ABSTRACT

Distress Investing: Corporate Bankruptcy Prediction

Erica Xuewei Jiang

Director: John D. Martin, PhD

Financial distress is a condition where a company has difficulty paying off its financial obligations to its creditors. Failing to relieve financial distress can lead to bankruptcy, which is very costly to the firm's creditors and investors, as well as its employees and citizens of the community where it is located. As a result, corporate bankruptcy predication is of great interest to both academics and practitioners. While a number of bankruptcy prediction models have been published, Altman's Z-score model is still commonly used by practitioners as a standard to evaluate a firm's financial health after almost fifty years since it was published in 1968. In this thesis, I explore the use of artificial neural networks to improve upon existing predictive techniques. My model shows strong discriminating power in a Receiver Operating Characteristics analysis, and outperforms the hazard model and the distance-to-default model in detecting bankruptcies.

APPROVED BY DIRECTOR OF HONORS THESIS:

____

Dr. John D. Martin, Department of Finance, Insurance & Real Estate

APPROVED BY THE HONORS PROGRAM

____

Dr. Andrew Wisely, Director

DATE: ___________________________

DISTRESS INVESTING: CORPORATE BANKRUPTCY PREDICTION

A Thesis Submitted to the Faculty of

Baylor University

In Partial Fulfillment of the Requirements for the

Honors Program

By

Erica Xuewei Jiang

Waco, Texas

May 2015

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor, Dr. John Martin, for the continuous support of my independent reading and research, for his patience, motivation, enthusiasm, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. Without his guidance and persistent help this thesis would not have been possible.

I would like to thank my committee members, Dr. James West and Dr. Greg Hamerly, for their encouragement, insightful comments, and hard questions.

My sincere thanks also go to Dr. Jeanne Hill, Dr. Van Pham, Dr. Scott Cunningham, and Dr. David VanHoose for their valuable inputs to my research.

Last but not least, my thanks and appreciations go to my friends and people who have willingly helped me out with their abilities.

CHAPTER ONE

Introduction

Financial distress is a condition where a company has difficulty paying off its financial obligations to its creditors. Failing to relieve financial distress can lead to bankruptcy, which is very costly to the financially distressed company. Knowing the probability of bankruptcy helps companies evaluate their distress risk and allows companies to prevent bankruptcy from happening. Other parties like banks, hedge funds and creditors are also concerned about the distress risk of a company. Furthermore, the stocks of financially distressed companies tend to move together, and thereby their risk cannot be diversified away. The premium for bearing such risk cannot be captured by the Capital Asset Pricing Model (CAPM) in many circumstances (Campbell et al. 2008). In this case creating a measure of distress risk is necessary to explain patterns that are anomalies in the CAPM. As a result, corporate bankruptcy predication is of great interest to both academics and practitioners.

A vast body of literature is devoted to assessing the risk that a firm will go bankrupt. Bankruptcy prediction models began to develop with Beaver's (1966) univariate study. Two years later, Altman published the first multivariate study. He combined several measures into a predictive model using multiple discriminant analysis and developed the well-known Z-score model, becoming one of the most important milestones in the bankruptcy prediction history. Bankruptcy prediction study turned to focusing on multivariate models and has continued to evolve since then.

In the later thirty years, the number and complexity of bankruptcy prediction models increased dramatically. Researchers tried to improve the accuracy of the prediction over a long time span. The logistic model was a major multivariate model being used during that time period. But the logistic model is a static model: it ignores the time variation of firms. To improve upon static models, many studies including Queen and Roll (1987), Theodossiou (1993), Denis et al. (1997) and Pagano et al (1998), began to use dynamic forecasting models. In 2001, Shumway proposed a simple hazard model: multi-period logit model. Unlike a logistic model, this model is dynamic and can explicitly account for time. Shumway's publication was another milestone in the bankruptcy prediction history, marking the beginning of a new era in this area.

It has been over ten years since Shumway's milestone publication. During this period of time, the cumulative number of corporate defaults has grown dramatically from 1990s. While the situation has been mitigated these years, the loss associated with corporate defaults is still not trivial. In Moody's 27th default study, 66 Moody's-rated corporate issuers defaulted in 2013, up from 63 in 2012, and half of the defaulted issuers were from North America. Defaulted debt in 2013 consisted of $36.7 billion of bonds and $17.8 billion of loans, increasing from $31.7 billion of bonds and $17.6 billion of loans that defaulted in 2012. Moreover, although a number of models have been published, all of which are claimed to be more or less improved upon past models, Altman's Z-score is still commonly used by practitioners as a standard to evaluate a firm's financial healthiness after almost fifty years since it was published. The question turns out to be: can the performance of bankruptcy prediction models be improved anymore? If so, can this new model be easily adopted by the real world? If it exists, such a model may become a new milestone in the area of bankruptcy prediction.

An exploration of artificial neural networks (ANNs) may answer these questions. Neural networks have been commonly used in various areas as computation technology develops. Some researchers, such as Zhang et al (1997), O'Leary (1998), Virag (2005) and Santos et al (2006), have applied it to bankruptcy prediction. By well programing, the neural network that learned to predict bankruptcy can be developed into a mature application that returns a company's risk of default given a set of specified ratios. Comparing with conventional statistical models, it does not require users (practitioners or researchers) to have significant background knowledge to play with it, and its prediction is easy to understand. Another advantage of a neural network lies under its learning mechanism. People do not need to worry that it might be obsolete because it can be trained with the most up-to-date data periodically to make it updated, which makes it stable over a long time span. Therefore, a well-trained neural network is expected to not only be more accurate and stable but also have practical meaning.

I study the ability of neural networks in bankruptcy prediction in two steps. In the first step, I design and train neural networks to identify bankrupt companies; and in the second step, I evaluate the performance of trained neural networks and compare the predictive power of our "best model" with that of the Shumway (2001) model and the distance-to-default model. My best model performs incredibly well in a Receiver Operating Characteristics analysis, and also outperforms its leading alternatives.

By studying the predictive ability of neural networks in corporate bankruptcy, I am confident to be an advocate for neural networks. By virtue of its powerful learning ability, the artificial neural network can most efficiently utilize publically available information and is able to become next milestone in the area of bankruptcy prediction.

The rest of the paper is structured as follows. Section II contains literature review; Section III specifies the methodology used in this study; Section IV and V present empirical results and assess the predictive power of the model; Section VI compares my best model with two leading alternatives; and Section VII summarizes my conclusions and provides suggestions for future research.

CHAPTER TWO

Previous Research

Modern day corporate bankruptcy prediction models can be traced back to Beaver's (1966) univariate study, where he tested the individual financial ratios' predictive abilities in classifying failed and non-failed firms. This study laid the groundwork for the studies that followed. In 1968, Altman constructed the well-known Z-score model. He employed discriminant analysis on a dataset containing thirty-three failed firms and thirty-three non-failed firms in the manufacturing industry and generated an accounting data based model for bankruptcy prediction. The Z-score model swiftly gained wide acceptance as a standard of financial healthiness, and even until today it remains popular among practitioners.

Despite the widespread use of the Z-score model, two limitations made researchers continue developing new models. First, the Z-score model could only be applied to firms in the manufacturing industry. Also, its underlying statistical method, discriminant analysis, has been heavily criticized for its restrictive assumptions. In more recent times the logit model [1] has become the primary multivariate model used in developing bankruptcy prediction models. Examples include Ohlson (1980), Zmijewski (1984) and Lennox (1999) who all achieved improved predictive accuracy.

While the logit model has less restrictive assumptions than discriminant analysis, it, like discriminant analysis, is a static model. Static models ignore the time series variation of a firm, and therefore can lead to biased and inconsistent estimates. To

[1] In this paper, I regard the logit model and the probit model as the same kind of models. They are slightly different from each other in a way that the logit model has slightly fatter tails.

improve upon static models, starting in 1987, many authors[2] began to explore dynamic forecasting models to predict corporate bankruptcies. In 2001, Shumway proposed a simple hazard model, the multi-period logit model, which can explicitly account for time. The introduction of the multi-period logit model then became one of the most important milestones in corporate bankruptcy prediction.

Shumway's (2001) publication marked the beginning of a new era in the study of corporate bankruptcy prediction. Researchers began to focus on dynamic models, especially the hazard model. In 2005, Beaver et al. extended Shumway (2001) by examining secular changes in the ability of accounting data to predict bankruptcy using a hazard model. Later on, Campbell et al. (2008) expanded the existing hazard models to include more market information. Campbell also tested the predictive power for different horizons. The results indicated that market capitalization, volatility, and the market-to-book ratio became more important relative to net income, leverage, and recent equity returns as predicting further into the future. Campbell's (2008) study was very influential in the area of financial distress and had been largely cited in other research.

While developing statistical models have been the mainstream of study in bankruptcy prediction, starting in 1990 a number of researchers began to apply ANN to bankruptcy prediction.[3] Many studies have showed that ANNs are more powerful than conventional statistical methods because of their nonlinear, nonparametric adaptive-learning properties. The first attempt to predict bankruptcies using ANNs was made by Odom and Sharda (1990), who constructed a three-layer feed-forward network. The

---

[2] Queen et al (1987), Theodossiou (1993), Denis et al (1997) and Pagano et al (1998) implemented dynamic models in bankruptcy prediction.

[3] Odom and Sharda (1990), Coleman et al. (1991), Coats and Fant (1993), Altman et al. (1994), Wilson and Shardra (1994), Zhang et al. (1997), O'Leary (1998), Atiya (2001), Virag (2005) and Santos et al (2006)

explanatory variables in their model were the same as the five variables in Altman's Z-score model. The same set of data was also analyzed using discriminant analysis. By comparing the predictive ability of the resulting ANN and the statistical model, they concluded that the ANN was applicable to bankruptcy prediction. Subsequently, a number of studies have investigated the application of ANNs in bankruptcy prediction. Coleman et al. (1991) extended Odom and Sharda's (1990) research using ANNs with different architectures, and improved upon the predictive accuracy of the Odom and Sharda's study.

Altman et al.'s (1994) study compared the results achieved with ANNs with that achieved by discriminant analysis, and provided a detailed discussion of the strengths and limitations of ANNs. Their experiments showed that the ANNs obtained results that were near or superior to the results obtained from discriminant analysis. Nonetheless, due to several limitations of ANNs, especially the inconsistent behaviors that made it impossible to interpret the model's operating logic on the basis of the coefficients, discriminant analysis was deemed to be better than the ANNs in their experiments.

Zhang et al. (1997) provided a comprehensive review of neural network applications to bankruptcy prediction and demonstrated the connection between ANNs and the traditional statistical classification methods. Zhang et al. illustrated that neural network outputs were the estimates of Bayesian posterior probabilities. Based on a matched sample of 220 firms they found that ANNs were significantly better than logit models.

Although a number of ANN studies comparing the performance of ANNs with conventional statistical models have been done, no one had done a comparative analysis of variable selection methods until du Jardin (2010). Du Jardin studied five selection techniques for three different models: discriminant analysis, logit models, and ANNs, and showed that the sets of variables that were selected with a likelihood criterion reached fairly high accuracy with a neural network.

CHAPTER THREE

Methodology

I first estimate a multi-period logit model with a pool of 36 variables that rely primarily on the use of accounting and stock information. Then, to study the power of neural networks in predicting corporate bankruptcies, I do an initial screening of the 36 variables to select the robust ones. The variables that pass my screening tests are then used in the development of a neural network. According to du Jardin (2010), using such a method to select variables for neural networks can reach fairly high accuracy. After determining the explanatory variables for developing neural networks, I decide the neural network architecture through a series of experiments in order to find a neural network with high predictive power in an out-of-sample test.

1. Variable Selection

A variety of corporate information has been used to evaluate a firm's risk of default. Accounting information[4] and market information[5] are two major sources that have been used in the previous literature. I first select a set of 33 initial explanatory variables that rely primarily on the use of accounting and stock information. The variables are divided into five categories: profitability, leverage, liquidity, solvency and stock information. I add two new variables, the net debt to invested capital ratio and the net debt to enterprise value ratio. In addition, I use annual GDP growth as a general economic indicator.

[4] Beaver (1966), Altman (1968), Ohlson (1980), Zmijewski (1984), Shumway (2001), etc.
[5] Shumway (2001), Campbell (2008), etc.

Table 1
List of Variables

This table lists the variables in five categories.

PROFITABILITY RATIOS	**LIQUIDITY RATIOS**
Sales to Total Assets	Working Capital to Total Assets
Net Income to Total Assets	Working Capital to Book Equity
Net Income to Market Assets	Working Capital to Market Equity
EBIT to Total Assets	Current Assets to Total Assets
Sales to Inventory	Cash to Total Assets
Net Income to Sales	Cash to Market Assets
Net Income to Equity	Cash and Equivalents to Total Assets
EBITDA to Total Liabilities	
Operating Income Dummy	
LEVERAGE RATIOS	**SOLVENCY RATIOS (SHORT/LONGTERM)**
Total Liabilities to Total Assets	Current Assets to Current Liabilities
Total Liabilities to Market Assets	(Current Assets-Inventory) to Current Liabilities
(Invested Assets+Inventory) to Equity	Cash to Total Liabilities
Retained Earnings to Total Assets	Book Equity to Total Liabilities
Long-term Debt to Invested Assets	Market Equity to Total Liabilities
Current Liabilities to Current Assets	**MARKET**
Current Liabilities to Total Assets	Annual Excess Return
Total Debt to Total Assets	Sigma
Net Debt to Invested Capital	Log of Market Capital
Net Debt to Enterprise Value	Market Equity to Book Equity

To test if the predictive power of these variables and baseline risks vary across industries in our sample, a Likelihood Ratio test (LR test) is used. Specifically, I test how many times more likely our data are under the unrestricted model compared to the restricted model and test if the difference is significant. The unrestricted model is defined as follows:[6]

$$Risk\ of\ default = \begin{cases} \dfrac{1}{1 + EXP(\beta_0 + \beta_1 x_1 + \cdots + \beta_{28} x_{28} + \beta_{29} \Delta GDP)} & if\ \gamma \\ \dfrac{1}{1 + EXP(\beta'_0 + \beta'_1 x_1 + \cdots + \beta'_{28} x_{28} + \beta'_{29} \Delta GDP)} & if\ \delta \end{cases}$$

[6] A multi-period logit model is defined by Shumway (2001) as a logit model that is estimated with data on each firm in each year of its existence as if each firm-year were an independent observation. The dependent variable is set equal to one only in the year in which a failure occurs.

The restricted model is the following:

$$Risk\ of\ default = \frac{1}{1 + EXP(\beta_0 + \beta_1 x_1 + \cdots + \beta_{28} x_{28} + \beta_{29} \Delta GDP + \beta_{30} Capital)}$$

$x_1 \dots x_{28}$: Variables about business and stock performance listed in Table 1
γ: Companies are in the goods-producing industry
δ: Companies are in the service industry
$Capital$: Capital-intensive industry dummy
(1 indicates the firm belongs to the capital-intensive industry, and 0 otherwise)

The restricted model assumes that baseline risks and the impact of each variable on the risk of default are consistent across industries in our sample, while the unrestricted model allows the baseline risks and the impact to vary across industries. One important thing to notice is that the unrestricted model is a combination of two models, where one is for goods-producing industries and the other is for service industries. Therefore, the essential argument is whether I need to estimate two distinct models to better fit the data of companies from different industries. If the test shows that the unrestricted model is significantly better than the restricted model, then baseline risks and/or the impact of each variable on the risk of default are different across industries, and therefore I need to develop two different models to better fit the data.

While all variables listed in Table 3 have been shown to be a good predictor, including all of them in a neural network is dangerous because doing so may lead to the problem of over-fitting, and thus influences the ability of generalization of our model. Normally, a neural network learns from training examples. During the learning process, the network adjusts itself to fit training data by minimizing the estimation errors. Ideally, we want to train a neural network to learn the “true” information from training examples and ignore the noise, so that the network can generalize to new situations. If we provide too much information of training examples, the network is likely to learn many noises inherent in the training data. As a result, the trained network is over-fitted the training data: it has only memorized the training examples, but it has not learned to generalize to new situations.

To avoid the problem of over-fitting, I do an initial screening of variables based on their performance in a univariate statistical test. The variables that can pass my initial screening are restricted to the ones that show significance in a single-variable multi-period logit model,

$$Risk\ to\ default = \frac{1}{1 + EXP\ (\beta_0 + \beta_1 x)}$$

where x is the variable being tested.

2. *Neural Network Architecture*

The architecture design is a major part of the development of a neural network. A complete architecture design specifies the number of layers, the number of neurons per layer, and interconnection patterns between layers.

Using ANNs to study the relationship between the likelihood of corporate bankruptcy and the selected explanatory variables, I need to address one important question: what is the appropriate neural network design? The design of a neural network includes many factors, such as selection of training methodology, architecture design, and specification of activation function. Logistic activation function is specified in my networks due to the fact that the probability of bankruptcy is between 0 and 1. Since back-propagation is the most commonly used training algorithm, my neural networks are designed to implement back-propagation learning algorithm. Back-propagation is an abbreviation for backward propagation of errors. It calculates the gradient of a loss function with respects to all the weights in the network. The gradient is then fed to the optimization method, which in turn uses it to update the weights, in an attempt to minimize the loss function.

The commonly used loss function is the squared error function defined as

$$L = \sum_i (\hat{y}_i - y_i)^2$$

where $\hat{y}$ is the estimated risk of default and y, a binary indicator of bankruptcy, is the real status of a company. The commonly used loss function treats the error from misclassifying a healthy company and the error from misclassifying a bankrupt company equally. Yet, if I take the investor's perspective, the cost of misclassifying a failed company is higher than the cost of misclassifying a company that does not fail, and therefore, the differential misclassification costs cannot be ignored. To make a model that well reflects this fact, I revise the squared error loss function to assign larger penalty to the misclassification of a failed company. Specifically, the loss function in our neural networks is specified to be

$$L = \sum_i 100(\hat{y}_i - y_i)^2 (y_i) + (\hat{y}_i - y_i)^2(1 - y_i).$$

The revised loss function magnifies the error from misclassifying a failed company by 100 times while keeps the error from misclassifying a healthy company unchanged. For example, if a bankrupt company is estimated to have a 40 percent chance of failure, the error will be treated as 60 instead of 0.6; and if a healthy firm is estimated to have a 60 percent chancee of failure, the error will be treated as 0.6, which is the same the true error. This modification to the loss function can make neural networks detect bankruptcy in a more conservative way.

Currently, there are no concrete principles to guide the architecture design. Network architecture refers to the number of layers and the number of neurons in each layer. In my models to distinguish failed companies from healthy companies, the number of input neurons is the number of selected explanatory variables. The number of hidden layers and the number of neurons in each hidden layer are not easy to be determined a priori. I decide them through experiments so that the predicting accuracy is maximized in an out-of-sample test.

CHAPTER FOUR

Results

1. Data Description

My measure of default risk is the probability of bankruptcy. I define bankruptcy to be a deletion from Compustat under the reason for deletion code 02-Bankruptcy and 03-Liquidation. The bankruptcy indicator equals one in a year prior to the year in which bankruptcy or liquidation occurs to a firm, and zero otherwise; in particular, the indicator is zero if the firm is deleted from Compustat for some reason other than bankruptcy and liquidation such as acquisition and no longer filing with SEC but pricing continues. I collect annual bankruptcy data that run from 1990 to 2013.

I use macroeconomic condition, accounting and market-based measures to forecast failure. My data are obtained from the Wharton Research Data Services (WRDS) database and the U.S. Bureau of Economic Analysis (BEA). Specifically, I construct accounting-based measures using annual financial data from Compustat, and construct market-based measures using monthly equity data from the Center for Research in Security Prices (CRSP).

Table 2
Number of Bankruptcies per Year
This table lists the total number of firms and bankruptcies for every year of our sample period before and after eliminating incomplete data.

Year	Bankruptcies (Before)	Firms (Before)	%	Bankruptcies (After)	Firms (After)	%
1990	0	1983	0.00%	0	1,574	0.00%
1991	0	2043	0.00%	0	1,636	0.00%
1992	15	2231	0.67%	10	1,768	0.57%
1993	69	2524	2.73%	57	2,063	2.76%
1994	43	2749	1.56%	30	2,315	1.30%
1995	51	2959	1.72%	31	2,510	1.24%
1996	42	3238	1.30%	34	2,785	1.22%
1997	38	3427	1.11%	30	2,962	1.01%
1998	49	3455	1.42%	40	3,011	1.33%
1999	61	3581	1.70%	45	3,121	1.44%
2000	87	3718	2.34%	58	3,223	1.80%
2001	63	3723	1.72%	48	3,178	1.51%
2002	34	3727	0.91%	26	3,159	0.82%
2003	23	3779	0.61%	17	3,143	0.54%
2004	30	3963	0.76%	17	3,274	0.52%
2005	19	4203	0.45%	8	3,429	0.23%
2006	16	4576	0.35%	7	3,626	0.19%
2007	12	5028	0.24%	4	3,812	0.10%
2008	43	5126	0.84%	12	3,794	0.32%
2009	86	5168	1.66%	29	3,772	0.77%
2010	74	5474	1.35%	21	3,936	0.53%
2011	35	5807	0.60%	11	4,068	0.27%
2012	104	6001	1.73%	14	4,201	0.33%
2013	69	5227	1.32%	13	3,421	0.38%

Table 2 summarizes the properties of our bankruptcy indicators. In order to construct both accounting-data based and market-based explanatory variables at the individual firm level, I need to combine annual accounting data from COMPUSTAT with monthly equity market data from CRSP. Therefore, only companies whose data are available in both databases are included in my sample.

While the initial sample whose summary statistics are reported in the first three columns in Table 2 may better reflect population, it contains many observations with incomplete data. After eliminating about 20,000 observations for incomplete data, I get a new sample. The last three columns in Table 2 summarize the new sample. My final sample contains 562 bankrupt firms and 73,781 data points for companies that stay healthy at the end of that year, compared with 1,063 bankrupt firms and 93,710 data points for companies that stay healthy at the end of that year in the initial sample.

As can be seen, most observations being eliminated from the initial sample are those in years 2004-2013. The average annual bankruptcy rate of 1.17 percent over years 1990-2003 in the new sample is very close to that of 1.33 percent in the initial sample. However, the average annual bankruptcy rate over years 2004-2013 drops to 0.36 percent from 0.96 percent in the new sample. In order to make the sample for model estimation close to the initial sample, I choose the subsample of data from 1990 to 2003 to be the training sample, and hold back the subsample of data from 2004 to 2013 for out-of-sample tests.

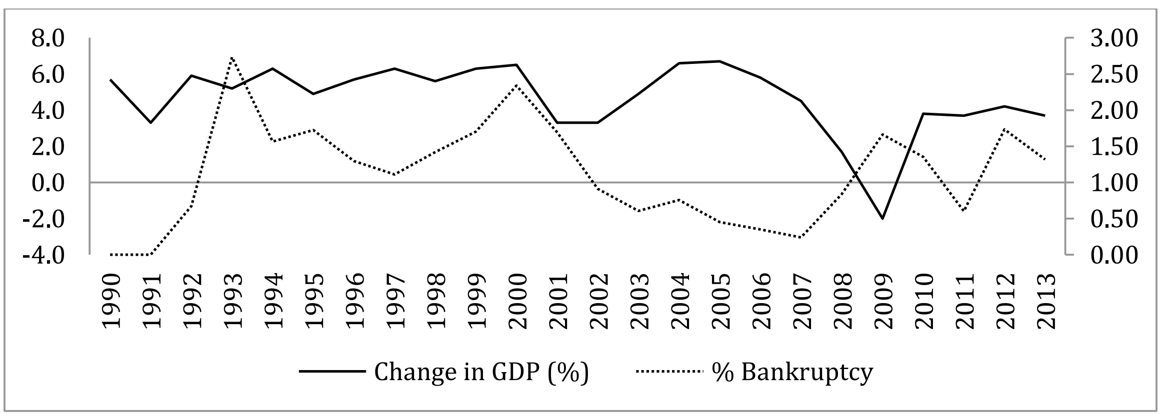

Figure 2. Change in GDP and Annual Bankruptcy Rate: this chart plots the GDP changes and the annual aggregate bankruptcy rate from 1990 to 2013.

To see if bankruptcy rate is correlated with the domestic economy, I plot the annual bankruptcy rate - the number of bankruptcies to the number of companies that stay healthy at the end of the previous year in our initial sample - and the GDP changes during the years 1990 to 2013 in Figure 2. It appears that the bankruptcy rate is negatively correlated with changes in GDP, especially during the four years 1994 to 1997 and the eight years from 2003 to 2010. The correlation was -0.63 in the first time period from 1994 to 1997 and was -0.76 in the second time period from 2003 to 2010.

Table 3
Number of Bankruptcies by Industries from 1990-2003
This table lists the total number of firms and bankruptcies for every industry from 1990-2003.

#	SIC	Industries	Bankruptcies	Firms	%
1	01-09	Agriculture, Forestry, and Fishing	0	15	0.00%
2	10-14	Mining	18	181	9.94%
3	15-17	Construction	7	56	12.50%
4	20-39	Manufacturing	149	1,829	8.15%
5	40-49	Transportation & Public Utilities	39	417	9.35%
		Goods-Producing Industry	213	2498	8.53%
6	50-51	Wholesale trade	23	159	14.47%
7	52-59	Retail trade	64	311	20.58%
8	60-67	Finance, insurance, & real estate	48	793	6.05%
9	70-89	Services	74	706	10.48%
10	91-97	Public administration	0	0	-
11	99	Nonclassifiable establishments	4	49	8.16%
		Service Industry	213	2018	10.56%
		Total	426	4516	9.43%

Table 3 describes bankruptcies in my training sample by industry. Companies are divided into 11 industries according to the first two digits of their SIC codes. Column 4 shows the number of bankruptcies in each industry. Column 5 shows the number of firms in each industry over my sample period. I categorize my sample of firms into two industries: the goods-producing industry and the service industry. The goods-producing industry contains firms that have SIC codes with the first two digits less than 50, and the service industry otherwise. Specifically, the goods-producing industry includes industries like agriculture, forestry, & fishing, mining, construction, manufacturing, and transportation & public utilities; and the service industry includes industries like wholesale trade, retail trade, finance, insurance, and real estate, services, public administration, and non-classifiable establishments. As a result, 50% firms in my bankruptcy sample are categorized into service industry.

2. Findings

2.1 Likelihood Ratio Test Results

Table 4 reports the Likelihood Ratio test results of the restricted and unrestricted multi-period models. Note that the restricted model defined in the previous chapter assumes that baseline risks and the impact of each variable on the risk of default are consistent across industries, while the unrestricted model says that baseline risks and the impact may vary across industries. In absolute value, the log-likelihood of the unrestricted model is smaller than that of the restricted model, showing that my data are more likely under the unrestricted model, which makes sense given that I have fewer restrictions in the unrestricted model. The question now turns to be whether it is worth developing a model for each industry separately. In other words, is the discrepancy significantly large? If not, I may still want to develop an "omnipotent" model that can predict bankruptcies in both industries. The test results in Table 4 indicate that the difference is statistically significant at a 1 percent level of significance. It implies that in my sample, the impact of the variables listed in Table 3 and/or baseline risks change(s) across industries.

Table 4
Likelihood Ratio Test Results

	Log-likelihood
Unrestricted Model	-1724
Restricted Model	-1856
D statistic	265
Prob. > D-stat	0.00

2.2 Multi-period Logit Models

Based on the LR test results, the unrestricted multi-period logit model is estimated using 36 variables for comparison purpose. Specifically, to estimate the unrestricted multi-period logit model, I split my sample into two sub-samples. One sample contains goods-producing companies, and the other contains service companies.

Table 5 reports the estimation results. Since high multicollinearity is a problem inherent in models that use financial ratios as explanatory variables, the estimated standard errors for the coefficients of our variables are inflated, and therefore some variables that have strong predictive power are reported to be not significant. I report variance inflation factors (VIFs) in Table 6. The square root of the VIF of an variable tells us how many times the standard error for the coefficient of this variable is inflated from what it would be if it were not correlated with the other explanatory variables.

Table 5
Results by Samples

Variable	Goods-producing	Services
Total Liabilities/Total Assets	0.3204	-0.1188
Total Liabilities/Market Assets	-0.0040***	0.0468**
Working Capital/Book Equity	0.0014	-0.0062
Working Capital/Market Equity	-0.0017**	-0.0003***
Annual Excess Return	-0.3603**	-1.2015***
Net Debt/Invested Capital	0.0004	0.0021
Net Debt/Enterprise Value	-0.0004	-0.0028**
Retained Earnings/Total Assets	0.0584	0.0132
Working Capital/Total Assets	1.5977**	1.2167***
EBIT/Total Assets	-0.2691	0.1085
Sales/Total Assets	0.1726*	0.1208***
Net Income/Total Assets	-0.0723	-0.1912***
Net Income/Market Assets	-0.0673***	-0.0138
Current Assets/Current Liabilities	-0.1523	-0.0071
(Current Assets-Inventory)/Current Liabilities	-0.4054*	-0.2842
Cash/Total Assets	-3.1539**	-1.7412*
Cash/Market Assets	0.0370***	-0.1215**
Cash/Total Liabilities	0.9793***	0.0219**
(Invested Assets+Inventory)/Equity	0.0009	0.0087**
Current Assets/Total Assets	-0.3491	-0.2300
Sales/Inventory	0.0003	-0.0003
Long-term Debt/Invested Assets	-0.0096	0.0023
Current Liabilities/Current Assets	-0.0178	-0.0426
Net Income/Sales	0.0007	0.0027
Net Income/Total Equity	0.0034	0.0132**
Cash and Equivalents/Total Assets	0.2946	1.3828**
Current Liabilities/Total Assets	1.4531**	0.7291*
Total Debt/Total Assets	0.3747	0.6803**
EBITDA/Total Liabilities	0.0498	0.0088*
Sigma	1.6201***	1.0891***
Operating Income Dummy	-1.6662***	-1.8884***
Log of Market Cap	-0.0811***	-0.0744***
Book Equity/Total Liabilities	-0.0158	0.0021
Market Equity/Total Liabilities	-0.0039*	-0.0027**
Market Equity/Book Equity	0.0009	-0.0033
GDP Growth	0.2979***	0.2817***

-*** 1% significance level before adjusted for multicollinearity. -** 5% significance level before adjusted multicollinearity. ***- 1% significance level after adjusted for multicollinearity. **- 5% significance level after adjusted for multicollinearity.

Table 6
Variance Inflation Factors

Variable	Goods-producing	Services
Total Liabilities/Total Assets	7.76	3.03
Total Liabilities/Market Assets	2.26	1.47
Working Capital/Book Equity	40.56*	1.13
Working Capital/Market Equity	1.01	1.01
Annual Excess Return	1.06	1.06
Net Debt/Invested Capital	1.03	1.02
Net Debt/Enterprise Value	1.00	1.00
Retained Earnings/Total Assets	2.40	24.76*
Working Capital/Total Assets	5.78	4.22
EBIT/Total Assets	7.33	28.11*
Sales/Total Assets	2.13	1.66
Net Income/Total Assets	5.29	15.28*
Net Income/Market Assets	2.06	1.38
Current Assets/Current Liabilities	71.68*	386.67*
(Current Assets-Inventory)/Current Liabilities	68.28*	533.80*
Cash/Total Assets	4.38	3.33
Cash/Market Assets	1.79	1.89
Cash/Total Liabilities	5.24	21.22*
(Invested Assets+Inventory)/Equity	986.07*	1.54
Current Assets/Total Assets	7.24	136.95*
Sales/Inventory	1.04	1.00
Long-term Debt/Invested Assets	1.01	1.02
Current Liabilities/Current Assets	1.19	1.00
Net Income/Sales	1.02	1.00
Net Income/Total Equity	746.90*	3.06
Cash and Equivalents/Total Assets	6.42	3.74
Current Liabilities/Total Assets	3.76	3.12
Total Debt/Total Assets	7.76	1.64
EBITDA/Total Liabilities	2.61	5.94
Sigma	1.29	1.31
Operating Income Dummy	1.80	1.53
Log of Market Cap	1.29	1.12
Book Equity/Total Liabilities	2.35	5.79
Market Equity/Total Liabilities	1.67	4.43
Market Equity/Book Equity	3.18	2.83

2.3 Variable Selection Results

Table 7 and 8 report the variables that are significant in the single-variable models and their summary statistics. There are 14 robust predictors for both goods-producing industries and service industries. In addition, GDP growth is a significant predictor in all models.

Table 7 Goods-producing Industry

Variable	Mean	
	Bankrupt	Healthy
Profitability		
NI/TA	-0.53	-0.06
S/TA	1.28	1.00
EBIT/TA	-0.34	-0.01
EBITDA/TL	-0.35	-0.06
RE/TA	-1.81	-0.30
OPDUM	0.26	0.73
Leverage		
TL/TA	0.86	0.49
LL/IA	-0.21	0.29
TD/TA	0.46	0.24
Liquidity		
WC/TA	-0.01	0.26
CHE/TA	0.12	0.17
Solvency		
ME/TL	1.75	7.96
Market		
LnFSize	14.62	18.43
Sigma	0.23	0.14

Table 8 Services Industry

Variable	Mean	
	Bankrupt	Healthy
Profitability		
S/TA	1.58	1.07
NI/MA	-0.60	-0.01
OPDUM	0.31	0.79
Leverage		
TL/TA	0.71	0.58
CL/TA	0.40	0.21
TL/MA	0.77	0.46
ND/ME	-3.42	7.49
Liquidity		
WC/TA	0.05	0.16
WC/ME	-0.23	3.11
Solvency		
CA/CL	1.28	2.22
(CA-INV)/CL	0.92	1.85
Market		
Excess Return	-0.13	-0.01
LnFSize	14.91	18.13
Sigma	0.21	0.13

As shown in Tables 7 and 8, bankrupt companies are less profitable, more highly leveraged, less liquid, less solvent, and have poorer stock performance than healthy companies on average. The discrepancy in solvency is more apparent between goods-producing industries, but less apparent between services companies. The discrepancy in liquidity is more obvious between services companies, but less obvious between goods-producing companies. The discrepancy in profitability and stock performance is clear both between goods-producing companies and between services companies.

2.4 Neural Network Results and Evaluation

Note that the LR test shows that baseline risks and the impact of each explanatory variable on predicting bankruptcy vary across industries. Based on this result, I develop two neural networks, one for goods-producing industries and the other for service industries.

As mentioned earlier, my sample is split into two subsamples: data in years 1990 to 2003 are included in the training sample for training neural networks; and data in years 2004 to 2013 are included in the holdout sample. I use data in years 2004 to 2008 from this holdout sample to do the out-of-sample test for selecting the best architectures, and this sample will be referred to as the "validation sample;" I use the rest in the holdout sample to evaluate the ability of generalization of our neural networks, and this sample will be referred to as the "testing sample."

After a number of experiments, the best performing model is chosen for each industry. The best performing model is defined to be the model that achieves the highest accuracy in detecting bankruptcies among all of my experiments for the same industry in an out-of-sample test using the validation sample. While I do not intentionally keep the two architectures consistent, both best performing models have two hidden layers with three neurons in the first hidden layer and two neurons in the second hidden layer. Figure 3 and 4 are the best performing neural networks.

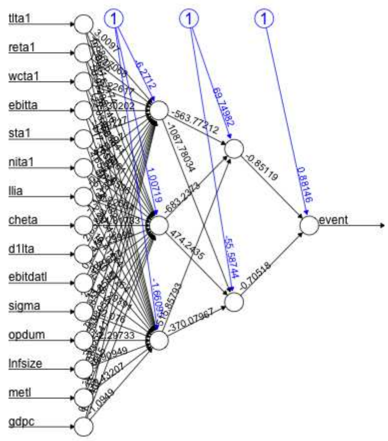

Figure 3. Goods-producing Industry

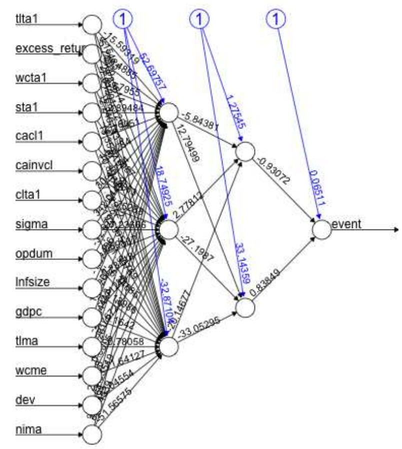

Figure 4 Service Industry

Finally, to evaluate the performance of our models, I construct confusion matrices that summarize the training results in Table 9. The neural network for goods-producing industries achieved overall accuracy of 86 percent, where its true positive rate is 94 percent and true negative rate is 86 percent: it classifies 94 percent bankrupt companies into the correct category and classifies 86 percent healthy companies into the correct category. However, as it is designed to be conservative in order to avoid misclassification of bankrupt companies as far as possible, the model has a low precision, only 7 percent. The precision of 7 percent indicates that, among all the companies that my model classified as bankrupt companies, only 7 percent of them are actual default companies. The neural network for service industries performed as satisfactorily as the model for goods-producing industries. It achieved overall accuracy of 77 percent, where it correctly identifies 97 percent bankrupt companies and 77 percent healthy companies. Similarly, it has a low precision, 6.3 percent.

Table 9 – Confusion Matrix – Training

Panel A – Goods-producing Industries

# Company-year	Actual Bankruptcies	Actual Healthy Companies
Predicted Bankruptcies	200	2726
Predicted Healthy Companies	13	16883

Panel B – Service Industries

# Company-year	Actual Bankruptcies	Actual Healthy Companies
Predicted Bankruptcies	206	3065
Predicted Healthy Companies	6	10139

The confusion matrices for the out-of-sample test are reported in Table 10. In the out-of-sample test, I use my models to predict bankruptcies that happened in years 2004 to 2013. The neural network for goods-producing industries achieved overall accuracy of 94 percent, where its true positive rate is 55 percent and true negative rate is 94 percent, and reached the precision of 2 percent. The neural network for service industries achieved overall accuracy of 88 percent, where it correctly identifies 56 percent bankrupt companies and 88 percent healthy companies. Its precision reached 2 percent.

Table 10 – Confusion Matrix – Out-of-sample Test

Panel A – Goods-producing Industries

# Company-year	Actual Bankruptcies	Actual Healthy Companies
Predicted Bankruptcies	27	1379
Predicted Healthy Companies	22	21106

Panel B – Service Industries

# Company-year	Actual Bankruptcies	Actual Healthy Companies
Predicted Bankruptcies	50	2175
Predicted Healthy Companies	40	15702

I also consider the ability of my model to explain variation in the aggregate bankruptcy rate over time. As I noted in our discussion of Table 1 in the previous section, there is considerable variation in the bankruptcy rate over time. How well does my model fit this pattern? Figure 5 plots the actual bankruptcy rate and the predicted bankruptcy rate. The predicted aggregate bankruptcy rate is the average of the predicted probability of bankruptcy for each company using our best model (Campbell, 2008). As shown in the graph, while my model closely follows the pattern, the predicted aggregate bankruptcy rate is about 10 times as high as the actual bankruptcy rate in each year. The result is not surprising because I started with the assumption that the cost of misclassifying a bankrupt firm is much high than that of misclassifying a healthy firm, and thereby the model is designed to be conservative in a sense that it may provide higher-than-reality bankruptcy rate.

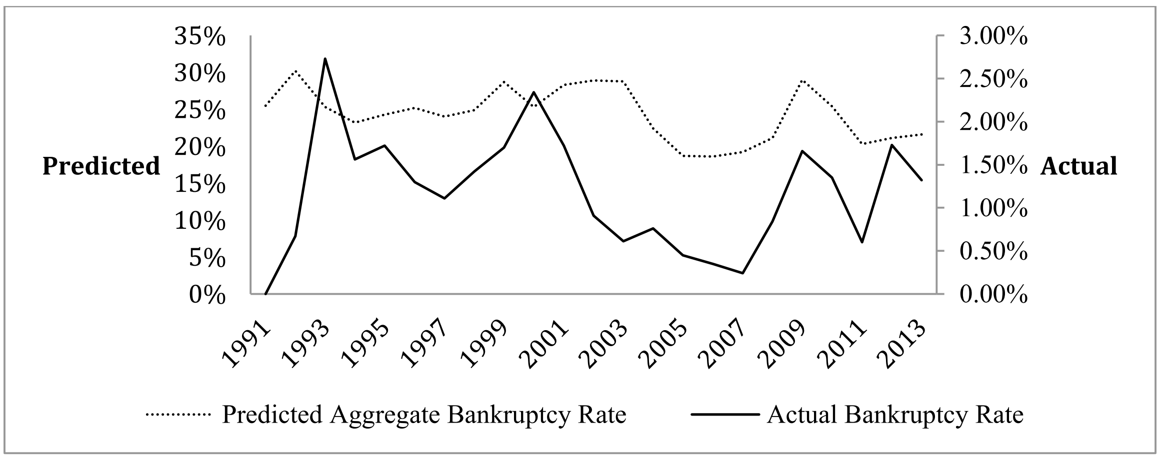

Figure 5. Predicted aggregate bankruptcy rate versus actual bankruptcy rate

CHAPTER FIVE

Comparing ANNs to Alternative Models

I next compare the performance of my model with that of the multi-period logit model estimated in Chapter Four. I also compare my model with a common alternative that is used especially by practitioners: distance-to-default.

The distance-to-default model is one of the most popular approaches to forecast bankruptcy. It is based on Merton's (1974) bond pricing model, where the equity of the firm is regarded as a call option on the underlying value of the firm with a strike price equal to the face value of the firm's debt. The model specifies that the probability of bankruptcy is the normal cumulative density function of a distance-to-default score that depends on the firm's underlying value, the firm's asset volatility and the face value of the firm's debt.

However, neither the firm's underlying value nor its asset volatility is directly observable. Theoretically, under the model's assumptions, both can be inferred from the market value of the equity, the stock volatility, the risk free rate, and the face value of debt by solving two equations simultaneously. Campbell et al. (2008) solved the two equations numerically in order to find values for the underlying assets and the asset volatility. Yet, the procedure is complicated and cannot be easily implemented.

Bharath and Shumway (2008) proposed a naïve alternative that does not require complex numerical methods. They approximated the value of underlying asset as the market value of equity plus the face value of debt, and approximated the asset volatility as a weighted average of the equity volatility and the debt volatility,

$$Naive\ \sigma_A = \frac{E}{E+D}\sigma_E + \frac{D}{E+D}Naive\ \sigma_D,$$

where E is the market value of equity, D is the face value of debt, σ_E is the volatility of equity and $Naive\ \sigma_D$ is the debt volatility approximated as

$$Naive\ \sigma_D = 0.05 + 0.25 * \sigma_E.$$

Bharath and Shumway showed that the predictive power of the naïve distance-to-default predictor constructed with the approximated firm's underlying value and the asset volatility is comparable to that of the KMV distance-to-default predictor.

Since the naïve predictor is extremely easy to calculate and is tested to be comparable to the true distance-to-default predictor, I choose to use the naïve approach to construct the distance-to-default model. The model used in my comparison is

$$DD = \frac{\ln\left[\frac{E+D}{D}\right] + \left(r_f - 0.5 * Naive\ \sigma_A^2\right)T}{Naive\ \sigma_A\sqrt{T}}.$$

I assume T=1 and use the book value of short term debt plus one half the book value of long term debt to proxy for D, the face value of debt (Vassalou and Xing (2004)). Table 11 reports summary statistics.

Table 11: Summary Statistics

Table 11 reports summary statistics for all the variables used in the naïve distance-to-default model. E is the market value of equity in millions of dollars and is taken from CRSP as the product of share price at the end of the year and the number of shares outstanding. D is the approximation of the face value of debt in millions of dollars (computed as short term debt plus one half long term debt). rrf is the risk-free rate measured as the ten year Treasury-bill rate. $Naive\ \sigma_A$ is the approximated asset volatility. *Naïve* DD is the naïve distance-to-default score. Probability is the estimated probability of default. Our sample spans 1990 through 2013.

	Mean	Std. Dev.	Min	Max
E (Million)	100.00	456.00	0.13	9410.00
	3060.00	14700.00	0.02	602000.00
D (Million)	49.50	165.00	0.00	2090.00
	1040.00	9900.00	0.00	534000.00
rrf	0.05	0.02	0.02	0.08
$Naive\ \sigma_A$	0.17	0.14	0.00	2.23
	0.11	0.09	0.00	3.80
Naïve DD	536.00	7035.00	-7.18	128196.00
	1573.00	125428.00	-5.96	25700000.00
Probability	0.14	0.29	0.00	1.00
	0.02	0.11	0.00	1.00

I compare my model with the multi-period logit model estimated in the previous chapter and the naïve distance-to-default model. Table 12 and 13 report the confusion matrices for these two alternative models.

The distance-to-default model performs well compared to multi-period logit model. It is able to correctly classify 34 percent of bankrupt companies and 78 percent of healthy companies in the sample of data during years 1990-2013, and is able to correctly predict 55 percent of bankruptcies in the holdout sample only. While the multi-period logit model can achieve very high overall accuracy and relatively high precision, I can hardly say that the model has discriminating power due to its low true positive rate. My neural network model is more powerful in detecting bankruptcies than both the alternatives. The shortcoming of my model is that its precision is only 2 percent in the out-of-sample test, but it still outperforms the distance-to-default model whose precision is less than 1 percent in the out-of-sample test.

Table 12 – Confusion Matrix – 1990 to 2013 (Entire Sample)
Panel A – Multi-period Logit Model

# Company-year	Actual Bankruptcies	Actual Healthy Companies
Predicted Bankruptcies	12	72
Predicted Healthy Companies	553	73144

Panel B – Distance to Default Model

# Company-year	Actual Bankruptcies	Actual Healthy Companies
Predicted Bankruptcies	190	16447
Predicted Healthy Companies	375	56769

Panel C – Artificial Neural Network

# Company-year	Actual Bankruptcies	Actual Healthy Companies
Predicted Bankruptcies	440	9345
Predicted Healthy Companies	125	63830

Table 13 – Confusion Matrix – 2004-2013 (Out-of-sample Test)
Panel A – Multi-period Logit Model

# Company-year	Actual Bankruptcies	Actual Healthy Companies
Predicted Bankruptcies	2	45
Predicted Healthy Companies	137	40339

Panel B – Distance-to-Default Model

# Company-year	Actual Bankruptcies	Actual Healthy Companies
Predicted Bankruptcies	70	8796
Predicted Healthy Companies	52	28449

Panel C – Artificial Neural Network

# Company-year	Actual Bankruptcies	Actual Healthy Companies
Predicted Bankruptcies	77	3554
Predicted Healthy Companies	62	36808

To further evaluate my model's discriminant performance compared with other models, I do a Receiver Operating Characteristics (ROC) analysis of all three models. In ROC analysis, the predicated classification of each individual firm is compared to its true status, bankrupt or healthy. The area under an ROC curve is an indicator of the discriminating performance. The area is equal to the probability that the model can correctly identify the failed firms when given the information of a random chosen pair of firms where one firm is failed and the other is not. If the model could perfectly discriminate between failed and non-failed firms, then the ROC curve would consist of two straight line segments encompassing the entire unit square, and thus the area under the curve is one, which would be interpreted as the model having a one hundred percent probability of correctly classifying a random pair containing one failed firm and one non-failed firm.   Figure 6 plots the ROC curves of all three models on the holdout sample, which contains data during years 2004-2013. The green one is the ROC curve of my neural network, the red one is the curve of distance-to-default model, and the blue one is the curve of multi-period logit model. The area under the ANN's ROC curve (AUC) is 0.82, under the distance-to-default's ROC curve is 0.80, and under the multi-period logit model's ROC curve is 0.63. The plot shows that my model has very strong discriminating power and greatly outperforms the multi-period logit model while slightly outperforms the distance to default model.

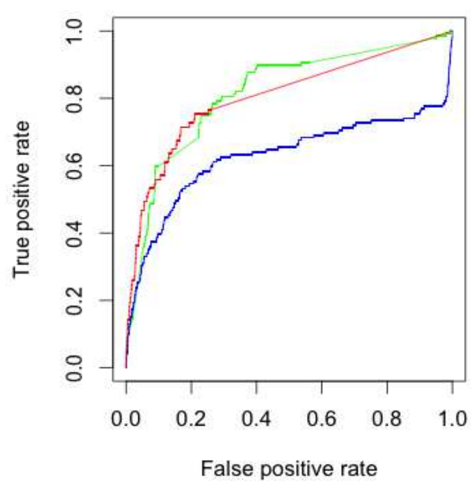

Figure 6 – ROC

CHAPTER SIX

Conclusion

In the light of comparisons in my study, the neural network is a very powerful model in predicting bankruptcy. The networks trained on my sample have shown significant predictive power, with out-of-sample test results that are superior to multi-period logit model and the well-known distance-to-default model. My study finds that a neural network with revised loss function that gives more penalties to the misclassification of failed companies increases the power of detecting bankruptcies upon a normal neural network, while the overall predicting accuracy may decline. The exceptional discriminating power of my networks is also demonstrated in a ROC analysis, which shows that 82 percent bankrupt companies are predicted to have a higher risk of default than a randomly chosen healthy company.

The strong learning ability of neural networks is a very attractive characteristic. This characteristic is especially appealing in discriminating potential bankrupt companies from healthy companies. Although the topic I focus on is known as bankruptcy prediction, "discriminating" or "detecting bankruptcies" would be a more accurate terminology because what we try to do is to detect potential bankruptcies from a pool of companies using our knowledge. The knowledge is based on historical examples. While every company is financially distressed in a distinct way, some shared patterns must more or less exist, which may be highly twisted and not be easily captured by simple statistical models. Learning from massive examples enables a neural network to recognize these patterns and give comparably accurate prediction.

Since neural networks are superior to conventional statistical models in terms of predictive accuracy, it is a promising and powerful tool in practice. For entities like hedge funds and vulture investors who seek to predict the probability of bankruptcy accurately, neural networks are more promising than other models, such as Altman's Z-score model, the distance-to-default model, and the Shumway model, which may perform well in evaluating a company's financial risk, but are not as accurate as neural networks in predicting bankruptcy. Given the superiority of neural networks in predicting bankruptcies, the next question is, can this new model be easily adapted to business practice? The resulting networks can be developed into software like Altman's Z-score Calculator, where users can get to know a company's risk of default simply by entering the required company information. By virtue of its reliable learning mechanism and demonstrated predictive power, the neural network is worth further development by software engineers, which can turn it into a powerful tool that can be easily adopted by the real world.

Furthermore, successfully predicting bankruptcies does not only benefit entities that actively invest in financially distressed companies but also benefits entities that try to avoid potentially bankrupt companies, such as banks, bondholders and stockholders. As shown in Campbell et al. (2008), stocks of financially distressed companies significantly underperform the S&P500 and are very risky. These stocks are shown to be mispriced. Moreover, even large companies with higher than average analyst coverage and with high levels of institutional holdings can face the problem of financial distress, leading to significant underperformance of stocks. I present a model that provides a reliable measure of financial distress for investors. Although bond rating agencies, like Moody's, Standard and Poor's, and Fitch, assess the creditworthiness of debt securities and their issuers, our model gives more precise evaluation specifically aimed at a company's potential for declaring bankruptcy. It assists investors to avoid distressed companies when constructing a diversified portfolio. In addition, it allows investors that do not want to miss the opportunity of investing in companies that appears to be distressed but will not go bankrupt at a huge discount price. The model can also help with evaluating the likelihood of a seemingly distressed company to recover its financial healthiness.

While the model closely follows the pattern, my model cannot explain the variation in the aggregate bankruptcy rate over time. The higher-than-actual predicted aggregate bankruptcy rate is a side-effect of magnifying the loss of misclassifying a bankrupt company. In fact, several studies have been done on the distribution of loan/corporate defaults. One of the major factors that have been considered to have significant impact on the distribution of defaults is the interdependency between companies.

Defaults have been shown to not occur independently over time because of the dependency between obligors. According to Elliott et al. (2014), in a networked market, financial contagions of failures among organizations are linked through a network of financial interdependencies that has been created by organizations' holdings of each other's obligations like shares and debts. Such interdependencies can lead to widespread failures when the "central company" of a highly networked market defaults.

To improve upon our best model, future work can focus on this interdependency and study if incorporating the measure of interdependency can better fit the distribution of bankruptcies as well as further increase the model accuracy. In addition, a more specific bankruptcy prediction model can be developed merely for those "central companies" of different regional financial networks. If we can correctly locate the central companies and develop a model to evaluate the financial distress risk of those companies, such a model can be more powerful: the predicted risk will not only show the default risk of a particular company but also approximate the default risk of a group of companies. Moreover, such a model can also be used to predict financial crisis. If central companies simultaneously melt down, the likelihood of financial crisis soars. Therefore, such a model could be very powerful.

REFERENCES

Altman, Edward I. “Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy.” *The Journal of Finance* 23.4 (1968): 589–609.

Altman, Edward I., Giancarlo Marco, and Franco Varetto. “Corporate Distress Diagnosis: Comparisons Using Linear Discriminant Analysis and Neural Networks (the Italian Experience).” *Journal of Banking & Finance* 18.3 (1994): 505–529.

Atiya, A.F. “Bankruptcy Prediction for Credit Risk Using Neural Networks: A Survey and New Results.” *IEEE Transactions on Neural Networks* 12.4 (2001): 929–935.

Balcaen, Sofie, and Hubert Ooghe. “35 Years of Studies on Business Failure: An Overview of the Classic Statistical Methodologies and Their Related Problems.” *British Accounting Review* 38 (2006): 63–93.

Bauer, Julian, and Vineet Agarwal. “Are Hazard Models Superior to Traditional Bankruptcy Prediction Approaches? A Comprehensive Test.” *Journal of Banking & Finance* 40 (2014): 432–442.

Beaver, William H. “Financial Ratios As Predictors of Failure.” *Journal of Accounting Research* 4 (1966): 71.

Beaver, William H., Maureen F. McNichols, and Jung-Wu Rhie. “Have Financial Statements Become Less Informative? Evidence from the Ability of Financial Ratios to Predict Bankruptcy.” *Review of Accounting Studies* 10.1 (2005): 93–122.

Campbell, John Y., Jens Hilscher, and Jan Szilagyi. “In Search of Distress Risk.” *The Journal of Finance* 63.6 (2008): 2899–2939.

Coats, Pamela K., and L. Franklin Fant. “Recognizing Financial Distress Patterns Using a Neural Network Tool.” *Financial Management* 22.3 (1993): 142.

Coleman, Kevin G., Timothy J. Graettinger, and William F. Lawrence. “Neural Networks for Bankruptcy Prediction: The Power to Solve Financial Problems.” *AI Rev.* 4 (1991): 48–50.

Denis, David J., Diane K. Denis, and Atulya Sarin. “Ownership Structure and Top Executive Turnover.” *Journal of Financial Economics* 45.2 (1997): 193–221.

Du Jardin, Philippe. "Predicting Bankruptcy Using Neural Networks and Other Classification Methods: The Influence of Variable Selection Techniques on Model Accuracy." *Neurocomputing* 73.10–12 (2010): 2047–2060.

Eng, John. "Receiver Operating Characteristic Analysis." *Academic Radiology* 12.7 (2005): 909–916.

Gershenson, Carlos. "Artificial Neural Networks for Beginners." *arXiv:cs/0308031* (2003): n. pag.

Lennox, Clive S. "The Accuracy and Incremental Information Content of Audit Reports in Predicting Bankruptcy." *Journal of Business Finance & Accounting* 26.5-6 (1999): 757–778.

"Moody's: Low Global Corporate Default Rate Holds Steady in 2013, Likely to Drop Further in 2014." *Moodys.com*.

Odom, M.D., and R. Sharda. "A Neural Network Model for Bankruptcy Prediction." *, 1990 IJCNN International Joint Conference on Neural Networks, 1990*. N.p., 1990. 163–168 vol.2.

Ohlson, James A. "Financial Ratios and the Probabilistic Prediction of Bankruptcy." *Journal of Accounting Research* 18.1 (1980): 109.

O'Leary, Daniel E. "Using Neural Networks to Predict Corporate Failure." *Intelligent Systems in Accounting, Finance & Management* 7.3 (1998): 187–197.

Pagano, Marco, Fabio Panetta, and Luigi Zingales. *Why Do Companies Go Public? An Empirical Analysis*. National Bureau of Economic Research, Inc, 1995. *RePEc - IDEAS*. Web. 19 Dec. 2014.

Queen, Maggie, and Richard Roll. "Firm Mortality: Using Market Indicators to Predict Survival." *Financial Analysts Journal* 43.3 (1987): 9–26. *cfapubs.org (Atypon)*.

Santos, M. F. et al. "Corporate Bankruptcy Prediction Using Data Mining Techniques." Vol. 1. WIT Press, 2006. 349–357.

Shumway, Tyler. "Forecasting Bankruptcy More Accurately: A Simple Hazard Model." *The Journal of Business* 74.1 (2001): 101–24.

Sreedhar T. Bharath, Tyler Shumway. "Forecasting Default with the KMV-Merton Model." *differences* (2004): n. pag.

Sule, Ahmed. "Predicting Financial Distress and the Performance of Distressed Stocks." *CFA Digest* 41.4 (2011): 7–9. *cfapubs.org (Atypon)*.

Theodossiou, Panayiotis T. “Predicting Shifts in the Mean of a Multivariate Time Series Process: An Application in Predicting Business Failures.” *Journal of the American Statistical Association* 88.422 (1993): 441.

Virág, Miklós, and Tamás Kristóf. “Neural Networks in Bankruptcy Prediction - A Comparative Study on the Basis of the First Hungarian Bankruptcy Model.” *Acta Oeconomica* 55.4 (2005): 403–426.

Zhang, Guoqiang et al. “Artificial Neural Networks in Bankruptcy Prediction: General Framework and Cross-Validation Analysis.” *European Journal of Operational Research* 116.1 (1999): 16–32.

Zmijewski, Mark E. “Methodological Issues Related to the Estimation of Financial Distress Prediction Models.” *Journal of Accounting Research* 22 (1984): 59.