VALIDATING BANKRUPTCY PREDICTION BY USING BAYESIAN NETWORK MODEL: A CASE FROM MALAYSIAN FIRM

By

MUHAMMAD ZUHAIRI BIN ABD HAMID

Thesis Submitted to
Thesis/Dissertation Submitted to the
Othman Yeop Abdullah Graduate School
of Business, Universiti Utara Malaysia,
in Fulfillment of the Requirement for the Degree of
Doctor of Philosophy /Doctoral of Business
Administration/ Master of Science

Othman Yeop Abdullah
Graduate School of Business

Universiti Utara Malaysia

PERAKUAN KERJA KERTAS PROJEK
(Certification of Project Paper)

Saya, mengaku bertandatangan, memperakukan bahawa
(I, the undersigned, certified that)
MUHAMMAD ZUHAIRI BIN ABD HAMID (812905)

Calon untuk Ijazah Sarjana
(Candidate for the degree of) **MASTER OF SCIENCE (FINANCE)**

telah mengemukakan kertas projek yang bertajuk
(has presented his/her project paper of the following title)

VALIDATING BANKRUPCY PREDICTION BY USING BAYESIAN NETWORK MODEL: A CASE FROM MALAYSIAN FIRM

Seperti yang tercatat di muka surat tajuk dan kulit kertas project
(as it appears on the title page and front cover of the project paper)

Bahawa kertas projek tersebut boleh diterima dari segi bentuk serta kandungan dan meliputi bidang ilmu dengan memuaskan.
(that the project paper acceptable in the form and content and that a satisfactory knowledge of the field is covered by the project paper).

Nama Penyelia *(Name of Supervisor)* : **DR. NORSHAFIZAH BINTI HANAFI**

Tandatangan *(Signature)* :

Tarikh *(Date)* : **21 DECEMBER 2014**

PERMISSION TO USE

In presenting this dissertation / project paper in partial fulfillment of requirement for a post graduate Master from Universiti Utara Malaysia, I agree that the library of this university may make it freely available for inspection. I further agree that permission for copying this dissertation / project paper in any manner in whole or in part, for scholarly purposes may be granted by my supervisor(s) or in their absence, by Dean of Othman Yeop Abdullah Graduate School of Business where I did my dissertation / project paper. It is understood that any copying or publication or use of this dissertation / project paper part of it financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the UUM in any scholarly use which may be made of any material in my dissertation / project paper.

Request for permission to copy or to make other use of material in this dissertation / project paper in whole or in part should be address:

Dean of Othman Yeop Abdullah Graduate School of Business

Universiti Utara Malaysia

06010 UUM Sintok

Kedah Darul Aman

Abstract

This paper provides operational guidance for validating Naïve Bayes model for bankruptcy prediction. First, researcher suggests heuristic methods that guide the selection of bankruptcy potential variables. Correlations analyses were used to eliminate variables that provide little or no additional information beyond that subsumed by the remaining variables. A Naïve Bayes model was developed using the proposed heuristic method and it performed well based on logistic regression, which is used for validation analysis. The developed Naïve Bayes model consists of three first-order variables and seven second-order variables. The results show that the model's performance is best when the method of *enter* is used in logistic regression which is percentage of correct is 90%. Finally, the results of this study could also be applicable to businesses and investors in decision making, besides validating bankruptcy prediction.

Keywords: Bankruptcy prediction, financial distress, Naïve Bayes model, Variables selection, Logistic regression

Abstrak

Karya ini memberi panduan operasi untuk mengesahkan model naive Bayes untuk ramalan muflis. Pertama, penyelidik mencadangkan kaedah heuristik yang membimbing pemilihan muflis potensi bolehubah. Berdasarkan korelasi dan korelasi separa antara pemboleh ubah, matlamat kaedah ini adalah untuk menghapuskan pembolehubah yang memberikan maklumat tambahan sedikit atau tidak lebih dari itu digolongkan oleh pembolehubah yang tinggal. Model Bayes naif dibangunkan dengan menggunakan kaedah heuristik yang dicadangkan dan didapati prestasi yang baik berdasarkan regresi logistik yang digunakan untuk analisis pengesahan. Model Bayes naif dibangunkan terdiri daripada tiga pembolehubah yang mula- perintah dan tujuh pembolehubah tertib kedua . Keputusan kami menunjukkan bahawa prestasi model adalah yang terbaik apabila kaedah memasukkan digunakan dalam regresi logistik yang merupakan hasil daripada peratusan yang betul ialah 90 %. Akhir sekali, hasil kajian ini juga boleh diguna pakai untuk membuat perniagaan dan pelabur keputusan konteks yang lain daripada mengesahkan ramalan muflis

Kata kunci: ramalan Kebankrapan, tekanan kewangan, model naif Bayes, pemilihan Pembolehubah, regresi logistic.

Acknowledgement

In the name of Allah, most Gracious, most Merciful.

First and foremost, I would like to thank Universiti Utara Malaysia's Dean of Othman Yeop Abdullah Graduate School of Business for giving me this chance to carry out this project paper. It has been an invaluable experience for me and I have succeeded these far thanks to his support.

I would like to express my appreciation and thankful to my supervisor, Dr. Norshafizah bt. Hanafi for her support and guidance. I would also like to thank examiner and everyone evaluating my project paper.

Lastly, I would like to thank my parents Mr. Abd Hamid Bin Musa and Mrs. Nor Rashidah bt Shamsuddin for supporting my project paper. Without them, I would not have been able to complete this project paper.

Thanks to all of you.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

1.0 Introduction

This chapter will provide an explained of the research background, problem statement, scope, objectives and the significance of the study. The research questions and the organization of the dissertation are also presented, while in operational definitions used in this study are described.

1.1 Background of Studies

Unpredicted performance in a dynamic economic and monetary stability nowadays has become one of the reasons that the number of bankruptcies continued to rise significantly from time to time. This scenario became worst when financial experts failed to make accurate judgment as Hopwood (1994) states, "Even auditors, who have good knowledge of firms' situations, often fail to make an accurate judgment on firms' going-concern conditions". Bankruptcy prediction is a key importance for companies to ensure that the cost of bankruptcy is zero and the interests of stakeholders are protected. According to Salehi and Abedini (2009), "One of the bankruptcy factors is lack of existing control by different claimants". In the corporate management, shareholders have the right to control the management of the company and use the various types of operation to avoid bankruptcy. They agreed that the factors that led to bankruptcy are the management of a company failed to maximize shareholders wealth and cause financial problems occur.

The purpose of this study is to model bankruptcy prediction of listed companies in Kuala Lumpur Composite Index (KLCI) under PN17 and GN3 companies. Bayesian Network (BN) would be used as a bankruptcy prediction model due to its ability to solve complex problems. Pearl (1988) stated "Criteria and structure of Bayesian network such as probabilistic graphical models".

Among the reasons why BN is widely used is the ability of this model to interpret the output easily and problems can be structured in the form of graphs. The authors states, "The basic concept in BN is conditional probability that measure how our beliefs in certain propositions are changed by the introduction of related knowledge" Naslmosavi, Chadegani and Mehri (2011). After all data and information needed have been collected, next is to develop bankruptcy prediction model for Malaysian's firm. At first, this study have to consider 20 predictor variables of 18 companies under PN17 and GN3 for 10 years from 2001 until 2010 including liquidity ratios, leverage ratios, profitability ratios and other factors like firm's size and auditor's opinion. Thereafter this study used correlation and heuristic method to choose variables. In view of the above, Naïve Bayes model would be developed and all performances from the model would be compared by using regression model.

Finally result shows that operating cash flow over total liability (X2), depreciation and amortization over total assets (X7), current ratio (X12), working capital divided by total assets (X1), total liabilities over total assets (X4), long term debts over total assets (X5) sales over total assets (X6), earnings before interest, tax, net income over sales (X9), current assets over total assets (X10), and natural log of total assets (X14) are factors of bankruptcy based on Naïve Bayes model. This study concluded that Naïve Bayesian network model is a simple and easy model to implement and at the most important thing is it performs well and proven by fact of 90% accuracy of prediction bankruptcy for Malaysian firms by using this model. After conclusion stated, few recommendations are suggested for the outcome.

1.2 Problem Statement

The main objective of this study is to measure the possibility to predict financial problems using Bayesian model on companies in Malaysia. In addition, it may be evidence that the company's financial statements contain information that can be applied because it is the main source for the study to be performed. Altman Z-score is one of the popular bankruptcy predictions model and often used by researchers around the world in an effort to predict bankruptcy. This model was published in 1968 by Edward I. Altman and according to Zmijewski (1983), “Some research shows that this model has certain limitation such as samples were small and not proportional to actual bankruptcy rates.

According to Scott (1981), another limitation of Altman Z-score is hypothetical bias in section technique of the variables. He would like to consider a multitude of variables because of lack of bankruptcy theory and then deduction of original set to the utmost accurate subset was proposed.

Several decades ago, many studies have been carried out by experts from the field of accounting and finance, to predict the company's financial problems by using information from the financial statements of the company. Beaver (1966) states, “Business ratio is the best sole predictor and this statement”. He has conducted study by using financial ratios as predictor variables. Finally he concluded that the best bankruptcy predictor is income to debt ratio. Altman (1968) decided to use Multiple Discriminant Analysis (MDA) in his effort to find a bankruptcy prediction model. Sample data form 33 random manufacturing firms were selected in between year 1945 and 1946.

In 1990’s era, the latest bankruptcy prediction model is Neural Networks technique and the first researchers who are applied are Odom and Sharda. Millennium year 2000, BN models was developed by authors in order to detect early warning of bank failure (Sarkar and Siram, 2001). They managed to conclude that Naïve BN model and a composite BN model has comparable performance to the well-known induced decision tree classification algorithm. After all these techniques being implemented, there are other bankruptcy prediction techniques introduced in next era such as Rough Ser Theory (McKee 1998), Discrete Hazard models (Shumway, 2001) and Genetic Programming (McKee and Lensberg, 2002).

1.3 Research Questions

This study examines three research questions using recent sample data:

a. Can Bayesian model be a tool to validate bankruptcy prediction?

b. Is Bayesian model a good method to select potential variables?

c. Can the Naïve Bayesian model used as bankruptcy prediction model for Malaysian firms?

1.4 Research Objectives

The main objectives of this project paper are:

a. To examine the validity of Bayesian Model as bankruptcy prediction tool.

b. To provide a good method to guide the selection of variables in naïve Bayes models.

c. To propose the naïve Bayesian model as a bankruptcy prediction model for the Malaysian firms.

1.5 Significance of the Study

The first significance of the study is to act as a benchmark for most investors to prevent them to invest in companies with financial problems. By using Naive Bayesian model, investors can identify early warning of financial problems faced by the company before making any investment decisions.

Another significance of the study is Bayesian network model can help Bursa Malaysia to produce new software on bankruptcy prediction to prevent bankruptcy. The purpose is to reduce a number of companies that listed under PN17 and GN3.

In addition, the Naive Bayesian can be used as an awareness tool for companies to prevent bankruptcy. Therefore companies can be prepared an action plan to deal with in future bankruptcy.

Finally, the significance of this study is to contribute something new in the modeling system to forecast an intelligent bankruptcy for the firm as a practical contribution to the industry and at the same time will bring the world's financial industry to a more forward.

1.6 Scope of Study

The scope of this study is focusing only in PN17 and GN3 companies which listed on Kuala Lumpur Composite Index to predict bankruptcy. The data are collected for the period of 10 years from 2001 until 2010. Based on Bursa Malaysia, a PN17 is a listed company that having financial distress and has failed to meet minimum capital or equity (not less than 25% of the paid up capital). A PN17 company must submit to Bursa Malaysia their plan on how to regularize or face possible delisting. Meanwhile a GN3 is a company designated as an Affected Listed Company because its poor or adverse fiscal condition and level of operations fall into the criteria described in Bursa Malaysia's Guidance Note 3.

1.7 Organization of the Dissertation

In Chapter 1, the introductory section starts with a research objective, together with background information of why and how the study was initiated will also ne stated. The body of the report, the literature survey, the theoretical framework and the hypotheses are furnished will be done in Chapter 2. Meanwhile in Chapter 3 this study will discuss about design details such as sampling and data collection method, as well as the nature and type of study, the time horizon, the field setting and the unit of analysis will be described. The details of the type of data analyses done to be test the hypotheses and the finding will be provided next in Chapter 4. The final part of the report or Chapter 5 will contain the conclusion and recommendation drawn from the finding.

1.8 Summary

Chapter 1 has provided an explained of the research work, problem statement, scope, objective and significant of studies. The research questions and organizations of dissertation are also presented while operational terms used in this study also have been explained. In order to develop bankruptcy prediction model, this study consider 20 predictor variables including liquidity ratios, leverage ratios, profitability ratios and other factors like firm's size and auditor's opinion and then this study use two methods for choosing variables. Therefore the aim of this study is to model Financial Distress prediction of listed companies in Kuala Lumpur Composite Index (KLCI) under PN17 and GN3 companies by using Bayesian Network.

CHAPTER TWO

LITERATURE REVIEW

2.0 Introduction

This chapter present a review of previous study related to this topic. It provides the background knowledge on the research work undertaken by previous researchers. In this chapter it also covers a review of analyses on the analysis of other bankruptcy prediction model that have been developed in old times.

2.1 Previous Studies

Bankruptcy prediction model has become one of the famous topics for researchers and they tried to investigate the factors since the past few decades have seen a dramatic rise. Highly leveraged and changes in governance could be a reason of bankruptcies. Halpern, Kieschnick and Rotenberg (2008) stated "The core issue is whether highly levered transactions avoid financial distress or bankruptcy because of the way their governance is changed or because of the way they are financed?" They concluded that it is the composition of the debt used to finance highly levered transactions rather than how its governance is restructured that matters most to the likelihood of it facing financial distress or bankruptcy. The evidence was supported by Chemmanur and Fulghieri (1994) who proposed banks appear to reduce 'inefficient' liquidation, which suggests that reputation effects are important to banks.

Tan and Dihardjo (2001) have extended on previous study written by Tan (1996, 1997) regarding comparison between artificial neural models with probit model. They reconstructed the data set that was failed before and they also did an enhancement on Type II Error by redefining a correct prediction tools. They found that Type II Error was reduced significantly with redefinition of correct prediction and they duly confirmed that the usefulness of the artificial neural models.

In order to have a deep understanding of bankruptcy prediction model, this study has decided to conduct comprehensive review regarding other techniques of bankruptcy prediction done by other researchers. Grice and Ingram (2001) examined 3 research questions and first is "Is Altman's original model as useful for predicting bankruptcy in recent periods as it was for the periods in which it was developed and tested by Altman?" Second question is "Is the model as useful for predicting bankruptcy of non-manufacturing firms as it is for predicting bankruptcy of manufacturing firms?" and last but not least, "Is the model as useful for predicting financial stress conditions other than bankruptcy as it is for predicting bankruptcy" . They found that Altman's accuracy was decline year after year. They concluded that Altman's model was not sensitive to financial distress.

Suarez (2004) predicted failure of construction companies by using neural network model. Sample data consist of 16 healthy and bankrupt companies and he found that 7 ratios are the most relevant indicators to predict bankruptcy. Sample data was entered to Altman's model and he concluded that neural network models have more accuracy compared than Altman's model. Ahn and Kim (2009) introduced a new hybrid case based on reasoning model that used genetic algorithms. They concluded that this new model give new opportunities to other researchers to use as a bankruptcy prediction's tool.

Etemadi, Rostam and Dehkordi (2009) proposed to investigate application of genetic programming for bankruptcy prediction model. Bankruptcy and non-bankruptcy Iranian listed firms applied as sample data in order to predict bankruptcy by using genetic programming then multiple discriminant analysis used as benchmarking tool. Results show that genetic programming has 94% and 90% accurate on training and holdout meanwhile multiple discriminant analysis has 77% and 73% accuracy. Finally they concluded that genetic programming model is more accurate compared than traditional multiple discriminant analysis. Again Chen and Du (2009) used neural network and data mining in their research in order to predict bankruptcy. Finally they found that artificial neural network approach obtains better prediction accuracy than the multiple discriminant analysis approach.

The prediction of bankruptcy has been extensively studied using various statistical techniques and one of them is Bayesian model. According to Daciana, Dominic and Crina (2010), using Bayesian and AgenaRisk tool, it is possible to show all of the implication and results of a complex Bayesian argument without requiring and understanding of the underlying theory of mathematics. They structured 4 types of nodes that consist of sample, probability, result and assumption nodes The initially model was built and all 4 sample nodes were parents for the result node. They introduced the 2

Mediate Nodes in order to reduce the number of parents' node and of the calculation time. This study was moldings the Anghel Prediction Model for bankruptcy risk using the Bayesian probability. Again, Sun and Shenoy (2006) stated, Naïve Bayes model have an average prediction accuracy of 81.12% for the bankruptcy sample and 81.85% for the non-bankruptcy sample. Besides that, nature of bankruptcy prediction is comparable to results reported by some other studies (e.g. Ohlson, 1980; Hopwood et al., 1994; McKee and Greenstein, 2000; McKee and Lensberg, 2002).

Many studies have been conducted by researchers to predict bankruptcy using financial ratio such as Beaver and Altman. Previous study has been used to support this study that using financial ratios as a sample data. Holmen (1988) compared performance of several classical bankruptcy prediction models for bankruptcies occurring between 1977 and 1984 and out of 84 bankrupt firms only 58 companies were selected as sample data. Data were collected from financial statements for each firm from various industries. The main purpose of study is to compare ability of univariate model proposed by Beaver (1966) and multivariate model proposed by Altman (1968). He used 3 samples to measures error rate for both model and he found that cash flow total debt ratio that used in inivariate model have fewer errors than multivariate model.

Zeytinoglu and Akarim (2013) calculated 20 financial ratios that included of liquidity, operation, debt management and profitability were used to predict the financial failure of firms and develop the most reliable model by analyzing these ratios statistically. They employed discriminant analysis by using financial ratios of 115 firms and Altman Z-score have been used to determine the financial distress of firms. They concluded that sales over fixed assets, sales over inventories and turnover of receivable is efficient to determine the financial failure of the firms and at the same time important for financial analysts, investors and other company officials.

Salehi and Abedeni (2009) proposed the ability of financial ratios to predict financial distress of the listed companies in Tehran Stock Exchange (TES). Data were divided into two groups. The first group contained 30 companies which did not have financial distress, and the second group, contained 30 companies which have had financial distress. The presented model was according to 5 ratios, namely; ratios indicate liquidity, profitability, managing of debt and managing of property. They found that results yielding from the current study indicate that the accounting data can be predicted the financial cases that have a high predicting power.

Hypothesis is a statement or claim made about the value of a population parameter. It is theory or idea used as the basis of a statistical test and guide investigation in the light of established facts. Sun and Shenoy (2006) on their study have been provided operational guidance for building Naive Bayes Bayesian network models for bankruptcy prediction. They suggested a heuristic method to guide the selection of bankruptcy predictors that based on the correlations and partial correlations among variables and finally develop a Naive Bayes model. According to Koller and Sahami (1996), the purpose of the propose method is to identify key predictors and eliminating redundant or irrelevant ones. They adapted the extended Pearson–Tukey (EP–T) method (Keefer and Bodily, 1983) and results show that the proposed heuristic method for variable selection is simple to implement and performs well. Therefore, it is possible to approve this Naïve Bayes model can provide a good method in selection of variables models.

Aghaie and Saeedi (2009) stated "Naïve bayes model is best performance of predicting bankruptcy among these 3 models". The results supported by fact that this model has an average prediction accuracy of 94% for the bankruptcy sample and 92% for the non-bankruptcy sample. They concluded that it is possible to predict financial distress using Bayesian models. They also examined several important methodological issues related to the use of Naïve Bayes Bayesian Network models to predict bankruptcy.

Study by Sarkar and Sriram (2001) has demonstrated how probabilistic models may be used to provide early warnings for bank failure. The automatic system examines the financial ratios as predictors of bank performance and assesses the posterior probability of bank financial health. Both models are able to make accurate predictions with the help of historical data to estimate the required probabilities. In particular, the more complex model is found to be very well calibrated in its probability estimates. Two different probability models were examined in this context, the Naïve Bayes classification model and a composite attributes model. Although out of two models, the Naïve Bayes model appeared sharper while the composite attributes model was better calibrated.

2.2 Summary

At hand, few journals that have been reviewed in term of authors, year published, and method that they are using also finally their findings. Finally this study hopes that the finding of previous studies could also applicable and assist to develop bankruptcy prediction model by using Naïve Bayes model.

CHAPTER 3

DATA AND METHODOLOGY

3.0 Introduction

This section present in details the data and method that are being adopted in this study. The method used in this study includes conditional correlation, conditional likelihood, a Naïve Bayes Bayesian Network model and regression.

3.1 Theoretical Framework

Theories are formulated to explain, predict, and understand phenomena and, in many cases, to challenge and extend existing knowledge, within the limits of the critical bounding assumptions. The theoretical framework is the structure that can hold or support a theory of a research study. The theoretical framework introduces and describes the theory which explains why the research problem under study exists.

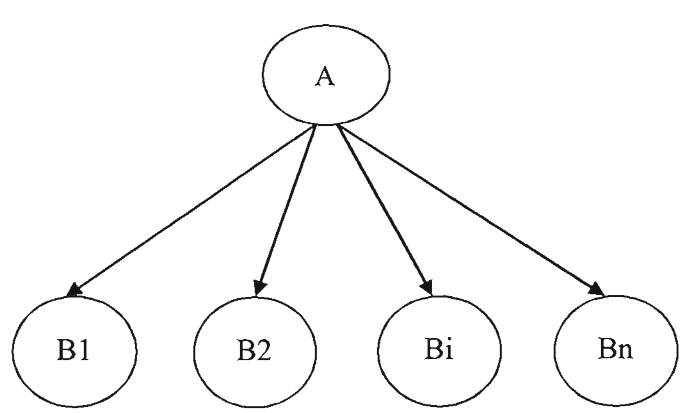

Figure 3.1 Naïve Bayes Model Nodes

The naïve Bayes model is named by Titterington et al (1981) because of its simplicity and Figure 3.1 presents a graphical representation of naïve Bayes model where A represents the bankruptcy variable or independence variable and B1, B2, Bi and Bn represent bankruptcy predictor variables or dependent variables (Sun and Shenoy, 2006).

3.2 Hypothesis Development

For the hypothesis part, where hypothesis is a statement or claim made about the value of a population parameter. It is theory or idea used as the basis of a statistical test. The two types of hypothesis are null hypothesis states that there is no different between the parameters and alternative hypothesis states that the opposite of null hypothesis that is; there exist a difference between the parameters.

According to Sun and Shenoy (2006) naïve Bayes model correctly predicts 81.57% of bankruptcies, and 81.78% of non-bankruptcies.

H1: It is possible to validate a bankruptcy prediction by using Bayesian Model.

According to Sun and Shenoy (2006), heuristic method is very easy to implement and proves to be effective by the empirical results. Besides that, only variables that have significant correlations with the variable of interest are selected.

H2: It is possible to approve this Naïve Bayes model can provide a good method in selection of variables models.

Aghaie and Saeedi (2009) stated “Naïve bayes model is the best performance of predicting bankruptcy among these three models”. The results show that it is possible to predict financial distress using Bayesian models.

H3: It is possible to predict financial distress of firm listed in Kuala Lumpur Composite Index using Bayesian networks (Aghaie and Saeedi, 2009)

3.3 Data Collection

The sources of data for this study were taken from Bursa Malaysia and the official website of the company. Data derived from the annual financial statements of companies that comprise the liquidity ratios, leverage ratios, profitability ratios and other factors such as the size of the firm and the auditor's opinion. Data consisted of 20 predictor variables of 18 companies under PN17 and GN3 for 10 years from 2001 to 2010. Data were restructured and streamlined with 20 variables that have been proposed in the Microsoft Office software. Out of the 20 predictor variables, 15 variables were identified and used in this study. The data consists of 18 companies sorted by year for every 15 predictor variables. The last step is to transfer all the data into SPSS software to conduct a study.

3.4 Operational Terms

The main terms used in this study are defined below:

i. Financial distress: Term in corporate finance used to indicate a condition when promises to creditors of a company are broken or honored with difficulty. If financial distress cannot be relieved, it can lead to bankruptcy (Wikipedia, the free encyclopedia).

ii. Bankruptcy: A legal proceeding involving a person or business that is unable to repay outstanding debts (Investopedia).

iii. Bursa Malaysia Securities Berhad Market: Malaysian Stock Exchange regulated and supervised by the Malaysian Securities Commission under the Malaysia Securities Act 1993.

iv. PN17 Company: Company that triggered any of the criteria pursuant to Practice Note 17 of the Main Market Listing Requirements of Malaysian Stock Exchange are said to be reprimanded under the PN17 list as financial distress companies.

v. GN3 Company: GN3 companies are a company designated as an Affected Listed Company because its poor or adverse financial condition and level of operations fall into the criteria described in Bursa Malaysia's Guidance Note 3.

vi. Bayesian networks: Probabilistic graphical models that represent a set of random variables for a given problem, and the probabilistic relationships between them.

vii. Altman Z - score: Is a linear combination of five common business ratios, the coefficients were estimated by identifying a set of firms which had declared bankruptcy and then collecting a sample of firms which had survived, with matching by industry (Wikipedia, the free encyclopedia).

3.5 Methodology

a. In order to develop a bankruptcy prediction model, this study have considered 20 predictor variables including liquidity ratios, leverage ratios, profitability ratios and other factors like firm's size and auditor's opinion (Sun and Shenoy, 2006).

Table 3.1Definitions of potential predictor variables

(Sun and Shenoy, 2006)

Variables	Definition
X1	Natural log of (total Assets/GNP Index)
X2	(Current Assets-Current Liabilities)/Total Assets
X3	Current Assets/Current Liabilities
X4	Operating Cash Flow/Total Liabilities
X5	Current Assets/Total Assets
X6	Cash/Total Assets
X7	Total Liabilities/Total Assets
X8	Long Term Debt/Total Assets
X9	Sales/Total Assets
X10	Current Assets/Sales
X11	Earnings before Interest and Taxes/Total Assets
X12	Net Income/Total Assets
X13	One if net income was negative for the last two years, else zero
X14	Retained Earnings/Total Assets

X15	(Net income in year t – net income in year t-1)/(Absolute net income in year t + absolute net income in year t-1)
X16	Natural log of total assets
X17	Zero if auditors' opinion is unqualified otherwise one
X18	Net income/Sales
X19	Retained Earnings/Total owner's equity
X20	Quick assets/total assets

From 20 predictor variable, only 15 are identified as potential bankruptcy predictors. The predictor variables not available for this study are working capital over total assets, one if net income was negative for the last two years, net income in year t minus net income in year t-1 divide to absolute net income in year t plus absolute net income in year t-1, zero if auditors' opinion is unqualified otherwise one and retained earnings over total owner's equity.

b. There are two methods used to choose the variables. The first is using correlations and partial correlations to measures correlation among variables. Secondly this study used heuristic methods to measures first and second order variables that are connected with bankruptcy.

i. Correlations and partial correlations.

Based on the correlations and partial correlations among variables, the method aims at eliminating redundant and less relevant variables (Sun and Shenoy, 2006).

ii. Heuristic methods.

Koller and Sahami (1996) stated, "One purpose of this paper is to provide a heuristic method to guide the selection of variables in Naıve Bayes models". Their goal is to eliminate variables that provide little or no additional information beyond that subsumed by the remaining variables. To achieve the goal, the proposed heuristic relies on correlations and partial correlations among variables and based on the assumption that the dependence between every pair of variables was measured by the correlation coefficient.

c. Once variables have been selected through proposed heuristic method, this study has started to develop the naïve Bayes model consists of first and second order variables

i. A Naïve Bayes BN Model

In a naïve Bayes model as shows in Figure 3.1, the node of interest has to be the root node, which means, it has no parent nodes. In a bankruptcy prediction situation, A represents the bankruptcy variable. B1, B2, Bi and Bn represent n bankruptcy predictor variables. The naïve Bayes model assumes the following conditional independence:

$Bi \perp \{B1, B2 \ldots Bi - 1, Bi{+}1 \ldots Bn\} \mid A$, for $i = 1,2,\ldots,n$.

The above assumption says that predictors, B1, B2...Bn are conditionally mutually independent given the state of bankruptcy (Sun and Shenoy, 2006).

ii. Regression model.

In this section, this study has compared the performance of the Naive Bayes model in with that of logistic regression, a widely used bankruptcy prediction tool. The goal of logistic regression is to find the best fitting and most parsimonious, yet biologically reasonable model to describe the relationship between outcomes variable (dependent) and set of independent variables. Logistic regression allows one to predict a discrete outcome, such as group membership, from a set of variables that may be continuous, discrete, dichotomous, or a mix of any of these (Wiley, Sons, Menard and Scott, 1995).

3.6 Summary

At the end of this chapter, this study has provided an explained of the research framework besides dependent and independent variables were determined. This paper considered 20 predictor variables of 18 companies under PN17 and GN3 for the period of 10 years from 2001 until 2010 as sample data by using proposed methodologies and finally two hypotheses were developed. The findings of this study will be discussed in the next chapter.

CHAPTER FOUR

RESULTS

4.0 Introduction

This section reports the results of tests used to confirm the Naive Bayes model as a predictive tool for bankruptcy. The decision is divided into four main categories. The first is using correlation and partial correlation to calculate the correlation between variables. Hence this study using heuristics method to select first and second order variables that related to bankruptcy. Therefore, this study develops a Naive Bayes model after heuristics method completed. Finally, Naive Bayes model performance was measured by using logistic regression.

4.1 Correlations and Partial Correlations

First, this study obtains the correlations among all variables, including 15 potential predictors and the variable of interest, firms' bankruptcy status (Altman Z-score). Variables that have significant correlations (Pearson correlation coefficient ≤ 0.05) are assumed to be dependent and therefore related. This study uses the cutoff of 0.05 to help identify a small subset of extremely important predictors while excluding the unimportant ones.

Three predictors, operating cash flow over total liabilities, earnings before interest, depreciation and amortization over total assets and current ratio are connected with Altman Z-score and assumed as first order variables. This study used Altman Z-score to measure either the companies consider as bankrupt or non-bankrupt based on Altman Z-score benchmark which are Z $\geq$2.99 consider as safe zone, 1.81 $\leq$Z $\leq$2.99 as grey zone and Z $\leq$1.81 consider as distress (Wikipedia).

Among the first order variables, only two pairs of variables have dependency (correlations P $\leq$0.05) as shown in Panel A of Appendix 4.1. To avoid double counting information, this study analyze whether one variable is dependent with Altman Z-score given the other variable in the pair by examining the partial correlations between that variable and Altman Z-score, while controlling the other variable in the pair.

Next this study discusses how to identify second order variables that have significant correlations with first order variables. To select second order variables, this study used same method (partial correlation) to select first order variables. The major difference is that now this study considers each first order variable instead of Altman Z-score as a root variable. To select second order for operating cash flow over total liabilities, this study identifies those non-first order variables that are connected to say variable in Figure 4.3. Variables that have significant correlations with operating cash flow over total liabilities are working capital over total assets, total liabilities over total assets, long term debt over total assets, sales over total assets, net income over sales, current assets over total assets and natural log of total assets. Pairs are assuming as significant if the P value less than 0.05 and has bidirectional significant relationship by using same variable but different controlling variable. For that reason, this study used partial correlation between one variable with operating cash flow over total liabilities after controlling for the other variable in the pair. Based on Panel B.1 of Appendix 4.1, only 18 pairs have bidirectional significant relationships with operating cash flow over total. The remaining pairs in Panel B.1 of Appendix 4.1 considered as not significant because of P value more than 0.05 and they do not have two ways significant relationship.

Same goes to Panel B.2 of Appendix 4.1; this study identifies those non-first order variables that are connected with earnings before interest, depreciation and amortization over total assets in Figure 4.3. Variables that have bidirectional significant correlations are working capital over total assets, total liabilities over total assets, long term debt over total assets, sales over total assets, net income over sales, current assets over total assets and natural log of total assets. The results showed that out of the 66 pairs tested, only 40 pairs have a significant relationship and 26 pairs were vice versa.

This study identifies those non-first order variables that are connected to current ratio in Figure 4.3. Variables that have significant correlations with current ratio are cash over total assets, long term debt over total assets, sales over total assets, current assets over total assets and market value over total liabilities. Only 10 pairs were tested to analyze their significant relationship but none of the pairs have significant relationship as shows in Panel B.3 of Appendix 4.1.

4.2 Heuristic Method

In order to describe how the proposed heuristic works, initially this study has obtained the correlations among all variables, including 15 potential predictors and the variable of interest, Altman Z-score. To avoid double counting information, this study analyze whether one variable is dependent with Altman Z-score given the other variable in the pair by examining the partial correlations between that variable and Altman Z-score, while controlling the other variable in the pair.

One purpose of this paper is to provide a heuristic method to guide the selection of variables in Naïve Bayes models. Koller and Sahami (1996) stated, "The goal of heuristic model is to eliminate variables that provide little or no additional information beyond that subsumed by the remaining variables". To achieve the goal, the proposed heuristic method relies on correlations and partial correlations among variables.

Heuristic method is based on the assumption that the dependency between every pair of variables is linear and measured by the correlation coefficient. Result shows that only 3 out of 15 variables are connected with Altman Z-score which are operating cash flow over total liabilities, earnings before interest, tax and depreciation amortization over total assets and current ratio because of p-value is less than 0.05 as shows in Figure 4.2.

Meanwhile, there are 7 variables selected as second order variables including working capital over total assets, total liabilities over total assets, long term debt over total assets, sales over total assets, net income over sales, current assets over total assets and natural log of total assets. As shown in Figure 4.3, natural log of total assets is not connected with any variables in second order variables because the p-value is greater than 0.05 and prove they do not has bidirectional significant relationship even though natural log of total assets is connected with Altman Z-score.

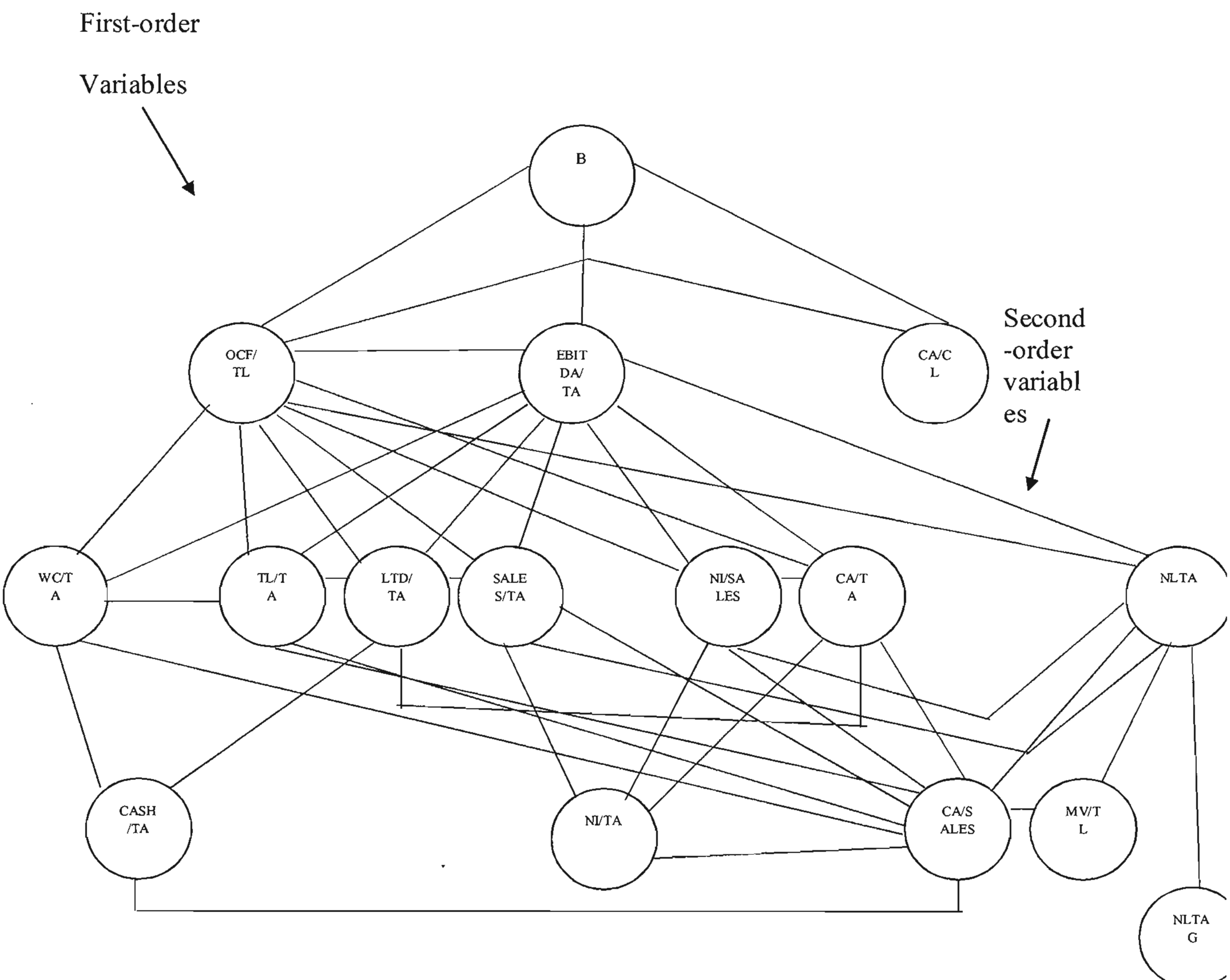

Figure 4.1: Correlation among the variables

Figure 4.1 shows that correlation between every pair of variables is linear and measured by the correlation coefficient by using heuristic method. Heuristic method guides us to select predictor variables from a pool of probable variables. Under this method, only variables that have significant correlations with the variable of interest, the status of bankruptcy are selected.

Definitions:

B	= Altman Z-score
WC/TA (X1)	= (Current Assets-Current Liabilities)/Total Assets.
OCF/TL (X2)	= Operating Cash Flow/Total Liabilities.
Cash/TA (X3)	= Cash / Total Assets.
TL/TA (X4)	= Total Liabilities / Total Assets.
LTD/TA (X5)	= Long Term Debt / Total Assets.
Sales/TA (X6)	= Sales / Total Assets.
EBITDA/ TA (X7)	= Earnings before Interest, Depreciation and Amortization / Total Assets.
NI/TA (X8)	= Net Income / Total Assets.
NI/Sales (X9)	= Net Income / Sales.
CA/TA (X10)	= Current Assets / Total Assets.
CA/Sales (X11)	= Current Assets / Sales.
CA/CL (X12)	= Current Assets / Current Liabilities.
MV/TL (X13)	= Market Value / Total Liabilities.
NLTA (X14)	= Natural Log of Total Assets.
NLTAG (X15)	= Natural Log of (Total Assets / GNP Index)

4.3 Naïve Bayes Model

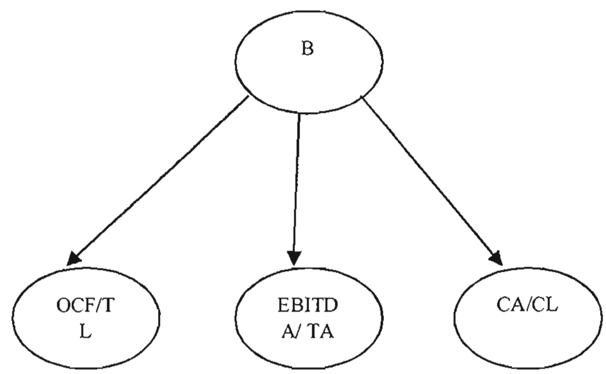

Figure 4.2 The Structure of the Naïve Bayes Model

Using Bivariate correlation, only 3 out of 15 variables are connected with Altman Z-score which is operating cash flow over total liabilities, earnings before interest, tax and depreciation amortization over total assets and current ratio as shown in Figure 4.2. According to Sun and Shenoy (2006) those variables are directly having significant correlation (P ≤ 0.05) with bankruptcy status will be assumed as first order variables.

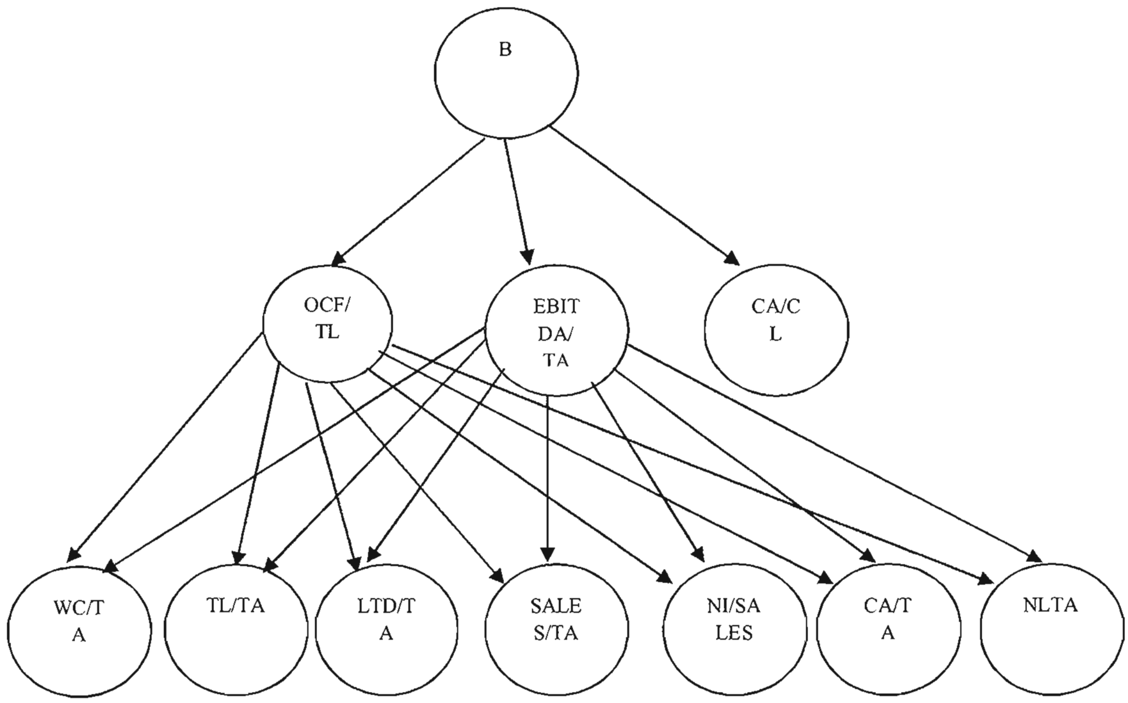

Figure 4.3 Naïve Bayes Model

Figure 4.3 shows first and second order variables for naïve Bayes model that have been developed. The first order variables for Naïve Bayes model including operating cash flow over total liabilities, earnings before interest, tax and depreciation amortization over total assets and current ratio. Meanwhile, there 7 variables were selected as second order variables which is working capital over total assets, total liabilities over total assets, long term debt over total assets, sales over total assets, net income over sales, current assets over total assets and natural log of total assets. As shown in Figure 4.3, current ratio is not connected with any variables in second order variables because the p-value is greater than 0.05 and prove they do not have bidirectional significant relationship even though it is connected with Altman Z-score.

According to Sun and Shenoy (2006), those variables are directly having significant correlation ($P \leq 0.05$) with bankruptcy status will be first order variable and the remaining non-first order variable that have positively significant relationship with first order variable are considered as second order variable. Only variable that has significant bidirectional relationship will be stay in second order variable. Finally this study can formulate the naïve Bayes model as follow:

Naïve Bayes model = X2 + X7 + X12 + X1 + X4 + X5 + X6 + X9 + X10 + X14

Whereas:

X2 = Operating Cash Flow/Total Liabilities

X7 = EBITDA/Total Assets

X12 = Current ratio

X1 = (Current Assets-Current Liabilities)/Total Assets

X4 = Total Liabilities/Total Assets

X5 = Long Term Debt/Total Assets

X6 = Sales/Total Assets

X9 = Net income/Sales

X10 = Current Assets/Total Assets

X14 = Natural log of total assets

4.4 Logistic Regression

Logistic regression is part of a category of statistical models called generalized linear models. Generally, the dependent or response variable is dichotomous, such as presence or absence and success or failure (A. Agresti, 1996). The goal of logistic regression is to find the best fitting and truly parsimonious, yet biologically reasonable model to describe the relationship between outcomes variable (dependent) and set of independent variables. These independent variables are often called covariates. What distinguish a logistic regression model from the linear regression model is that the outcome variable in logistic regression is binary or dichotomous (Hosmer D.W). Therefore this study explained whether the logistic regression result is significant or not with our null hypothesis which is it is not possible to predict financial distress of firm listed in Bursa Malaysia using Bayesian networks. This study would like to reject the null hypothesis base on p-value ($p \leq 0.05$).

Table 4.1 Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	88.095[a]	.206	.402

a. Estimation terminated at iteration number 12 because parameter estimates changed by less than .001.

Table 4.1 above shows the model summary that includes three type of parameter which are 2 Log likelihood, Cox & Snell R Square and Nagelkerke R Square. 2 Log likelihood is the -2 log likelihood for the final model. By itself, this number is not very informative. However, it can be used to compare nested (reduced) models. Meanwhile, Cox & Snell R Square and Nagelkerke R Square are pseudo R-squares. Logistic regression does not have an equivalent to the R-squared that is found in OLS regression. However, many people have tried to come up with one. There are a wide variety of pseudo-R-square statistics (these are only two of them). Since this statistic does not mean what R-squared means in OLS regression (the proportion of variance explained by the predictors), this study suggest interpreting this statistics with great caution.

Table 4.2 Hosmer and Lemeshow Test

Step	Chi-square	Df	Sig.
1	4.921	8	.766

Table 4.3 Hosmer and Lemeshow Test's Contingency

		B or NB = NB		B or NB = B		Total
		Observed	Expected	Observed	Expected	
Step 1	1	12	10.432	6	7.568	18
	2	2	3.364	16	14.636	18
	3	2	2.458	16	15.542	18
	4	1	1.455	17	16.545	18
	5	2	1.163	16	16.837	18
	6	2	.880	16	17.120	18
	7	0	.603	18	17.397	18
	8	0	.471	18	17.529	18
	9	0	.174	18	17.826	18
	10	0	.000	18	18.000	18

The Hosmer–Lemeshow test is a statistical testing for goodness of fit for logistic regression models. It is used frequently in risk prediction model. The test assesses whether or not the observed event rates match the expected event rates in subgroups of the model population. The Hosmer–Lemeshow test specifically identifies subgroups as the deciles of fitted risk values. Models for which expected and observed event rates in subgroups are similar are called well calibrated. The test is similar to a χ^2 goodness of fit test and has the advantage of partitioning the observations into groups of approximately equal size, and therefore near are less likely to be groups with very low observed and expected frequencies (Wikipedia). The observations are grouped into deciles based on the predicted probabilities. The statistical testing were calculated as per above using the observed and expected counts for both the bankrupt and non-bankrupt, and has an approximate χ^2 distribution with 8 degrees of freedom as shown in Table 4.2. Calibration results for the model from the example data are shown in Table 4.3. The Hosmer–Lemeshow test (P = 0.766) indicates that the numbers of bankruptcy are not significantly different from those predicted by the model.

Table 4.4 Model Classification

Observed			Predicted		
			B or NB		Percentage Correct
			NB	B	
Step 1	B or NB	NB	6	15	28.6
		B	3	156	98.1
	Overall Percentage				90.0

a. The cut value is .500

Table 4.4 above shows the estimation for the full model with the independents as well as the constant. The observed column indicates the number of B's and NB's that are observed in the dependent variable, the predicted column shows the predicted values of the dependent variable based on the full logistic regression model. This table shows how many cases are correctly predicted (6 cases are observed to be NB and are correctly predicted to be NB, 156 cases are observed to be B and are correctly predicted to be B), and how many cases are not correctly predicted (15 cases are observed to be NB but are predicted to be B, 3 cases are observed to be B but are predicted to be NB). The overall percentage column gives the overall percentage of cases that are correctly predicted by the model (in this case, the full model that this study specified). The percentage has increased from 88.3 for the null model to 90 for the full model.

4.5 Summary

After study has been done, 10 predictor variables were identified as the factors of bankruptcy based on naïve Bayes model. The variables are operating cash flow over total liability, depreciation and amortization over total assets, current ratio, working capital divided by total assets, total liabilities over total assets, long term debts over total assets sales over total assets, earnings before interest, tax, depreciation and amortization over total assets, net income over sales, current assets over total assets, and natural log of total assets. Key point is the results shows that naïve Bayes models can be used in predicting bankruptcy in Malaysia. It can be concluded that the proposed Naïve Bayesian network model is an easy to implement and supported by 90% accuracy of bankruptcy prediction models.

CHAPTER 5

CONCLUSION AND RECOMMENDATION

5.0 Introduction

After analyzing the results from the Multiple Regression Analysis, conclusion and recommendations have been made to show the relationship between capital structure and corporate performance. From the conclusion, few recommendations are suggested.

5.1 Finding for Demography and Hypotheses

Every single company in the world has possibility to face bankruptcy problem whether they want or not. Therefore a lot of bankruptcy prediction models have been introduced such as logit and probit models (Ohlson, 1980; Zmijewski, 1984), neural network models (Tam and Kiang, 1992), Rough Ser Theory (McKee, 1998), Discrete Hazard models (Shumway, 2001) and Genetic Programming (McKee and Lensberg, 2002).

First, this study provides a heuristic method that guides the selection of predictor variables from a pool of probable variables. Under this method, only variables that have significant correlations with the variable of interest, the status of bankruptcy are selected. Next three variables are selected as first order variables and seven variables are selected as second order variables. Next this study test the validity of naïve Bayes models in Figure 4.1 by using logistic regression that widely used as bankruptcy prediction tool. The overall percentages of cases that are correctly predicted by the model were increased from 88.3 for the null model to 90 for the full model. The ability of naïve Bayes models in predicts bankruptcy also reported by some other reports such as Sarkar and Sriram, 2001 and McKee and Lensberg, 2002.

This study identified 10 predictor variables to predict bankruptcy and to begin with operating cash flow over total liability. If the total liabilities of the company increased, the company's operating income will decrease and will result in the company being unable to expand their businesses because cash was used to finance liabilities of the company. Next factor is earning before interest, tax, depreciation and amortization over total assets. It shows how much of gross income is generated by total assets.

Coming factor of bankruptcy is current ratio. Current ratio measure the firm ability to meet its day to day operating expenses and satisfy its short term obligation. If a company has lower rate of current ratio, it shows that the company is in a poor condition and will contribute to financial distress. Another factor is working capital over by total assets. The working capital to total assets ratio quantifies a company's ability to cover its short term financial obligations (total current liabilities) by comparing its total current assets to its total assets. This ratio can provide some insight as to the liquidity of the company, since this ratio can uncover the percentage of remaining liquid assets (with total current liabilities subtracted out) compared to the company's total assets.

Total liabilities over total assets also can be factor of bankruptcy problem. This ratio indicates proportion of asset is finance by debt. If companies always increase their assets by using debt, they will increase their obligation to make periodically repayment of debt included interest rate expenses. Meanwhile, long term debts over total assets also contribute to bankruptcy. Ratio shows how much debt of company is using to support it total assets and the higher ratio indicates the firm will face more risk. Furthermore, sales over total assets also became indicator of firm's effectiveness in uses it assets to generate sales. If the ratio is lower, its shows that the company is inefficient or not well enough in managing their assets.

Following factor is net income over sales and measures the amount of worth earned from sales. If amount of company's sales drop, company's earnings also drop and will produce losses. Thereafter factor is current assets over total assets. This factor evaluates the liquidation of company from its total assets. If the ratio lowers, its shows that company will fail to face any emergency cases. The last variable considered as bankruptcy factor is natural log of total assets. Total assets are listed on a company's balance sheet based on their level of liquidity, which is based on the speed in which they can be exchanged for cash. The liquidity of assets can be found toward the top of a financial statement. Therefore it can used when unexpected event happen in company such as bankruptcy.

More notable, the above reported results show that Naïve Bayes models can be considered one of the tools in predicting bankruptcy in Malaysia. In completing this research, certain limitations encountered. The primary problem is choosing predictor variables. Out of 20 predictor variables suggested by Sun and Shenoy (2006), only 15 variables are provided by Bursa Malaysia. Following problems is a number of companies selected as sample data. In attendance, only 18 companies under the PN17 and GN3 have data for the period of ten years. This study can be concluded that the proposed heuristic method for variables selection is simple to implement and performs well according to the overall percent of cases that are correctly predicted by the model were increased from 88.3 for the null model to 90 for the full model.

5.2 Theoretical and Practical Contributions of the Study

One contribution of this study is the availability of naïve Bayesian network to predict bankruptcy prediction for Malaysia firms. This study run a logistic regression model and results show that it is possible to predict financial distress using Bayesian models. Also, because this prediction is based on the information that existed in financial statements of companies, it can be evidence that the financial statements of companies have information content. This study concluded that the model is a simple and easy to implement and at the most important thing is it performs well and proven by fact of 90% accuracy of prediction bankruptcy for Malaysian firms by using this model.

Another contribution is the heuristic methods as guidance to select bankruptcy potential variables. This study examines several important methodological issues related to the use of naïve Bayes Bayesian Network models to predict bankruptcy. First, this study provides two different methods that guide the selection of predictor variables from a pool of potential variables. Under the first method, only variables that have significant correlations with the variable of interest, the Altman Z-score, are selected. The purpose of this method is to eliminate variables that provide little or no additional information using correlations and partial correlations among the 20 predictor variables. This study found that heuristic method was performed well based on logistic regression which is used to validation analysis.

5.3 Recommendation and Future Research

The investors should be aware on the trend of economic cycles and company condition before they invest their money in certain company. Therefore Naïve Bayes models can be used as evaluation tool before investors make investment decision in certain company and to avoid any in future losses.

This study also recommends Bursa Malaysia to use Naïve Bayes model to evaluate the listed companies. This is crucial to ensure only qualified companies can be listed in Bursa Malaysia index and help investors to make a right choice in investing decision. Naïve Bayes model not only useful to investors and Bursa Malaysia but also financial institutions and banks. Banks can use Naïve Bayes model for loan making decisions. The purpose is to avoid any default repayment of loan disbursement.

Further this study would like to recommend for future research, instead using Naïve Bayes models, researcher can use other prediction bankruptcy tools such as neural network, genetic algorithm, computerized programming and logit or probit regression. Lastly this study recommend testing bankruptcy by using other model such as Fulmer approach to improvise other bankruptcy tools.

5.4 Conclusion

This chapter provides operational guidance for validating naïve Bayes model for bankruptcy prediction. Firstly, this study suggests heuristic methods that guide the selection of bankruptcy potential variables. Eliminate variables that provide little or no additional information was identified by using correlations and partial correlations among the variables. Thus a Naïve Bayes model is developed by using the proposed heuristic method and is found to perform well based on logistic regression which is used to validate the analysis. Results show that the model's performance is best when the method of *enter* is used in logistic regression which is the result of percentage correct is 90%. Finally, the results of this study could also be applicable to businesses and investor's decision making other than validating bankruptcy prediction.

REFERENCES

Ahn, H & Kim, K.J. (2009). Bankruptcy Prediction Modeling with Hybrid Case-Based Reasoning and Genetic Algorithms Approach. *Applied Soft Computing 9*, Pp 599- 607.

Altman, Edward I. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *The Journal of Finance*, Pp. 589-609.

Beaver, William, H., (1966). Financial Ratios as Predictors of Failure. *Journal of Accounting Research, Supplement, Empirical Research in Accounting: Selected Studies*, pp. 71-111.

Chen, W.S & Du, Y.K. (2009). Using Neural Networks and Data Mining Techniques for the Financial Distress Prediction Model.*Expert Systems with Applications 36,* Pp 4075-4086.

Drugan, M.M. and Wiering, M.A. (2010). Feature Selection for Bayesian Network Classifiers Using the MDL-FS Score. *International Journal of Approximate Reasoning*, Vol. 51, Pp 695–717.

Etemadi, H, Rostamy, A.A. &Dehkordi, H.F. (2009). A Genetic Programming Model for Bankruptcy Prediction: Empirical Evidence from Iran. *Expert Systems Applications 36*, Pp 3199-3207.

Estevam R. Hruschka, E.R. and Ebecken, N.F.F. (2007). Towards Efficient Variables Ordering for Bayesian Networks Classifier.*Data & Knowledge Engineering*, Vol. 63, Pp 258–269.

Grice, J.S & Ingram, R.W. (2001). Test of the Generalizability of Altman's Bankruptcy Prediction Model. *Journal of Business Research*, Pp 53-77.

Halperna, P, Kieschnick, R &Rotenberga, W. (2009). Determinants of Financial Distress and Bankruptcy in Highly Levered Transaction.*The Quarterly Review of Economics and Finance 49*, Pp 772-783.

Holmen, J. S. (1988). Using financial ratios to predict bankruptcy: An evaluation classic models using recent evidence. *Akron Business and Economic Revie,* Pp 52 – 62.

Hopwood, W.S., Mckeown, J.C., Mutchler, J.P. (1989). A test of the incremental explanatory power of opinions qualified for consistency and uncertainty. *The Accounting Review 64,* Pp 28 - 48.

Jay, S.H. (1988). Using Financial Ratios to Predict Bankruptcy: An Evaluation. Akron Business and Economic Review, Pp 52.

Naslmosavi, S, Chadegani, A.A., and Mehri, M. (2011). Comparing the Ability of Bayesian Networks and Adaboost for Predicting Financial Distress of Firms Listed on Tehran Stock Exchange (TSE). *Australian Journal of Basic and Applied Sciences,* Pp 629 – 634.

Pearl, J. (1988). Probabilistic Reasoning in Intelligent Systems. Network of Plausible Inference.

Scott, J. (1981). The Probability of Bankruptcy: A comparison of Empirical Prediction and Theoretical Models. *Journal of Banking and Finance, Vol 5*, Pp 317 – 344.

Salehi, M &Abedini, B. (2009). Financial Distress Prediction in Emerging Market:Empirical Evidences from Iran. *Journal of Contemporary Research in Business,* Vol. 1(1), Pp 6-26.

Sun, L &Shenoy, P. (2007).Using Bayesian Networks for Bankruptcy Prediction.*European Journal of operational Research,* Vol. 180, Issue 2, Pp 738-753.

Sarkar, S &Sriram, R. (2001).Bayesian Models for Early Warning of Bank F Failures. M*anagement Science*, Pp 1457-1475.

Thomas E. McKee, T.E. and Lensberg, T. (2001). Genetic programming and Rough sets: A hybrid approach to bankruptcy classification. *European Journal of Operational Research*, Pp 436–451.

Tan, C.N.T & Dihardjo, H. (2001).A Study on Using Artificial Neural Networks to Develop an Early Warning Predictor for Credit Union Financial Distress with Comparison to the Probit Model.*Managerial Finance*, Pp 56-61.

Zhang, G, Hu, M.Y, Patuwo, B.E &Indro, D.C. (1999). Artificial Neural Networks in Bankruptcy Prediction: General Framework and Cross-Validation Analysis. *European Journal of Operational Research 116*, Pp16-32.

Zeytinoglu, E and Akarim, Y.D. (2013).Financial Failure Prediction Using Financial Ratio: AnEmpirical Application on IstanbulStockExchange. *Journal of Applied Finance & Banking*, Pp107 – 116.

http://www.bnm.gov.my/

http://en.wikipedia.org/wiki/Bayesian_network

http://www.google.com.my/

http://www.bursamalaysia.com/website/bm/

http://www.investopedia.com/