings and loan institutions”, Decision Sciences, Volume: 20, Issue: 2 (1989), pp. 320--330

[17] Boritz, J. and Kennedy D., Effectiveness of neural networks types for prediction of business failure, Expert Systems with Applications, Volume: 9, pp. 503—512, 1995.

[18] Charalambous, C., Charitous, A. and Kaourou, F., “Comparative analysis of

artificial neural network models: application in bankruptcy prediction", Annals of Operations Research, Volume: 99, pp. 403—425, 2000.

[19] Charnes A., Cooper W. W. and Rhodes E., "Measuring the efficiency of decision making units", European Journal of Operational Research, 30 (9), pp. 1078-1092, 1984.

[20] Charnes, A., Roussea J. J. and Semple J. H., "Sensitivity and stability of efficiency classifications in data envelopment analysis", Journal of Productivity Analysis 7 (1): 5-18 MAR 1996.

[21] Chen, S. H. and Yeh, C. H., "Using genetic programming to model volatility in financial time series" In: Koza, J., et al., editors. Proceedings of the Second Annual Conference on Genetic Programming, pp. 58–63, 1997.

[22] Chung, W. W. C., Tam, M. M. C. and Saxena, K. B. C., "Evaluation of DSS use in Hong-Kong Manufacturing-Industries", Computers in Industry 21 (3): 307-324 APR 1993.

[23] Cielen, A., Peeters, L. and Vanhoof, K., "Bankruptcy prediction using a data envelopment analysis", European Journal of Operational Research, 154 (2), pp. 526-532, 2004.

[24] Curry, B., "Rough sets: current and future developments", The International Journal of Knowledge Engineering and Neural Networks, 20(5), pp. 247-250, 2003.

[25] Courtis, J. K., "Modeling a financial ratios categoric framework", Journal of Business Finance and Accounting, 5(4), pp. 371-386, 1978.

[26] Deakin B. E., "A discriminant analysis of predictors of business failure", Journal of Accounting Research, pp. 167-179, 1976.

[27] Dimitras, A. I., Zanakis, S. H. and Zopounidis, C., "A survey of business

failures with an emphasis on prediction methods and industrial applications", European Journal of Operational Research, Volume: 90, Issue: 3, pp. 487-513, 1996.

[28] Dimitras A. I., Slowinski R., Susmaga R. and Zopounidis C., "Business failure using rough set", European Journal of Operational Research, Volume: 114, Issue: 2 , pp. 263-280, 2003.

[29] Dimitras, A. I., Slowinski, R., Susmaga, R. and Zopounidis, C. , "Business failure prediction using rough sets", European Journal of Operational Research, 114(2), pp. 263-28, 1999.

[30] Doumpos, M. and Zopounidis, C., "Assessing financial risks using a multicriteria sorting procedure: the case of country risk assessment", Omega: International Journal of Management Science, 29 (1): 97-109 FEB 2001.

[31] Doumpos, M. and Zopounidis, C., "Rough sets and multivariate statistical classification: a simulation study" Computational Economics, 19(3), pp. 287-301, 2002.

[32] Etheridge, H. and Sriram, R., "A comparison of the relative costs of financial distress models: artificial neural networks, logit and multivariate discriminant analysis" , Intelligent Systems in Accounting, Finance and Management, Volume: 6, (1997), pp. 235--248

[33] Fletcher D. and Goss E., "Forecasting with neural networks: an application using bankruptcy data", Information and Management, Volume: 24, Issue: 3, pp. 159—167, 1993.

[34] Goh, C. and Law, R., "Incorporating the rough sets theory into travel demand analysis", Tourism Management, 24(5), 511-517, 2003.

[35] Goldberg, D.E., Genetic algorithms in search, optimization and machine learning, Addison-Wesley, Reading, MA., 1989.

[36] Greco, S., Matarazzo, B. and Slowinski, R., “A new rough set approach to evaluation of bankruptcy risk” In New Operational Tools in the Management of Financial Risks, Zopounidis C (ed.). Kluwer Academic: Dordrecht; pp.121-136, 1998.

[37] Greco, S., Matarazzo, B. and Slowinski, R., “Extension of the rough set approach to multicriteria decision support”, INFOR 38 (3): 161-195 AUG 2000.

[38] Greco, S., Matarazzo, B. and Slowinski, R., “Rough sets methodology for sorting problems in presence of multiple attributes and criteria”, European Journal of Operational Research, 138(2), pp. 247-259, 2002.

[39] Grice, S. J. and Dugan, T. M., “The limitations of bankruptcy prediction models: some cautions for the researcher”, Review of Quantitative Finance and Accounting, Volume: 17, pp. 151--166, 2001.

[40] Grzymala-Busse, J. and Zuo, X., "Classification strategies using certain and possible rules”, Lecture Notes in Artificial Intelligence, *RSCTC´98*, (1998), pp. 37-44 Springer.

[41] Hansen, J. V. and Messier, Jr. W. F., “Artificial neural networks: foundations and application to a decision problem”, Expert Systems With Applications, 3, pp. 135-141, 1991.

[42] Hashemi, R. R., Le Blanc, L. A. and Rucks, C. T., et al. ”A hybrid intelligent system for predicting bank holding structures”, European Journal of Operational Research ,109 (2), pp. 390-402, 1998.

[43] Hertz, J., Krogh, A. and Lautrup, B., et al., ”Nonlinear backpropagation:

Doing backpropagation without derivatives of the activation function", IEEE Transactions on Neural Networks , 8 (6), pp. 1321-1327, 1997.

[44] Jensen, R. and Shen, Q., "Semantics-preserving dimensionality reduction: Rough and fuzzy-rough-based approaches ", IEEE Transactions on knowledge and data engineering, 16 (12): 1457-1471 DEC 2004

[45] Jo, H., Han, I. and Lee, H., "Bankruptcy prediction using case-based reasoning, neural networks, and discriminant analysis", Expert Systems With Applications, Volume: 13, Issue: 2, pp. 97—108, 1997.

[46]. Keasey, K. and Watson, R., "Consensus and accuracy in accounting studies of decision-making - a note on a new measure of consensus", Accounting Organizations and Society 14 (4): 337-345 1989.

[47] Keasey, K. and Watson, R., "Financial distress prediction models: A review of their usefulness", British Journal of Management, Vol 2, pp.89-102, 1991.

[48] Kittler R. and Wang W. D., "The emerging role for data mining", Solid State Technology, 42 (11): 45, NOV 1999.

[49] Koh H. C. and Tan S. S., "A neural network approach to the prediction of going concern status", Accounting and Business Research, 29(3), pp. 211-216, 1999.

[50] Koza, J., Genetic programming, MIT Press, Cambridge, MA, 1992.

[51] Koza, J., "Genetic programming for economic modeling", Statistics and Computing, Volume: 4, Issue: 2, pp. 187-197, 1994.

[52] Kusiak, A, Kern J. A. and Kernstine, K. H., et al., "Autonomous decision-making: A data mining approach", IEEE Transactions on Information Technology in Biomedicine, 4 (4), pp. 274-284, 2000.

[53] Kusiak, A., "Rough set theory: A data mining tool for semiconductor

manufacturing”, IEEE Transactions on Electronics Packaging Manufacturing, 24 (1), pp. 44-50, JAN 2001.

[54] Lee, K. C., Han, I. G. and Kwon, Y., “Hybrid neural network models for bankruptcy predictions”, Decision Support Systems, Volume: 18, pp. 63-72, 1996.

[55] Leshno M., Spector Y., “Neural network prediction analysis: the bankruptcy case”, Neurocomputing, Volume: 10, pp. 125—247, 1996.

[56] Lin TY, “A rough logic formalism for fuzzy controllers: A hard and soft computing view”, International Journal of Approximate Reasoning 15 (4): 395-414 NOV 1996.

[57] Liu LC, Lee C, Tzeng GH, “Dea approach for the current and the cross period efficiency for evaluating the vocational education” , International Journal of Information Technology & Decision Making, 3 (2): 353-374, 2004

[58] Mak, B. and Munakata T., "Rule extraction from expert heuritics: a comparative study of rough sets with neural networks and ID3", European Journal of Operational Research, Volume: 136, pp. 212-229, 2002.

[59] McKee, T. E., “Developing a bankruptcy prediction model via rough sets theory”, International Journal of Intelligent Systems in Accounting, Finance & Management, 9, pp. 159-173, 2000.

[60] McKee, T. E., Greenstein M. “Predicting bankruptcy using recursive partitioning and a realistically proportioned data set”, Journal of Forecasting, 19, pp. 219-230, 2000.

[61] McKee, T. E., Lensberg T. “Genetic programming and rough sets: a hybrid approach to bankruptcy classification”, European Journal of Operational

Research 138: 436-451. 2002.

[62] McKee, T. E., "Rough sets bankruptcy prediction models versus auditor signalling rates", Journal of Forecasting, 22 (8), pp. 569-586, 2003.

[64] Messier, W. and Hansen J., "Inducing rules for expert system development: an example using default and bankruptcy data", Management Science, Volume: 34, Issue: 12 (1998), pp. 1403--1415

[65] Mi, J. S., Wu, W. Z. and Zhang, W. X., "Approaches to knowledge reduction based on variable precision rough set model", Information Sciences, 159(3-4), pp. 255-272, 2004.

[66] Michalewicz Z., Genetic algorithms + data structures = evolution programs, Springer, Berlin, 1996.

[67] Min, Jae H. and Lee, Young-Chan, "Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters", Expert Systems With Applications , 28(4), pp. 603-614, May, 2005.

[68] Mitra, S., Pal, S. K. and Mitra, P., "Data mining in soft computing framework: A survey ", IEEE Transactions on Neural Networks, 13 (1): 3-14 JAN 2002.

[69] Morris, R., Early Warning Indicators of Corporate Failure, Ashgate Publishing Ltd, Aldershot England, 1997.

[70] Odom, M. and Sharda, R., "A neural networks model for bankruptcy prediction", Proceedings of the IEEE International Conference on Neural Network, Volume: 2, (1990), pp. 163--168

[71] Ohlson, J., "Financial ratios and the probabilistic prediction of bankruptcy", Journal of Accounting Research, Volume: 18, Issue: 1, pp. 109—131, 1980.

[72] Ong, C. S., Huang, J. J. and Tzeng, G. H., "Using rough set theory for detecting the interaction terms in a generalized Logit model", Lecture Notes in Artificial Intelligence, 3066, pp.624-629, 2004

[73] Ong, C. S., Huang, J. J. and Tzeng, G. H., "Building credit scoring models using genetic programming",Expert Systems With Applications, 29 (1): 41-47 JUL, 2005.

[74] Pal, S. K., Mitra, S. and Mitra, P., "Rough-fuzzy MLP: Modular evolution, rule generation, and evaluation", IEEE Transactions on Knowledge and Data Engineering, 15 (1): 14-25 JAN-FEB 2003

[75] Pal S.K., Mitra P, "Case generation using rough sets with fuzzy representation", IEEE Transactions on Knowledge and Data Engineering, 16 (3): 292-300 MAR 2004

[76] Peel MJ, Wilson N, "The Liquidation merger alternative – some results for the UK corporate sector", Managerial and Decision Economics, 10 (3): 209-220 SEP 1989.

[77] Pendharkar P.C., "A potential use of data envelopment analysis for the inverse classification problem" Omega: International Journal of Management Science, pp. 243-248, 2002.

[78] Paradi, J.C., Asmild, M., Simak, P.C., "Using DEA and worst practice DEA in credit risk evaluation", Journal of Productivity Analysis, 21 (2), pp. 153–165, 2004.

[79] Paul Simak, Inverse and Negative DEA and their Application to Credit Risk Evaluation, Centre for Management of Technology and Entrepreneurship, Ph.D. Dissertations, University of Toronto, Canada, Advisor: J.C. Paradi, 2000.

[80] Pawlak, Z., "Rough sets", International Journal of Computer and Information Science", 11(5), 341-356, 1982.

[81] Pawlak, Z., "Rough Classification", International Journal of Man-Machine Studies, 20(5), 469-483, 1984.

[82] Pawlak Z., Rough sets. Theoretical aspects of reasoning about data, Kluwer Academic Publishers, London, 1991.

[83] Pawlak, Z., "Rough Sets" In: Lin, T. Y. & Cercone, N. (Eds.). Rough Sets and Data Mining: Analysis for Imprecise Data. Norwell, MA: Kluwer Academic Publishers, 1997.

[84] Pawlak, Z., "Rough sets, decision algorithms and Bayes' theorem", European Journal of Operational Research, 136(1), pp. 181-189, 2002.

[85] Pendharkar P. C., "A potential use of data envelopment analysis for the inverse classification problem", Omega: International Journal of Management Science, pp. 243-248, 2002.

[86] Predki, B., Slowinski, R., Stefanowski, J., Susmaga, R. and Wilk, S., "ROSE: software implementation of the rough set theory. Rough sets and current trends in computing", Lecture Notes in Artificial Intelligence, vol. 1424. Berlin: Springer, pp. 605–608, 1998.

[87] Predki, B. and Wilk, S., "Rough set based data exploration using ROSE system. Foundations of intelligent systems", Lecture Notes in Artificial Intelligence, vol. 1609. Berlin: Springer, pp. 172–80, 1999.

[88] Salchenberger, L., Cinar, E. and Lash, N., "Neural networks: a new tool for predicting thrift failures", Decision Sciences, Volume: 23, pp. 899-916, 1992.

[89] Seiford, Lawrence M. and Zhu, Joe: An acceptance system decision rule

with data envelopment analysis. Computers & Operations Research, 25 (4) pp. 329-332, 1998.

[90] Seiford, L. M. and Zhu J., “Context-dependent data envelopment analysis: measuring attractiveness and progress”, Omega: International Journal of Management Science, Volume: 31, Issue: 5, pp. 397-408, 2003.

[91] Shaw, M. and Gentry, J., “Inductive learning for risk classification”, IEEE Expert, pp. 47-53, 1990.

[92] Skowron, A. and Rauszer, C., "The discernibility matrices and functions in information systems", Slowinski R., (Ed.) Intelligent Decision Support—Handbook of Applications and Advances of the Rough Sets Theory, pp. 331-362 Kluwer Academic Publishers, 1992.

[93] Slowinski, R. and Zopounidis, C., “Business failure prediction using rough sets” International Journal of Intelligent Systems in Accounting, Finance & Management , 4 (1), pp.27-41, 1995.

[94] Slowinski, R., Zopounidis, C. and Dimitras, A. I., "Prediction of company acquisition in Greece by means of the rough set approach", European Journal of Operational Research, Volume: 100, pp. 1-15, 1997.

[95] Slowinski, R and Vanderpooten, D, “A generalized definition of rough approximations based on similarity”, IEEE Transactions on Knowledge and Data Engineering, 12 (2): 331-336 MAR-APR 2000

[96] Stefanowski, J., "On rough set based approaches to induction of decision rules", Polkowski J., Skowron A., (Ed.) Rough Sets in Knowledge Discovery, Volume: vol. 1, pp. 500-529, Physica-Verlag, Heidelberg, 1998.

[97] Suzuki, T, “Credit channel of monetary policy in Japan: resolving the supply versus demand puzzle, Applied Economics 36 (21), pp. 2385-2396,

DEC 10 2004.

[98] Tam, K. Y., "Neural network models and the prediction of bankruptcy", Omega: International Journal of Management Science, Volume: 19, Issue: 5, pp. 429-445, 1991.

[99] Tam, K. Y. and Kiang, M. Y., "Managerial applications of neural networks: the case of bank failure predictions", Management Science, Volume: 38, Issue: 7, pp. 926-947, 1992.

[100] Tay, F. E. H. and Shen, L. “Economic and financial prediction using rough sets model”, European Journal of Operational Research, 141(3), pp. 641–659, 2002.

[101] Theodossiou, P., Kahya, E., Saidi, R. and Philippatos, G., "Financial distress and corporate acquisitions: further empirical evidence", Journal of Business Finance, Volume: 23, Issue: 5–6, pp. 699-719, 1996.

[102] Troutt, M. D., Rai, Arun and Zhang, Aimao, “The potential use of DEA for credit applicant acceptance systems”, Computers & Operations Research, 23 (4), pp. 405-408, 1996.

[103] Tversky, A. and Simonson, I., "Context-dependent preferences" Management Sciences, pp. 1179-1189, 1993.

[104] Walczak, B. and Massart, D., “Rough sets theory”, Chemometrics and Intelligent Laboratory Systems, 47(1), pp.1-16, 1999.

[105] Wilson, R. L. and Sharda, R., "Bankruptcy prediction using neural networks", Decision Support Systems, Volume: 11, pp. 545-557, 1994.

[106] Wright, M. E. and Davidson, R. A., “The effect of auditor attestation and tolerance for ambiguity on commercial lending decisions”, Auditing - A Journal of Practice & Theory 19 (2): 67-81 FALL 2000.

[107] Yao, Y. Y., “Two views of the theory of rough sets in finite universes “, International Journal of Approximate Reasoning, 15 (4): 291-317, NOV 1996.

[108] Yao, Y. Y., “Information granulation and rough set approximation “, International Journal of Intelligent Systems 16 (1): 87-104, JAN 2001

[109] Yasdi, R.,"Combining Rough Sets Learning-Method and Neural Learning method to deal with uncertain and imprecise information”, Neurocomputing 7 (1), pp. 61-84, 1995.

[110] Zhang, G., Hu Y. M., Patuwo, E. B. and Indro, C. D., “Artificial neural networks in bankruptcy prediction: general framework and cross-validation analysis”, European Journal of Operational Research, Volume: 116, pp. 16-32, 1999.

[111] Ziarko, W., "Variable precision rough set model", Journal of Computer and System Sciences, Volume: 46, pp. 39-59, 1993.

[112] Ziarko, W., "Analysis of uncertain information in the framework of variable precision rough sets", Foundations of Computing and Decision Sciences, Volume: 18, Issue: 3–4, pp. 381-396, 1993.

[113] Zopounidis, C, Doumpos M and Matsatsinis N. F., “On the use of knowledge-based decision support systems in financial management: A survey “, Decision Support Systems 20 (3): 259-277, JUL 1997

[114] Zopounidis, C and Dimitras, A. I., Multicriteria Decision Aid Methods For The Prediction of Business Failure, Kluwer Academic: Dordrecht, 1998.

[115] Zopounidis, Constantin and Doumpos, Michael, "A preference disaggregation decision support system for financial classification problems", European Journal of Operational Research, pp. 402-413, 2001.

[116] Zopounidis, C. and Doumpos, M., "Multicriteria classification and sorting methods: A literature review", European Journal of Operational Research, 138 (2): 229-246 APR 16 2002.

Appendix A Financial Information in DEA Analysis

Co.		Output				Input					
ID	Liabilities/assets ratio (%)	Average collection days	Average inventory days	Current ratio (%)	Quick ratio (%)	Return on shareholders' equity (%)	Operating income to paid-in capital (%)	Profit before tax to paid-in capital (%)	Worst Practice DEA	Result	Failure
2301	40.27	71.15	37.86	133.07	109.97	98.25	45.76	99.65	0.207		
2302	17.24	226.7	246.62	192.19	67.02	58.88	35.18	40.54	1	X	
2303	26.87	44.95	43.45	413.66	359.62	82.72	42.22	80.16	0.12		
2305	45	78.15	84.29	112	66	81.58	43.13	77.58	0.325		
2308	39.15	85.08	12.15	95.44	90.56	94.4	43.33	108.69	0.204		
2311	42.5	72.27	28.14	101.73	80.51	79.9	51.92	73.62	0.294		
2312	23.89	69.52	7.21	191.19	173.05	89.78	46.65	88.78	0.121		
2313	41.37	69.25	45.68	302.62	229.97	72.32	38.01	55.59	0.257		
2314	45.89	83.52	163.67	250.93	185.77	65.83	43.74	58.07	0.325		
2315	39.57	97.85	38.91	152.88	113.36	85	44.8	84.01	0.229		
2316	36.6	86.08	23.77	143.87	119.38	86.92	49.34	86.55	0.205		
2317	48.12	56.5	21.79	132.87	101.47	106.54	78.56	168.45	0.171		
2318	57.83	528.98	44.4	69.3	66.65	11.04	22.28	25	1	V	V
2321	43.79	89.68	63.03	158.75	94.19	81.14	44.59	76.64	0.285		
2323	44.68	118.89	31.96	137.25	113.87	79.11	49.72	74.27	0.282		
2324	48.29	58.68	19.05	258.02	230.6	97.83	67.45	108.27	0.196		
2325	43.28	67.34	28.18	197.12	172.44	81.18	45.07	77.05	0.239		
2326	78.49	276.51	514.08	261.69	215.01	18.79	21.46	12.86	1	V	V
2327	37.48	114.06	98.64	80.47	68.8	67.48	45.01	55.16	0.457		
2328	15.32	78.66	18.22	338.68	319.75	82.43	46.62	79.81	0.077		
2329	64.18	43.71	22.69	48.64	38.92	50.06	23.76	50.48	0.704		
2330	20.04	40.19	31.54	304.07	264.11	87.12	59.28	89.19	0.089		
2331	59.04	70.32	21.23	184.89	156.57	107.91	131.62	156.21	0.205		
2332	42	96.81	31.87	264	235	88.58	61.13	94.58	0.187		
2333	44.57	91.25	172.98	109.34	53.18	70.91	34.36	62.1	0.661		
2336	42.54	121.66	12.16	155	141	82.01	46.66	78.75	0.249		
2337	53.48	62.6	228.12	126.63	86.68	48.62	24.63	44.81	0.79		
2338	28.29	94.07	27.48	405.6	381.71	91.12	57.95	90.56	0.116		
2340	50.08	115.87	128.97	204.9	152.04	81.18	45.38	73.28	0.315		
2341	52.36	87.52	100.27	248.19	158.66	63.97	34.31	53.12	0.382		
2342	64.9	41.28	79	40.01	22.94	1	10.3	58.06	1	V	V

Co.	Output				Input					
ID	Liabilities/assets ratio (%)	Average collection days	Average inventory days	Current ratio (%)	Quick ratio (%)	Return on shareholders' equity (%)	Operating income to paid-in capital (%)	Profit before tax to paid-in capital (%)	Worst Practice DEA	Result	Failure
2343	19	66.36	58.4	209	165	85.58	46.13	85.58	0.125		
2344	19.11	38.62	72.13	207.84	151.05	73.03	38.07	64.34	0.155		
2345	21.32	90.34	47.83	327.3	266.97	89.73	65.45	96.74	0.092		
2347	34.01	50.06	27.94	186.21	132.76	94.72	65.51	118.39	0.15		
2348	72.6	280.76	44.51	29.2	17.2	88.18	85.03	96.08	1	V	V
2349	40.59	180.69	66	202.39	186.9	73.35	36.48	54.2	0.289		
2350	52.76	74.94	30.41	117.07	90.53	82.83	51.4	79.59	0.334		
2351	46.25	82.39	134.19	126.12	71.4	78.31	45.18	73.25	0.431		
2352	41	70.87	16.5	137	121	100.58	80.13	126.58	0.176		
2353	26.02	111.62	17.6	169.68	155.14	93.35	46.44	119.95	0.112		
2354	29.48	82.76	9.35	288.76	270.25	88.98	45.68	87.73	0.137		
2355	35	89.46	18.75	358	342	88.58	74.13	100.58	0.137		
2356	39.44	70.05	17.79	174.4	141.66	93.09	52.43	96.43	0.196		
2357	14.7	40.46	40.82	418.89	348.4	94.77	78.65	129.41	0.05		
2358	73.78	39.97	6.6	73.97	73.93	27.59	30.65	47.26	0.803		
2359	57	104.58	28.07	62	51	54.58	42.13	46.58	0.627	X	V
2360	31.77	105.18	79.52	205.51	156.17	91.21	60.71	95.71	0.178		
2361	27.6	55.3	19.51	285.9	239	93.98	60.43	89.38	0.125		
2362	28.56	28.71	41.52	235.76	161.93	82.68	47.92	78.99	0.159		
2363	55.93	77	105.49	139.1	94.91	53.99	26.15	37.13	0.606		
2364	58.54	92.17	96.05	158.04	111.72	61.57	33.95	63.16	0.43		
2365	31.32	43.71	6.85	191.04	183.11	93.14	68.03	105.72	0.142		
2366	51.83	57.12	51.55	175.14	103.13	97.05	69.78	101.14	0.265		
2367	35.34	74.79	16.67	94.65	85.9	75.44	57.46	71.5	0.254		
2368	51.81	114.42	43.71	98.21	84.98	40.9	23.71	8.84	1	X	
2369	25.68	89.68	25.32	144.96	125.29	70.03	40.62	61.59	0.18		
2370	16.07	111.62	106.41	388.14	256.55	81.15	41.66	77.64	0.163		
2371	59.44	80.39	48.86	88.82	59.28	66.92	39.27	63.49	0.495		
2373	58.92	56.58	51.33	90.28	68.95	97.32	51.39	91.82	0.355		
2374	35.17	87.11	36.57	81.26	65.71	88.19	43.52	87.2	0.226		
2375	18.13	229.55	82.02	280.79	224.32	82.99	39.06	80.63	0.238		

Co.	Output					Input					
ID	Liabilities/assets ratio (%)	Average collection days	Average inventory days	Current ratio (%)	Quick ratio (%)	Return on shareholders' equity (%)	Operating income to paid-in capital (%)	Profit before tax to paid-in capital (%)	Worst Practice DEA	Result	Failure
2376	35.91	42.19	44.24	209.05	180.78	97.03	93.3	131.54	0.138		
2377	49.69	46.2	47.34	192.7	135.8	102.13	104.97	142.27	0.19		
2379	13.15	56.94	66.36	911	865.44	105.35	99.55	136.74	0.037		
2380	21.51	62.82	55.98	215.2	163.82	88.41	49.78	88.76	0.119		
2381	43	48.53	30.04	219	174	82.58	48.13	77.58	0.234		
2382	51.39	87.74	27.32	156.38	132.44	104.05	71.78	121.17	0.221		
2383	22.6	107.66	103.69	230.29	158.69	73.33	43.35	66.83	0.234		
2384	41.9	76.04	125.86	206.7	122.4	82.18	49.43	80.28	0.295		
2385	35.39	60.73	44.84	136.9	91.99	113.73	54.53	118.53	0.164		
2387	27.45	76.51	55.47	185.24	154.26	100.3	53.85	98.51	0.139		
2388	32.86	66.84	95.54	281.63	212.82	81.33	66.36	78.69	0.177		
2389	30.31	140.92	36.57	127.4	112.39	77.12	45.46	71.62	0.207		
2391	17.58	64.37	46.97	337.17	284.36	100.72	80.67	119.7	0.067		
2392	27.66	84.88	48.08	211.19	168.09	94.2	99.78	128.28	0.112		
2393	26	80.21	71.15	185	138	93.58	68.13	103.58	0.159	X	V
2394	34.62	60.73	52.51	291.47	194.74	95.58	69.13	105.58	0.152		
2395	32.58	70.19	54.07	369.38	303.02	105.99	96.91	130.43	0.11		
2396	46.03	199.45	55.21	117.85	100.17	76	46.41	66.02	0.341		
2397	21	74.48	36.5	461	420	83.58	52.13	85.58	0.086		
2398	49	192.1	68.73	342.61	302.16	81.58	61.13	79.58	0.226	X	V
2399	60.93	88.16	52.66	136.22	101.23	92.04	62.61	89.86	0.343		
2401	9.22	53.44	53.91	769.59	683.98	96.26	68.13	105.72	0.042		
2403	25.79	47.83	28.92	312.68	223.57	90.4	82.68	108.36	0.107		
2404	53.21	70.73	480.26	152.33	117.61	98.15	87.69	111.57	1	X	
2405	41.37	77.65	21.62	160.16	139.08	92.91	72.38	101.67	0.201		
2406	33.41	148.37	52.89	143.74	122.5	64.48	26.46	51.8	0.28		
2407	40.25	107.66	29.57	146.82	87.7	61.45	35.43	53.72	0.344	X	V
2408	48.6	43.5	96.56	161.82	99.92	85.74	53.37	82.14	0.309		
2409	39	46.43	49.59	208	173	89.58	63.13	90.58	0.194		
2410	45.75	68.73	79.34	109.27	64.95	68.39	24.37	68.8	0.361		
2412	16.68	39.41	34.14	83.4	79.91	91.37	99.54	133.33	0.105		

Co. ID	Output: Liabilities/assets ratio (%)	Average collection days	Average inventory days	Current ratio (%)	Quick ratio (%)	Input: Return on shareholders' equity (%)	Operating income to paid-in capital (%)	Profit before tax to paid-in capital (%)	Worst Practice DEA	Result	Failure
2413	36.56	87.52	131.29	197.54	130.34	83.54	51.86	81.86	0.287		
2414	53.5	77.49	29.31	165.42	120.32	87.07	53.53	86.55	0.299		
2415	36.93	69.92	36.06	133.12	120.83	90.74	57.49	99.26	0.191		
2416	52.08	57.03	45.85	182.88	109.77	88.79	78.2	103.33	0.263		
2417	31.09	58.02	53.28	340.94	300.45	92.55	57.97	95.16	0.13		
2418	36.12	91.93	84.88	225.66	141.79	97.51	78.72	111.63	0.178		
2419	60.27	160.08	77	150.75	119.01	81.31	44.44	75.37	0.372		
2420	32.33	84.29	96.56	163.01	103.41	93.53	67.41	96.77	0.239		
2421	34.48	91.7	109.6	182.99	109.63	84.28	52.02	83.36	0.274		
2422	43.79	102.81	203.91	247.26	158.66	88.5	53.46	84.95	0.372		
2423	22.7	94.07	102.24	206.78	141.93	89.48	52.98	88.65	0.222		
2424	30.4	60.43	23.82	130.36	98.29	72.6	35.09	66.66	0.215		
2426	45	144.26	215.97	197	124	84.58	48.13	80.58	0.485		
2427	38.46	131.76	101.67	191.27	123.64	82.46	50.62	81.09	0.274		
2428	21	143.7	95.3	333	283	91.1	60.73	96.84	0.149		
2429	50.84	173.8	75.88	92.02	76.21	51.32	25.51	37.71	0.597		
2430	53.16	30.77	53.05	151.21	89.67	96.97	72.69	125.81	0.238		
2432	18.3	72.13	122.07	79.58	53.63	77.36	55.42	72.6	0.487		
2433	29.72	72.56	133.69	90.09	65.26	83.58	48.65	77.46	0.472		
2434	43	146	260.71	906	529	83.48	48.73	80.28	0.205		
2435	75.79	128.97	51.12	73.26	49.78	41.52	26.41	58.09	0.723		
2436	8.52	84.29	57.84	833.05	699.34	93.79	61.55	94.07	0.067		
2437	8.45	218.56	91.93	971.61	911.86	84.29	50.04	88.7	0.188		
2438	36.07	93.58	30.67	121.68	92.84	69.38	47.91	62.79	0.272	X	V
2439	32.15	83.71	52.59	184.53	176.32	87.86	53.17	94.78	0.162		
2440	57.6	101.38	110.6	164.7	93.4	81.78	50.73	82.98	0.367		
2441	36.48	92.63	25.56	213.44	179.14	97.34	74.16	103.05	0.164		
2442	39.79	52.36	1.04	197.82	195.21	76.47	38.32	68.48	0.24		
2443	36.56	116.24	58.87	73.48	47.41	69.39	33.85	64.52	0.349		
2444	53.43	57.38	86.9	323.42	225.85	85.38	51.2	85	0.258		
2445	53.52	154.1	73.44	96.16	12.78	55.17	30.79	58.09	1	V	V

Co. ID	Output			Input							
	Liabilities/assets ratio (%)	Average collection days	Average inventory days	Current ratio (%)	Quick ratio (%)	Return on shareholders' equity (%)	Operating income to paid-in capital (%)	Profit before tax to paid-in capital (%)	Worst Practice DEA	Result	Failure
2446	38.71	90.34	44.24	140.19	108.56	72.38	35.81	66.03	0.269		
2447	66.18	118.12	13.36	132.63	124.53	97.62	75.3	101.66	0.335		
2448	37.38	87.74	158	199.51	147.93	100.87	83.9	114.04	0.293		
2449	49.09	73.29	4.14	88.95	87.21	76.21	38.14	60.44	0.404		
2450	49.38	38.25	41.33	266.18	158.84	99.95	72.76	107.55	0.221		
2451	15.91	41.8	46.61	473.5	360.92	114.43	105.32	146.76	0.049		
2452	11.43	91.25	100.55	471.84	336.51	93.24	61.79	101.82	0.113		
2453	25.5	140.38	96.05	316.8	237	78.78	37.93	71.08	0.18		
2454	21.57	43.81	29.03	351.23	325.25	145.68	318.44	354.17	0.039		
2455	36.28	68.73	144.84	125.68	76.21	62.13	9.68	40.35	0.592		
2456	44.19	148.97	120.46	334.9	262.48	87.63	48.39	86.31	0.208		
2457	47.58	64.37	50.41	146.5	123.6	92.44	57.24	100.67	0.238		
2458	27.62	33.98	70.46	485.95	394.69	92.61	65.66	91.11	0.106		
2459	48.95	58.87	9.65	159.18	146.92	103.57	62.26	112.19	0.221		
2460	45.55	99.45	91.93	224.22	157.69	98.16	78.2	107.69	0.21		
2461	60.87	74.48	42.34	211.23	201	95.61	58.09	102.73	0.272		
2462	25.33	109.6	8.59	280.65	267.18	83.64	56.56	83.49	0.121		
2463	33.33	49.52	87.32	354.87	226.13	83.54	50.35	80.93	0.17		
2464	47.39	85.68	143.7	180.07	85.61	98.16	65.58	101.69	0.351		
2466	42.76	88.59	189.11	437.16	268.71	80.07	44.3	71.83	0.252		
2467	23.72	106.41	133.69	264.23	189.73	86.86	53.32	87.84	0.232		
2468	48	109.93	29.34	242	212	85.58	54.13	86.58	0.232		
2469	33.36	114.77	61.03	555.64	439.6	96.8	49.67	104.92	0.118		
2470	61.21	74.03	50.48	237.31	152.64	89.98	72.88	98.43	0.297		
2471	15.03	125.86	89.68	612.72	559.36	73.46	31.94	63.09	0.181		
2472	34.37	135.68	52.74	175.61	146.54	88.5	62.66	96.17	0.177		
2474	27.12	67.21	59.25	642	564.73	90.72	76.85	104.09	0.087		
2475	44.27	63.14	43.34	144.79	113.59	72.62	44.24	69.21	0.296		
2477	36.73	56.5	4.95	144.79	113.59	106.79	69	138.5	0.147		
2478	11.05	152.08	98.91	566.25	497.5	87.57	57.5	92.47	0.126		
2479	49.59	153.36	51.04	152.95	136.08	83.86	36.67	79.75	0.288		

Co.	Output					Input					
ID	Liabilities/assets ratio (%)	Average collection days	Average inventory days	Current ratio (%)	Quick ratio (%)	Return on shareholders' equity (%)	Operating income to paid-in capital (%)	Profit before tax to paid-in capital (%)	Worst Practice DEA	Result	Failure
2480	48.44	136.19	102.81	265.79	184.8	101.82	89.41	119.68	0.202		
2481	53.28	106.41	138.78	186.58	107.5	87.87	70.48	92.38	0.34		
2482	21.69	67.97	87.32	291.64	208.44	89.74	48.5	87.07	0.141		
2483	6.07	69.52	132.24	733.83	629.53	86.68	57.06	91.69	0.09		
2484	30.02	100.82	158	165.47	107.38	77.33	42.82	68.38	0.413		
2486	54.61	110.27	61.55	123.87	104.64	84.38	49.83	82.19	0.329		
2487	7	120.46	58.4	1,488.00	1,438.00	80.93	53.13	78.58	1		
2488	33.13	30.77	35.54	317.12	262.89	98.11	54.44	103.86	0.135		
2489	45.71	77.33	22.67	420.48	391.78	92.17	60.26	99.51	0.175		
2490	37.33	179.8	140.92	427.32	300.01	31.37	35.13	1	1	V	V
2491	25.04	84.1	41.24	158.71	150.07	81.33	52.01	78.64	0.151	X	V
2492	35.56	138.25	110.27	314.85	249.96	82.38	54.83	81.44	0.184		
2493	29.03	158.69	85.68	295.65	244.04	82.15	43.89	78.1	0.168		
2494	35.28	52.14	45.51	249	161	74.47	37.13	67.65	0.218	X	V
2495	10.86	63.69	160.08	867.86	708.2	100.93	71.99	102.51	0.09		
2496	20.01	105.49	456.25	348.16	136.47	57.4	28.72	53.38	1	X	
2497	23.9	74.94	129.89	217.44	140.94	93.01	61.81	94.94	0.253		
2498	54.24	64.26	38.66	133.14	95.08	118.95	172.32	168.26	0.189		
2499	33.52	194.14	113.35	304.7	217.88	89.59	58.18	96.18	0.218		
3002	22.69	96.05	12.35	267.86	258.24	88.81	47.51	88.93	0.107		
3003	16.58	48.08	110.27	424.41	311.48	98.66	75.43	112.44	0.121		
3004	54.73	197.29	160.79	245.02	187.24	81.83	60.91	79.91	0.329	X	V
3005	49.57	47.77	32.47	176.25	122.64	100.44	63.77	110.5	0.23		
3006	23.85	40.69	71.85	375.74	291.72	111.46	105.05	160.77	0.081		
3007	30.86	66.72	22.44	176.59	151.33	100.47	60.5	113.28	0.136		
3008	16.32	77.65	23.51	472	455	106.46	90.49	183.22	0.042		
3009	59.82	48.73	61.34	107.66	67.04	96.7	71.73	99.9	0.338		
3010	44.57	97.33	52.82	182.85	139.55	102.2	93.01	132.93	0.178		
3011	44.48	131.29	80.04	266.02	226.36	91.83	59.08	96.3	0.199		
3012	57.16	55.98	65.41	107.14	63.51	78.68	43.85	72.88	0.421		
3013	46	98.64	14.44	131	115	96.58	66.13	104.58	0.229		
3014	23.56	82.76	56.5	338.51	302.63	91.71	60.75	92.98	0.101		
3015	50	112.65	47.65	164	130	108.58	109.13	140.58	0.193		
3016	16.94	55.81	22.04	238.56	204.99	86.94	50.66	83.13	0.085		
3017	50.1	73	37.86	136.32	107.35	102.24	69.2	112.23	0.235		
3018	22.3	104.58	131.29	379.17	365.75	90.98	62.05	97.73	0.158		
3019	45.7	59.64	14.47	221.64	209.77	113.75	61.35	181.61	0.137		
3020	64.8	97.59	24.79	170.04	139.65	93.26	75.72	104.1	0.308		
3021	59.02	105.18	77.99	331.84	267.38	84.47	67.36	87.12	0.264	X	V
5336	67.8	70.32	33.21	345.52	311.12	50.92	33.32	42.34	0.459	X	V
5347	54.64	45.85	45.45	123.55	94.34	53.96	22.97	59.47	0.438	X	V
5385	90.42	180.69	133.7	34.22	24.74	38.45	1	54.03	1	V	V
6157	38.34	106.72	201.65	179.47	76.41	78.95	49.45	75.26	0.582	X	V
8011	63.72	96.56	244.96	110.47	43.42	65.87	38.07	55.84	1	V	V
8012	12.75	146	116.61	522.13	370.92	79.49	43.54	75.9	0.164	X	V

Appendix B Risk Level Prediction Obtained through Hybrid Model

Highest Risk	High Risk	Medium Risk
First Level	**Second Level**	**Third Level**
2302	2363	2440
2318	2429	2410
2326	2455	2373
2342	6157	2464
2348	2371	2443
2368	2432	2407
2404	2426	2399
2445	2433	2396
2487	5336	2481
2490	2327	3009
2496	5347	2447
5385	2351	2350
8011	2364	2486
2358	3012	3004
2337	2484	2305
2435	2449	2314
2329	2341	2340
2333	2419	2408
2359	2422	3020
2302	2363	2440
2318	2429	2410
2326	2455	2373
2342	6157	2464
2348	2371	2443
2368	2432	2407
2404	2426	2399
2445	2433	2396
2487	5336	2481

Appendix C Distress Related Regulations in Taiwan

我國法律上財務危機定義－台灣證券交易所

台灣證券交易所股份有限公司營業細則		
第４９條	第５０條	第５１條
變更交易方法為全額交割	停止買賣	終止其上市
１．其依證交法第三十六條規定公告並申報之最近期財務報告顯示淨值已低於實收資本額二分之一者。 ２．未於營業終了後六個月內加開股東常會。 ３．經會計師簽發保留意見者。 ４．違反上市公司重大訊息查證暨公開相關章則規定，經通知補行辦理公開程度，未依限期辦理且個案情節重大者。 ５．依公司法第２８２條規定向法院聲請重整者。 ６．其他原因。	１．其淨值仍為負數。 ２．公司營運全面停頓，暫時無法恢復者。	１．裁定解散公司之經營有顯著困難或重大損害。 ２．公司法規定解散事由包括破產、合併及解散等。 ３．法院裁定宣告破已確定者。 ４．公司營業範圍有重大變更，不宜繼續上市買賣者。

我國法律上財務危機定義－公司法、銀行法

公司法		銀行逾期放款、催收款、及呆帳處理辦法		
第２１１條	第２８２條	第二條	第三條	第四條
宣告破產	公司重整	逾期	催收款	呆帳
１.公司虧損達實收資本額二分之一。 ２.公司資產顯有不足抵償其負債，得依解散辦理者外，董事會應即聲請宣告破產。	公開發行股票或公司債之公司有： (1)財務困難 (2)暫停營業 (3)或有停業之虞者法院得依準申請，裁定准定予重整。	已屆清償期而未受清償之各項放款及其他授信款項。	１.系指經轉入「催收款項」科目之各項放款及授信款項。 ２.凡逾期放款應於清償期屆滿六個月內轉入催收款項。	１.債務人因故致債權人全部或一部份不能收回者。 ２.擔保品已無法受償者。 ３.催收款逾清償期二年經催收未收回

資料來源：整理自我國法律上財務危機的定義

Appendix D Lingo Code for Decision Rule Induction

```
Model:

Sets:
    DMU/1..200/:object, d_value;    ! there are 200 DMUs;
    DMU_class/1..200/:DC;         ! each DMU is assigned to class;
    criteria_A/1..6/:CC;          ! there are 6 atributes;
    Criteria_level/1..2/:A_level;   ! each attribute has 2
classification levels;
    jl(criteria_A, Criteria_level):g;
    a_jl(DMU,criteria_A,criteria_level):av; ! attribute value for each
DMU;
    classification_level/1..2/:cl;      ! each object has 4 values for ;
                             ! 1. group indx, 2. starting indx,
3.ending indx;
    group_index/1..2/:temp_cg;
    company_group(classification_level, group_index):cg;

    criteria_control(DMU,DMU):v, u, pn, pq, check_pn, check_pq,
check_final;
                    ! u: covered by classification rule;
                    ! v: not covered by classification rule;
                    ! pn is discernibility index (between different
groups);
                    ! pq is undiscernibility index (same group);

    d_jl(DMU,criteria_A,criteria_level):d;      ! decision rule;
    rule_control(DMU): accuracy, coverage;      ! sum of rule d ;
    filnal_control(classification_level,criteria_A,criteria_level):fi
nal_d;
Endsets

Data:
    av=@OLE(c:\jjs\VPRS3.xls); ! attribute value of companies;
    dc=@OLE(c:\jjs\VPRS3.xls); ! group number of companies;
      @OLE(c:\jjs\VPRS3.xls)=d;
```

```
        @OLE(c:\jjs\VPRS3.xls)=v;
        @OLE(c:\jjs\VPRS3.xls)=u;

        @OLE(c:\jjs\VPRS3.xls)=pq;
        @OLE(c:\jjs\VPRS3.xls)=check_pn;
        @OLE(c:\jjs\VPRS3.xls)=check_pq;

        @OLE(c:\jjs\VPRS3.xls)=accuracy;
        @OLE(c:\jjs\VPRS3.xls)=coverage;
Enddata

min=@sum(d_jl(p,j,l):d(p,j,l));

! ------------  test decision rule -------------;
@for(DMU(i):d_value(i)=@sum(d_jl(p,j,l)|p #eq# i:d(p,j,l)));

m=9999;
e=0.001;

! group range of companies;
    cg(1,1)=1;  ! starting indx;
    cg(1,2)=3;  ! ending indx;
    cg(2,1)=4;
    cg(2,2)=7;

! ----------------- discernibility with big M -------;

@for(DMU(p)|(p #ge# cg(1,1)) #and# (p #le# cg(1,2)):
    @for(DMU(n)|n #ge# cg(2,1):
        ! discernible attributes;
        pn(p,n)=@sum(a_jl(p,j,l):av(p,j,l)*(1-av(n,j,l)));
        pn(p,n)>= 1+ M*(v(p,n)-1); pn(p,n)<= M*v(p,n);
        ! create discernibility ;
        check_pn(p,n)=@sum(d_jl(p,j,l)|(av(p,j,l) #eq# 1) #and#
(av(n,j,l) #eq# 0):d(p,j,l)*(1-av(n,j,l)));
        check_pn(p,n)>= 1 + M*(v(p,n)-1);check_pn(p,n)<= M*v(p,n);
    ));
```

```
@for(DMU(p)|(p #ge# cg(2,1)) #and# (p #le# cg(2,2)):
    @for(DMU(n)|n #le# cg(1,2):
        pn(p,n)=@sum(a_jl(p,j,l):av(p,j,l)*(1-av(n,j,l)));
        pn(p,n)>= 1+ M*(v(p,n)-1); pn(p,n)<= M*v(p,n);
        check_pn(p,n)=@sum(d_jl(p,j,l)|(av(p,j,l) #eq# 1) #and#
(av(n,j,l) #eq# 0):d(p,j,l)*(1-av(n,j,l)));
        check_pn(p,n)>= 1 + M*(v(p,n)-1); check_pn(p,n)<= M*v(p,n);
    ));

! --- undercinibility with big M ;
@for(criteria_control(p,q)|(p #ge# cg(1,1)) #and# (p #le# cg(1,2)):
        ! undiscernibility attributes;
        pq(p,q)=@sum(d_jl(p,j,l):d(p,j,l)*(1-av(q,j,l)));
        pq(p,q) >= e - M*u(p,q); pq(p,q) <= e+ M*(1-u(p,q));
        ! create undiscernibility;
        check_pq(p,q)=@sum(d_jl(p,j,l)| av(p,j,l) #eq# 1 :
d(p,j,l)*(1-av(q,j,l)));
        check_pq(p,q)>= 1 - M*u(p,q); check_pq(p,q)<= 1+ M*(1-u(p,q));
 );

@for(criteria_control(p,q)|(p #ge# cg(2,1)) #and# (p #le# cg(2,2)):
        ! undiscernibility attributes;
        pq(p,q)=@sum(d_jl(p,j,l):d(p,j,l)*(1-av(q,j,l)));
        pq(p,q) >= e - M*u(p,q); pq(p,q) <= e+ M*(1-u(p,q));
        ! create undiscernibility;
        check_pq(p,q)=@sum(d_jl(p,j,l)| av(p,j,l) #eq# 1 :
d(p,j,l)*(1-av(q,j,l)));
        check_pq(p,q)>= 1 - M*u(p,q); check_pq(p,q)<= 1+ M*(1-u(p,q));
 );

!----------  Accuracy ----------------------;
@for(rule_control(p)|(p #ge# cg(1,1)) #and# (p #le# cg(1,2)):
    accuracy(p)=(@sum(criteria_control(p,i)|(i #ge# cg(1,1)) #and# (i
#le# cg(1,2)):u(p,i))+
             @sum(criteria_control(p,j) |(j #ge# cg(2,1)) : v(p,j)))
             /(cg(2,2)-cg(1,1)+1));

@for(rule_control(p)|(p #ge# cg(2,1)) #and# (p #le# cg(2,2)):
```

```
    accuracy(p)=(@sum(criteria_control(p,i)|(i #ge# cg(2,1)) #and# (i
#le# cg(2,2)):u(p,i))+
             @sum(criteria_control(p,j) |(j #le# cg(1,2)) : v(p,j)))
             /(cg(2,2)-cg(1,1)+1));

@for(rule_control(p):accuracy(p) <=1);
!@for(rule_control(p):accuracy(p)>=0.3);
!accuracy(1)>=0.7;

! ------------------ coverage---------------;
@for(rule_control(p)|(p #ge# cg(1,1)) #and# (p #le# cg(1,2)):
    coverage(p)=@sum(rule_control(i)|(i #ge# cg(1,1)) #and# (i #le#
cg(1,2)):u(p,i))/(cg(1,2)-cg(1,1)+1));

@for(rule_control(p)|(p #ge# cg(2,1)) #and# (p #le# cg(2,2)):
    coverage(p)=@sum(rule_control(i)|(i #ge# cg(2,1)) #and# (i #le#
cg(2,2)):u(p,i))/(cg(2,2)-cg(2,1)+1));

!@for(rule_control(p):coverage(p)>=0.2);
@for(rule_control(p):coverage(p)<=1);
coverage(1)>=0.6;
@for(criteria_control(p,n):@gin(check_pn(p,n)));
@for(criteria_control(p,q):@gin(check_pq(p,q)));
@for(criteria_A(i):@bin(cc(i)));
@for(a_jl(i,j,l):@bin(av(i,j,l)));
@for(d_jl(k,j,l):@bin(d(k,j,l)));
@for(criteria_control(k,i):@bin(u(k,i)); @bin(v(k,i)));
```

簡歷

姓名： 帥嘉珍 (Jia-Jane Shuai)

學歷： 國立交通大學 資訊管理所 博士，2005

Master of Computer Science, University of Illinois at Urbana-Champaign, U.S.A., 1993

Master of Accountancy, University of Illinois at Urbana-Champaign, U.S.A., 1991

國立台灣大學 會計學系 學士，1986

經歷：

明新科技大學 資訊管理系 專任講師，1995 迄今

大華技術學院 資訊管理系 專任講師，1994 - 1995

安達信顧問公司(Andersen Consulting & Co.) 系統分析師，1988 – 1990

勤業會計師事務所 (Arthur Andersen) 查帳員，1986 – 1988

證照： 美國會計師 (Certified Public Accountant)，1991

【期刊論文】

1. Jia-Jane Shuai and Han-Lin Li, “Using Rough Set and Worst Practice DEA in Business Failure Prediction”, accepted by *Lecture Notes in Computer Science,* (SCI).

2. Han-Lin Li and Jia-Jane Shuai, “Time and Money Trade-off in Competence Set Expansion”, *International Journal of Information Technology and Decision Making,* Vol. 3, No. 2, (2004) pp. 251-263, (SCI).

3. Jia-Jane Shuai, Gwo-Hshiung Tzeng and Han-Lin Li, “The Multi-source Fabless-Foundry Partnership Selection Model”, *International Journal of Manufacturing Technology and Management,* Vol. 6, No. 1/2, (2004) pp. 137-154.

【研討會論文】

1. Jia-Jane Shuai and Han-Lin Li (2005) Using Rough Set and Worst Practice DEA in Business Failure Prediction. *The Tenth International Conference on Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing World Conference*, Regina, Canada.
2. Jia-Jane Shuai and Han-Lin Li (2002) Multi Source Fabless-Foundry partnership Decision Model. *Proceedings of the 16th Multiple Criteria Decision Making World Conference*, Semmering, Austria.
3. Han-Ying Kao, Jia-Jane Shuai, and Han-Lin Li (2000) Probabilistic Reasoning Models in Supply Chain Alliance. *Proceedings of the 5th Annual International Conference on Industrial Engineering - Theory, Applications and Practice*, HsinChu, Taiwan.

【專書論文】

1. Han-Lin Li and Jia-Jane Shuai (2001), The Three Vs of global competitiveness. In: *Made by Taiwan*. C-Y Chang and P-L Yu Eds. *World Scientific Publishing*, Singapore, pp. 27-42.

2. Han-Lin Li and Jia-Jane Shuai (2000), A Decision Support System for Incorporating Competencies of a Software Company. In: *New Frontiers of Decision Making for the Information Technology Era*. Yong Shi and Milan Zeleny Eds. *World Scientific Publishing*, New Jersey, NY, pp. 31-47.