YAŞAR UNIVERSITY

GRADUATE SCHOOL OF SOCIAL SCIENCES

BUSINESS ADMINISTRATION PROGRAMME

PHD THESIS

PREDICTING FRAUDULENT FINANCIAL ACTIVITIES THROUGH NEURAL NETWORK ALGORITHMS

MUSTAFA REHA OKUR

THESIS ADVISORS

ASSOC. PROF. DR. F. DİLVİN TAŞKIN YEŞİLOVA

ASSOC. PROF. DR. YASEMİN ZENGİN KARAİBRAHİMOĞLU

2019, İZMİR

DOKTORA/SANATTA YETERLİK TEZİ JÜRİ ONAY SAYFASI

Bu tezi okuduğumu ve görüşüme göre doktora/sanatta yeterlik derecesi için bir tez olarak kapsam ve nitelik açısından tam olarak yeterli olduğunu onaylarım.

Danışman Doç.Dr. F.Dilvin Taşkın Yeşilova
06.08.2019

Eş-Danışman Doç. Dr. Yasemin Karaibrahimoğlu

Bu tezi okuduğumu ve görüşüme göre doktora/sanatta yeterlik derecesi için bir tez olarak kapsam ve nitelik açısından tam olarak yeterli olduğunu onaylarım.

Dr. Öğr. Üyesi Özen Ece Aşar

Bu tezi okuduğumu ve görüşüme göre doktora/sanatta yeterlik derecesi için bir tez olarak kapsam ve nitelik açısından tam olarak yeterli olduğunu onaylarım.

Doç. Dr. Sevinç Güler Özçalık
06.08.2019

Bu tezi okuduğumu ve görüşüme göre doktora/sanatta yeterlik derecesi için bir tez olarak kapsam ve nitelik açısından tam olarak yeterli olduğunu onaylarım.

Prof. Dr. A. Fatih Dalkılıç
06.08.2019

Bu tezi okuduğumu ve görüşüme göre doktora/sanatta yeterlik derecesi için bir tez olarak kapsam ve nitelik açısından tam olarak yeterli olduğunu onaylarım.

Dr. Öğr. Üye. M. Gürol Durak
06.08.2019

Bu tezi okuduğumu ve görüşüme göre doktora/sanatta yeterlik derecesi için bir tez olarak kapsam ve nitelik açısından tam olarak yeterli olduğunu onaylarım.

Doç.Dr. Çağrı Bulut
SOSYAL BİLİMLER ENSTİTÜ MÜDÜRÜ

ABSTRACT

PREDICTING FRAUDULENT FINANCIAL ACTIVITIES THROUGH NEURAL NETWORK ALGORITHMS

Mustafa Reha OKUR

Ph.D., Business Administration

Advisors: Assoc. Prof. Dr. F. Dilvin Taskin / Assoc. Prof. Dr. Yasemin Zengin Karaibrahimoglu

Despite worldwide regulatory efforts (e.g., Sarbanes – Oxley Act, Financial Security Law of France, Fraud Act 2006 of the United Kingdom), fraud is still a major concern of today's capital markets. This study aims to forecast the risk of fraudulent financial activities of cross-listed companies in US stock exchanges (NYSE, NASDAQ) by employing a Neural Network based algorithm. Data of financial fraud filings, financial statements, corporate governance variables, and macroeconomic indicators are collected to construct a comprehensive study. By this method, this study tries to develop a broader framework on fraud detection that does not focus only on firm-specific aspects, instead of covering a more comprehensive dataset, which incorporates country-specific institutional factors into consideration. This study employs four machine learning based classification algorithms. Random Forest and C4.5 algorithm outperformed others with superior classification power. Moreover, this study mostly exceeds the classification ability of the previous literature.

Keywords: *Financial Fraud, Accounting Fraud, Neural Network, Machine Learning, Artificial Neural Network, Decision Trees, Forecasting*

ÖZ

HİLELİ FİNANSAL AKTİVİTELERİN SİNİR AĞLARI ALGORTİMALARI İLE ÖNGÖRÜLMESİ

Mustafa Reha OKUR

Doktora, İşletme

Danışmanlar: Doç. Dr. F. Dilvin Taşkın / Doç. Dr. Yasemin Zengin Karaibrahimoğlu

Dünya çapında yasal düzenlemelere rağmen (Sarbanes-Oxley Yasası, Fransa Finansal Güvenlik Yasası, Birleşik Krallık 2006 yılı Hile Yasası) finansal hileler bugünün sermaye piyasaları için hala ana sorunlardan birisidir. Bu çalışma, Amerika Birleşik Devletleri borsalarında (NYSE ve NASDAQ) çapraz listelenen firmaların hileli finansal aktivite risklerinin Sinir Ağları temelli algoritmalar kullanılarak tahminlenmesini amaçlamaktadır. Bu çalışmada, kapsamlı bir veriseti oluşturabilmek için finansal hile davaları, finansal tablo verileri, kurumsal yönetim verileri ve makroekonomik gösterge verileri toplanmıştır. Bu yöntem sayesinde bu çalışma hile tespitinde sadece firmaya özgü yönlere odaklanmak yerine ülkelere özgü kurumsal etmenleri de kapsayan oldukça geniş çaplı bir çerçeve geliştirmeye çalışmaktadır. Bu çalışma makine öğrenme temelli dört sınıflandırma algoritmasını kullanmaktadır. Rassal Orman ve C4.5 algoritmaları diğer kullanılan algoritmalardan daha iyi sonuçlar elde etmiştir. Dahası, bu çalışma literatürdeki önceki çalışmaların sınıflandırma performanslarından daha iyi sonuçlara ulaşmıştır.

Anahtar Kelimeler: *Finansal Hile, Muhasebe Hilesi, Sinir Ağları, Makine Öğrenme, Yapay Sinir Ağları, Karar Ağaçları, Tahminleme*

Acknowledgment

First and foremost, I would like to acknowledge the utmost efforts of my supervisors, Dr. Fatma Dilvin Taşkın Yeşilova from Yaşar University and Dr. Yasemin Zengin Karaibrahimoğlu from the University of Groningen. This study could not be completed without their enormous effort. They illuminated my Ph.D. journey through their wisdom and extensive knowledge. Dr. Yeşilova has allowed me full academic freedom to explore new fields and choose my direction in academia. She is a reference for me about hardworking, and scientific curiosity. Dr. Karaibrahimoğlu has enhanced this study's overall structure through her extensive scientific knowledge on the field. Her valuable contribution not only effects this study, but it also extensively shape my academic perspective and knowledge. Additionally, I would also like to thank her guidance and invaluable help during my research visit at the University of Groningen.

Additionally, I also would like to thank my thesis committee, who are Dr. A. Fatih Dalkılıç, Dr. Sevinç Güler Özçalık, Dr. Özen Ece Acar, and Dr. Mustafa Gürol Durak, for their valuable comments, contribution and support.

I want to thank several other valuable people who influence this study and my scientific development during the Ph.D. journey. First of all, I want to thank Dr. Çağrı Bulut, who is a mentor and a guiding spirit for my academic and personal life. Secondly, I would like to thank Dr. Vlad Andrei Porumb for his critical suggestions about the sample selection procedure of this thesis. Thirdly, I would like to thank Dr. İge Pırnar for her support to my research visit allowance and for her guidance. Lastly, I thank colleagues from the University of Groningen for the warm welcome and scientific support during my research visit.

Special thanks must go to my parents and sisters for their unmeasurable and unconditional support through all my life. This thesis can not be completed without their help and encouragement.

Finally, I would like to thank my beloved wife, Sinem. I always feel her support during my graduate studies. My Ph.D. journey will be more stressful and pessimistic without her optimism, cheer and confidence on me.

TEXT OF OATH

I declare and honestly confirm that my study, titled "PREDICTING FRAUDULENT FINANCIAL ACTIVITIES THROUGH NEURAL NETWORK ALGORITHMS" and presented as a Ph.D. Thesis, has been written without applying to any assistance inconsistent with scientific ethics and traditions. I declare, to the best of my knowledge and belief, that all content and ideas drawn directly or indirectly from external sources are indicated in the text and listed in the list of references.

Mustafa Reha OKUR

August 6, 2019

TABLE OF CONTENT

Thesis Approval Page

Abstract i

Acknowledgments iii

Text of Oath iv

List of Tables viii

List of Figures ix

1. Introduction 1

1.1. Overview 1

1.2. Motivations and Contributions 6

2. History of Financial Fraud and Literature Review 9

2.1. Historical Background 9

2.2. Main Theories Related to Fraud and Management Commitment 13

2.2.1. Fraud Triangle and Fraud Diamond 15

2.2.2. Agency Theory 17

2.3. Legal Regulations 19

2.3.1. OECD 19

2.3.2. Sarbanes-Oxley Act (2002) 21

2.3.3. International Financial Reporting Standards 23

2.3.4. International Standards on Auditing 24

2.3.5. PCAOB Auditing Standards 25

3. Research Methodology 27

3.1. Artificial Neural Network 28

3.1.1. Artificial Neural Networks in Accounting and Finance 30
3.1.2. Unsupervised Learning 31
3.1.3. Semi-supervised Learning 32
3.1.4. Supervised Learning 32
3.2. Utilized Algorithms 33
3.2.1. Multilayer Perceptron 33
3.2.2. Logistic Regression 33
3.2.3. Decision Trees 33
3.2.3.1. C4.5. (J48) Algorithm 34
3.2.3.2. Random Forest 35
4. Data and Sample Selection 36
4.1. Data 36
4.1.1. Data of the Financial Fraud 37
4.1.2. Data of the Financial Statement 39
4.1.3. Macroeconomic Indicators 41
4.1.4. Feature Selection 41
4.1.4.1. Gain Ratio Attribute Evaluation 42
4.2. Sample 43
4.2.1. Cross-listing 43
4.2.2. Bonding Hypothesis 45
4.2.3. American Depository Receipts (ADRs) 45
5. Results 47
5.1. Feature Selection 47
5.2. ROC Curves 48
5.3. Multilayer Perceptron 49
5.4. Logistic Regression 50
5.5. C4.5 Algorithm 52
5.6. Random Forest 53

6. Conclusion 55

References 57

Appendices 74

Appendix A. Multilayer Perceptron Sigmoid Node Weights 74

Appendix B. Odds Ratios of the Logistic Function 91

Appendix C. Employed pillars and subpillars of the Global Competitiveness Index of World Economic Forum 92

Appendix D. Visualized Decision Tree 94

List of Tables

Table 1. Triggering Events for Recent Fraud Regulations

Table 2. Top Ten Largest Settlements in the SCAC Database

Table 3. List of Financial Variables and Financial Ratios

Table 4. List of Corporate Governance and Nonfinancial Variables

Table 5. Pillars of the Global Competitiveness Index

List of Figures

Figure 1. Revenue prediction of the firms that develop Artificial Intelligence applications for enterprises, from 2016 to 2025

Figure 2. Share prices of the Bank of England, the East India Company, the South Sea Company during the South Sea Bubble

Figure 3. Single layer neural network model

"Some clandestine companies combine,
Erect new stocks to trade beyond the line:
With air and empty names beguile the town,
And raise new credits first, then dry'em down:
Divide the empty nothing into shares,
To set the town together by the ears."

Daniel Defoe, 1703

1. INTRODUCTION

1.1. Overview

Fraudulent activities of a firm are enormously investigated topic, yet no clear answer has found about this complex and chaotic phenomena. Stock markets are full of bubbles, frauds, and crises due to lack of regulations until the 20th century. Financial manipulation in organized markets started with the manipulations in the Tulip Mania period in Holland. Investors hysterically rushed and paid house equivalent prices for a single tulip bulb (Chancellor, 2000. However, the fraudulent actions of managers are not related to the triggering point of this Mania. Maximizing self – interest and excessive gain over a transaction are always attractive motivators for individuals (Wang, Malhotra, & Murnighan, 2011). Tulip Mania can be the best example to understand human's greediness and losing collective wisdom for easy earned money. These disorganized and lawless periods revealed the first signs of a strong relationship between fraud and financial markets.

The early era of stock markets was quite different from the 2000s. The first dividend based lending mechanism was among the investors and the Dutch, British, and French East Indian Trade Companies at the end of the 16th century (Petram, 2014). In that period, investors mostly focused and invested in uncertain news about company activities in distant lands. First signs of excess return from stock-based investments cause great attention to primitive stock markets. This interest caused the first financial market bubbles and stock market collapse.

Fraud was mostly seen only as a criminal activity between outlaws and investors in the infant era of stock markets. Many management perspectives, theories, and regulations developed to cope with fraud activities in the following centuries. Especially, the second half of the 20th century was a key period for the fight between fraud and regulations. From that time, expanding literature over finance and management areas turned spotlights to the backstage of fraud. Many research focused on the psychological side of managers' actions to understand their attitudes and behaviors on particular events. White Collar Crime (Sutherland, 1940), Theory of Fraud Triangle (Cressey, 1950), Agency Theory (Berle and Means, 1932; Jensen and Meckling, 1976), Stewardship Theory (Donaldson and Davis, 1991), Stakeholder Theory (Freeman, 1994), Fraud Diamond (Wolfe and Hermanson, 2004) and many other theories developed or adopted to explain managers' actions on company – related decisions. Some other studies focused on the costs and effects of financial fraud over stakeholders (Rezaee, 2005; Farber, 2005; Karpoff et al., 2008; Ugrin & Odom, 2010; Wells, 2017).

Abovementioned theories and regulations have one common aim, to prevent shareholders' (or stakeholders') financial loss that occurs due to corporate fraud activities. The endless efforts of researchers identify that there are cultural, psychological, behavioral, country-specific, judicial, and managerial reasons behind fraud activities. Nevertheless, the combined effort of countries and researchers cannot hinder the greediness of top managers.

This study aims to forecast the risk of fraudulent financial activities of cross-listed companies in US stock exchanges (NYSE, NASDAQ) by employing an Artificial Neural Network (hereafter, ANN) model. ANNs are the base of many machine learning algorithms that have the ability of prediction. This study tries to develop an ANN-based semi-supervised model to construct a proactive fraud detection tool by employing an inclusive data set including both systematic and unsystematic risk, by combining key financial ratios, corporate governance variables, and country-specific macroeconomic and institutional indicators. By this method, this study tries to develop a broader framework on fraud detection that does not focus only on firm-specific aspects, instead of covering a more comprehensive dataset, which incorporates country-specific institutional factors into consideration.

In legal perspective, managers or agents are the people who delegated to manage a company by the name of shareholders (Drucker, 2008). In most cases, the interests of shareholders and managers are not overlapping. Managers tend to value their interests above shareholders' (Jensen & Meckling, 1976). Such circumstances cause great damage to shareholders' wealth and the company. Moreover, due to highly integrated economic systems, a manager can affect the entire economy by his/her fraudulent decision. Because of integrated economic systems, countries experience agency costs with a multiplier effect.

The vast amount of studies (Beasley, 1996; Summers & Sweeney, 1998; James, 2003; Skousen & Wright; 2006; Schrand & Zechman, 2012; Donelson et al., 2017) employ conventional methods to understand fraud and their effects. Most of them focus on ex – post effects, a lot fewer focus on ex – ante events. In the modern world, newly emerging methods can handle the most complex issues. The epoch that we live in is the golden age of computers and computer – based artificial intelligence[1]. In today's world, computer – based systems can also predict the most complicated creature in the world: human. Even a tiny mobile phone can predict our daily routines and adapt itself. Moreover, %60 of stock market transactions based on the decisions that are made by machines in today's financial world[2].

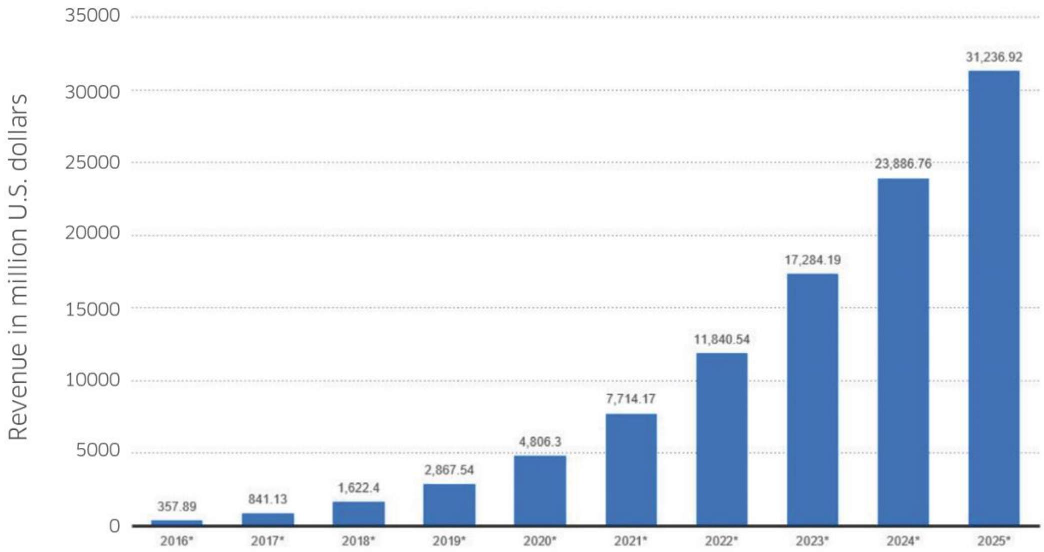


Figure 1. Revenue prediction of the firms that develop Artificial Intelligence applications for enterprises, from 2016 to 2025 (Source: Artificial Intelligence: Industry Report and Investment Case, Nasdaq (2019)).

Technological developments offered new opportunities for researchers who try to understand the occurrence of financial fraud. The benefits of forecasting financial fraud via neural network algorithms will be discussed in the following sections in detail. In a nutshell, a robust algorithm can reduce the risk exposures of investors and stakeholders due to financial fraud. Moreover, the risk of financial misstatements will be minimized due to the continuous evaluation of companies. Adoption of such an algorithm to the financial markets can be beneficial for regulatory bodies and beneficial for other stakeholders like banks, individual investors, investment funds, and companies.

Origin of artificial neural network (hereafter, ANN) based on the article of McCulloch and Pitts (1942). They developed a mathematical model, and that model

1 See OECD Report on Private Equity Investment in Artificial Intelligence (December, 2018)

2 See "Artificial Intelligence: Industry Report and Investment Case (2019)" of Nasdaq to better understand the economic consequences of the Artificial Intelligence in the finance sector.

triggers two distinctive neural network research area. One side focused on brain related research topics, and the other side focused on employing neural networks for artificial intelligence. ANNs are inspired by human brain activities and mimics the pattern classification and pattern recognition ability of it (Zhang, Patuwo, & Hu, 1998). ANNs do not need input assumptions, can also learn from previous knowledge and can generalize it (Bahrammirzaee, 2010). ANN applications in the accounting and finance area began with the article of Tam and Kiang (1990). Their main aim was to predict bank failure by using neural networks. Later on, neural networks attract many researchers in the field of accounting and finance[3]. Most of them focused on bankruptcy prediction. Fanning et al. (1995) published first fraud related research that employed neural network. Their research consists of prediction power comparison between Bell et al.'s (1993) cascaded logit model and artificial neural network. According to their results, artificial neural network outperforms cascaded logit model in accounting fraud prediction. This first bullet drew many researchers attention and many research (Green & Choi, 1997; Fanning & Cogger, 1998; Lin, Hwang, & Becker, 2003; Kirkos, Spathis, & Manolopoulos, 2007; Ngai, Hu, Wong, Chen, & Sun, 2011; Lin, Chiu, Huang, & Yen, 2015) published on this topic since that time.

Financial ratios were the key variables instead of raw financial statement data for the aforementioned researches. However, most of them employ variables different from each other. Additionally, fraud related studies in neural network and fraud related fields cover very few variables. Some other researches focused on corporate governance related variables (Chen et al., 2006; Lin et al., 2015; Chen, 2016) to forecast fraud with neural networks. In other words, there was no consensus about key fraud indicators. Neural networks gave us the freedom of variable selection to have a holistic approach. From that perspective, this study tries to combine variables from previous literature.

Financial fraud is not only related to the internal environment of a company. It affects not only the internal environment but also the external environment. Besides, companies have a strong connection with the external environment and economy. Financial trouble may trigger managers of a company to commit fraud (Kirkos, Spathis, & Manolopoulos, 2007). Prior researches mostly focused on the financial indicators of a company. Nevertheless, a company's economic condition cannot solely

[3] 59 articles published between the years of 1990 – 1995 in the field of accounting and finance that employed neural network as a method (Wong, Bodnovich, & Selvi, 1997).

be described by its financial data. Economic conditions of a country or global economic circumstances have a strong effect on a company's economic situation. Liu (2004) identified that interest rates are a critical factor on the company's overall financial health. Effects of interest rates can be easily observable through the financial statements of a company. Moreover, Birz and Lott (2011) found that GDP and unemployment rate of a country influence the company's stock returns. Prior literature on financial fraud pays insufficient attention to macroeconomic indicators. Additionally, as far as we know, there is no research published yet that combines macroeconomic indicators, financial ratios, corporate governance variables, and ANNs.

Researchers face some tricky points when they observe fraud activities. Difference between legal systems in countries is the key challenging point of fraud related research corpus (Coffee, 2005). Besides, a sharp divergence between common law and civil law decrease the generalizability of fraud related researches (Reese & Weisbach, 2002). Cultural differences, ownership structures, restrictive regulations, historical differences have a significant effect on fraud related research. Those vulnerable points make it hard to construct a reliable cross – country research by researchers. For this reason, a vast amount of research (Huijgen & Lubberink, 2005; Leuz, 2006; Chang & Sun, 2009; Berger, Li, & Wong, 2011; Hope, Kang, & Kim, 2013) in different fraud related areas focuses on cross – listed companies to eliminate this complex issue. US cross – listed companies listed on stock exchanges that established outside of US and non-US cross – listed companies that listed on selected US stock exchanges are included in this research. Reaching mutual legal ground for companies from different countries and cultures is the main reason for choosing cross – listed companies[4].

In the last decades, financial fraud attracts a great deal of attention from many researchers. Managerial theories, psychological methods, surveys, statistical models have developed to understand the back backstage of financial fraud. Additionally, many legal regulations put into practice by governments to monitor companies and financial markets. The common purpose of this joint effort is to prevent the costs of financial fraud on society. This study tries to develop a combined approach on financial fraud by employing accounting data, corporate governance data,

[4] Corporate Governance Principles of the OECD, Sarbanes-Oxley Act, International Financial Reporting Standards, International Auditing Standards will be discussed in following sections.

macroeconomic indicators and artificial neural. Furthermore, data from fraudulent and non-fraudulent companies will be used to understand fraud commitment of firms better and to train the neural network algorithm. The main purpose of this study is to forecast (predict) financial fraud beforehand by employing an artificial neural network model. I anticipate that this study will help to overcome the costly, time consuming and imprecise nature of financial fraud detection.

1.2. Motivations and Contributions of the Research

Despite worldwide regulatory efforts (e.g., Sarbanes – Oxley Act, Financial Security Law of France, Fraud Act 2006 of the United Kingdom), fraud is still a major concern of today's capital markets. Association of Certified Fraud Examiners (ACFE) reports in their review "*2018 Global Study on Occupational Fraud and Abuse*" that the yearly cost of fraud to the countries is approximately USD 4 trillion.[5] Numerous academic studies examine the reasons and consequences of managers' fraudulent actions on company-related decisions. According to theory and previous studies, there are cultural, psychological, behavioral, country-specific, judicial, and managerial reasons behind fraudulent activities. To avoid fraud, theoretical and empirical studies and regulations point out that instead of focusing on ex – post consequences, ex – ante events outside or inside the organization should be taken into consideration. Furthermore, considering worldwide technological developments, instead of applying traditional analysis, alternative data analytics technique which is also used by audit companies may advance our understanding to detect the probability of fraud before it occurs.[2]

This study aims to forecast the risk of fraudulent financial activities of cross-listed companies in US stock exchanges (NYSE, NASDAQ, CBOE) by employing an artificial neural network model. ANNs are the base of many machine learning algorithms that have the ability of prediction. Such algorithms can be beneficial for the risk management approaches of companies' with the prediction ability (Wu et al., 2014). The vast amount of studies (e.g., Fanning et al., 1995; Kirkos et al., 2007; Ngai et al., 2011; Niaki & Hoseinzade, 2013; Zhao et al., 2015) from business-related fields

[5] See https://www.acfe.com/report-to-the-nations/2018/

[2] See https://www2.deloitte.com/content/dam/Deloitte/in/Documents/finance/Forensic-Proactive-services/in-fa-frm-noexp.pdf

employed ANNs as a prediction tool for their research. However, especially in fraud forecasting literature, employed data sets are narrowly scoped and lack of being comprehensive (e.g., Kirkos et al., 2007; Zhao et al., 2015). Thus, this study tries to develop an ANN-based algorithm to construct a proactive fraud detection tool by employing an extensive data set including both systematic and unsystematic risk, by combining key financial ratios, corporate governance variables, and country-specific macroeconomic and institutional indicators. By this method, this study tries to develop a broader framework on fraud detection that does not focus only on firm-specific aspects, instead of covering a more comprehensive dataset, which incorporates country-specific institutional factors into consideration.

I believe that using such a prediction model, which incorporates macro-economic and institutional factors and institutional differences into consideration, has significant contributions and implications for academic literature and regulators. First, a more comprehensive prediction model will cater to the needs of investors on potential fraud risks in a better way. Prior researches mostly focused on the financial indicators of a company. Nevertheless, a company's economic condition cannot solely be described by its financial data. Economic conditions of a country or global economic circumstances have a substantial effect on the company's economic situation. However, macro-economic factors and institutional differences are more likely to trigger managers of a company to commit fraud (Kirkos et al., 2007).

Consequently, we argue that the omission of macroeconomic and institutional factors are more likely to cause failures in ex-ante fraud prediction. Second, I believe that findings of this study will advance previous studies and provide insightful findings into the understanding of regulators and capital market participants to provide high-quality financial numbers that help users to make more informed decisions through the signals that are produced by the algorithm. Considering the fact that, from an equity market perspective, a more precise prediction model is more likely to minimize the potential costs of fraudulent financial activities for stakeholders, mitigate the use of managers' and companies' fraudulent activities and enhance trust on capital markets via continuous fraud risk assessment of companies based on advanced machine learning.

Companies, investors or stakeholders face several risk exposures during financial activities as a nature of the investment. Currency risk, market risk, political risk, liquidity risk, default risk, the risk of material misstatement are only a few of them.

ANN algorithms can be beneficial for the risk management approaches of companies' with predictive ability (Wu et al., 2014). Our research will provide insights into the prediction of financial fraud risk of companies. A robust algorithm can reduce the risk exposures of investors and stakeholders due to financial fraud.

Moreover, continuous evaluation of companies can minimize the risk of financial misstatements. Adoption of such an algorithm to the financial markets can be beneficial for regulatory bodies and beneficial for other stakeholders like banks, individual investors, investment funds, and companies. Commercial banks are started to develop several ANN based algorithms for credit risk evaluation (Angelini et al., 2008). Audit companies can also benefit from the developed algorithm as auditor's decision aid tool. In general, auditing firms adopt a strategic systems approach or transaction focused approach to evaluate the risk of material misstatement (Schultz et al., 2010). Our research will enlarge the audit companies' evaluation procedures for the risk of material misstatement. Additionally, auditor's trust-based relationship with company managers can affect managerial fraud evaluation (Kerler & Killough, 2009). An emotionally indifferent algorithm will reduce the risk of biased fraud assessment.

Development of an artificial neural network based prediction algorithm can also be beneficial for the academic corpus. Researchers from other fields of finance and accounting can be encouraged for using different methods and big data sets.

2. HISTORY OF FINANCIAL FRAUD AND LITERATURE REVIEW

2.1. Historical Background

Commerce was the critical element of social development among different nations throughout the history. Nations started to exchange goods on the base of barter at the beginning of trade activities. Later on, long distance trade activities had developed, and precious metals were used as the instrument of payment. There are no written records found, but tricks, deceive, and rip off were always an issue of trade activities in the early trade era. Greek merchant Hegestratos, who lived 300 B.C., was the first fraudster of known history according to many historians. He had a deal with the lender to transport corn by his boat. In exchange for that, the lender gave him money to finance this operation. Hegestratos will pay his debt when the duty is fulfilled. However, he decided to intentionally sink his empty boat, sell the corn secretly and never pay back his debt. Unfortunately, the plan went wrong and he lost his life with his sunken boat (Johnstone, 1998).

Dante Alighieri (1320) reserved the deepest dungeons of hell to the individuals who act fraudulently to people who connected with love and trust in his famous long narrative poem called Divine Comedy. The main idea of this book did not cover financial fraud. Nevertheless, in essence, it punishes people who deceive others based on their fiduciary relationship.

Many societies condemned fraud or related activities culturally. Legal regulations based on religions had been developed to prevent or punish fraud on trade activities. Those cultural curses and social oppression never restrain fraudsters from criminal activities. They found new ways to delude societies, investors, and counterparties with the development of the economic environment of the world. At the beginning of commercial activities, they only targeted individuals or small groups. Effects of their fraud activities had impacts only on small environment. The impact of their fraud actions enlarged with the development of economic systems. New trade markets, developed economic systems, booming trade opportunities with new trade routes, capital gains through newly developed lending mechanisms (Tracy, 1993), the newly born wealthy upper class who had excess capital provoke fraudsters to perform their job.

In classical efficient market perspective, investors are fully informed, have rational expectations, and markets are efficient on some levels (Malkiel & Fama, 1970). Although none of the market bubbles can be explained if the market conditions are in perfect balance. A typical investor in a market mostly misprice stocks or securities (Fama, 1965). On the other hand, in most cases, investors have incomplete information. Additionally, most of the stock market “manias” and underlying fraudulent acts are triggered by irrational behaviors of investors (Kindleberger & Aliber, 2011). Whether we accept such kind of definition or the contrary, none of them can fully explain the effects of fraudulent activities.

The Netherlands’ Tulip Mania was the first market collapse that investors faced in known financial market history (Gisler & Sornette, 2010). During the great prosperity times of the Netherlands among 1585 – 1650, surprising commodity contracts has risen as a new guaranteed way of profit (Sornette, 2003). Tulip bulbs were the underlying security of such future contracts. A house equivalent price paid for a bulb until the collapse of the market in 1637 (Malkiel, 2012). In this period, probably many fraudsters appeared in the market. However, fraud-based behavior was not the triggering point of this collapse. Maximizing self – interest and excessive gain over a transaction are always attractive motivators for individuals (Wang, Malhotra, & Murnighan, 2011). Tulip Mania can be the best example to understand human’s greediness for easy earned money. An absurd commodity can charm the whole society, and this human characteristic can easily be manipulated by an expert.

Later on, the East Indian Companies era has started. The first known monopolistic company to trade commodities from the Far East has been established in Russia in 1553 (Baskin, 1988). However, the most important and well-known one was the British East India Company (hereafter, EIC) which was founded in 1600 to reduce the debt of British Empire and to reach commodities which were produced in East Indies (Chaudhuri, 1999). EIC cannot only be seen as a developed version of merchant unions or an early version of corporations. It is much more complicated and, indeed, it has much power than any other competing company in that period. Besides, EIC has a massive impact on the evolution of British economic philosophy and development (Erikson, 2014). Furthermore, EIC was also the dominant player in trade activities and had an enormous influence on the trade income of Britain for nearly 200 years (Ward, 1994; Broadberry & Gupta, 2009). EIC is not only crucial for the economic history of Britain; it is also a symbol for the rise of shareholder capitalism (Lawson, 1993). On the other hand, EIC's agents frequently trade for their accounts. The company's financial structure highly effected by agency problems and unhealthy principal-agent relationship results in financial failure (Hejeebu, 2005). Moreover, EIC is also known as the first publicly traded companies that committed financial fraud (Dorminey et al., 2012). Adam Smith's (1776) modern corporation criticism highly influenced by the fraudulent activities of the EIC and shareholder wealth decrease because of those fraudulent activities.

The potential and power of EIC forced Dutch states to reconstruct their marine trade power. They paired up their naval trade merchants under one joint-stock company that called as Dutch United East India Company (Verenigde Oostindische Compagnie – hereafter, VOC) in 1602. VOC was a very well established company as a result of well-developed Dutch stock markets. It was also the first established publicly traded joint-stock company and has ten times equity then EIC (Robins, 2017). Moreover, VOC outperformed EIC in terms of the voyage numbers until the 1780s. By the years of the 1780s, EIC's voyage number sprang to 318 and outnumbered VOC's 297 voyages (Erikson, 2014). Dutch economic history highly influenced by VOC's sea trade activities. Potential gain through capital investments attract investors and improve Amsterdam capital markets. This private finance market even became more favorable and stronger than public finance options after the year of 1609 by the significant influence of the VOC (Gelderblom & Jonker, 2004). Such a developed structure of Dutch capital markets and highly profitable companies induce managers for fraudulent activities. The powerful and profitable status of the VOC makes its managers self-centered, greedy and they act like a tyrant. This improper and fraudulent activities results first generally accepted shareholder lawsuit by the letter of the investor Isaac Le Maire on 1609 (Koppell, 2011).

Harsh competitive environment among those rival companies and nations was presumably the biggest commercial competition until that time. Those companies' fraudulent activities involve managerial and financial fraud together. However, especially in EIC case, fraudsters also triggered because of those two great companies (Chancellor, 2000). Capital gains of investors who invested voyages of EIC and VOC, and the booming capital markets charmed the behindhand investors. They started a hunt to invest new companies and fraudsters were there with their paper companies.

A while later (nearly a century), fraudsters were rubbing their hands in glee on the other side of the world. Fierce competition in the East Indies, high public debt and capital need forced the British Empire to find new trade routes and colonies. For this reason, the South Sea Company was founded in 1711 with the trading rights from South America. On the beginning of 1720, the stock price of the South Sea Company was 130 pound (Kleer, 2012). The stock price was multiplied seven times in just a couple of months. Additionally, the market capitalization of the company reached 164 million pounds and that was five times higher than the tangible assets of the company. That situation was the biggest madness for a stock until that time and even the greatest mind of that century had deceived. Sir Isaac Newton had lost a significant amount of money during the South Sea Bubble (Temin & Voth, 2004). Moreover, the company has no significant trading activities to generate income (Dale et al., 2005). During this bubble, first insider trading activities occurred and some parliament members and investors gain unearned income (Hoppit, 2002).

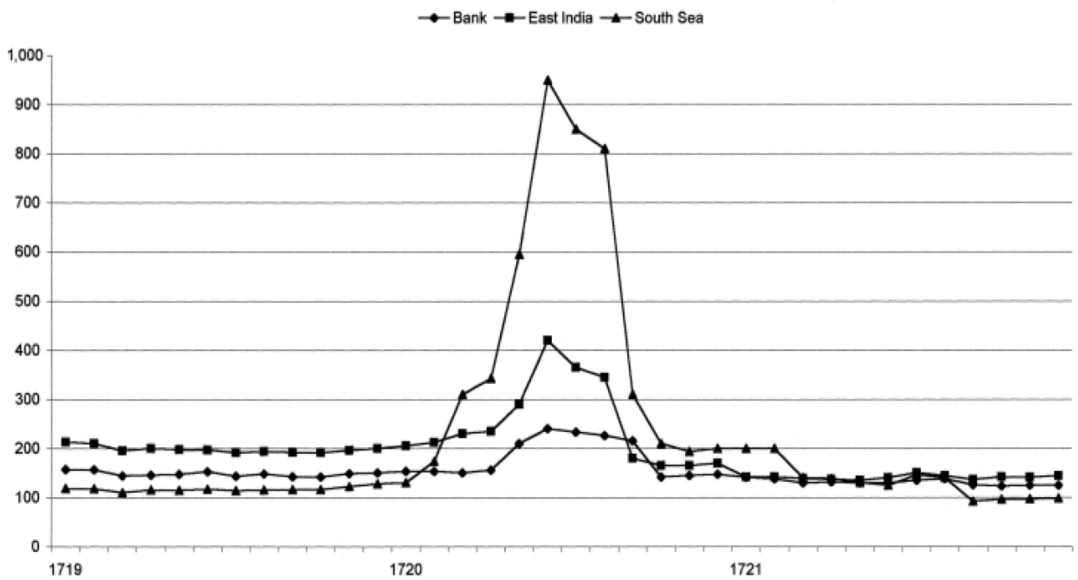


Figure 2. Share prices of the Bank of England, the East India Company, the South Sea Company during the South Sea Bubble (Source: The Myths of the South Sea Bubble, Hoppit (2002))

Company	Year	Description
Enron Corporation	2001	Misappropriate use of special-purpose entities.
Worldcom	2002	Accounting fraud through improper expenses.
Tyco	2002	Issues with merger related accounting practices.

Fannie Mae	2004	Excessive executive payments through fraudulent accounting activities.
AIG	2005	Poor corporate governance and internal control practices
Subprime Mortgage Loans	2007	Mortgage backed toxic securities.

Table 1. Triggering Events for Recent Fraud Regulations (Source: Giroux, 2008)

Newspapers or news agencies play a critical role and become watchdog against accounting fraud in such circumstances (Miller, 2006). Market regulators and employees of the company also play a vital role to determine fraudulent financial activity in addition to journalists (Dyck et al., 2010).

2.2. Main Theories Related to Fraud and Management Commitment

Fraud is a highly debated topic from the beginning of the first trade activities. Modern fraud literature begins with the influential work, “White-Collar Criminality”, of Edwin H. Sutherland in 1940. It is a milestone in the fraud literature because, starting with his study, criminologists have started to acknowledge that the criminal activities are not only associated with the actions of immigrants or poor people but also with the actions of rich and powerful people (Coleman, 1987). In short, Sutherland’s (1940) interdisciplinary article combines the perspectives of economists and criminologists to identify business related criminal activities. Researchers can also use the combination of terms fraud, white-collar crime and financial crime (Pickett & Pickett, 2002).

Several theories, Theory of Fraud Triangle (Cressey, 1950), Agency Theory (Berle and Means, 1932; Jensen and Meckling, 1976), Stewardship Theory (Donaldson and Davis, 1991), Fraud Diamond (Wolfe and Hermanson, 2004), have had a significant impact on financial fraud literature. Those theories mostly focus on the relationship between financial fraud and managerial commitment. They also influence and shape fraud literature throughout the 20th century.

First, fraudulent financial activities should be separated from some other technical terms in accounting and finance literature. Creative accounting (Breton & Taffler, 1995; Gowthorpe & Amat, 2005; Jones, 2011), earnings management (Dechow et al., 1995; Leuz et al., 2003; Cohen et al., 2008; Chen et al., 2015), income smoothing (Tucker & Zarowin, 2006; Grant et al., 2009; Acharya & Lambrecht, 2015) terms can be erroneously used alone in some cases. Above all, financial fraud or accounting fraud are separate from all of those. Furthermore, there is even a distinction between the terms of fraud and error. The intentional act of material misstatement is the critical point that separates financial statement fraud from error. The American Institute of Certified Public Accountants (AICPA) defines financial fraud in SAS No. 99 as the intentional activities that result in materially misstated financial statements. However, the definition of fraud or using the right term for misstatements is not clear for each case. Heated discussion ongoing to define financial statement fraud correctly. Financial misreporting, financial misrepresentation, financial fraud or financial misconduct can bear the same meaning (Amiram et al., 2018). ISA 240 (IAASB, 2018) defines fraud as the intentional act of one or more professionals that aim to deceive the shareholders or group of stakeholders to gain an advantage. Country specific definitions of fraud might be different from each other, although nearly all of the fraud definitions focused on the violation of laws and regulations (Jones, 2011).

Several perspectives in prior financial fraud literature highlight that the analysts' forecasts and market expectations put significant pressure on managers shoulder to meet expectations of them. That kind of situation force managers to employ some earning management methods and even incentivize fraudulent financial activities (Burgstahler & Eames, 2006). Perols and Lougee (2011) also supports this idea and propose that fraud firms tend to manage earnings before the fraud year. They also assert that the analysts' expectations are also the primary reason behind those activities. Wells (2017) explains the actions behind the managerial commitment to the overstatement of financial statements;

- To financially meet the market expectations

- To raise the potential financing options

- To meet parent company's performance criteria

- To meet personal goals and performance

- To support company backed securities and stock price for potential stock exchange for M&As.

Gottschalk (2010) defines fraud as an intentional act of deceiving some others for seizing their belongings or legal rights. Rijsenbilt and Commandeur (2013) define managerial fraud as the intentional financial misstatement activities of managers that mislead the shareholders and related parties. We cannot analyze most of the financial crimes without a fraud perspective. Therefore, fraud, financial crime and financial fraud or accounting fraud bear the same meaning in this and the following sections. This study does not strictly categorize fraud related studies on a terminological basis. The categorization of this study mostly based on the context and the essence of those studies.

Beasley (1996) limits the definition of financial statement fraud in two different categories for his research. The first category covers the managers' intentional financial statement misrepresentations to the stakeholders. The second category involves senior managements' intentional exploitation of company assets for their self-interest. The common ground of both categories is the intentional act of top management.

Rezaee (2005) defines financial statement fraud as the intentional act of firms to manipulate creditors and potential investors by producing misreported financial statements.

Jones (2011) defines fraud as fabricated accounting transactions that are contrary to broadly accepted accounting principles and penalized by related courts or enforcement bodies. Murphy and Dacin (2011) define fraud as dishonest or illegal actions that are intentionally committed by the employee.

We can also define financial statement fraud as managers' (corporations') intentional manipulation of financial statements to deceive market participants for improper benefit by taking advantage of information asymmetry.

2.2.1. Fraud Triangle and Fraud Diamond

Modern fraud literature mostly based on the research of Sutherland (1940) who tried to explain corporate managers' fraudulent actions against stockholders. He developed the theory of "White-Collar Crime" and derived the term to describe the illegal actions of companies and managers (Choo & Tan, 2007). Following Sutherland's work, in 1953, Daniel Cressey, one of his students and a well-known criminologist, developed several hypotheses to understand what triggers people to commit financial fraud. He conducted interviews with 250 prisoners who accused violation of financial trust. The findings of his study document that among all other factors, (1) perceived pressure, (2) perceived opportunity and (3) rationalization are the key motivators of fraudsters which are lately named as the elements of Fraud Triangle. Although, Cressey's work has been criticized by many aspects such as the ignorance of major white-collar crimes – collective fraud and tax evasion, highlighted in Sutherland's (1940) research-, the sample selection procedures, and the lacking angles of Cressey's theory research (Trompeter et al. 2013; Morales, Gendron, & Guénin-Paracini, 2014). The background of this theory developed by Cressey, however, the name of the theory was given by other researchers. The roots of Fraud Triangle Theory (hereafter, FTT) grounded on the findings of Cressey (1953).

After Cressey's research, many other researchers have focused on this area to understand the patterns of fraud. The FTT argued that managers would commit fraud if there is incentive/pressure to commit fraud, weak control mechanisms within the company (which lowers possibility of being caught) and perpetrator can legitimize his/her fraudulent actions (Mui & Mailley, 2015). FTT not only focuses on the behavioral or managerial side of fraud, but also it connects the links between accounting, risk management, auditing, and organizational deviance (Morales, Gendron, & Guénin-Paracini, 2014).

Another bunch of studies commented heavily on the lacking angles of Cressey's theory. FTT is lack of culture related perspective and ignores fraudsters' capability about the profession (Rubasundram, 2015). Wolfe and Hermanson (2004) argue that the position of a manager in the organization, competences and psychological attributes has an interlinked connection about the perpetuator's ability to identify potential fraud and realize it. Furthermore, Dellaportas (2013) adds that, fraud appears when the man in charge is the right person with suitable capabilities and Donegan and Ganon (2008) support abovementioned views and argued that there is no empirical basis to implement Cressey's theory as explanatory model for fraud in American Institute of Certified Public Accountants' (AICPA) SAS No. 99. Therefore, following the critiques on FTT, Wolfe and Hermanson (2004) developed the Fraud Diamond (hereafter, FD) to extend the scope of the FTT. While FTT argues that the fraudster has three thought steps before committing fraud; incentive, opportunity, and rationalization. FD considers the idea that a fraudster should also have the ability to recognize potential fraud opportunity and realize the fraudulent activity, which is the fourth angle and named as Capability. This additional angle is valuable because, without the necessary abilities, a fraudster cannot realize the incentivized and rationalized fraud opportunity (Kapp & Heslop, 2011). Additionally, capabilities angle not only covers the ability to do the job but also covers the position within the organization, intelligence, self-confidence/ego, pressure, effective lying and resistance to stress (Wolfe & Hermanson, 2004). Those characteristics play an important role when fraudulent activity consists of large sums and continue over the long run (Dorminey et al., 2012). This additional perspective directly affects the fraud decision procedures of the FTT. Boyle, DeZoort and Hermanson (2015) investigated 89 auditors' fraud decision aid type. They found that FD practice aid results in more conservative fraud risk assessments than FTT.

2.2.2. Agency Theory

Modern management approaches (Berle & Means, 1932; Fama and Jensen, 1983; Claessens et al., 2000) bring forward that the shareholders' role and the managers' role should be separated from each other. Some others (Demsetz & Lehn, 1985; La Porta et al., 1999) raised opposing views against the idea of separation of the ownership. However, in most cases, shareholders hire talented professionals to manage the daily activities of their company. Shareholders (principals) delegate their managerial duties to the professional managers with this employment procedure. This employment procedure brings out agency problems (Jensen & Meckling, 1976). Spence and Zeckhauser (1971) mention agency related issues from the individual perspective that are limited monitoring capability of companies and the utility function maximization of the individuals. Alchian and Demsetz (1972) mention contractual issues and monitoring cost of individuals within the organization. Centuries long accumulated knowledge about the principal-agent relationship leads perspectives to solid principal-agent theory. Nevertheless, Stephen A. Ross (1973) had proposed the first integrated perspective about the agency theory and followed by the research of Barry M. Mitnick (1975). Besides, Jensen and Meckling (1976) develop a perspective on the agency theory that explains the complicated relationship among the shareholders, managers and third-party stakeholders. Additionally, agency theory mainly focuses on constructing the most effective contract to overcome the conflicts and manage the relationship between shareholders and managers (Eisenhardt, 1989). The vast amount of empirical studies (Ross, 1973; Fama, 1980; Fama & Jensen, 1983; Hill & Jones, 1992; Berger & di Patti, 2006; Hypko et al., 2010; Pepper & Gore, 2015) focus on the principal-agent relationship to improve the managerial performance of organizations.

The underlying reason of agency problem is that the parties of this relationship seek their self interest in most cases. Adam Smith (1776) highlighted this problem centuries ago in his famous book. He argues that the managers of a company never treat shareholders' money as their own and it should not be expected. Of course, there are several other reasons for agency problems. Duration of principal-agent relationship, the organizational structure of the company, industry specific features and organizational climate (Shapiro, 2005) can be the reasons behind the agency problems.

In some cases, agents try to manipulate inputs of financial statements to maximize their interests and this circumstance results failed firm value maximization (Berger & di Patti, 2006). Cost of those and similar conflict of interests among parties called as agency costs. Agency costs can be calculated as the sum of the amount of principals' spending to monitor the agents, the bonding expenses of the agents, and loss of principals' income as a result of the conflict of interest between related parties (Jensen & Meckling, 1976). Jensen and Meckling (1976) also assumed zero agency cost situation within the companies that entirely owned and managed by a single person. However, it is only possible for nonpublic companies and not practical or possible for publicly traded companies. Effects of ownership structures on agency costs are highly investigated topic (Pagano & Röell, 1998; Fleming et al., 2005; McKnight & Weir; 2009; Rashid, 2016). We are currently in a stock market system that the publicly traded companies' ownership structures mostly consist of institutional investors. Such an environment results in a phenomenon called "the agency costs of agency capitalism" (Gilson & Gordon, 2013). Various researches focus on measuring the amount of agency cost (Demsetz & Lehn, 1985; Zhou, 2001; Huang et al., 2011; Songini & Gnan, 2015).

Agency theory also focuses on compensation plans of managers within the organization. Behavioral side of the agency theory tries to explain the executive's behavior against risky situations of compensation plans (Shi et al., 2017). Fixed compensation plans and performance based compensation plans are the primary distinctions of agency theory for the compensation contracts (Christen et al., 2006). In recent years, performance based compensation contracts are more popular than fixed compensation contracts. However, there is a debate about the compensation contracts effects on triggering fraud related managerial activities (Crutchley et al., 2007; Crocker & Slemrod, 2007; Conyon & He, 2016). Coffee (2005) highlighted that there is a distinction between the US related fraud cases and Europe related fraud cases based on compensation contracts. He proposed that a higher option based contract environment in the US results in more fraud cases than Europe's less equity based contracts. Bruner et al. (2008) identified that the managers' fraud related activities are positively correlated with the amount of performance related equity. Efendi et al. (2007) find that CEOs tend to misreport the financial statements when they have a large amount of in the money stock options. Thus, financial fraud can be related to the compensation contracts which are resultant of the principal-agent relationship.

Corporate governance theory and practices are profoundly affected by agency theory (Lan & Heracleous, 2010). This impact is highly influential especially in the infant era of the corporate governance practices (Shleifer & Vishny, 1997; Dalton et al., 1998). Protecting shareholder rights is one of the primary objectives of corporate governance practices. La Porta et al. (2000) broadly define corporate governance as regulations which protect investments of outside stakeholders from who has access to insider information. In other words, corporate governance regulations mostly care about to guarantee the return of investment of the shareholders who invest in companies (Shleifer & Vishny, 1997).

2.3. Legal Regulations

2.3.1. OECD Corporate Governance Principles

Corporate Governance regulations had arisen in accordance with the need for regulations that can adopt country to country. In a broad perspective, corporate governance practices are the fullest extent of regulations that balance the relationship between a company and the society. OECD was the key institution, which had published the OECD Corporate Governance Principles in 1999 and revised it in 2002 and 2004. Before that time, several countries have designed their corporate governance regulations. Adaptability power was the essence of OECD principles. Most of the institutional regulations about corporate governance are based on four pillars; fairness, transparency, accountability, and responsibility.

Corporate governance is a key framework to understand fraud related regulations. Corporate governance regulations firstly developed in the United Kingdom with the reports of Cadbury (1992), Greenbury (1995), and Hampel (1998). However, the most important and inclusive one was published by the Organisation for Economic Cooperation and Development (OECD) in 1999 as a recommendation. Later on, those principles are revised in 2002, 2004, and 2015. The revision in 2015 was different from others because it was published under the mutual authority of G20 and the OECD. The reason behind the revisions is to meet the new requirements because of worldwide corporate scandals (Jesover & Kirkpatrick, 2005). OECD is an organization that aims to promote and improve economic conditions around the world. From that point of view, OECD corporate governance principles is a guide that can be adapted for each country's particular economic conditions (OECD, 2012).

OECD corporate governance principles focused on following main areas; *the rights of shareholders and equal treatment to them, stock markets and intermediaries, the role of stakeholders, disclosure and transparency, and the responsibilities of the boards* (OECD, 2015). All of those subjects have direct or indirect effects on the management of companies. Applying corporate governance principles do not guarantee the efficient management of a company, but, it will contribute it and shareholders can be protected from managerial malice.

Sound corporate governance system is vital to establish a sustainable financial market that has potential to growth (Claessens, 2006). Numerous researches find a positive relationship between the application of corporate governance practices and the firm performance (Brown & Caylor, 2004; Brown & Caylor, 2009; Agrawal & Knoeber, 2012), and market valuation (Bauer et al., 2004; Beiner et al., 2006; Cheung et al., 2010). Several other research finds contrary or no relationship between corporate governance practices and firm performance (Arora & Sharma, 2016), and market valuation (Peni & Vahamaa, 2012).

Rightminded nature of OECD corporate governance principles did not result positively in every case. OECD principles especially criticized when implementing those principles in underdeveloped or emerging countries (Chen et al., 2011; Peters et al., 2011; Siems & Alvarez-Macotela, 2014). The weak legal system and powerless institutions of underdeveloped or emerging economies lead to an unstable environment for OECD principles (Klapper & Love, 2002).

Nature of corporate governance activities can be associated with fraud because of the relationship between managerial activities and corporate governance. Shi et al. (2017) claimed that the external corporate governance regulations force managers to act fair and truthful. Additionally, corporate governance practices regulate the role of independent directors in the board of directors and CEO duality (separation of the CEO and the chairperson of the board of directors) to avoid uncontrolled decision-making process (Sharma, 2004). Without those regulations, uncontrolled decision making procedure will encourage fraud related activities of managers. Corporate governance practices also regulate the organizational structure of companies (Dalton & Dalton, 2011; Carcello et al., 2011). Chen et al. (2006) highlight that the number of outside directors, CEO tenure, and the total number of board meetings is also linked with fraud related activities. Farber (2005) investigated 87 firms that participated in fraudulent activities through manipulating financial statements. He finds that the companies that fraudulently misreport financial statements tend to have poor governance the year before fraud detection. Agrawal and Anup (2005) find that financial reporting restatements are lower in the companies that have experienced outside directors.

2.3.2. Sarbanes-Oxley Act (2002)

Sarbanes-Oxley Act was prepared to overcome the company related fraud and accounting cases and enacted on July 30, 2002. The main idea of the Sarbanes-Oxley Act is to protect shareholders and overcome conflicts among shareholders and companies by improving the precision and the correctness of companies' announcements (Li et al., 2008). Officially, the Corporate and Auditing Accountability, Responsibility, and Transparency Act of 2002 is the name of the Sarbanes-Oxley Act. Later on, it was titled as the Sarbanes-Oxley Act after the U.S. Senator Paul Sarbanes and the U.S. Senator Michael Garver Oxley. Banking Committee of U.S. Senate participated nine days of trial to hear from former SEC employees, former SEC chairpersons, five major accounting profession representatives, and accounting professionals, academics, attorneys and investors before the enactment of Sarbanes-Oxley Act.

Enron scandal was the most illicit accounting fraud case in recent economic history and it was also the primary motivator behind the Sarbanes-Oxley Act. Enron Corporation was one of the fastest-growing energy giants in the United States at the end of the 1990s. Enron has had nearly 30,000 employees and titled as the most innovative company six years in a row by Fortune, the influential magazine, in the

United States[6]. Additionally, Enron published the income of $101 billion in 2000[7]. The stock price of Enron Corporation reached the maximum level of $90.75 per share on 23 August 2000. In the following couple of months, the Enron's brand value and share price collapsed and its value plummeted to less than $1. Positive public opinion in large corporations was at its lowest level (%20) in 2002 with respect to the previous decade (Romano, 2005). The reason behind Enron's failure was that the top executives of Enron Corporation set their self-interests before shareholders' rights and benefit. They hide billions of dollars worth debt by manipulating financial reports skillfully and also by establishing hundreds of special purpose entities for fraudulent transactions. Additionally, they also cooperated with Arthur Andersen, the infamous audit company, to hide their fraudulent transactions (Linthicum et al., 2010). Thirty-four former employees of Enron were penalized to pay $163 million to the victims of fraud activities after the detection of the fraud scheme (Sun & Zhang, 2006). Besides these, SEC had canceled the auditing license of the Arthur Andersen and closed it. As a result, Enron Corporation scandal had a huge impact on fraud related legislation in the United States.

The Sarbanes-Oxley Act has several positive impacts on the legislative environment of the US economy. New enforcement extensively affects the board structures of companies. Corporate boards become much more independent after the Sarbanes-Oxley Act (Linck et al., 2009). However, this perspective is still contentious among scholars. Dah et al. (2014) proposed that a considerable amount of companies reduced their independent director number to fulfill the %50 requirement of the new legislation. Adoption of the Sarbanes-Oxley Act lowered the fraudulent financial activities (Patterson & Smith, 2007). Additionally, the adoption of the Sarbanes-Oxley Act also reduces the risk-taking level of the listed companies (Bargeron et al., 2010). Companies that have agency related problems lobbied heavily against the implementation of the Sarbanes-Oxley Act. Nevertheless, the Act had come into force and decreased the agency costs of lobbying and non-lobbying companies (Hochberg et al., 2009). The Act also has positive and significant effects on liquidity by improving the quality of financial reports, and market related factors (Jain et al., 2008).

On the other hand, the Sarbanes-Oxley Act highly criticized in the early application period. Economic consequences of the implementation of the Sarbanes-Oxley Act is

---

[6] Between the years 1995-2000.

[7] See Enron Annual Report of 2000 (http://picker.uchicago.edu/Enron/EnronAnnualReport2000.pdf).

highly debated topic in the literature (Zhang, 2007; Leuz, 2007; Linck et al., 2009; Gao et al., 2009). The annual cost of applying the Sarbanes-Oxley Act for smaller firms vary from $6 million to $39 million for larger corporations (Ahmed et al., 2010). That cost is not only related to the stock price reactions but also related to the total assets and the cash flows. On the other hand, the Sarbanes-Oxley Act has had adverse effects on the firm values that are listed on worldwide markets (Bianconi et al., 2013). The audit fees also increased after the implementation of the Sarbanes-Oxley Act (Griffin & Lont, 2007; Asthana et al., 2009).

2.3.3. International Financial Reporting Standards

International Financial Reporting Standards (hereafter, IFRS) are proposed to set top-notch reporting standards for companies. IFRS standards aim to construct transparent, accountable and efficient financial markets (IFRS, 2019). Conceptual framework of IFRS firstly published in 1989 and updated in 2010 and 2018. In June 2003, IFRS 1 issued to regulate the first time adoption of companies. It highlights the procedures that a company should follow when adopting GAAP based financial statements to IFRS ones (Deloitte, 2018). Following two years are voluntary adoption periods. In 2005, a huge milestone reached and IFRS became mandatory first time. The regulation of that enormous auditing adoption passed from the European Parliament in 2002. This adoption process was one of the most significant reporting change in recent history (Armstrong et al., 2010). Countries have two options to adopt IFRS; voluntary or mandatory adoption. Scholars hugely investigate the effects, impacts, and results of voluntary or mandatory adoption in early stages of IFRS (Soderstrom & Sun, 2007; Daske et al., 2008; Byard et al., 2011; Christensen et al., 2013; DeFond et al., 2015). In the early adoption period, voluntary adoption of IFRS seen as an improvement of the company's accounting quality because of the principal based nature of the IFRS (Carmona & Trombetta, 2008). On the other hand, disparities among local GAAPs and IFRS are also attracted scholars attention (Tendeloo & Vanstraelen, 2005; Jeanjean & Stolowy, 2008; Horton et al., 2013; DeFond et al., 2015). Overall, IFRS has a massive impact on the economic environments of countries. These impacts also have cross-national effects that are specific to cross-listed companies. Those companies are listed on the stock exchanges of IFRS adopted countries and the exchanges of local GAAP applying countries at the same time.

The most significant disagreement about the settlement of generally accepted international reporting standards is between the U.S. GAAP and the IFRS. The divergence of U.S. GAAP and IFRS are based on the essence of the regulations. U.S. GAAP is classified as rules based; however, IFRS are classified as principals based accounting regimes (Karim & Jamal, 2010; Agoglia et al., 2011). This divergence among the perspectives of accounting standards causes huge differences in financial statements and reporting practices of firms. Especially, foreign-based companies that are planning to list on U.S. stock exchanges or already cross listed on U.S. stock exchanges face dramatic financial statement volatilities when they are adopting their financial statements according to U.S. GAAP (Bradshaw & Miller, 2008; Sun et al., 2011; Burnett et al., 2015). There are some efforts to overcome those issues however, no consistent solution has found yet.

The coverage area of the IFRS is increasing day by day. Principal based nature of the IFRS makes it easy to adapt according to different economies. For this reason, if the political, economic and cultural obstacles between U.S. and European countries tackled, IFRS will become the basis of the ultimate financial reporting standards of the world economy.

2.3.4. International Standards on Auditing

International Standards on Auditing (hereafter, ISAs) are published by the International Auditing and Assurance Standards Board (hereafter, IAASB) of the International Federation of Accountants. Those published standards comprise 36 single standards. Each ISA addresses the introduction and purpose of the standard, definitions and requirements of the related terms and mentions the application procedures (IFAC, 2019). Additionally, one standard of International Standard on Quality Control (ISQC) is published. This section focus on financial fraud related standards. "ISA 240: The Auditor's Responsibilities Relating to Fraud in an Audit of Financial Statements" and "ISA 315: Identifying and Assessing the Risks of Material Misstatement through Understanding the Entity and Its Environment" covered in this section. Additionally, this section based on the 2018 edition of the "Handbook of International Quality Control, Auditing, Review, Other Assurance, and Related Services Pronouncements (IAASB, 2018)".

ISA 240 regulates the auditor's liabilities concerning fraudulent financial activities. ISA 240 splits fraudulent financial activities or misstatements into two categories; intentional misstatements count as fraudulent activity, and unintentional misstatements count as errors (IAASB, 2018). This distinction results in great differences before legislative bodies and laws. Additionally, ISA 240 discusses and highlights the importance of professional skepticism in auditing and lays a burden on auditors (Quadackers et al., 2014). According to ISA 240, auditors should assess financial statements to highlight the potential material misstatements and test them for fraud.

ISA 315 regulates the auditor's responsibility on recognizing materially misstated financial statements through understanding internal control practices and economic environment of the company. Additionally, the auditor is responsible for the assessment of the firm's risk evaluation procedures. Moreover, auditors should recognize material misstatements on financial statements, account balances, transactions and disclosures level (IAASB, 2018).

2.3.5. PCAOB Auditing Standards:

Compared to the majority of countries following ISAs set by IAASB, in the U.S., Public Company Accounting Oversight Board (hereafter, PCAOB) oversees the auditing procedures of public companies and SEC registered financial markets brokers. It established after the enactment of Sarbanes-Oxley Act in 2002. In other words, PCAOB authorized by the Congress of the United States. PCAOB promotes superior auditing and directs the auditing practices of professionals and accurately preparation of independent reports of the public companies (PCAOB, Standards, 2019) and PCAOB publishes. Among all ASs[8], from a fraudulent financial reporting perspective, this section covers "AS 2110: Identifying and Assessing Risks of Material Misstatement" and "AS 2401: Consideration of Fraud in a Financial Statement Audit" because of the significant relation and importance for the topic.

AS 2110, which is effective for fiscal year-ending on and after December 15, 2010, aims to identify material misstatements from the gathered information and define the auditors' responsibilities on assessments of such cases (PCAOB, 2010). AS 2110 links material misstatements to fraud triangle (Albrecht et al., 2018). This perspective based on the idea that a fraudster should have the necessary capabilities to operationalize the

[8] The American Institute of Certified Public Accountants (AICPA) is responsible for the national accounting standard setting and rulemaking of the United States and represents the certified public accountants profession. The Auditing Standard Board committee of the AICPA is responsible for the Statements on Accounting Standards (SASs) in the United States. Those standards are specific to the United States and guide accounting professionals to the auditing of nonpublic companies. In general, ISAs and SASs are similar to each other based on texting style, scope and intention (Trotman et al., 2009).

fraudulent activity. For this reason, PCAOB highlights that auditors' should also be skeptical throughout the potential misstatement cases (Nolder & Kadous, 2018).

Furthermore, AS 2401, effective since April 25, 2003, highlights the responsibility of the auditors on financial statements' error and fraud. It also covers the detailed description and features of fraud and stresses the importance of professional skepticism of the auditor. Moreover, it also states how to response fraudulent activities and how to verbalize them to the management and its audit committees (PCAOB, 2010).

Overall, IAASB's standard and PCAOB's standard are the counterparts of each other. The main difference is that the PCAOB standards are custom tailored for U.S. specific conditions. However, as a result, both target to improve the quality of auditing standards and set robust audit practices.

3. RESEARCH METHODOLOGY

Vast amount of studies (Fama, 1965; Altman, 1968; Fama, 1992; Dimitras et al., 1996; Chava & Jarrow, 2004; Agarwal & Taffler, 2008; Demyanyk & Hasan, 2010; Smith, 2012) in the finance and accounting literature try to classify and predict several issues in order to understand the financial markets, accounting related issues and the surrounding financial climate of the organizations. Abovementioned studies and many others employed traditional methods to explain the outlying phenomenon of their study. Most of them employed the appropriate method for their research and performed very well. A few other studies underperformed because of employing the wrong method. Nevertheless, the main point is that thousands of studies already employed traditional methods to explain finance and accounting related issues.

This study employs several machine learning based methods that are not common in the scientific background of the scholars in the finance and accounting field. One of the reasons for employing such methods is to draw attention to the paradigm shift (Kuhn, 1962) in scientific progress. Recently, machine learning based methods are a crucial propulsive force behind the scientific advancements in different disciplines. Additionally, the development of an artificial neural network based prediction algorithm can also be beneficial for the finance and accounting related academic corpus. Researchers from other fields of finance and accounting can be encouraged for using different methods and big data sets.

We can resemble the fraud literature to the root system of a tree. Fraud literature ramifies several disciplines to understand the background of the fraudulent activities of managers. Cultural, psychological, managerial, historical and legal sides of fraudulent financial activities are highlighted in the previous sections of this study. Motives behind model selection procedures and the methodology of this study will be discussed in the following subsections. To avoid potential misrepresentations and for the sake of clarity, this study is not going to discuss the mathematical proofing of the machine learning based methods. This study only covers the machine learning methods that overlap with the field of interest of this study. Additionally, image recognition and image classification area of the machine learning literature is ignored due to irrelevance.

Predicting human nature is more laborious than predicting organizations because of the unique nature of individuals. The characteristics of a fraudster can be different in each fraud cases (Crain et al., 2015). In line with this perspective, this study focuses on the organizational aspects of fraudulent activities instead of the fraud cases of individuals.

3.1. Artificial Neural Network

Origin of Artificial Neural Network (hereafter, ANN) research based on the article of McCulloch and Pitts (1943). They developed a mathematical model that triggered two distinctive neural network research area. One side focuses on brain-related research topics, and the other side focuses on employing neural networks for artificial intelligence. Besides the breakthrough in the field, the early version of neural networks could not learn. However, enormous developments have been occurring in the neural network corpus since the early adaptors of neural networks (Schmidhuber, 2015).

The epoch that we live in is the golden age of computers and computer-based artificial intelligence. ANN is a statistical model that comprises simple units to process the data or the information and can generalize the existing situation to future events (Gencay et al., 2002). ANN is a model that inspires the human brain’s neuron interaction (Chen, 2016). It imitates the human brain’s processing steps to perceive a large number of inputs and produce outputs based on those interactions. This allows us to construct complex models and to identify prior unanswered questions. The power of ANNs comes from analogous processing of the data and ANNs do not need predetermined assumptions for model construction (Khashei & Bijari, 2010).

We can explain the construction of an ANN model in six phases:

1- Defining and presenting the data to the ANN model as patterns of input variables.
2- The dataset should split into two groups as training or test set. The training set is for the learning procedures of the ANN model, and test set is for the validation of the predictive ability of the model.
3- The number of hidden layers and the neurons of them are decided to determine the architecture of the ANNs.
4- ANN parameters are decided before starting the training procedure.
5- The algorithm is trained by employing input data to predict the output variable.
6- The last step is the evaluation of the predictive ability of the ANN model. This procedure can be repeatable to increase the performance of the model.

A simple neural network model composed of three layers[9]. The input layer consists of explanatory (independent) variables. The second layer named as the hidden layer and hidden layer cannot be directly detected. Hidden layer emerged through the multiplication of inputs and the connection strengths. This procedure produces hidden units (logistic functions). The linear combinations of inputs and connection strengths are produced and converted into a value between 0 – 1 through activation functions. In the end, those values are multiplied by the weights to generate the output (layer). In this simple model, information flows only in one direction, from inputs to outputs. Additionally, there is only one hidden layer in the model. Such kind of simple ANN model called a single hidden layer feedforward network model as a result of model characteristics. Mathematical explanations of the neural network models are well explained in the literature (White, 1992; Franses & van Dijk, 2000).

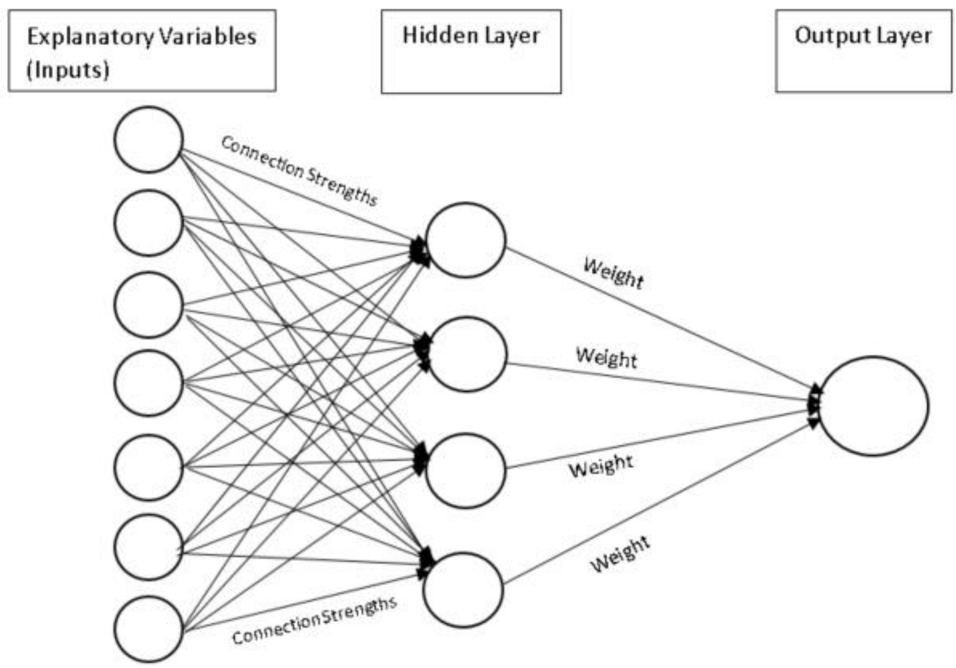


Figure 3. Single layer neural network model.

ANNs are also superior with several capabilities. Prior analyses (Ngai et al., 2011; Ticknor, 2013) revealed that ANNs detection performance is superior with large datasets. Such a characteristic allows clustering large and varied data set. ANNs do

[9] Some researchers in the ANN field call layers as slabs.

not need input assumptions, can also learn from previous knowledge and can generalize the learned knowledge for future predictions (Bahrammirzaee, 2010). Consequently, the use of ANNs also allows the modeling of complicated operations to solve intricate, nonlinear or stochastic problems (Graupe, 2013). An erroneous cell (or variable) cannot affect the overall performance of the algorithm. This feature allows ANNs to be accurate in predictions that have an uncertain environment or potential to change with the time (Gençay et al., 2002). In addition to that, the ANN model can quickly adopt new environment without any predefinition or preprogramming (Yegnanarayana, 2006). Except that, constructing the data set will be time consuming depending on the chosen learning method. ANNs are inspired by human brain activities and mimics the pattern classification and pattern recognition ability of it (Zhang et al., 1998). On the other hand, nonparametric ANNs do not need distribution requirements of traditional parametric statistical models (Coakley & Brown, 2000).

Naturally, ANNs have not only superior capabilities but also have inferior features compared to other statistical methods. The Achilles heel of ANNs is the overfitting problem of the model. Srivastava et al. (2014) define overfitting as producing complicated relationships through sample noising that even not exist in the original data. Several methods (Guresen et al., 2011; Ticknor, 2013; Srivastava et al., 2014) are developed to overcome the overfitting problem. Application of neural networks on financial markets contains high overfitting probability due to the noisy nature of financial data (Ticknor, 2013; Krauss et al., 2017). Possibility of such problem force scholars, who deal with financial statements, to be cautious against overfitting and overtraining of the model.

3.1.1. Artificial Neural Networks in Accounting and Finance

ANN applications in the accounting and finance area began with the article of Tam and Kiang (1990). In the infant era, ANN models mostly employed to predict the risk of bankruptcy (Odom & Sharda, 1990; Tam, 1991; Wilson & Sharda, 1994; Tsai & Wu, 2008). In addition to that, ANNs came into prominence among scholars to forecast the economic time series data (Kaastra & Boyd, 1996; Thawornwong & Enke, 2004). Later on, ANNs are applied on several finance related topics such as stock market index predictions (Guresen et al., 2011; Niaki & Hoseinzade, 2013), exchange rate predictions (Adhikari & Agrawal, 2014; Galeshchuk, 2016), credit risk predictions (Bekhet & Eletter; 2014; Zhao et al., 2015). Fanning et al. (1995) published the first fraud-related research that employed ANN. Since their seminal work, the ANN-based fraud prediction topic drew the attention of many researchers (e.g., Green & Choi, 1997; Fanning & Cogger, 1998; Lin et al., 2003; Kirkos et al., 2007; Ngai et al., 2011; Lin et al., 2015).

The vast amount of previous studies employ financial ratios as input variables to forecast the risk of financial fraud. Few types of research (Fanning & Cogger, 1998; Lin et al., 2015; Chen, 2016) focused on the combination of financial and nonfinancial data when employing ANNs. Hajek and Henriques (2017) employed financial statement data and managerial comments on annual reports to detect financial statement fraud. They find that the more comprehensive data set should be constructed to develop an algorithm that has prediction ability. Focusing only on financial data unintentionally damages the efforts of developing a holistic approach to fraud detection.

Perols (2011) compares six machine learning algorithms to detect financial statement fraud under different conditions. He analyzed 42 independent variables as fraud predictors for detection. Six variables coherently selected by employed algorithms. Besides, support vector machines and logistic regression outperforms other algorithms.

3.1.2. Supervised Learning

Supervised learning's roots are based on the early era of neural computation literature. The idea behind supervised learning is to train the algorithm based on the instructions that were given by the supervisor (teacher) (Basu et al., 2010). A goal of supervised neural network training is to find the weights that have the minimum error and minimize the total error of the model (Schmidhuber, 2015). This minimum error based approach will increase the generalizability power of the model in the later phases.

Ponulak and Kasinski (2010) define the underlying assumption of the supervised learning in ANNs aims to minimize the error between the actual and the predicted results by modifying the changeable parameters of the given neuron.

Caruana and Niculescu-Mizil (2006) evaluate ten supervised learning methods to compare the classification performances of them. They classify the performance of those methods according to eight performance metrics.

A groundbreaking article (Silver et al., 2016) had published in the field of neural network based machine learning in 2016. Scholars from Google DeepMind team develop an algorithm for the ancient game of the GO. Researchers suggest that GO is the most complex game for artificial intelligence. They produce an algorithm that can play GO on an expert level and has beaten the world GO champion five times in a row. They train their algorithm by using a supervised learning approach. Specifically, their model is a type of semi-supervised learning algorithm.

3.1.3. Semi-supervised Learning

Semi-supervised learning is in between supervised and unsupervised learning methods. Combination of supervised information and unlabeled data employed in semi-supervised learning models (Chapelle et al., 2006). Traditional learning methods need labeled input data to train the algorithm. Semi-supervised algorithms developed to deal with the time consuming and costly nature of labeled data. A semi-supervised algorithm can be trained by employing the combination of a huge amount of unlabeled data and a small amount of labeled data (Zhu et al., 2003; Zhu, 2005).

Semi-supervised learning algorithms are capable of dealing with the classification problem when there is only a small group of observations that have matching labels (Kingma et al., 2014). Naturally, unlabeled data carry less information than labeled data. Such a characteristic force semi-supervised learning algorithms to have a massive amount of data to increase the prediction power of the model (Chapelle et al., 2006).

3.1.4. Unsupervised Learning

There is no supervisor (teacher) in unsupervised learning for the training of the algorithm. The unsupervised learning aims to find answers to the questions without having correct answers beforehand (Hastie et al., 2008). In another saying, there are n number of cases and n number of potential circumstances in the training processes of the unsupervised learning algorithms (Zhu & Goldberg, 2009). The goal of unsupervised learning is to find new relations that are already there but hidden in the data set (Chapelle et al., 2006). In unsupervised learning models, we cannot construct a linear regression model because of the absence of the dependent variable. On the other hand, allows us to cluster our undefined information to find patterns and build supervised models (James et al., 2013).

3.2. Utilized Algorithms

3.2.1. Multilayer Perceptron

Artificial neural networks or neural networks, generally based on multilayer perceptron networks (Panchal et al., 2011). The multilayer perceptron is a backpropagation based classifier to learn and classify instances. A typical network of perceptrons comprises of input layer (independent variables), hidden layers and output layer (dependent variable). Each connection between layers has its weight. The algorithm lowers the error of the weights through the method of the steepest descent (Battiti, 1992) and by iterative repetitions. In other saying, multilayer perceptrons can learn and update the weight of the connections during the training (Pal & Mitra, 1992). Multilayer perceptrons are slower among other algorithms but perform well especially with large datasets (Witten, 2019). Moreover, multilayer perceptron based models are popular for financial predictions (Tsai & Wu, 2008).

Multilayer perceptron tool of the Weka has employed as one of the benchmark methods for this research.

3.2.2. Logistic Regression

Logistic regression based classifier is employed because of the binary[10] character of the dependent variable. Several other research (Lin et al., 2003; Koh & Low, 2004; Lin et al., 2015) also employ logistic regression based neural networks as a classifier in fraud detection literature. Logistic class under the classification section in Weka employs a modified version of the logistic regression model with a ridge estimator. Additionally, the logistic algorithm under the classification section of Weka is modified because the logistic regression models do not work with instance weights intrinsically.

3.2.3. Decision Trees

Decision trees are models that have the ability of classification or regression based prediction. In decision trees, decision-maker aims to reach the best possible scenario (Rokach & Maimon, 2015). The term for specified decision tree model based on whether it is employed for classification or regression. A typical decision tree is very

[10] y_i=0 or 1 for all n cases

similar to a tree in nature and formed through three main components (Amor et al., 2004):

1- a node,
2- a branch,
3- a leaf.

Abovementioned components determine an attribute (variable), an attribute outcome, and an answer based on nodes and branches, respectively. Decision trees became favorable among the researchers because of ease of use and understandability. Additionally, decision trees also can classify categorical and numerical data, except that the dependent variable (or the output attribute) must be categorical (Zhao & Zhang, 2008). From that point, this study employs the undermentioned decision tree algorithms.

Pruning has vital importance in decision tree based research corpus. We can conceptualize pruning in decision trees same as pruning in horticulture. In Japanese Bonsai art, one of the reasons for pruning a healthy bonsai tree is to distribute the energy of the tree in itself efficiently[11]. In machine learning, the main motivation behind pruning is to construct an efficient (minimized error) decision tree based model. Bratko and Bohanec (1994) summarize pruning as “trading accuracy for simplicity”. Pruning can be applied to the branches or leaves of a single decision tree based model or trees within a random forest model (Kulkarni & Sinha, 2012). Pruning can continue until the stopping criterion is fulfilled. Those stopping criteria can be attaining maximum tree depth, or the most outperformed splitting criterion is lower than a specific threshold (Rokach & Maimon, 2015).

3.2.3.1. C4.5 (J48) Algorithm

C4.5 is a simple software that is based on decision trees and developed by J. Ross Quinlan (1993). He indicates that C4.5 is a developed form of ID3[12]. Quinlan has a strong impact on classification based machine learning application with his serial publications in the 1980s and 1990s (Quinlan, 1983; Quinlan 1986; Quinlan, 1987; Quinlan, 1993). In broad terms, C4.5 is an algorithm that is based on the principles of decision trees. Main advantages of the C4.5 algorithm, when it is compared with other classification algorithms, are it has higher accuracy in classification problems and it is

[11] http://bonsai4me.com/Basics/Basics%20Bonsai%20Continual%20trimming.htm
[12] A decision tree construction algorithm which is based on entropy (information) measure.

faster than many other algorithms in data mining and machine learning applications (Ruggieri, 2002). As a summary, the C4.5 algorithm starts with a root node at the top part of the tree, which analyzes the information of the whole sample and transfers it to the branch node. In this step, the algorithm produces rules depending on information measure of subsamples. In the final step, C4.5 produces huge tree comprises all attributes (variables) and produce the final decision after the pruning (Ali & Smith, 2006).

C4.5 algorithm is named as J48 in Weka software terminology. For this reason, J48 results stand for C4.5 algorithm results in this thesis results.

3.2.3.2. Random Forest

The random forest method can be sub-classified under supervised machine learning methods. Leo Breiman (2001) proposes a random forest method as a combination of decision trees that depend on equally distributed and separately sampled vectors. According to his article, the random forest algorithm outperformed the well-known and widely used AdaBoost (Ratsch et al., 2001; Schapire, 2013) algorithm. He also notes that the algorithm overcomes the overfitting problem of classification models. Additionally, a random forest is a decision tree based method that can efficiently handle large datasets in real life examples (Oshiro et al., 2012). In another saying, the random forest contains a bunch of decision trees that represent the identically distributed random vectors (Kulkarni & Sinha, 2012). They individually contribute to the ultimate result. Each decision tree contains different subsample and feature set to decrease the error of the final model.

Random Forest tool of the Weka has employed as one of the benchmark methods for this research.

4. DATA AND SAMPLE

This section discusses the sample structure and the dataset of this study. Firstly, the sample of this study covers the data of the companies that have headquarter in the United States but cross-listed on stock exchanges outside of the US and companies that have headquarter in outside of the United States but listed on the National Association of Securities Dealers Automated Quotations (hereafter, NASDAQ) and New York Stock Exchange (NYSE).

4.1. Data

Data of fraudulent and nonfraudulent companies collected for this research. Fraudulent and nonfraudulent companies' data set allow us to identify the key variables that have direct or indirect effects on fraud. Furthermore, the training algorithm trained much more consistent by combining the data from fraudulent and nonfraudulent companies. The sample of this study covers the corporate frauds that are securities class action lawsuits and filed between the periods of January 2002 – December 2017. Post-Sarbanes – Oxley Act period has chosen because; I would like to evaluate the fraud-related cases in a single legal basis. Data regarding financial ratios and corporate governance variables are retrieved through the Compustat, BoardEx and ORBIS databases. Additional country-specific macroeconomic and institutional variables are imported from the historical dataset of the Global Competitiveness Index of World Economic Forum.

Collecting fraud filing data is complicated because of the hidden nature of financial fraud data. Furthermore, the language barrier for reaching filing results is a serious obstacle against collecting multi-country data. This research covers the fraud data from the U.S. because of data availability. We intend to overcome the single market and generalizability issues by including foreign companies through cross-listing data. Choosing cross-listed companies allow us to reach mutual legal ground for companies from different countries. Additionally, this allows us to understand the country-specific financial statement fraud characteristics of foreign countries. Cross-listed companies are highlighted through the American Depository Receipts (ADR)[13].

[13] "A depositary receipt (DR) is a physical, negotiable certificate that represents ownership of shares in an overseas company that is held in custody in the issuer's home market. The structure of a depositary receipt includes a ratio, which correlates the amount of underlying shares to the receipt, a well as other general terms and conditions applicable to holders. A depositary receipt can be cancelled for its underlying shares at any time. An American Depositary Receipt ("ADR") references DRs that are

The dataset of this research collected through multiple databases and in several different file formats. However, Weka, the employed knowledge analysis software for this research, can only read Attribute-Relation File Format (ARFF). For this reason, final dataset constructed as an ARFF file.

4.1.1. Data of the Financial Fraud

The most challenging part in financial fraud research is the hand collection the data of fraud filings. In this research, fraud filings data set is hand collected through Securities Class Action Clearinghouse (hereafter, SCAC). SCAC is an online database of Stanford Law School and covers "a database of more than 4,000 securities class action lawsuits filed since passage of the Private Securities Litigation Reform Act of 1995." (SCAC, 2019). Employing securities class action lawsuits data is well accepted in the financial fraud research corpus (Choi, 2007; Dyck et al., 2010; Karpoff et al., 2017). Furthermore, collecting financial fraud data from the SCAC database is preferable because of the match up with the financial definition of the database and the included financial fraud data. There are huge differences among fraud definitions of the databases and the included data. In an example, Karpoff et al. (2017) indicate that there are 4155 observations in the SCAC database and 100% of them are fraud observations according to fraud definition of the database. However, the same ratios for other financial fraud filings databases, such as Government Accountability Office, Center for Financial Reporting and Management and Audit Analytics, are 26,4%, 31,3%, and 1,7% respectively.

The financial fraud dataset of this study contains 3337 securities class action lawsuits between 04.01.2002 – 29.12.2017. Company names and tickers are collected as a company identifier. Moreover, filing date, date of the final order, listed stock exchange, district court of the filing, industry and sector, headquarter data are collected. The fraud data of this study is refined in several steps. First, the data of privately owned companies are excluded and 3203 filings left in the dataset. Second, to solely cover the companies that are listed on NASDAQ and NYSE, the data of the listed exchange is processed and as a result, 2887 lawsuits left. Later on, the crucial part of the data collection had come and each case is identified according to the result of the lawsuit. To omit frivolous cases, we categorized filing results in three categories

available in the U.S. The terms ADR and ADS (or DR and DS) are often used interchangeably." (https://www.adr.com/Investors/Glossary) named and labelled as 1: Ongoing, 2: Dismissed and 3: Settled and 4: Remanded[14]. There are 195 cases labelled as Ongoing[15], 1491 cases labelled as Dismissed, 1177 cases labelled as Settled in the dataset. 23 Remanded and one blank case are ignored for the sake of clarity. In the literature (Coffee Jr., 2006; Dyck et al., 2010; Arena & Julio, 2015), settled class action lawsuits in the SCAC database are generally accepted as the existence of fraudulent activities. For this reason, I identify fraudulent companies through the results of the lawsuits.

Company Name	**Settlement Amount($)**
Enron Corporation	7.227.390.000
WorldCom, Inc.	6.133.000.000
Tyco International Ltd.	3.200.000.000
Cendant Corporation	3.186.500.000
Petroleo Brasileiro S.A. - Petrobras: American Depository Shares	3.000.000.000
Nortel Networks Corporation (Nortel I & II)	2.935.901.451
AOL Time Warner, Inc.	2.500.000.000
Bank of America Corporation: Merger with Merrill Lynch	2.425.000.000
Household International, Inc.	1.576.500.000
Koninklijke Ahold NV: Royal Ahold Corporation Securities on the United States and European Stock Exchanges	1.100.000.000

Table 2. Top Ten Largest Settlements in the SCAC Database

I have to clarify that, in the SCAC database, companies are identified through official company names or tickers. Using a company name is useless for this study. On the other hand, using tickers as company identifiers can be useful in several cases; however, it is ineffective. Additionally, tickers in SCAC database are not overlapping with other databases in a notable amount of cases. For example, 654 filings[16] ticker is updated to combine the SCAC data with the BoardEx data effectively. Firstly, ticker

[14] Ongoing term indicates that the lawsuit has started but not decided yet, Dismissed term indicates that the lawsuit is dismissed because of no sign of criminal activity or voluntarily dismissal, Settled term indicates that there is a settlement between plaintiffs and the company.
[15] The oldest and still ongoing lawsuit case is on trial since October 9, 2009.
[16] The ticker of 304 settled filings is updated.

matching had targeted. Later on, matching between the official company names are controlled for double-checking.

Ultimately, I match companies with their International Securities Identification Numbers (ISIN) and 916 fraudulent companies left.

Derived fraud variable from settled cases is dichotomous. Companies that have settled fraud filing(s) in the SCAC database are going to have the value of 1, otherwise 0. By this way, a clear distinction between fraudulent and nonfraudulent companies had been made.

4.1.2. Data of the Financial Statement

Raw financial data, financial ratios and corporate governance variables are retrieved through the Compustat, the BoardEx and ORBIS databases. The financial variables of the dataset are combined through the literature (Persons, 1995; Kirkos et al., 2007; Lin et al., 2015). Corporate governance variables are also gathered through the combination of the datasets of similar studies (Chen et al., 2006; Lin et al., 2015; Chen, 2016). Abovementioned researchers mostly employ normalized financial data in the financial fraud detection literature. This research mutually employs raw and normalized financial data. Employed feature selection methods allow us to construct a comprehensive dataset.

Table 3. List of Financial Variables and Financial Ratios

Current Assets - Total	Pro Forma Net Sales - Prior Year
Assets - Other	Stockholders Equity - Parent
Accounts Payable - Trade	Stockholders Equity - Total
Assets - Total	Unearned Income
Book Value Per Share	Working Capital (Balance Sheet)
Cash	Operating Expenses - Total
Cost of Goods Sold	Prepaid Expenses
Dividends - Total	Stock Exchange Code
Earnings Before Interest and Taxes	Active/Inactive Status Marker
Earnings Before Interest	Current ISO Country Code - Incorporation
Earnings Per Share (Basic) - Including Extraordinary Items	Market Value - Total - Fiscal
Goodwill	Auditor
Gross Profit (Loss)	Auditor Opinion
Invested Capital - Total	Chief Executive Officer SOX Certification
Intangible Assets - Total	Chief Financial Officer SOX Certification
Inventories - Finished Goods	Current Ratio
Inventories - Other	Quick Ratio

Inventory/Stock - Other	Cash Ratio
Inventories - Total	Operating Cash Flow
Investment Securities -Total	Debt Ratio
Current Liabilities - Total	Debt to Equity
Liabilities - Other - Total	Gross Margin
Liabilities - Total	Operating Margin Ratio
Net Income (Loss)	Return on Assets
Net Interest Margin	Return on Equity
Operating Activities - Net Cash Flow	Operating Income/Total Assets
Operating Income Before Depreciation	Asset Turnover Ratio
Pretax Income	Book Value Per Share
Retained Earnings	Earnings Per Share (Basic) - Including Extraordinary Items
Retained Earnings - Restatement	Cash to Total Assets
Receivables - Total	Current Liab/Total Assets
Revenue - Total	Net Profit/Total Assets
Sales/Turnover (Net)	Working Capital/Total assets
Pro Forma Net Sales - Current Year	Net Profit/Net Sales

Table 4. List of Corporate Governance and Nonfinancial Variables

Total Number of Board Members	Liquid Wealth ED Average
Total Number of EDs	Liquid Wealth ED Total
Total Number of NEDs	Liquid Wealth NED Average
CEO and Chairman Roles are combined on the Board	Liquid Wealth NED Total
Average Salary EDs	Average time in role for EDs
Average Salary NEDs	Average time in role for NEDs
Average Bonus EDs	Average years on Other Quoted Boards EDs
Average Bonus NEDs	Average years on Other Quoted Boards NEDs
Average Total Direct Compensation for EDs	Average Age EDs
Average Total Direct Compensation for NEDs	Average Age NEDs
Average Total Equity-Linked Compensation for EDs	Average Number of Education EDs
Average Total Equity-Linked Compensation for NEDs	Average Number of Education NED
Average Wealth Shares EDs	Gender (% Male) EDs
ED Total Wealth Shares	Gender (% Male) NED
Average Wealth Shares NEDs	Nationality Mix ED
NED Total Wealth Shares	Nationality Mix NED

4.1.3. Macroeconomic Indicators

There are several online and official databases[17] for macroeconomic indicators. Even countries publish their statistics for the attention of the public. However, this study covers the data of the Global Competitiveness Index[18] (hereafter, GCI) of the World Economic Forum for the period among 2002-2017. The aim of employing the pillars of the GCI dataset is to capture more complex information with a less numerical approach. The GCI dataset covers 114 indicators under 12 pillars. For the sake of clarity, this study covers only nine of the 12 pillars. Those pillars are;

1st pillar	Institutions
3rd pillar	Macroeconomic environment
5th pillar	Higher education and training
6th pillar	Goods market efficiency
7th pillar	Labor market efficiency
8th pillar	Financial market development
9th pillar	Technological readiness
10th pillar	Market size
11th pillar	Business sophistication

Table 5. Pillars of the Global Competitiveness Index

In an example, the 10th pillar consists of the indicators of domestic market size index, foreign market size index, gross domestic product valued at purchasing power parity (GDP) and exports as a percentage of GDP. All of the pillar values are matched with the companies according to the headquarter of the company.

4.1.4. Feature Selection

Feature term has very similar meaning with the variable. Choosing the correct term for the same thing depends on the field of study. Feature selection is a crucial step for data mining applications and knowledge analysis. Piramuthu (2003) indicates that the %80 of resources in data mining applications (mostly, time) spend for the preprocessing and cleaning of the dataset. This period mostly focuses on constructing well-structured dataset by employing several feature selection methods. If the feature selection

[17] https://data.worldbank.org/ , http://data.un.org/

[18] Visit https://www.weforum.org/reports/the-global-competitiveness-report-2017-2018 to reach the insights about the report.

procedures applied poorly, the results of data mining applications could be highly unstable (Ravisankar et al., 2011).

Feature selection procedures increase the overall performance, increase the prediction power, decrease the noise of the data and lower the calculation time of machine learning application by dropping out unimportant features (Chandrashekar & Sahin, 2014). In another perspective, feature selection is a process of exchanging the explanatory power of the model with efficiency. There are tens of feature selection methods in the literature. To implement the perspective in the previous line, this study employs the Gain Ratio Attribute Evaluation method as a feature selection method.

4.1.4.1. Gain Ratio Attribute Evaluation:

Gain Ratio Attribute Evaluation method assesses the value of an attribute by calculating the gain ratio through the corresponding class (Hall, Class InfoGainAttributeEval, 2019). This method is superior to the decision-making process when there is a large number of attributes. On the other hand, this perspective mostly deals with the uncertainty through manipulating, processing and evaluating the available information (Ghahramani, 2006). The attribute with the highest ranking value in the gain ratio method will be chosen as a splitting attribute (Karegowda et al., 2010).

The effect of the attribute on the entropy (information gain) of the class can be formulated as (Hall & Holmes, 2002);

$$\mathrm{H(Y)} = -\sum_{y \in Y} p(y) \log_2 p(y)$$

$$\mathrm{H(Y|X)} = -\sum_{x \in X} p(x) \sum_{y \in Y} p(y|x) \log_2 p(y|x)$$

Gain Ratio can be formulated as (Frank & Witten, 2004);

$$Gain\ Ratio = \frac{Information\ Gain\ (Attribute)}{Intrinsic\ Information\ (Attribute)}$$

In Weka, each attribute ranked to decide the selected features according to the following formula;

$$GainRatio(Y|X) = \frac{(H(Y) - H(Y|X))}{H(X)}$$

where Y represents class and X represents attributes.

4.2.Sample

The sample of this study focuses the data of the companies that have headquarter in the United States but cross-listed on stock exchanges outside of the US and companies that have headquarter in outside of the United States but listed on the National Association of Securities Dealers Automated Quotations (hereafter, NASDAQ) and New York Stock Exchange (NYSE).

In total, 916 identifiable fraud cases data collected from the SCAC database. However, I apply another filter to capture the data of cross-listed companies solely. After that, 74 fraudulent companies are identified. Those companies are matched with 416 cross-listed companies that are identified as nonfraudulent to construct a dataset for neural network training.

4.2.1. Cross – Listing

Highly competitive and capital intensive structure of economies increases the outside financing requirement of companies. Liberalization of economies, developments in financial systems, lowered barriers for capital moves, new opportunities for reaching financial capital in different countries, excessive dependency on financial capital for new investment projects create a sophisticated and fragile economic environment. Such an economic environment force governments and regulatory bodies to handle complicated economic issues.

Companies can reach funds by raising debt or equity. Many research (Modigliani & Miller, 1958; Myers, 1984; Rajan & Zingales, 1995; Frank & Goyal, 2003; Fan, Titman, & Twite, 2012) focus on the optimal capital structure to identify the most convenient financing option for a company. Modigliani and Miller (1958) proposed that equity or debt financing options have the same cost to the company in efficient markets. However, Baker and Wurgler (2002) identified that firms could gain advantage from equity issuing if they examine the market conditions and timing in detail. In other words, a company can decrease the cost of capital through financial markets. Cross-listing on different financial markets allow companies to reach different capital resources. Coffee (2002) proposes that companies from countries that have poor legal environment tend to list on stock exchanges that have higher legal standards. By this way, they aim to increase the disclosure standards of the company voluntarily and reach potential investors. Besides, cross-listing on U.S. stock exchanges reduces the cost of capital (Lambert et al., 2007) and this reduction results in higher firm valuation (Hail & Leuz, 2009). On the other hand, cross-listing on major stock exchanges in the U.S. rather than over the counter markets results in higher valued stocks (Hope et al., 2007).

Difference between legal systems in countries is the crucial challenging point of fraud related research corpus (Coffee, 2005). Besides, a sharp divergence between common law and civil law mitigates the generalizability of fraud related researches (Reese & Weisbach, 2002). For this reason, a vast amount of research (Huijgen & Lubberink, 2005; Leuz, 2006; Chang & Sun, 2009; Berger et al., 2011; Hope et al., 2013) in various fraud related areas focus on cross-listed companies to eliminate this complex issue. This research contains the data of US cross-listed companies listed on stock exchanges that established outside of US and non-US cross-listed companies that listed on selected US stock exchanges. Rule 10b-5 allows US investors to sue cross-listed companies due to their fraudulent financial activities whether it has occurred in the US or it has occurred in another country (Reese & Weisbach, 2002). Only US investors can benefit under this rule, but all investors can benefit from it in practice. Reaching mutual legal ground for companies from different countries and cultures is the main reason for choosing cross-listed companies. Additionally, it allows us to observe multiple countries instead of a single country. That allows us to observe the country-specific fraud environment.

4.2.2. Bonding Hypothesis

This subsection is motivated by the reasons behind cross-listing that has direct or indirect effects on financial statement fraud. It mostly focuses on the Bonding Hypothesis proposed by Coffee (1999) and Stulz (1999).

Firms have various motivations to be cross-listed on different stock exchanges. One of them is to give signals about the company's perception of investor protection. A foreign-based company, from a country that has lower shareholder protection, can increase its value by cross-listing on strictly regulated stock exchange regimes (Coffee, 1999; Stulz, 1999). By this way, a company signals to the market and shareholders about its positive views and respect to the shareholders' protection. Additionally, a company also bonds itself to a more strict law environment, disclosure rules, auditing standards, and enforcements through cross-listing. According to Shleifer and Vishny (1997), legal systems play two critical roles. First, it limits the managers to steal from investors' wealth. Second, shareholders can monitor managers and protect their rights through the mechanisms of inclusive legal systems.

The effects of the bonding hypothesis can be clearly seen on corporate governance related issues. Charitou et al. (2007) investigate the relationship between the cross-listing and corporate governance activities of Canadian companies that are cross-listed on U.S. exchanges. They find that those cross-listed companies have more independent board structures and audit committees after the cross-listing.

Loureiro (2010) proposes an interesting finding about cross-listing on U.S. stock exchanges. Foreign firms cross-listed on U.S. stock exchanges and started with IPO procedures are more likely to hire underwriters that are more prestigious if they are from countries that have weak shareholder protection. By this way, companies have done their window dressing and have higher valued shares.

4.2.3. American Depository Receipts (ADRs)

Companies can either choose directly cross list on stock exchanges or with depository receipts. Depository receipts are the physical certificates that bear the ownership rights of oversea companies and hold under custody in the issuer's home market. American Depository Receipts (ADRs) have three levels and all of them come with different legal responsibilities. Level I ADRs have no additional reporting requirements for companies and mostly traded on over the counter at the pink sheet markets. However, Level II and Level III ADRs come with SEC regulations and US GAAP (Huijgen & Lubberink, 2005). Companies listed with Level II ADRs can be traded on secondary markets. Meanwhile, companies that listed with Level III ADRs can be traded both on primary and secondary markets.

ADR data of this study is collected through two different way for double-checking. First, ADR data is downloaded from the web site of www.adr.com that is the DR database of the J.P.Morgan. Second, ADR data also collected from the COMPUSTAT database. However, we cannot directly identify the companies that cross-listed through ADRs in COMPUSTAT. Two methods suggested by Wharton Database[19] to identify the ADR issued cross-listed companies. I follow those instructions to identify ADRs in the COMPUSTAT database.

[19] https://wrds-www.wharton.upenn.edu/pages/support/research-wrds/research-guides/guide-adrs-and-research/

5. RESULTS

5.1. Feature Selection

There are 48 variables left for the final training and testing procedure after applying Gain Ratio feature selection method. Threshold set to the 0.005 level to increase the captured information. On the other hand, raw financial data and normalized financial variables have the same gain ratio value. However, none of them omitted because of potential contribution to the neural network training stage.

=== Run information === **Evaluator: weka.attributeSelection.GainRatioAttributeEval** **Search: weka.attributeSelection.Ranker -T 0.005 -N -1** **Relation: fraud-weka.filters.unsupervised.attribute.** **Instances: 5250** **Attributes: 166** **[list of attributes omitted]** **Evaluation mode: evaluate on all training data** **=== Attribute Selection on all input data ===** **Search Method:** **Attribute ranking.** **Threshold for discarding attributes: 0.005** **Attribute Evaluator (supervised, Class (nominal): 166 fraud):**	
Gain Ratio feature evaluator	
 Ranked attributes:	
0.14441	Pillar10
0.12484	Pillar1
0.09409	NationalityMixED
0.06998	GenderMaleEDs
0.06033	NationalityMixNED
0.05958	Pillar11
0.05867	TotalNumberofBoardMembers
0.05201	Pillar5
0.04672	AccountingStandard
0.04541	Pillar7
0.03694	TotalNumberofEDs
0.03592	Pillar3
0.03289	TotalNumberofNEDs
0.03232	Pillar6
0.02899	Pillar9
0.02713	Pillar8
0.023	AverageAgeEDs
0.01615	ChiefFinancialOfficerSOXCert

0.01615	ChiefExecutiveOfficerSOXCert
0.01491	AveragetimeinroleforEDs
0.01415	AverageNumberofEducationEDs
0.01331	InventoriesOther
0.01331	normalizedinvoth
0.01174	GenderMaleNED
0.00783	WorkingCapitalBalanceSheet
0.00783	normalizedworkcap
0.00779	AverageyearsonOtherQuotedBo
0.00694	AveragetimeinroleforNEDs
0.0066	normalizedcrrntliabtot
0.0066	CurrentLiabilitiesTotal
0.00651	Goodwill
0.00651	normalizedgoodwill
0.00613	LiquidWealthEDAverage
0.00612	LiabilitiesTotal
0.00612	normalizedliabtot
0.0061	normalizedintastot
0.0061	IntangibleAssetsTotal
0.00589	normalizedaccpaytra
0.00589	accountspayabletrade
0.00586	AverageTotalDirectCompensatio
0.00568	Fiscalyear
0.0056	normalizeddivtot
0.0056	dividendstotal
0.00541	AssetTurnoverRatio
0.00541	normalizedassetturnrat
0.00537	AverageSalaryNEDs
0.00533	CEOandChairmanRolesarecombi

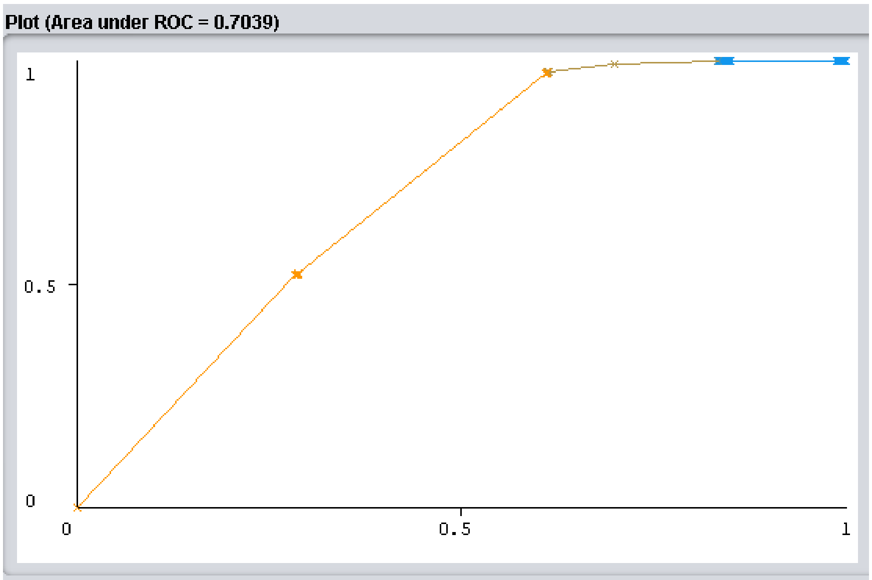
5.2. ROC Curves

Receiver Operator Characteristic (ROC), or Area Under the Curve (AUC), curves are plots that have the ability to show the diagnostic ability of binary classifiers. In the beginning, it was employed to classify the patients correctly. Later on, it mostly employs to understand the probability of correctly classifying a randomly chosen case (Bradley, 1997; Pencina et al., 2008). In another saying, it measures the performance of the classifier.

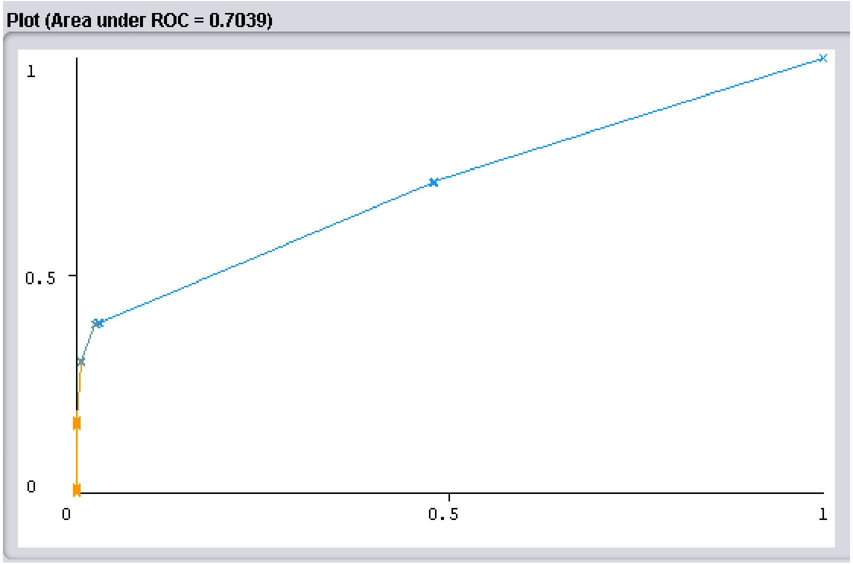
5.3. Multilayer Perceptron

Multilayer perceptron network has correctly classified the 88.96% of the instances (observations). In detail, it correctly classified 4670 instances of the full dataset and incorrectly classified 580 instances. Thirty-two of incorrectly classified instances consist of nonfraudulent cases.

```
=== Summary ===

Correctly Classified Instances        4670               88.9524 %
Incorrectly Classified Instances       580               11.0476 %
Kappa statistic                          0.4023
Mean absolute error                      0.162
Root mean squared error                  0.3097
Relative absolute error                 63.8067 %
Root relative squared error             86.9397 %
Total Number of Instances             5250

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0,993    0,700    0,890      0,993    0,939      0,475    0,704     0,906     0
                 0,300    0,007    0,880      0,300    0,448      0,475    0,704     0,458     1
Weighted Avg.    0,890    0,597    0,889      0,890    0,865      0,475    0,704     0,840

=== Confusion Matrix ===

    a    b   <-- classified as
 4435   32 |    a = 0
  548  235 |    b = 1
```



ROC Curve for Class 0 (Nonfraudulent)

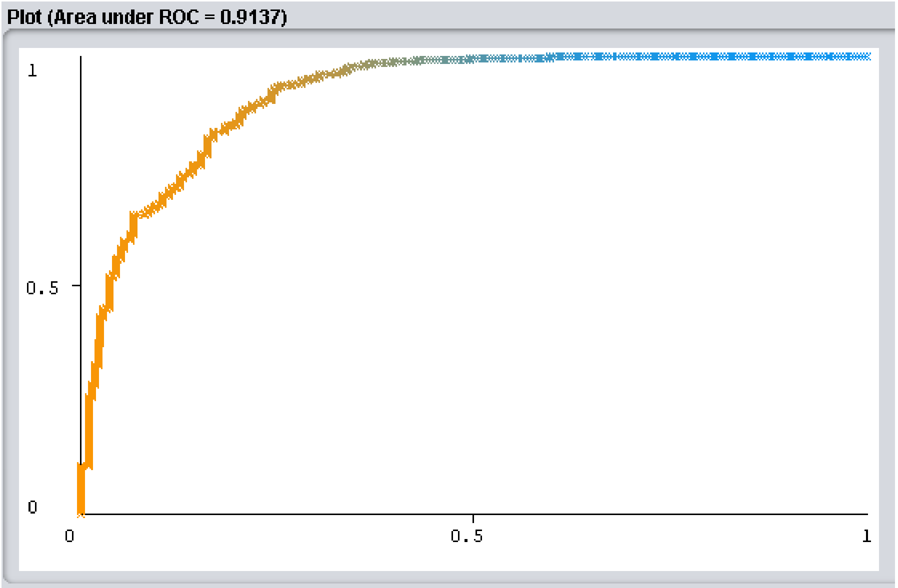


ROC Curve for Class 1 (Fraudulent)

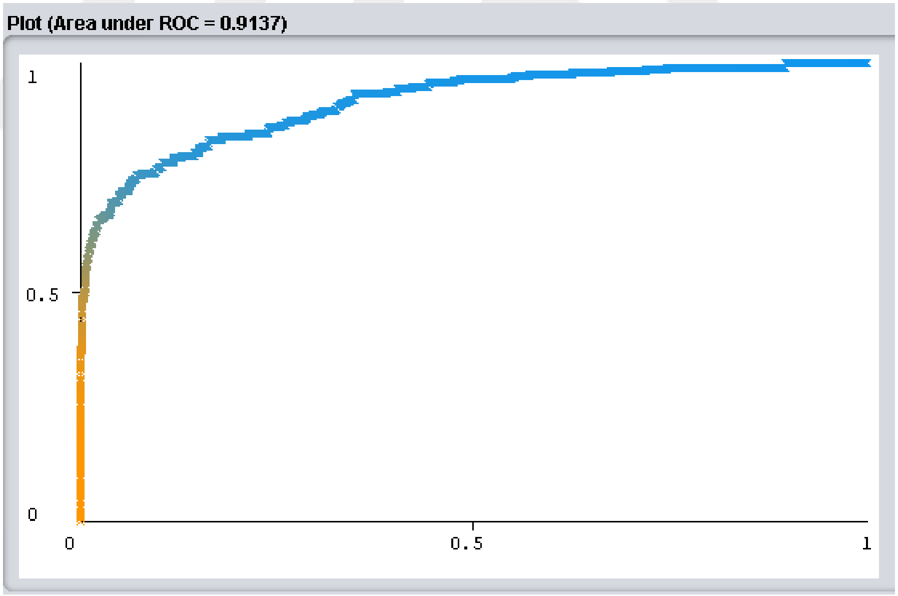
5.4. Logistic Regression

Multinomial logistic regression based classification model has correctly classified the 92.825% of the instances (observations). In detail, it correctly classified 1656 instances and incorrectly classified 128 instances. Twenty of incorrectly classified instances consist of nonfraudulent cases. 108 of incorrectly classified instances are consist of fraudulent instances. The ratio of incorrectly classified cases seems high in fraudulent cases, but the correctly classified fraudulent cases (158) number is still better than incorrectly classified instances.

```
=== Summary ===

Correctly Classified Instances        1656               92.8251 %
Incorrectly Classified Instances       128                7.1749 %
Kappa statistic                          0.6726
Mean absolute error                      0.1204
Root mean squared error                  0.243
Relative absolute error                 47.4561 %
Root relative squared error             68.2127 %
Total Number of Instances             1784

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0,987    0,406    0,933      0,987    0,959      0,690    0,914     0,979     0
                 0,594    0,013    0,888      0,594    0,712      0,690    0,914     0,794     1
Weighted Avg.    0,928    0,347    0,926      0,928    0,922      0,690    0,914     0,952

=== Confusion Matrix ===

    a    b   <-- classified as
 1498   20 |    a = 0
  108  158 |    b = 1
```



ROC Curve for Class 0 (Nonfraudulent)

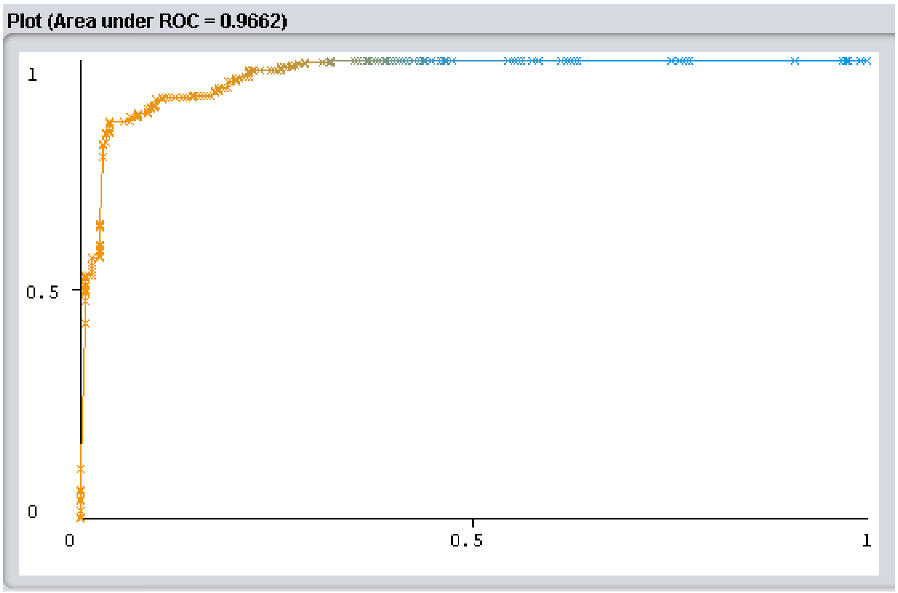


ROC Curve for Class 1 (Fraudulent)

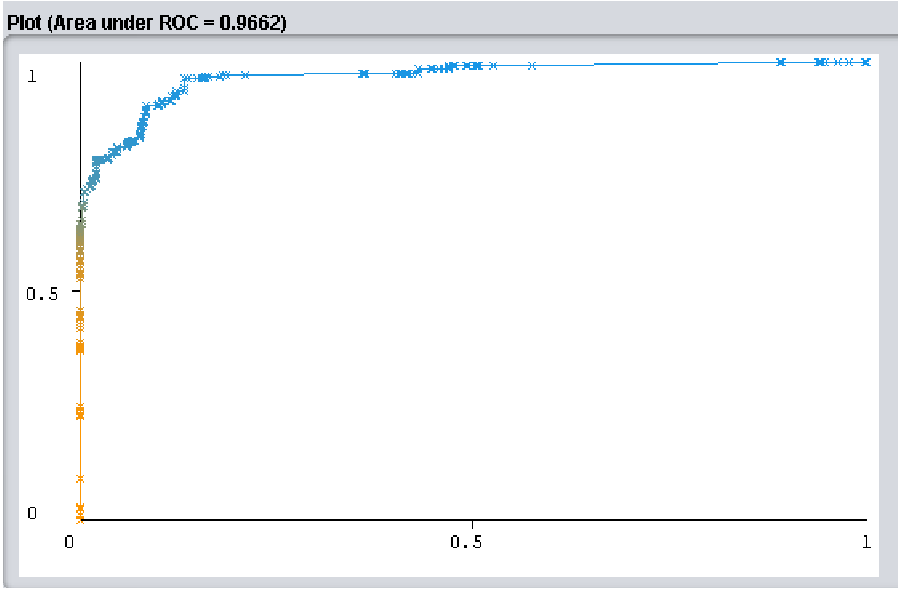
5.5. C4.5 Algorithm

C4.5 pruned decision tree has correctly classified the 94.51% of the instances (observations). In detail, it correctly classified 1686 instances and incorrectly classified 98 instances. Four of incorrectly classified instances consist of nonfraudulent cases. 94 of incorrectly classified instances are consist of fraudulent instances. The ratio of incorrectly classified cases seems high in fraudulent cases, but the correctly classified fraudulent cases (172) number is still better than incorrectly classified instances.

```
=== Summary ===

Correctly Classified Instances        1686               94.5067 %
Incorrectly Classified Instances        98                5.4933 %
Kappa statistic                          0.7484
Mean absolute error                      0.1015
Root mean squared error                  0.2098
Relative absolute error                 39.9824 %
Root relative squared error             58.8957 %
Total Number of Instances             1784

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0,997    0,353    0,942      0,997    0,969      0,769    0,966     0,992     0
                 0,647    0,003    0,977      0,647    0,778      0,769    0,966     0,895     1
Weighted Avg.    0,945    0,301    0,947      0,945    0,940      0,769    0,966     0,978

=== Confusion Matrix ===

    a    b   <-- classified as
 1514    4 |    a = 0
   94  172 |    b = 1
```



ROC Curve for Class 0 (Nonfraudulent)

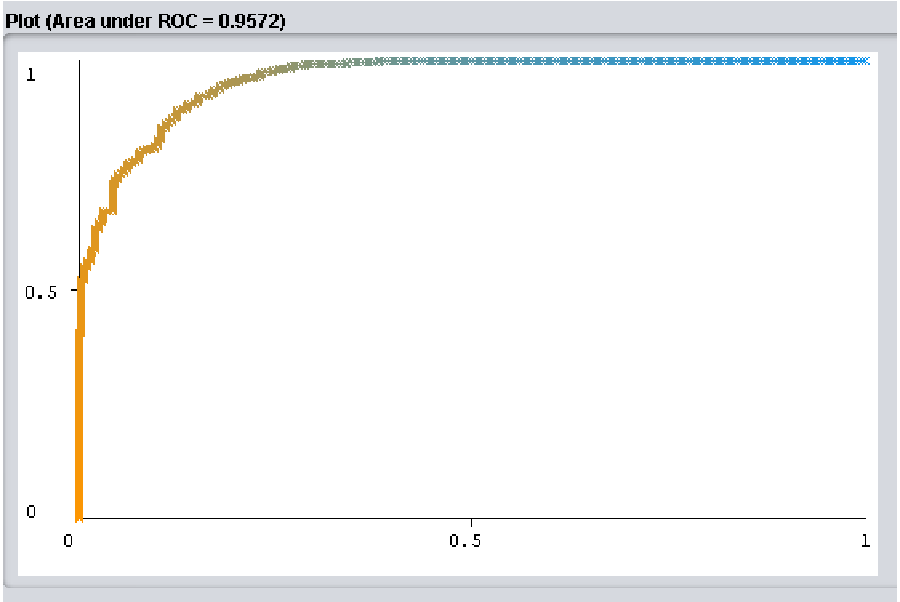


ROC Curve for Class 1 (Fraudulent)

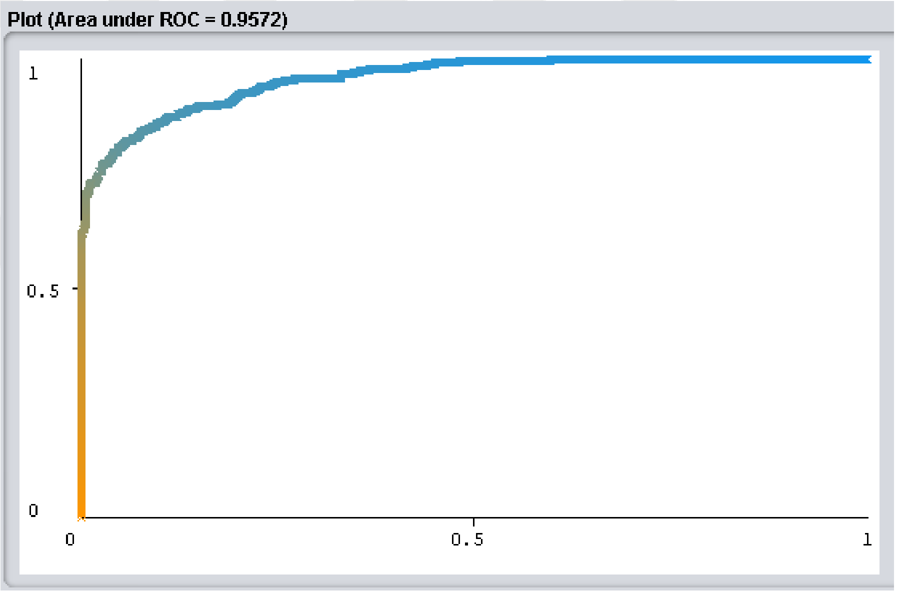
5.6. Random Forest:

Random forest decision tree based model has correctly classified the 94.57% of the instances (observations). In detail, it correctly classified 1688 instances and incorrectly classified 97 instances. 11 of incorrectly classified instances consist of nonfraudulent cases. 86 of incorrectly classified instances are consist of fraudulent instances. The ratio of incorrectly classified cases seems high in fraudulent cases, but the correctly classified fraudulent cases (192) number is still better than incorrectly classified instances.

```
=== Summary ===

Correctly Classified Instances        1688               94.5658 %
Incorrectly Classified Instances        97                5.4342 %
Kappa statistic                          0.7678
Mean absolute error                      0.1687
Root mean squared error                  0.236
Relative absolute error                 65.8506 %
Root relative squared error             65.0662 %
Total Number of Instances             1785

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0,993    0,309    0,946      0,993    0,969      0,780    0,957     0,991     0
                 0,691    0,007    0,946      0,691    0,798      0,780    0,957     0,881     1
Weighted Avg.    0,946    0,262    0,946      0,946    0,942      0,780    0,957     0,974

=== Confusion Matrix ===

    a    b   <-- classified as
 1496   11 |    a = 0
   86  192 |    b = 1
```



ROC Curve for Class 0 (Nonfraudulent)



ROC Curve for Class 1 (Fraudulent)

6. CONCLUSION

In a broad perspective, fraudulent activities of people play a significant role in the history of humanity. Countless negative events in human history have deceptive people behind it. Those devious human beings have been betraying civilizations, empires, and communities throughout history. In a narrow perspective, the modern economic system attracts people who have the same crooked characteristics as those in history. Nowadays, those deceptive individuals in history change their style and today, corporates host the most notorious fraudsters within the organization. Governments and various organizations try to protect citizens from the effects of such fraudsters. Because, in the modern economic world, economic boundaries are removed and economies are engaged with each other. A single company's bankruptcy can trigger a worldwide financial crisis and cause billions of dollar worth financial damage to the commonwealth of people. Besides, some of the previously mentioned financial crisis are triggered by corporate bankruptcies. These bankruptcies are not only related to economic struggles but mostly related to the fraudulent managerial decisions of top management teams. Stakeholders of the economic system try to understand the aspects of fraudulent activities.

Early age of financial fraud research mostly focused on the psychological side of fraudulent activities. Researchers try to understand fraud through surveys and interviews with fraudsters. Later on, statistical methods arouse and became the generally accepted method for fraud literature. However, I believe that the rapidly rising machine learning based algorithms have a promising future for financial fraud related literature.

This study tries to contribute to financial fraud and machine learning implementation literature. By combining these two fields, this study aims to help to minimize the risk exposures of investors and stakeholders due to fraudulent financial activities of companies. On the other hand, this study also aims to lower the risk of material misstatements through continuous evaluation. Four machine learning based classification algorithms are benchmarked to evaluate the most outperforming one. All of them have decent and consistent results, but, C4.5 and Random Forest algorithms outperform the other two algorithms.

Additionally, the first time in the literature, a comprehensive set of macroeconomic indicators are included in a machine learning based financial fraud prediction model.

Macroeconomic indicators are collected from the Global Competitiveness Index of the World Economic Forum to capture the essence of macroeconomic indicators in practice and in daily life. Nearly all of the macroeconomic indicators have a statistically significant effect on each algorithm. Moreover, the algorithm and dataset structure in this study classify fraudulent and nonfraudulent companies better than previous financial fraud prediction related corpus. This shows that the financial fraud prediction researchers should not only focus on company-specific data but also focus on macroeconomic indicators to construct robust and comprehensive prediction tool.

The developed model can be beneficial for regulatory bodies and beneficial for other stakeholders like banks, individual investors, investment funds, and companies. Commercial banks are started to develop several ANN based algorithms for credit risk evaluation (Angelini et al., 2008). Audit companies can also benefit from the developed algorithm as auditor’s decision aid tool. In general, auditing firms adopt a strategic systems approach or transaction focused approach to evaluate the risk of material misstatement (Schultz et al., 2010). This research will enlarge the audit companies’ evaluation procedures for the risk of material misstatement. Additionally, auditor’s trust-based relationship with company managers can affect managerial fraud evaluation (Kerler & Killough, 2009). An emotionally indifferent algorithm will reduce the risk of biased fraud assessment.

References

Acharya, V. V., & Lambrecht, B. M. (2015). A Theory of Income Smoothing When Insiders Know More Than Outsiders. *The Review of Financial Studies, 28*(9), 2534-2574.

Agarwal, V., & Taffler, R. (2008). Comparing the performance of market-based and accounting-based bankruptcy prediction models. *Journal of Banking & Finance, 32*(8), 1541-1551.

Agoglia, C. P., Doupnik, T. S., & Tsakumis, G. T. (2011). Principles-Based versus Rules-Based Accounting Standards: The Influence of Standard Precision and Audit Committee Strength on Financial Reporting Decisions. *The Accounting Review, 86*(3), 747-767.

Agrawal, A., & Chadha, S. (2005). Corporate Governance and Accounting Scandals. *The Journal of Law & Economics, 48*(2), 371-406.

Agrawal, A., & Knoeber, C. R. (2012). Corporate Governance and Firm Performance . In C. R. Thomas, & F. S. William, *Oxford Handbook in Managerial Economics.* New York: Oxford University Press.

Ahmed, A. S., McAnally, M. L., Rasmussen, S., & Weaver, C. D. (2010). How costly is the Sarbanes Oxley Act? Evidence on the effects of the Act on corporate profitability. *Journal of Corporate Finance*, *16*(3), 352-369.

Albrecht, A., Mauldin, E. G., & Newton, N. J. (2018). Do Auditors Recognize the Potential Dark Side of Executives’ Accounting Competence? *The Accounting Review, 93*(6), 1-28.

Alchian, A. A., & Demsetz, H. (1972). Production, Information Costs, and Economic Organization. *The American Economic Review, 62*(5), 777-795.

Ali, S., & Smith, K. A. (2006). On Learning Algorithm Selection for Classification. *Applied Soft Computing, 6*(2), 119-138.

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

Amiram, D., Bozanic, Z., Cox, J. D., Dupont, Q., Karpoff, J. M., & Sloan, R. (2018). Financial Reporting Fraud and Other Forms of Misconduct: A Multidisciplinary Review of the Literature. *Review of Accounting Studies, 23*(2), 732-783.

Amor, N. B., Benferhat, S., & Elouedi, Z. (2004). Naive Bayes vs Decision Trees in Intrusion Detection Systems. *Proceedings of the 2004 ACM symposium on Applied computing* , (pp. 420-424). Nicosia.

Angelini, E., Tollo, G. d., & Roli, A. (2008). A Neural Network Approach for Credit Risk Evaluation. *The Quarterly Review of Economics and Finance, 48*(4), 733-755.

Arena, M., & Julio, B. (2015). The Effects of Securities Class Action Litigation on Corporate Liquidity and Investment Policy. *Journal of Financial an Quantitative Analysis, 50*, 251-275.

Armstrong, C. S., Barth, M. E., Jagolinzer, A. D., & Riedl, E. (2010). Market Reaction to the Adoption of IFRS in Europe. *The Accounting Review, 85*(1), 31-61.

Arora, A., & Sharma, C. (2016). Corporate Governance and Firm Performance in Developing Countries: Evidence from India. *Corporate Governance, 16*(2), 420-436.

Asthana, S., Balsam, S., & Kim, S. (2009). The Effect of Enron, Andersen, and Sarbanes-Oxley on the US Market for Audit Services. *Accounting Research Journal, 22*(1), 4-26.

Bahrammirzaee, A. (2010). A Comparative Survey of Artificial Intelligence Applications in Finance: Artificial Neural Networks, Expert System and Hybrid Intelligent Systems. *Neural Computing and Applications, 19*(8), 1165-1195.

Baker, M., & Wurgler, J. (2002). Market Timing and Capital Structure. *The Journal of Finance, 57*(1), 1-32.

Bargeron, L. L., Lehn, K. M., & Zutter, C. J. (2010). Sarbanes-Oxley and corporate risk-taking. *Journal of Accounting and Economics, 49*, 34-52.

Baskin, J. B. (1988). The Development of Corporate Financial Markets in Britain and the United States, 1600-1914: Overcoming Asymmetric Information. *The Business History Review, 62*(2), 199-237.

Basu, J. K., Bhattacharyya, D., & Kim, T.-h. (2010). Use of Artificial Neural Network in Pattern Recognition. *International Journal of Software Engineering and Its Applications, 4*(2), 23-34.

Battiti, R. (1992). First- and Second-Order Methods for Learning: Between Steepest Descent and Newton's Method. *Neural Computation, 4*(2), 141-166.

Bauer, R., Guenster, N., & Otten, R. (2004). Empirical Evidence on Corporate Governance in Europe: The Effect on Stock Returns, Firm Value and Performance. *Journal of Asset Management, 5*(2), 91-104.

Beasley, M. S. (1996). An Empirical Analysis of the Relation Between the Board of Director Composition and Financial Statement Fraud. *The Accounting Review, 71*(4), 443-465.

Beiner, S., Drobetz, W., Schmid, M. M., & Zimmermann, H. (2006). An Integrated Framework of Corporate Governance and Firm Valuation. *European Financial Management, 12*(2), 249-283.

Bell, T., Szykowny, S., & Willingham, J. (1993, April). Assessing the Likelihood of Fraudulent Financial Reporting: A Cascaded Logit Approach. *Unpublished Manuscript*.

Berger, A. N., & di Patti, E. B. (2006). Capital Structure and Firm Performance: A New Approach to Testing Agency Theory and an Application to the Banking Industry. *Journal of Banking & Finance, 30*(4), 1065-1102.

Berger, P. G., Li, F., & Wong, M. H. (2011). *The Impact of Sarbanes-Oxley on Cross-listed Companies*. Retrieved January 2019, from https://pdfs.semanticscholar.org/3f53/6161daecb8702657076e8e9d78bfd4dc54e4.pdf

Berle, A. A., & Means, G. C. (1932). *The Modern Corporation and Private Property.* New york: Harcourt, Brace & World, Inc.

Bianconi, M., Chen, R., & Yoshino, J. A. (2013). Firm value, the Sarbanes-Oxley Act and cross-listing in the U.S., Germany and Hong Kong destinations. *North American Journal of Economics and Finance, 24*, 25-44.

Birz, G., & Lott Jr., J. R. (2011). The Effect of Macroeconomic News on Stock Returns: New Evidence from Newspaper Coverage. *Journal of Banking & Finance, 35*(11), 2791-2800.

Bohanec, M., & Bratko, I. (1994). Trading Accuracy for Simplicity in Decision Trees. *Machine Learning, 15*(3), 223-250.

Boyle, D. M., DeZoort, F. T., & Hermanson, D. R. (2015). The Effect of Alternative Fraud Model Use on Auditors' Fraud Risk Judgments. *Journal of Accounting and Public Policy, 34*(6), 578-596.

Bradley, A. P. (1997). The Use of the Area Under the ROC Curve in the Evaluation of Machine Learning Algorithms. *Pattern Recognition, 30*(7), 1145-1159.

Bradshaw, M. T., & Miller, G. S. (2008). Will Harmonizing Accounting Standards Really Harmonize Accounting? Evidence from Non-U.S. Firms Adopting U.S. GAAP. *Journal of Accounting, Auditing & Finance, 23*(2), 233-263.

Breiman, L. (2001). Random Forests. *Machine Learning, 45*, 5-32.

Breton, G., & Taffler, R. J. (1995). Creative Accounting and Investment Analyst Response. *Accounting and Business Research, 25*(98), 81-92.

Broadberry, S., & Gupta, B. (2009). Lancashire, India, and Shifting Competitive Advantage in Cotton Textiles,1700–1850: the Neglected Role of Factor Prices. *The Economic History Review, 62*(2), 279-305.

Brown, L. D., & Caylor, M. L. (2004, December 7). *Corporate Governance and Firm Performance*. Retrieved February 2019, from http://dx.doi.org/10.2139/ssrn.586423

Brown, L. D., & Caylor, M. L. (2009). Corporate Governance and Firm Operating Performance. *Review of Quantitative Finance and Accounting, 32*(2), 129-144.

Bruner, D., McKee, M., & Santore, R. (2008). Hand in the Cookie Jar: An Experimental Investigation of Equity-Based Compensation and Managerial Fraud. *Southern Economic Journal, 75*(1), 261-278.

Burgstahler, D., & Eames, M. (2006). Management of Earnings and Analysts' Forecasts to Achieve Zero and Small Positive Earnings Surprises. *Journal of Business Finance & Accounting, 33*(5&6), 633-652.

Burnett, B. M., Gordon, E. A., Jorgensen, B. N., & Linthicum, C. L. (2015). Earnings Quality: Evidence from Canadian Firms'Choice between IFRS and U.S. GAAP. *Accounting Perspectives, 14*(3), 212-249.

Byard, D., Li, Y., & Yu, Y. (2011). The Effect of Mandatory IFRS Adoption on Financial Analysts' Information Environment. *Journal of Accounting Research, 49*(1), 69-96.

Carmona, S., & Trombetta, M. (2008). On the Global Acceptance of IAS/IFRS Accounting Standards: The Logic and Implications of the Principles-Based System. *Journal of Accounting and Public Policy, 27*(6), 455-461.

Caruana, R., & Niculescu-Mizil, A. (2006). An Empirical Comparison of Supervised Learning Algorithms. *ICML '06 Proceedings of the 23rd international conference on Machine learning* (pp. 161-168). New York: Association for Computing Machinery.

Chancellor, E. (2000). *Devil Take the Hindmost: A History of Financial Speculation* (1st ed.). New York: Penguin Group.

Chandrashekar, G., & Sahin, F. (2014). A Survey on Feature Selection Methods. *Computers & Electrical Engineering, 40*(1), 16-28.

Chang, J. C., & Sun, H. L. (2009). Cross-Listed Foreign Firms' Earnings Informativeness, Earnings Management and Disclosures of Corporate Governance Information Under SOX. *The International Journal of Accounting, 44*(1), 1-32.

Chapelle, O., Schölkopf, B., & Zien, A. (2006). *Semi-Supervised Learning.* Cambridge: The MIT Press.

Chartiou, A., Louca, C., & Panayides, S. (2007). Cross-Listing, Bonding Hypothesis and Corporate Governance. *Journal of Business Finance & Accounting, 34*(7-8), 1281-1306.

Chaudhuri, K. N. (1999). *The English East India Company: The Study of an Early Joint-Stock Company 1600-1640* (2nd ed.). London: Routledge.

Chava, S., & Jarrow, R. A. (2004). Bankruptcy Prediction with Industry Effects. *Review of Finance, 8*, 537-569.

Chen, G., Firth, M., Gao, D. N., & Rui, O. M. (2006). Ownership Structure, Corporate Governance, and Fraud: Evidence from China. *Journal of Corporate Finance, 12*(3), 424-448.

Chen, S. (2016). Detection of Fraudulent Financial Statements Using the Hybrid Data Mining Approach. *SpringerPlus, 5*, 1-16.

Chen, V. Z., Li, J., & Saphiro, D. M. (2011). Are OECD-Prescribed “Good Corporate Governance Practices" Really Good in an Emerging Economy? *Asia Pacific Journal of Management*(28), 115-138.

Chen, X., Cheng, Q., & Wang, X. (2015). Does Increased Board Independence Reduce Earnings Management? Evidence from Recent Regulatory Reforms. *Review of Accounting Studies, 20*(2), 899-933.

Cheung, Y.-L., Jiang, P., Limpaphayom, P., & Lu, T. (2010). Corporate Governance in China: a Step Forward. *European Financial Management, 16*(1), 94-123.

Choi, S. J. (2007). Do the Merits Matter Less After the Private Securities Litigation Reform Act? *The Journal of Law, Economics, & Organization, 23*(3), 598-626.

Choo, F., & Tan, K. (2007). An “American Dream” Theory of Corporate Executive Fraud. *Accounting Forum, 31*, 203-215.

Christen, M., Iyer, G., & Soberman, D. (2006). Job Satisfaction, Job Performance, and Effort: A Reexamination Using Agency Theory. *Journal of Marketing, 70*(1), 137-150.

Christensen, H. B., Hail, L., & Leuz, C. (2013). Mandatory IFRS Reporting and Changes in Enforcement. *Journal of Accounting and Economics, 56*(2-3), 147-177.

Claessens, S. (2006). Corporate Governance and Development. *The World Bank Research Observer*, *21*(1), 91-122.

Claessens, S., Djankov, S., & Lang, L. H. (2000). The Separation of Ownership and Control in East Asian Corporations. *Journal of Financial Economics*(58), 81-112.

Coakley, J. R., & Brown, C. E. (2000). Artificial Neural Networks in Accounting and Finance: Modeling Issues. *International Journal of Intelligent Systems in Accounting, Finance & Management, 9*(2), 119-144.

Coffee Jr, J. C. (2002). Racing Towards the Top?: The Impact of Cross Listing and Stock Market Competition on International Corporate Governance. *Columbia Law Review, 102*(7), 1757-1831.

Coffee Jr, J. C. (2005). A Theory of Corporate Scandals: Why the USA and Europe Differ. *Oxford Review of Economic Policy, 21*(2), 198-211.

Coffee Jr, J. C. (2006). Reforming the Securities Class Action: An Essay on Deterrence and Its Implementation. *Columbia Law Review, 106*(7), 1534-1586.

Cohen, D. A., Dey, A., & Lys, T. Z. (2008). Real and Accrual-Based Earnings Management in the Pre- and Post-Sarbanes-Oxley Periods. *THE ACCOUNTING REVIEW, 83*(3), 757-787.

Coleman, J. W. (1987). Toward an Integrated Theory of White-Collar Crime. *American Journal of Sociology, 93*(2), 406-439.

Conyon, M. J., & He, L. (2016). Executive Compensation and Corporate Fraud in China. *Journal of Business Ethics, 134*, 669-691.

Crain, M. A., Hopwood, W. S., Pacini, C., & Young, G. R. (2015). *Essentials of Forensic Accounting.* New York: American Institute of Certified Public Accountants.

Cressey, D. (1953). *Other People's Money; A Study of the Social Psychology of Embezzlement.* New York: Free Press.

Cressey, D. R. (1950). The Criminal Violation of Financial Trust. *American Sociological Review, 15*(6), 738-743.

Crocker, K. J., & Slemrod, J. (2007). The Economics of Earnings Manipulation and Managerial Compensation. *The RAND Journal of Economics, 38*(3), 698-713.

Crutchley, C. E., Jensen, M. R., & Marshall, B. B. (2007). Climate for Scandal: Corporate Environments that Contribute to Accounting Fraud. *The Financial Review, 42*(1), 53-73.

Dah, M. A., Frye, M. B., & Hurst, M. (2014). Board Changes and CEO Turnover: The Unanticipated Effects of the Sarbanes–Oxley Act. *Journal of Banking & Finance, 41*, 97-108.

Daske, H., Hail, L., Leuz, C., & Verdi, R. (2008). Mandatory IFRS Reporting Around the World: Early Evidence on the Economic Consequences. *Journal of Accounting Research, 46*(5), 1085-1142.

Dechow, P. M., Sloan, R. G., & Sweeney, A. P. (1995). Detecting Earnings Management. *The Accounting Review, 70*(2), 193-225.

DeFond, M. L., Hung, M., Li, S., & Li, Y. (2015). Does Mandatory IFRS Adoption Affect Crash Risk? *The Accounting Review, 90*(1), 265-299.

Dellaportas, S. (2013). Conversations with Inmate Accountants: Motivation, Opportunity and the Fraud Triangle. *Accounting Forum, 37*(1), 29-39.

Deloitte. (2018, December). *Standards*. Retrieved from Iasplus Web site: https://www.iasplus.com/en/standards/ifrs/ifrs1

Demsetz, H., & Lehn, K. (1985). The Structure of Corporate Ownership: Causes and Consequences. *Journal of Political Economy, 95*(6), 1155-1177.

Demyanyk, Y., & Hasan, I. (2010). Financial Crises and Bank Failures: A Review of Prediction Methods. *Omega, 38*(5), 315-324.

Dimitras, A. I., Zanakis, S. H., & Zopounidis, C. (1996). A Survey of Business Failures with an Emphasis on Prediction Methods and Industrial Applications. *European Journal of Operational Research, 90*(3), 487-513.

Donaldson, L., & Davis, J. H. (1991). Stewardship Theory or Agency Theory: CEO Governance and Shareholder. *Australian Journal of Management, 16*(1), 49-64.

Donegan, J. J., & Ganon, M. W. (2008). Strain, Differential Association, and Coercion: Insights from the Criminology Literature on Causes of Accountant's Misconduct. *Accounting and the Public Interest, 8*(1), 1-20.

Donelson, D. C., Ege, M. S., & McInnis, J. M. (2017). Internal Control Weaknesses and Financial Reporting Fraud. *Auditing: A Journal of Practice & Theory, 36*(3), 45-69.

Dorminey, J., Fleming, A. S., Kranacher, M.-J., & Riley, R. A. (2012). The Evolution of Fraud Theory. *Issues in Accounting Education, 27*(2), 555-579.

Drucker, P. F. (2008). *Management - Revised Edition.* HarperBusiness.

Dyck, A., Morse, A., & Zingales, L. (2010). Who Blows the Whistle on Corporate Fraud? *The Journal of Finance, 65*(6), 2213-2253.

Efendi, J., Srivastava, A., & Swanson, E. P. (2007). Why Do Corporate Managers Misstate Financial Statements? The Role of Option Compensation and Other Factors. *Journal of Financial Economics, 85*(3), 667-708.

Eisenhardt, K. M. (1989). Agency Theory: An Assessment and Review. *The Academy of Management Review, 14*(1), 57-74.

Erikson, E. (2014). *Between Monopoly and Free Trade: The English East India Company, 1600-1757.* New Jersey: Princeton University Press.

Fama, E. F. (1965). The Behavior of Stock-Market Prices. *The Journal of Business, 38*(1), 34-105.

Fama, E. F. (1980). Agency Problems and the Theory of the Firm. *Journal of Political Economy, 88*(2), 288-307.

Fama, E. F., & French, K. R. (1992). The Cross-Section of Expected Stock Returns. *The Journal of Finance, 47*(2), 427-465.

Fama, E. F., & Jensen, M. C. (1983). Separation of Ownership and Control. *Journal of Law and Economics, 26*, 301-325.

Fan, J. P., Titman, S., & Twite, G. (2012). An International Comparison of Capital Structure and Debt Maturity Choices. *Journal of Financial and Quantitative Analysis, 47*(1), 23-56.

Fanning, K., & Cogger, K. O. (1998). Neural Networks Detection of Management Fraud Using Published Financial Data. *International Journal of Intelligent Systems in Accounting, Finance and Management, 7*(1), 21-41.

Fanning, K., Cogger, K. O., & Srivastava, R. (1995). Detection of Management Fraud: A Neural Network Approach. *Intelligent Systems in Accounting, Finance and Management, 4*(2), 113-126.

Farber, D. B. (2005). Restoring Trust after Fraud: Does Corporate Governance Matter? *The Accounting Review, 80*(2), 539-561.

Fleming, G., Heaney, R., McCosker, & Rochelle. (2005). Agency Costs and Ownership Structure in Australia. *Pacific-Basin Finance Journal, 13*(1), 29-52.

Frank, E., & Witten, I. H. (2004, March). *Data Mining.* Retrieved from http://phdies.ing.unisi.it/corsi/matdid/39_extra-slides-day-2.pdf

Frank, M. Z., & Goyal, V. K. (2003). Testing the Pecking Order Theory of Capital Structure. *Journal of Financial Economics, 67*(2), 217-248.

Franses, P. H., & van Dijk, D. (2000). *Nonlinear Time Series Models in Empirical Finance.* Cambridge: Cambridge University Press.

Freeman, R. E. (1994). The Politics of Stakeholder Theory: Some Future Directions. *Business Ethics Quarterly, 4*(4), 409-421.

Gao, F., Wu, J. S., & Zimmerman, J. (2009). Unintended Consequences of Granting Small Firms Exemptions from Securities Regulation: Evidence from the Sarbanes-Oxley Act. *Journal of Accounting Research, 47*(2), 459-506.

Gelderblom, O., & Jonker, J. (2004). Completing a Financial Revolution: The Finance of the Dutch East India Trade and the Rise of the Amsterdam Capital Market, 1595–1612. *The Journal of Economic History, 64*(3), 641-672.

Gençay, R., Selçuk, F., & Whitcher, B. (2002). *An Introduction to Wavelets and Other Filtering Methods in Finance and Economics.* San Diego: Academic Press.

Ghahramani, Z. (2006). Information Theory. In *Encyclopedia of Cognitive Science* (pp. 1-5). John Wiley & Sons, Ltd.

Gilson, R. J., & Gordon, J. N. (2013). The Agency Costs of Agency Capitalism: Activist Investors and the Revaluation of Governance Rights. *Columbia Law Review, 113*, 863-928.

Giroux, G. (2008). What Went Wrong? Accounting Fraud and Lessons from the Recent Scandals. *Social Research, 75*(4), 1205-1238.

Gisler, M., & Sornette, D. (2010). Bubbles Everywhere in Human Affairs. *Swiss Finance Institute Research Paper Series*. Swiss Finance Institute.

Gottschalk, P. (2010). Categories of Financial Crime. *Journal of Financial Crime, 17*(4), 441-458.

Gowthorpe, C., & Amat, O. (2005). Creative Accounting: Some Ethical Issues of Macro- and Micro-Manipulation. *Journal of Business Ethics, 57*(1), 55-64.

Grant, J., Markarian, G., & Parbonetti, A. (2009). CEO Risk-Related Incentives and Income Smoothing. *Contemporary Accounting Research, 26*(4), 1029-1065.

Graupe, D. (2013). *Principles of Artificial Neural Networks.* World Scientific.

Green, B. P., & Choi, J. H. (1997). Assessing the Risk of Management Fraud Through Neural Network Technology. *Auditing: A Journal of Practice & Theory, 16*(1), 14-28.

Griffin, P. A., & Lont, D. H. (2007). An Analysis of Audit Fees Following the Passage of Sarbanes-Oxley. *Asia-Pacific Journal of Accounting & Economics, 14*(2), 161-192.

Guresen, E., Kayakutlu, G., & Daim, T. U. (2011). Using Artificial Neural Network Models in Stock Market Index Prediction. *Expert Systems with Applications, 38*(8), 10389-10397.

Hail, L., & Leuz, C. (2009). Cost of Capital Effects and Changes in Growth Expectations Around U.S. Cross-Listings. *Journal of Financial Economics, 93*(3), 428-454.

Hajek, P., & Henriques, R. (2017). Mining Corporate Annual Reports for Intelligent Detection of Financial Statement Fraud – A Comparative Study of Machine Learning Methods. *Knowledge-Based Systems, 128*, 139-152.

Hall, M. (2019, January). *Class InfoGainAttributeEval*. Retrieved from http://weka.sourceforge.net/doc.dev/weka/attributeSelection/InfoGainAttributeEval.html

Hall, M., & Holmes, G. (2002). Benchmarking Attribute Selection Techniques for Discrete Class Data Mining. *Working Paper*.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data mining, inference, and prediction.* Springer.

Hejeebu, S. (2005). Contract Enforcement in the English East India Company. *The Journal of Economic History, 65*(2), 496-523.

Hill, C. W., & Jones, T. M. (1992). Stakeholder-Agency Theory. *Journal of Management Studies, 29*(2), 131-154.

Hochberg, Y. V., Sapienza, P., & Jorgensen-Vissing, A. (2009). A Lobbying Approach to Evaluatingthe Sarbanes-Oxley Act of 2002. *Journal of Accounting Research, 47*(2), 519-583.

Hope, O.-K., Kang, T., & Kim, J. W. (2013). Voluntary Disclosure Practices by Foreign Firms Cross-Listed in the United States. *Journal of Contemporary Accounting & Economics, 9*(1), 50-66.

Hope, O.-K., Kang, T., & Zang, Y. (2007). Bonding to the Improved Disclosure Environment in the US: Firms' Listing Choices and their Capital Market Consequences. *Journal of Contemporary Accounting & Economics, 3*(1), 1-33.

Hoppit, J. (2002). The Myths of the South Sea Bubble. *Transactions of the Royal Historical Society, 12*, 141-165.

Horton, J., Serafeim, G., & Serafeim, I. (2013). Does Mandatory IFRS Adoption Improve the Information Environment? *Contemporary Accounting Research, 30*(1), 388-423.

Huang, W., Jiang, F., Liu, Z., & Zhang, M. (2011). Agency Cost, Top Executives' Overconfidence, and Investment-Cash Flow Sensitivity — Evidence from Listed Companies in China. *Pacific-Basin Finance Journal, 19*(3), 261-277.

Huijgen, C., & Lubberink, M. (2005). Earnings Conservatism, Litigation and Contracting: The Case of Cross-Listed Firms. *Journal of Business Finance & Accounting, 32*(7&8), 1275-1309.

Hypko, P., Tilebein, M., & Gleich, R. (2010). Benefits and Uncertainties of Performance-Based Contracting in Manufacturing Industries: An Agency Theory Perspective. *Journal of Service Management, 21*(4), 460-489.

IAASB. (2018, November). *Publications.* Retrieved from Handbook of International Quality Control, Auditing, Review, Other Assurance, and Related Services Pronouncements: https://www.ifac.org/system/files/publications/files/IAASB-2018-HB-Vol-1.pdf

IFAC. (2019, February). *The Clarified Standards.* Retrieved from IFAC Web Site: https://www.iaasb.org/clarity-center/clarified-standards

IFRS. (2019, January). *About Us: IFRS*. Retrieved from IFRS Web site: https://www.ifrs.org/about-us/

Jain, P. K., Kim, J.-C., & Rezaee, Z. (2008). The Sarbanes-Oxley Act of 2002 and Market Liquidity. *The Financial Review*, *43*(3), 361-382.

Jamal, K., & Tan, H.-T. (2010). Joint Effects of Principles-Based versus Rules-Based Standards and Auditor Type in Constraining Financial Managers' Aggressive. *The Accounting Review, 85*(4), 1325-1346.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning with Applications in R.* Springer.

James, K. L. (2003). The Effects of Internal Audit Structure on Perceived Financial Statement Fraud Prevention. *Accounting Horizons, 17*(4), 315-327.

Jeanjean, T., & Stolowy, H. (2008). Do Accounting Standards Matter? An Exploratory Analysis of Earnings Management Before and After IFRS Adoption. *Journal of Accounting and Public Policy, 27*(6), 480-494.

Jensen, M. C., & Meckling, W. H. (1976). Theory of the Firm: Managerial Behavior, Agency Cost and Ownership Structure. *Journal of Financial Economics, 3*(4), 305-360.

Jesover, F., & Kirkpatrick, G. (2005). The Revised OECD Principles of Corporate Governance and their Relevance to Non-OECD Countries. *Corporate Governance: An International Review, 13*(2), 127-136.

Johnstone, P. (1998). Serious White Collar Fraud: Historical and Contemporary Perspectives. *Crime, Law and Social Change*, *30*(2), 107-130.

Jones, M. (2011). *Creative Accounting, Fraud and International Accounting Scandals.* Wiley.

Kaastra, I., & Boyd, M. (1996). Designing a Neural Network for Forecasting Financial and Economic Time Series. *Neurocomputing*, *10*(3), 215-236.

Kapp, L. A., & Heslop, G. (2011). Protecting Small Businesses from Fraud. *The CPA Journal, 81*(10), 62-67.

Karegowda, A. G., Manjunath, A. S., & Jayaram, M. A. (2010). Comparative Study of Attribute Selection Using Gain Ratio and Correlation Based Future Selection. *International Journal of Information Technology and Knowledge Management*, *2*(2), 271-277.

Karpoff, J. M., Koester, A., Lee, D. S., & Martin, G. S. (2017). Proxies and Databases in Financial Misconduct Research. *The Accounting Review, 92*(6), 129-163.

Karpoff, J. M., Lee, D. S., & Martin, G. S. (2008). The Cost to Firms of Cooking the Books. *Journal of Financial and Quantitative Analysis, 43*(3), 581-611.

Kerler, W. A., & Killough, L. N. (2009). The Effects of Satisfaction with a Client's Management During a Prior Audit Engagement, Trust, and Moral Reasoning on

Auditors’ Perceived Risk of Management Fraud. *Journal of Business Ethics*, *85*(2), 109-136.

Khashei, M., & Bijari, M. (2010). An Artificial Neural Network (p, d, q) Model for Timeseries Forecasting. *Expert Systems with Applications, 37*(1), 479-489.

Kindleberger, C. P., & Aliber, R. Z. (2011). *Manias, Panics and Crashes: A History of Financial Crises.* New York: Palgrave Macmillan.

Kingma, D. P., Mohamed, S., Rezende, D. J., & Welling, M. (2014). Semi-supervised Learning with Deep Generative Models. *Advances in Neural Information Processing Systems.*

Kirkos, E., Spathis, C., & Manolopoulos, Y. (2007). Data Mining Techniques for the Detection of Fraudulent Financial Statements. *Expert Systems with Applications, 32*(4), 995-1003.

Klapper, L. F., & Love, I. (2002). Working Paper. *Corporate Governance, Investor Protection, and Performance in Emerging Markets*. The World Bank.

Kleer, R. (2012). ‘The Folly of Particulars’: the Political Economy of the South Sea Bubble. *Financial History Review, 19*(2), 175-197.

Koh, H. C., & Low, C. K. (2004). Going Concern Prediction Using Data Mining Techniques. *Managerial Auditing Journal, 19*(3), 462-476.

Koppell, J. G. (2011). Shareholder Advocacy and the Development of the Corporation: The Timeless Dilemmas of an Age-old Solution. In J. G. Koppell, *Origins of Shareholder Advocacy* (pp. 1-28). New York: Palgrave Macmillan.

Krauss, C., Do, X. A., & Huck, N. (2017). Deep Neural Networks, Gradient-boosted Trees, Random Forests: Statistical Arbitrage on the S&P 500. *European Journal of Operational Research, 259*(2), 689-702.

Kuhn, T. S. (1962). *The Structure of Scientific Revolutions.* Chicago and London: University of Chicago Press.

Kulkarni, V. Y., & Sinha, P. K. (2012). Pruning of Random Forest classifiers: A survey and future directions. *2012 International Conference on Data Science & Engineering (ICDSE)* (pp. 64-68). IEEE.

La Porta, R., Lopez-De-Silanes, F., & Shleifer, A. (1999). Corporate Ownership Around the World. *The Journal of Finance, 54*(2), 471-517.

Lambert, R., Leuz, C., & Verrecchia, R. E. (2007). Accounting Information, Disclosure, and the Cost of Capital. *Journal of Accounting Research, 45*(2), 385-420.

Lan, L. L., & Heracleous, L. (2010). Rethinking Agency Theory: The View From Law. *The Academy of Management Review, 35*(2), 294-314.

Lawson, P. (1993). *The East India Company.* London: Routledge.

Leuz, C. (2006). Cross Listing, Bonding and Firms' Reporting Incentives: A Discussion of Lang, Raedy and Wilson (2006). *Journal of Accounting and Economics, 42*(1-2), 285-299.

Leuz, C. (2007). Was the Sarbanes–Oxley Act of 2002 Really This Costly? A Discussion of Evidence from Event Returns and Going-private Decisions. *Journal of Accounting and Economics, 44*(1-2), 146-165.

Leuz, C., Nanda, D., Wysocki, & D, P. (2003). Earnings Management and Investor Protection: An International Comparison. *Journal of Financial Economics, 69*(3), 505-527.

Li, H., Pincus, M., & Rego, S. O. (2008). Market Reaction to Events Surrounding the Sarbanes-Oxley Act of 2002 and Earnings Management. *The Journal of Law and Economics, 51*(1), 111-134.

Lin, C.-C., Chiu, A.-A., Huang, S. Y., & Yen, D. C. (2015). Detecting the Financial Statement Fraud: The Analysis of the Differences Between Data Mining Techniques and Experts' Judgments. *Knowledge-Based Systems, 89*(November), 459-470.

Lin, J. W., Hwang, M. I., & Becker, J. D. (2003). A Fuzzy Neural Network for Assessing the Risk of Fraudulent Financial Reporting. *Managerial Auditing Journal, 18*(8), 657-665.

Linck, J. S., Netter, J. M., & Yang, T. (2009). The Effects and Unintended Consequences of the Sarbanes-Oxley Act on the Supply and Demand for Directors. *The Review of Financial Studies, 22*(8), 3287-3328.

Linthicum, C., Reitenga, A. L., & Sanchez, J. M. (2010). Social Responsibility and Corporate Reputation: The Case of the Arthur Andersen Enron Audit Failure. *Journal of Accounting and Public Policy, 29*(2), 160-176.

Liu, J. (2004). Macroeconomic Determinants of Corporate Failures: Evidence from the UK. *Applied Economics, 36*(9), 939-945.

Loureiro, G. (2010). The Reputation of Underwriters: A Test of the Bonding Hypothesis. *Journal of Corporate Finance, 16*(4), 516-532.

Malkiel, B. G. (2012). Bubbles in Asset Prices. In D. C. Mueller, *The Oxford Handbook of Capitalism* (pp. 405-425). New York: Oxford University Press.

Malkiel, B. G., & Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance, 25*(2), 383-417.

McCulloch, W. S., & Pitts, W. (1943). A Logical Calculus of the Ideas Immanent in Nervous Activity. *The Bulletin of Mathematical Biophysics, 5*(4), 115-133.

McKnight, P. J., & Weir, C. (2009). Agency Costs, Corporate Governance Mechanisms and Ownership Structure in Large UK Publicly Quoted Companies: A Panel Data Analysis. *The Quarterly Review of Economics and Finance, 49*(2), 139-158.

Miller, G. S. (2006). The Press as a Watchdog for Accounting Fraud. *Journal of Accounting Research, 44*(5), 1001-1033.

Mitnick, B. M. (1975). The Theory of Agency: The Policing "Paradox" and Regulatory Behavior. *Public Choice, 24*(Winter), 27-42.

Modigliani, F., & Miller, M. H. (1958). The Cost of Capital, Corporation Finance and the Theory of Investment. *The American Economic Review, 48*(3), 261-297.

Morales, J., Gendron, Y., & Guénin-Paracini, H. (2014). The Construction of the Risky Individual and Vigilant Organization: A Genealogy of the Fraud Triangle. *Accounting, Organizations and Society, 39*(3), 170-194.

Mui, G., & Mailley, J. (2015). A Tale of Two Triangles: Comparing the Fraud Triangle with Criminology's Crime Triangle. *Accounting Research Journal, 28*(1), 45-58.

Murphy, P. R., & Dacin, M. T. (2011). Psychological Pathways to Fraud: Understanding and Preventing Fraud in Organizations. *Journal of Business Ethics, 101*(4), 601-618.

Myers, S. C. (1984). The Capital Structure Puzzle. *The Journal of Finance, 39*(3), 575-592.

Nasdaq. (2019). *Artificial Intelligence: Industry Report and Investment Case.* Retrieved from MarketInsite: https://business.nasdaq.com/marketinsite/2019/GIS/Artificial-Intelligence-Industry-Report-and-Investment-Case.html

Ngai, E. W., Hu, Y., Wong, Y. H., Chen, Y., & Sun, X. (2011). The Application of Data Mining Techniques in Financial Fraud Detection: A Classification Framework and an Academic Review of Literature. *Decision Support Systems*, *50*(3), 559-569.

Nolder, C. J., & Kadous, K. (2018). Grounding the Professional Skepticism Construct in Mindset and Attitude Theory: A Way Forward. *Accounting, Organizations and Society, 67*, 1-14.

Odom, M. D., & Sharda, R. (1990). A Neural Network Model for Bankruptcy Prediction. *International Joint Conference on Neural Networks,* (pp. 163-168). San Diego.

OECD. (2015). *G20/OECD Principles of Corporate Governance.* Ankara: OECD.

OECD. (2018). *Private Equity Investment in Artificial Intelligence.* OECD.

Oshiro, T. M., Perez, P. S., & Baranauskas, J. A. (2012). How many trees in a random forest? *International workshop on machine learning and data mining in pattern recognition* (pp. 154-168). Berlin: Springer.

Pagano, M., & Röell, A. (1998). The Choice of Stock Ownership Structure: Agency Costs, Monitoring, and the Decision to Go Public. *The Quarterly Journal of Economics, 113*(1), 187-225.

Pal, S. K., & Mitra, S. (1992). Multilayer Perceptron, Fuzzy Sets, and Classification. *Transactions on Neural Networks, 3*(5), 683-697.

Panchal, G., Ganatra, A., Kosta, Y. P., & Panchal, D. (2011). Behaviour Analysis of Multilayer Perceptrons with Multiple Hidden Neurons and Hidden Layers. *International Journal of Computer Theory and Engineering, 3*(2), 332-337.

Patterson, E. R., & Smith, J. R. (2007). The Effects of Sarbanes-Oxley on Auditing and Internal Control Strength. *The Accounting Review, 82*(2), 427-455.

PCAOB. (2018, January). *Rulemaking.* Retrieved from PCAOB Web site: https://pcaobus.org/Rulemaking/Docket%20026/Release_2010-004_Risk_Assessment.pdf

PCAOB. (2019, March). *Standards*. Retrieved from PCAOB Web site: https://pcaobus.org/Standards/Pages/default.aspx

Pencina, M. J., D' Agostino Sr, R. B., D' Agostino Jr, R. B., & Vasan, R. S. (2008). Evaluating the Added Predictive Ability of a New Marker: From Area Under the ROC Curve to Reclassification and Beyond. *Statistics in Medicine, 27*(2), 157-172.

Peni, E., & Vahamaa, S. (2012). Did Good Corporate Governance Improve Bank Performance during the Financial Crisis? *Journal of Financial Services Research, 41*(1-2), 19-35.

Pepper, A., & Gore, J. (2015). Behavioral Agency Theory: New Foundations for Theorizing About Executive Compensation. *Journal of Management, 41*(4), 1045-1068.

Perols, J. L., & Lougee, B. A. (2011). The Relation Between Earnings Management and Financial Statement Fraud. *Advances in Accounting, incorporating Advances in International Accounting, 27*(1), 39-53.

Peters, S., Miller, M., & Kusyk, S. (2011). How Relevant is Corporate Governance and Corporate Social Responsibility in Emerging Markets? *The International Journal of Business in Society, 11*(4), 429-445.

Petram, L. (2014). *The World's First Stock Exchange.* New York: Columbia University Press.

Pıckett, K. H., & Pickett, J. (2002). *Financial Crime Investigation and Control.* New York: John Wiley and Sons.

Piramuthu, S. (2003). On Learning to Predict Web Traffic. *Decision Support Systems, 35*(2), 213-229.

Ponulak, F., & Kasiński, A. (2010). Supervised Learning in Spiking Neural Networks with ReSuMe: Sequence Learning, Classification, and Spike Shifting. *Neural Computation, 22*(2), 467-510.

Porta, R. L., Lopez-de-Silanes, F., Shleifer, A., & Vishny, R. (2000). Investor Protection and Corporate Governance. *Journal of Financial Economics*(58), 3-27.

Quadackers, L., Groot, T., & Wright, A. (2014). Auditors' Professional Skepticism: Neutrality versus Presumptive Doubt. *Contemporary Accounting Research, 31*(3), 639-657.

Quinlan, J. R. (1983). Learning Efficient Classification Procedures and their Application to Chess End Games. In R. S. Michalski, J. G. Carbonell, & T. M. Mitchell, *Machine Learning: An Artificial Intelligence Approach* (pp. 463-481). Springer-Verlag Berlin Heidelberg GmbH.

Quinlan, J. R. (1986). Induction of Decision Trees. *Machine Learning, 1*(1), 81-106.

Quinlan, J. R. (1987). Simplifying Decision Trees. *International Journal of Man-Machine Studies, 27*(3), 221-234.

Quinlan, J. R. (1993). *C4.5: Programs for Machine Learning.* Morgan Kaufmann Publishers, Inc .

Rajan, R. G., & Zingales, L. (1995). What Do We Know About Capital Structure? Some Evidence From International Data. *The Journal of Finance, 50*(5), 1421-1460.

Ratsch, G., Onoda, T., & Muller, K.-R. (2001). Soft Margins for AdaBoost. *Machine Learning, 42*(3), 287-320.

Ravisankar, P., Ravi, V., Raghava Rao, G., & Bose, I. (2011). Detection of Financial Statement Fraud and Feature Selection Using Data Mining Techniques. *Decision Support Systems, 50,* 491-500.

Reese, W. A., & Weisbach, M. S. (2002). Protection of Minority Shareholder Interests, Cross-Listings in the United States, and Subsequent Equity Offerings. *Journal of Financial Economics, 66*(1), 65-104.

Reshid, A. (2016). Managerial Ownership and Agency Cost: Evidence from Bangladesh. *Journal of Business Ethics, 137*(3), 609-621.

Rezaee, Z. (2005). Causes, Consequences, and Deterrence of Financial Statement Fraud. *Critical Perspectives on Accounting, 16*(3), 277-298.

Rijsenbilt, A., & Commandeur, H. (2013). Narcissus Enters the Courtroom: CEO Narcissism and Fraud. *Journal of Business Ethics, 117*(2), 413-429.

Robins, N. (2017). *The Corporation That Change The World.* London: Pluto Press.

Rokach, L., & Maimon, O. (2015). *Data Mining with Decision Trees: Theory and Applications.* World Scientific Publishing Co. Pte. Ltd.

Romano, R. (2005). The Sarbanes-Oxley Act and the Making of Quack Corporate Governance. *The Yale Law Journal, 114*(7), 1521-1611.

Ross, S. A. (1973). The Economic Theory of Agency: The Principal's Problem. *The American Economic Review, 63*(2), 134-139.

Rubasundram, G. A. (2015). Perceived "Tone From The Top" During A Fraud Risk Assessment . *Procedia Economics and Finance* (pp. 102-106). Oxford: Elsevier.

Ruggieri, S. (2002). Efficient C4.5. *IEEE Transactions on Knowledge and Data Engineering, 14*(2), 438-444.

SCAC. (2019, January). *About Us*. Retrieved from Securities Class Action Clearinghouse: http://securities.stanford.edu/about-the-scac.html#about

Schapire, R. E. (2013). Explaining AdaBoost. In B. Schölkopf, Z. Luo, & V. Vovk, *Empirical Inference* (pp. 37-52). Berlin: Springer.

Schmidhuber, J. (2015). Deep Learning in Neural Networks: An Overview. *Neural Network, 61*(January), 85-117.

Schrand, C. M., & Zechman, S. L. (2012). Executive Overconfidence and the Slippery Slope to Financial Misreporting. *Journal of Accounting and Economics, 53*(1-2), 311-329.

Schultz Jr., J. J., Bierstaker, J. L., & O'Donnell, E. (2010). Integrating Business Risk into Auditor Judgment About the Risk of Material Misstatement: The Influence of a Strategic-Systems-Audit Approach. *Accounting, Organizations and Society, 35*(2), 238-251.

Shapiro, S. P. (2005). Agency Theory. *Annual Review of Sociology, 31*, 263-284.

Sharma, V. D. (2004). Board of Director Characteristics, Institutional Ownership, and Fraud: Evidence from Australia. *Auditing: A Journal of Practice & Theory, 23*(2), 105-117.

Shi, W., Connelly, B. L., & Hoskisson, R. E. (2017). External Corporate Governance and Financial Fraud: Cognitive Evaluation Theory Insights on Agency Theory Prescriptions. *Strategic Management Journal, 38,* 1268-1286.

Shleifer, A., & Vishny, R. W. (1997). A Survey of Corporate Governance. *The Journal of Finance, 52*, 737-783.

Siems, M. M., & Alvarez-Macotela, O. (2014). The OECD Principles of Corporate Governance in Emerging Markets: A Successful Example of Networked Governance? In M. Fenwick, S. V. Uytsel, & S. Wrbka, *Networked Governance, Transnational Business and the Law* (pp. 257-284). Springer.

Silver, D. H. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature, 529*, 484-489.

Skousen, C. J., & Wright, C. J. (2006, August 24). *Contemporaneous Risk Factors and the Prediction of Financial Statement Fraud.* Retrieved from SSRN: https://ssrn.com/abstract=938736

Smith, A. (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations.* London.

Smith, G. P. (2012). Google Internet Search Activity and Volatility Prediction in the Market for Foreign Currency. *Finance Research Letters, 9*(2), 103-110.

Soderstrom, N. S., & Jialin Sun, K. (2007). IFRS Adoption and Accounting Quality: A Review. *European Accounting Review, 16*(4), 675-702.

Songini, L., & Gnan, L. (2015). Family Involvement and Agency Cost Control Mechanisms in Family Small and Medium-Sized Enterprises. *Journal of Small Business Management, 53*(3), 748-779.

Sornette, D. (2003). Critical Market Crashes. *Physics Reports, 378*(1), 1-98.

Spence, M., & Zeckhauser, R. (1971). Insurance, Information, and Individual Action. *The American Economics Review, 61*(2), 380-387.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research, 15*, 1929-1958.

Summers, S. L., & Sweeney, J. T. (1998). Fraudulently Misstated Financial Statements and Insider Trading: An Empirical Analysis. *The Accounting Review, 73*(1), 131-146.

Sun, J., Cahan, S. F., & Emanuel, D. (2011). How Would the Mandatory Adoption of IFRS Affect the Earnings Quality of US Firms? Evidence from Cross-listed Firms in the US. *Accounting Horizons, 25*(4), 837-860.

Sun, P., & Zhang, Y. (2006). Is There Penalty For Crime? Corporate Scandal and Management Turnover in China. *EFA 2006 Zurich Meetings*, (pp. 1-54). Zurich.

Sutherlend, E. H. (1940). White-Collar Criminality. *American Sociological Review, 5*(1), 1-12.

Tam, K. Y., & Kiang, M. (1990). Predicting Bank Failures: A Neural Network Approach. *Applied Artificial Intelligence: An International Journal, 4*(4), 265-282.

Temin, P., & Voth, H.-J. (2004). Riding the South Sea Bubble. *The American Economic Review, 94*(5), 1654-1668.

Tendeloo, B. v., & Vanstraelen, A. (2005). Earnings Management under German GAAP versus IFRS. *Europen Accounting Review, 14*(1), 155-180.

Thawornwong, S., & Enke, D. (2004). The Adaptive Selection of Financial and Economic Variables for Use with Artificial Neural Networks. *Neurocomputing, 56*, 205-232.

Ticknor, J. L. (2013). A Bayesian Regularized Artificial Neural Network for Stock Market Forecasting. *Expert Systems with Applications, 40*(14), 5501-5506.

Tracy, J. D. (1993). *The Rise of Merchant Empires: Long Distance Trade in the Early Modern World 1350-1750.* Cambridge University Press.

Trompeter, G. M., Carpenter, T. D., Desai, N., Jones, K. L., & Riley, R. A. (2013). A Synthesis of Fraud-Related Research. *Auditing: A Journal of Practice & Theory, 32*(1), 287-321.

Trotman, K. T., Simnett, R., & Khalifa, A. (2009). Impact of the Type of Audit Team Discussions on Auditors' Generation of Material Frauds. *Contemporary Accounting Research, 26*(4), 1115-1142.

Tsai, C.-F., & Wu, J.-W. (2008). Using Neural Network Ensembles for Bankruptcy Prediction and Credit Scoring. *Expert Systems with Applications, 34*(4), 2639-2649.

Tucker, J. W., & Zarowin, P. A. (2006). Does Income Smoothing Improve Earnings Informativeness? *The Accounting Review, 81*(1), 251-270.

Ugrin, J. C., & Odom, M. D. (2010). Exploring Sarbanes–Oxley's Effect on Attitudes, Perceptions of Norms, and Intentions to Commit Financial Statement Fraud from a General Deterrence Perspective. *Journal of Accounting and Public Policy, 29*(5), 439-458.

Wang, L., Malhotra, D., & Murnighan, J. K. (2011). Economics Education and Greed. *Academy of Management Learning & Education, 10*(4), 643-660.

Ward, J. R. (1994). The Industrial Revolution and British Imperialism, 1750-1850. *The Economic History Review, 47*(1), 44-65.

Wells, J. T. (2017). *Corporate Fraud Handbook: Prevention and Detection* (5th ed.). New Jersey: John Wiley & Sons.

White, H. (1992). *Artificial Neural Networks : Approximation and Learning Theory.* Massachusetts: Blackwell.

Witten, I. H. (2019, March). *Simple Neural Networks.* Retrieved from More Data Mining with Weka: https://www.cs.waikato.ac.nz/ml/weka/mooc/moredataminingwithweka/slides/Class5-MoreDataMiningWithWeka-2014.pdf

Wolfe, D. T., & Hermanson, D. R. (2004). The Fraud Diamond: Considering the Four Elements of Fraud. *CPA Journal, 74*(12), 38-42.

Wong, B. K., Bodnovich, T. A., & Selvi, Y. (1997). Neural Network Applications in Business: A Review and Analysis of the Literature (1988-95). *Decision Support Systems, 19*(4), 301-320.

Yegnanarayana, B. (2006). *Artificial Neural Networks* (12 ed.). Prentice Hall of India.

Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with Artificial Neural Networks: The State of the Art. *International Journal of Forecasting, 14*, 35-62.

Zhang, I. X. (2007). Economic Consequences of the Sarbanes–Oxley Act of 2002. *Journal of Accounting and Economics, 44*(1-2), 74-115.

Zhao, Y., & Zhang, Y. (2008). Comparison of Decision Tree Methods for Finding Active Objects. *Advances in Space Research, 41*(12), 1955-1959.

Zhou, X. (2001). Understanding the Determinants of Managerial Ownership and the Link Between Ownership and Performance: Comment. *Journal of Financial Economics, 62*(3), 559-571.

Zhu, X. (2005). *Semi-Supervised Learning Literature Survey.* University of Wisconsin-Madison Department of Computer Sciences.

Zhu, X., & Goldberg, A. B. (2009). Introduction to Semi-Supervised Learning. In R. J. Brachman, & Dietterich, *Synthesis Lectures on Artificial Intelligence and Machine Learning* (pp. 1-130). Morgan & Claypool.

Zhu, X., Lafferty, J., & Ghahramani, Z. (2003). Combining Active Learning and Semi-Supervised Learning Using Gaussian Fields and Harmonic Functions. *ICML 2003* *workshop on The Continuum from Labeled to Unlabeled Data in Machine Learning and Data Mining.* Washington.

APPENDICES

Appendix A. Multilayer Perceptron Sigmoid Node Weights

Sigmoid Node 0
Inputs Weights
Threshold -1.2124651224235432
Node 12 1.9827366437573317
Node 13 0.5258560466694312
Node 14 2.0655077399644357
Node 15 -0.8110780482787602
Node 16 -2.049762422441353
Node 17 -0.6889039638035885
Node 18 -0.891560090002661
Node 19 -0.840705638125156
Node 20 -0.7898019014240588
Node 21 -2.7823914831736474
Node 22 -0.8664688343706386
Node 23 -0.7495052366421208
Node 24 -0.817877618200079
Node 25 -0.863327913197247
Node 26 -0.8662151317731168
Sigmoid Node 1
Inputs Weights
Threshold 1.2125233571095848
Node 12 -1.9829342065718465
Node 13 -0.5247482866759092
Node 14 -2.0654797334318866
Node 15 0.8325214510460163
Node 16 1.9841625333984052
Node 17 0.6680773950335065
Node 18 0.9128513447752149
Node 19 0.8045348572623877
Node 20 0.8073469202785369
Node 21 2.8249331416487298
Node 22 0.9231027279924459
Node 23 0.7669774695537129
Node 24 0.8280945926795423
Node 25 0.8418722638251046
Node 26 0.8107494464656568
Sigmoid Node 2
Inputs Weights
Threshold -0.19034672043417344
Attrib Pillar10 -0.09870454688555393
Attrib Pillar1 -0.10434166698955358
Attrib NationalityMixED -0.03943666711112736
Attrib GenderMaleEDs 0.4160293849990033
Attrib NationalityMixNED 0.028300416377402278
Attrib Pillar11 -0.029815741063218694

Attrib TotalNumberofBoardMembers 0.018301576426275976
Attrib Pillar5 -0.0707337469826625
Attrib AccountingStandard=DS 0.007199340371808176
Attrib AccountingStandard=DI 0.03712906408580637
Attrib AccountingStandard=US 0.015172163124062261
Attrib AccountingStandard=DU -0.02009413327801111
Attrib AccountingStandard=ND -0.04333636433755368
Attrib Pillar7 -0.06836680936932485
Attrib TotalNumberofEDs 0.006884839300042289
Attrib Pillar3 -0.0600067271745745
Attrib TotalNumberofNEDs 0.008285447830463775
Attrib Pillar6 -0.09682583452735395
Attrib Pillar9 -0.05107810075400473
Attrib Pillar8 -0.07830001180628436
Attrib AverageAgeEDs 0.3245634849006419
Attrib ChiefFinancialOfficerSOXCert=N -0.028706835852943733
Attrib ChiefFinancialOfficerSOXCert=Y -5.622353134691137E-4
Attrib ChiefFinancialOfficerSOXCert=E -0.02814673894994988
Attrib ChiefExecutiveOfficerSOXCert=N -0.01658516432480167
Attrib ChiefExecutiveOfficerSOXCert=Y 0.045792762394045644
Attrib ChiefExecutiveOfficerSOXCert=E 0.02196289989498144
Attrib AveragetimeinroleforEDs 0.04696169723936657
Attrib AverageNumberofEducationEDs 0.008330690167201955
Attrib InventoriesOther -0.026551898457404223
Attrib zinvoth -0.013050716470420632
Attrib GenderMaleNED 0.5150592367528356
Attrib WorkingCapitalBalanceSheet 0.5122745016617234
Attrib zworkcap -3.4253585875272885E-4
Attrib AverageyearsonOtherQuotedBo -0.007435615579479595
Attrib AveragetimeinroleforNEDs -0.017539745768781978
Attrib zcrrntliabtot -0.002730640370157984
Attrib CurrentLiabilitiesTotal 0.4072703524079606
Attrib Goodwill 0.17410595556144312
Attrib zgoodwill -0.04762047851616568
Attrib LiquidWealthEDAverage 1.2420101757310784
Attrib LiabilitiesTotal 0.8468240112351038
Attrib zliabtot 0.0014806952060337022
Attrib zintastot 0.014245131729747113
Attrib IntangibleAssetsTotal 0.3066974820802866
Attrib zaccpaytra 0.016515426860976578
Attrib accountspayabletrade 0.1046470491675615
Attrib AverageTotalDirectCompensatio 0.2992295363995753
Attrib Fiscalyear 0.36350549272336924
Attrib zdivtot -0.024100274392375536
Attrib dividendstotal 0.04929827255258368
Attrib AssetTurnoverRatio 0.025814900154626827
Attrib zassetturnrat -0.01003931828687698
Attrib AverageSalaryNEDs 0.09551472574844344

Attrib CEOandChairmanRolesarecombi=Yes -0.04754636990663684
Sigmoid Node 3
Inputs Weights
Threshold -0.7838915571425621
Attrib Pillar10 -0.08400954154618043
Attrib Pillar1 -0.09373654389474832
Attrib NationalityMixED -0.015479479279008673
Attrib GenderMaleEDs 0.665132130073045
Attrib NationalityMixNED 0.045139874935890355
Attrib Pillar11 -0.04576430555620222
Attrib TotalNumberofBoardMembers 0.02222324666118972
Attrib Pillar5 -0.043294994089097455
Attrib AccountingStandard=DS 0.03914830858728558
Attrib AccountingStandard=DI -0.037736306689375346
Attrib AccountingStandard=US -0.01583108142900666
Attrib AccountingStandard=DU 0.040648226014002815
Attrib AccountingStandard=ND -0.035606556914723114
Attrib Pillar7 -0.11772942799053948
Attrib TotalNumberofEDs 0.01735330776641683
Attrib Pillar3 -0.12667262557588216
Attrib TotalNumberofNEDs -0.014120191606468274
Attrib Pillar6 -0.090708797760893
Attrib Pillar9 -0.050816965941105206
Attrib Pillar8 -0.11925168602682179
Attrib AverageAgeEDs 0.11782405521085337
Attrib ChiefFinancialOfficerSOXCert=N -0.025614981853989617
Attrib ChiefFinancialOfficerSOXCert=Y 0.021140668468838678
Attrib ChiefFinancialOfficerSOXCert=E -0.009416461293480007
Attrib ChiefExecutiveOfficerSOXCert=N 0.01820026938095688
Attrib ChiefExecutiveOfficerSOXCert=Y -0.024194579182678252
Attrib ChiefExecutiveOfficerSOXCert=E -0.021264114785267758
Attrib AveragetimeinroleforEDs 0.08044384051862813
Attrib AverageNumberofEducationEDs -0.02194245025035344
Attrib InventoriesOther 0.01252210175523219
Attrib zinvoth 0.03592380124734448
Attrib GenderMaleNED 0.4721102820703647
Attrib WorkingCapitalBalanceSheet -0.10930122261034964
Attrib zworkcap 0.01868125343170005
Attrib AverageyearsonOtherQuotedBo -0.04824871570485169
Attrib AveragetimeinroleforNEDs -0.02021314038320458
Attrib zcrrntliabtot -0.038354768736209334
Attrib CurrentLiabilitiesTotal 1.9405872218061768
Attrib Goodwill 0.12669852170543353
Attrib zgoodwill 0.0229662283410009
Attrib LiquidWealthEDAverage 0.096024648525924
Attrib LiabilitiesTotal 2.876561300207215
Attrib zliabtot 0.016952161534192062
Attrib zintastot -0.048711039579586425

Attrib IntangibleAssetsTotal 0.36578866574841973
Attrib zaccpaytra -0.03405442781882393
Attrib accountspayabletrade 0.3746911749525989
Attrib AverageTotalDirectCompensatio 1.6645194277013406
Attrib Fiscalyear 2.4089917942218775
Attrib zdivtot -0.012745814726534079
Attrib dividendstotal 0.004591641818949581
Attrib AssetTurnoverRatio -5.9593949628357995E-5
Attrib zassetturnrat -0.011079279745239179
Attrib AverageSalaryNEDs -0.03412577745686157
Attrib CEOandChairmanRolesarecombi=Yes 0.04360236836534967
Sigmoid Node 4
Inputs Weights
Threshold -1.4436197175667547
Attrib Pillar10 -0.04572229778894843
Attrib Pillar1 -0.08159851595023528
Attrib NationalityMixED 0.019574719348232347
Attrib GenderMaleEDs -0.7117991840699391
Attrib NationalityMixNED 0.04143008381769154
Attrib Pillar11 -0.07168176847831578
Attrib TotalNumberofBoardMembers -0.14416962986309156
Attrib Pillar5 -0.09477799285953215
Attrib AccountingStandard=DS 0.020417643072365505
Attrib AccountingStandard=DI -0.02871232241617885
Attrib AccountingStandard=US -0.011879760994399169
Attrib AccountingStandard=DU 0.033110300531209655
Attrib AccountingStandard=ND -0.04177939437225711
Attrib Pillar7 -0.08993411385705961
Attrib TotalNumberofEDs -0.07384707951825312
Attrib Pillar3 -0.08598236119322991
Attrib TotalNumberofNEDs -0.04591972538466486
Attrib Pillar6 -0.038164275477662206
Attrib Pillar9 -0.06303094966505947
Attrib Pillar8 -0.11843539247017472
Attrib AverageAgeEDs -0.35984289499921934
Attrib ChiefFinancialOfficerSOXCert=N 0.03362373853885448
Attrib ChiefFinancialOfficerSOXCert=Y 0.05861768422252662
Attrib ChiefFinancialOfficerSOXCert=E 0.01270309443147864
Attrib ChiefExecutiveOfficerSOXCert=N 0.04980212528742511
Attrib ChiefExecutiveOfficerSOXCert=Y 0.018117953285551238
Attrib ChiefExecutiveOfficerSOXCert=E -0.04249546883814313
Attrib AveragetimeinroleforEDs -0.016128285910976717
Attrib AverageNumberofEducationEDs -0.011536146379696162
Attrib InventoriesOther -0.011631287344646288
Attrib zinvoth 0.01991869982934619
Attrib GenderMaleNED -1.3180849649950235
Attrib WorkingCapitalBalanceSheet 2.5236469362217155
Attrib zworkcap 0.022882325294681132

Attrib AverageyearsonOtherQuotedBo 0.05478429264106752
Attrib AveragetimeinroleforNEDs -0.08282436719155548
Attrib zcrrntliabtot -0.033911941824950366
Attrib CurrentLiabilitiesTotal 2.0460942697630062
Attrib Goodwill 2.1890642200594788
Attrib zgoodwill 0.01208465273512958
Attrib LiquidWealthEDAverage 3.0354346417992217
Attrib LiabilitiesTotal 3.4215909041084767
Attrib zliabtot -0.02794532698005283
Attrib zintastot 0.008335361560374281
Attrib IntangibleAssetsTotal 2.911585834536226
Attrib zaccpaytra -0.049412134564675636
Attrib accountspayabletrade 0.591459702590094
Attrib AverageTotalDirectCompensatio -11.822506851671873
Attrib Fiscalyear 0.25321160494391903
Attrib zdivtot 0.03691476441377346
Attrib dividendstotal 0.1749362068836662
Attrib AssetTurnoverRatio 0.004680507150133695
Attrib zassetturnrat 0.024469146750876223
Attrib AverageSalaryNEDs -4.314527401282892
Attrib CEOandChairmanRolesarecombi=Yes 0.008145371689108942
Sigmoid Node 5
Inputs Weights
Threshold -1.7782483324384646
Attrib Pillar10 -0.005960793269446717
Attrib Pillar1 -0.04587577099657712
Attrib NationalityMixED -0.026815471406646398
Attrib GenderMaleEDs -0.5623597933664661
Attrib NationalityMixNED 0.001667230718739245
Attrib Pillar11 -0.09042814180698287
Attrib TotalNumberofBoardMembers -0.023431751642887216
Attrib Pillar5 -0.09000711466769758
Attrib AccountingStandard=DS -0.06035852569344941
Attrib AccountingStandard=DI 0.03876965224161845
Attrib AccountingStandard=US 0.008600921330900641
Attrib AccountingStandard=DU 0.03207838030541521
Attrib AccountingStandard=ND 0.003674474393310166
Attrib Pillar7 -0.015348525702397628
Attrib TotalNumberofEDs -0.015765969011978423
Attrib Pillar3 -0.09230331496552727
Attrib TotalNumberofNEDs 0.005706244635796743
Attrib Pillar6 -0.06476435477966531
Attrib Pillar9 -0.07169302488572202
Attrib Pillar8 -0.0770692649901183
Attrib AverageAgeEDs -0.27265796430969996
Attrib ChiefFinancialOfficerSOXCert=N -0.0080764618739158
Attrib ChiefFinancialOfficerSOXCert=Y 0.00875646675582146
Attrib ChiefFinancialOfficerSOXCert=E -0.026572431434264403

Attrib ChiefExecutiveOfficerSOXCert=N   0.00937071757114231
Attrib ChiefExecutiveOfficerSOXCert=Y   -0.048331084875211976
Attrib ChiefExecutiveOfficerSOXCert=E   0.03339658529455468
Attrib AveragetimeinroleforEDs   -0.06606123194760473
Attrib AverageNumberofEducationEDs   0.019788935380496482
Attrib InventoriesOther   -0.011138088578764622
Attrib zinvoth   -0.04645300062326092
Attrib GenderMaleNED   -0.17812305974135925
Attrib WorkingCapitalBalanceSheet   6.2601253665131456
Attrib zworkcap   0.026155018216255883
Attrib AverageyearsonOtherQuotedBo   -0.021736120716767858
Attrib AveragetimeinroleforNEDs   -0.030634225413204194
Attrib zcrrntliabtot   0.02232680858489397
Attrib CurrentLiabilitiesTotal   45.353915786928056
Attrib Goodwill   32.50547547448447
Attrib zgoodwill   0.03343659740529102
Attrib LiquidWealthEDAverage   41.072979418718205
Attrib LiabilitiesTotal   79.07672807637098
Attrib zliabtot   -0.03320683652114876
Attrib zintastot   0.02923610365853324
Attrib IntangibleAssetsTotal   34.654783709791985
Attrib zaccpaytra   0.002911019822139807
Attrib accountspayabletrade   16.48642167914399
Attrib AverageTotalDirectCompensatio   5.062409696714559
Attrib Fiscalyear   10.686181953110856
Attrib zdivtot   0.046239024531529914
Attrib dividendstotal   1.1563383245031045
Attrib AssetTurnoverRatio   0.02699377380450031
Attrib zassetturnrat   -0.021728097527120134
Attrib AverageSalaryNEDs   0.6962458997067893
Attrib CEOandChairmanRolesarecombi=Yes   -0.02391449242229392
Sigmoid Node 6
Inputs   Weights
Threshold   1.0168226765685338
Attrib Pillar10   0.5427137963875748
Attrib Pillar1   0.48914467509752335
Attrib NationalityMixED   -0.027349455358110987
Attrib GenderMaleEDs   5.8110591119471575
Attrib NationalityMixNED   -0.012739312886701314
Attrib Pillar11   0.5271342146561253
Attrib TotalNumberofBoardMembers   0.7571717873426249
Attrib Pillar5   0.4997317346534101
Attrib AccountingStandard=DS   0.015798664215791625
Attrib AccountingStandard=DI   -0.04044343427160966
Attrib AccountingStandard=US   0.04781086288602196
Attrib AccountingStandard=DU   -0.021363718996473546
Attrib AccountingStandard=ND   -0.04423541281991686
Attrib Pillar7   0.47385126926684784

Attrib TotalNumberofEDs 0.08636779284507612
Attrib Pillar3 0.4373166202385501
Attrib TotalNumberofNEDs 0.6835973932647909
Attrib Pillar6 0.4898639299939558
Attrib Pillar9 0.5480341973419187
Attrib Pillar8 0.543967131471009
Attrib AverageAgeEDs 2.9716445634861275
Attrib ChiefFinancialOfficerSOXCert=N -0.04482322877984144
Attrib ChiefFinancialOfficerSOXCert=Y -0.022367844681928963
Attrib ChiefFinancialOfficerSOXCert=E 0.00682928017820151
Attrib ChiefExecutiveOfficerSOXCert=N 0.03663586534641565
Attrib ChiefExecutiveOfficerSOXCert=Y 0.016976713684588595
Attrib ChiefExecutiveOfficerSOXCert=E -0.023159349924654758
Attrib AveragetimeinroleforEDs -0.006832013227981406
Attrib AverageNumberofEducationEDs 0.10253806676314041
Attrib InventoriesOther 0.011586453488924899
Attrib zinvoth -0.015419937072157492
Attrib GenderMaleNED 7.597101457992561
Attrib WorkingCapitalBalanceSheet -0.03291684071069248
Attrib zworkcap 0.018569682542271232
Attrib AverageyearsonOtherQuotedBo 0.06972469949543322
Attrib AveragetimeinroleforNEDs 0.2341849431606104
Attrib zcrrntliabtot -0.008075124231221949
Attrib CurrentLiabilitiesTotal 0.45173464063092206
Attrib Goodwill 0.11607678359319484
Attrib zgoodwill -0.027620819317912028
Attrib LiquidWealthEDAverage 16.518814742631644
Attrib LiabilitiesTotal 0.48809954313579684
Attrib zliabtot 0.015569944396249993
Attrib zintastot 0.02167719365402172
Attrib IntangibleAssetsTotal 0.43736411048742757
Attrib zaccpaytra -0.023069270069011927
Attrib accountspayabletrade 0.3152086408083916
Attrib AverageTotalDirectCompensatio 21.11669569795502
Attrib Fiscalyear 116.29521578686234
Attrib zdivtot 0.012060226373161793
Attrib dividendstotal -0.2152057061599925
Attrib AssetTurnoverRatio -0.007650698621187134
Attrib zassetturnrat -9.088127070740167E-4
Attrib AverageSalaryNEDs 5.505647234373129
Attrib CEOandChairmanRolesarecombi=Yes -0.028372847324996023
Sigmoid Node 7
Inputs Weights
Threshold 0.17502378643114805
Attrib Pillar10 -0.015277820414687053
Attrib Pillar1 0.0024464126364633983
Attrib NationalityMixED 0.022461426411311484
Attrib GenderMaleEDs 0.6801928048891913

Attrib NationalityMixNED 0.03151858381345792
Attrib Pillar11 -0.055699119866551405
Attrib TotalNumberofBoardMembers 0.04386339219489322
Attrib Pillar5 -0.012233478316968508
Attrib AccountingStandard=DS 0.04803525458070823
Attrib AccountingStandard=DI -0.011216283517934791
Attrib AccountingStandard=US 0.04700651604612474
Attrib AccountingStandard=DU 0.04395209786915846
Attrib AccountingStandard=ND 0.017845252775572806
Attrib Pillar7 0.00724999947970915
Attrib TotalNumberofEDs 0.0013664012238111015
Attrib Pillar3 -0.015178778792700189
Attrib TotalNumberofNEDs -0.00793703174232298
Attrib Pillar6 -0.06015144292382283
Attrib Pillar9 -0.036243628133229976
Attrib Pillar8 -0.03094744283368303
Attrib AverageAgeEDs 0.36566736344570466
Attrib ChiefFinancialOfficerSOXCert=N 0.03826105496251318
Attrib ChiefFinancialOfficerSOXCert=Y 0.013494970675713524
Attrib ChiefFinancialOfficerSOXCert=E 0.01683173118050209
Attrib ChiefExecutiveOfficerSOXCert=N -0.011421174597212366
Attrib ChiefExecutiveOfficerSOXCert=Y 0.022815986207594775
Attrib ChiefExecutiveOfficerSOXCert=E 1.0371376138085779E-4
Attrib AveragetimeinroleforEDs 0.03524887309886683
Attrib AverageNumberofEducationEDs -0.027811637939826564
Attrib InventoriesOther -0.023351931776204295
Attrib zinvoth 0.02004168455136949
Attrib GenderMaleNED 0.4997743421636507
Attrib WorkingCapitalBalanceSheet 0.22265933378289346
Attrib zworkcap -0.01568563043331227
Attrib AverageyearsonOtherQuotedBo 0.06525204351660878
Attrib AveragetimeinroleforNEDs 0.053508280360542206
Attrib zcrrntliabtot -0.013019980224779211
Attrib CurrentLiabilitiesTotal 0.22792438014560765
Attrib Goodwill 0.18281425953623204
Attrib zgoodwill -0.01268923191573128
Attrib LiquidWealthEDAverage 0.21257428249057658
Attrib LiabilitiesTotal 1.0703039211574545
Attrib zliabtot -0.020039824608578328
Attrib zintastot 0.0038224201846993114
Attrib IntangibleAssetsTotal 0.24009936496296633
Attrib zaccpaytra 0.02137683617527702
Attrib accountspayabletrade 0.11297661275228063
Attrib AverageTotalDirectCompensatio 1.1479510202663918
Attrib Fiscalyear 0.2473097785371834
Attrib zdivtot 0.013347459703148361
Attrib dividendstotal 0.029395332357517782
Attrib AssetTurnoverRatio -0.003527638406176112

Attrib zassetturnrat -0.01318843594734462
Attrib AverageSalaryNEDs -0.11989965612296812
Attrib CEOandChairmanRolesarecombi=Yes -0.018346897169931632
Sigmoid Node 8
Inputs Weights
Threshold -0.14953330294241485
Attrib Pillar10 -0.11088510749777718
Attrib Pillar1 -0.03740128620095727
Attrib NationalityMixED 0.03655703645688137
Attrib GenderMaleEDs 0.33368450529715676
Attrib NationalityMixNED 0.02105171708672701
Attrib Pillar11 -0.04826569184193631
Attrib TotalNumberofBoardMembers -0.01065883284427155
Attrib Pillar5 -0.10563341648179818
Attrib AccountingStandard=DS 0.019116366634305642
Attrib AccountingStandard=DI 0.03521333906742878
Attrib AccountingStandard=US 0.012088062311327648
Attrib AccountingStandard=DU -0.0015639783794694814
Attrib AccountingStandard=ND -0.02992680088969403
Attrib Pillar7 -0.06393705341152096
Attrib TotalNumberofEDs 0.010334383905506356
Attrib Pillar3 -0.03396223361466837
Attrib TotalNumberofNEDs 0.03676513958724629
Attrib Pillar6 -0.021438494815874364
Attrib Pillar9 -0.05246014705107529
Attrib Pillar8 -0.07872130160400802
Attrib AverageAgeEDs 0.28906429123714805
Attrib ChiefFinancialOfficerSOXCert=N 0.02995302842440785
Attrib ChiefFinancialOfficerSOXCert=Y 0.03576721172942994
Attrib ChiefFinancialOfficerSOXCert=E 0.04139298891168714
Attrib ChiefExecutiveOfficerSOXCert=N 0.027717625481740255
Attrib ChiefExecutiveOfficerSOXCert=Y 0.028935248887204744
Attrib ChiefExecutiveOfficerSOXCert=E -0.006668495609480793
Attrib AveragetimeinroleforEDs 0.015834448889999016
Attrib AverageNumberofEducationEDs 0.036688343607174885
Attrib InventoriesOther 0.003735158095401797
Attrib zinvoth 0.04879322958355782
Attrib GenderMaleNED 0.3465571682304786
Attrib WorkingCapitalBalanceSheet 0.1989307662985592
Attrib zworkcap -0.011138429098752263
Attrib AverageyearsonOtherQuotedBo 0.006649404826163374
Attrib AveragetimeinroleforNEDs 0.03301916892181437
Attrib zcrrntliabtot -0.030695239691935876
Attrib CurrentLiabilitiesTotal 0.1962811781499879
Attrib Goodwill 0.0885182373346054
Attrib zgoodwill 0.03973107425354906
Attrib LiquidWealthEDAverage 0.8255405494978932
Attrib LiabilitiesTotal 0.7036212742040574

Attrib zliabtot 0.01094735415209215
Attrib zintastot -0.030311618453631207
Attrib IntangibleAssetsTotal 0.07727989153776492
Attrib zaccpaytra -0.033559025128393496
Attrib accountspayabletrade 0.03958084280735719
Attrib AverageTotalDirectCompensatio 0.13408593499684232
Attrib Fiscalyear 0.5320951707902527
Attrib zdivtot -0.049786016339032485
Attrib dividendstotal 0.06931986536464436
Attrib AssetTurnoverRatio -0.041712769943250134
Attrib zassetturnrat -0.035788934499033295
Attrib AverageSalaryNEDs -0.027352969530476924
Attrib CEOandChairmanRolesarecombi=Yes 0.028606469467754442
Sigmoid Node 9
Inputs Weights
Threshold 0.5218486566057465
Attrib Pillar10 0.07139521996683748
Attrib Pillar1 0.1288409329552222
Attrib NationalityMixED 0.019300198760550613
Attrib GenderMaleEDs -0.8870365390809082
Attrib NationalityMixNED -0.0022939925068031198
Attrib Pillar11 0.11946108161171295
Attrib TotalNumberofBoardMembers -0.05708671651729668
Attrib Pillar5 0.04146682748010383
Attrib AccountingStandard=DS -0.06026256546854449
Attrib AccountingStandard=DI 0.034485950720465446
Attrib AccountingStandard=US -0.003751743736154804
Attrib AccountingStandard=DU -0.014308660680269378
Attrib AccountingStandard=ND -0.041036322031244056
Attrib Pillar7 0.07200990675371874
Attrib TotalNumberofEDs -0.007060637138729899
Attrib Pillar3 0.03859351072289564
Attrib TotalNumberofNEDs 0.018979363674654193
Attrib Pillar6 0.06961402017386223
Attrib Pillar9 0.09879051319002168
Attrib Pillar8 0.03280335131754619
Attrib AverageAgeEDs -0.5036560473478837
Attrib ChiefFinancialOfficerSOXCert=N 0.010019492666645945
Attrib ChiefFinancialOfficerSOXCert=Y 0.02971577703087999
Attrib ChiefFinancialOfficerSOXCert=E -0.03599250669381422
Attrib ChiefExecutiveOfficerSOXCert=N 0.040535986001564266
Attrib ChiefExecutiveOfficerSOXCert=Y 0.03645441477286727
Attrib ChiefExecutiveOfficerSOXCert=E 0.004143208851682578
Attrib AveragetimeinroleforEDs -0.16589314622048765
Attrib AverageNumberofEducationEDs -0.056444978312033406
Attrib InventoriesOther -0.02589815850254736
Attrib zinvoth 7.606453061950773E-4
Attrib GenderMaleNED -0.8206400365634462

Attrib WorkingCapitalBalanceSheet -0.6540498052542358
Attrib zworkcap 0.05049440214808985
Attrib AverageyearsonOtherQuotedBo -0.04816452941474363
Attrib AveragetimeinroleforNEDs -0.023087253930027706
Attrib zcrrntliabtot 0.04972865086608123
Attrib CurrentLiabilitiesTotal -0.8092849853577468
Attrib Goodwill -0.35993388450341157
Attrib zgoodwill -0.026561795419797897
Attrib LiquidWealthEDAverage -4.420133868353998
Attrib LiabilitiesTotal -2.6318472825834935
Attrib zliabtot 0.002251603664184271
Attrib zintastot 0.03981340459438507
Attrib IntangibleAssetsTotal -0.6776953763272593
Attrib zaccpaytra 0.049946702805309724
Attrib accountspayabletrade -0.3080548221953625
Attrib AverageTotalDirectCompensatio 3.2137603869623628
Attrib Fiscalyear 3.136601239857178
Attrib zdivtot -0.022124092053374126
Attrib dividendstotal -0.09523806814952081
Attrib AssetTurnoverRatio -0.012307511925592592
Attrib zassetturnrat 0.04432020486590137
Attrib AverageSalaryNEDs 0.0595229230705361
Attrib CEOandChairmanRolesarecombi=Yes 0.022159532315907976
Sigmoid Node 10
Inputs Weights
Threshold -0.5906085499262801
Attrib Pillar10 -0.09655214892260149
Attrib Pillar1 -0.1190698919043798
Attrib NationalityMixED 0.020095108725900035
Attrib GenderMaleEDs -0.5073118749739036
Attrib NationalityMixNED 0.04244977252005433
Attrib Pillar11 -0.1115054275110019
Attrib TotalNumberofBoardMembers -0.20138674691290304
Attrib Pillar5 -0.13816577864987636
Attrib AccountingStandard=DS 0.021150311292731847
Attrib AccountingStandard=DI 0.01725832998898934
Attrib AccountingStandard=US -0.018331158577593286
Attrib AccountingStandard=DU -0.04215168146096565
Attrib AccountingStandard=ND 0.03313660928125245
Attrib Pillar7 -0.1444920092736756
Attrib TotalNumberofEDs -0.04065906966184702
Attrib Pillar3 -0.10886476817730988
Attrib TotalNumberofNEDs -0.1577027657610063
Attrib Pillar6 -0.16266091108791664
Attrib Pillar9 -0.12872633948131557
Attrib Pillar8 -0.089558337025764
Attrib AverageAgeEDs 0.1908612156758749
Attrib ChiefFinancialOfficerSOXCert=N -0.04709897782257354

Attrib ChiefFinancialOfficerSOXCert=Y 0.027347047616258585
Attrib ChiefFinancialOfficerSOXCert=E 0.01924923354403208
Attrib ChiefExecutiveOfficerSOXCert=N 0.04822333283969413
Attrib ChiefExecutiveOfficerSOXCert=Y -0.005686481064783499
Attrib ChiefExecutiveOfficerSOXCert=E 0.03078385097995332
Attrib AveragetimeinroleforEDs 0.04212530154991283
Attrib AverageNumberofEducationEDs 0.0026730806572135943
Attrib InventoriesOther 0.032654349934929806
Attrib zinvoth 0.01912886085695336
Attrib GenderMaleNED -0.6079837741186529
Attrib WorkingCapitalBalanceSheet 1.1322766728789988
Attrib zworkcap 0.04234041665934474
Attrib AverageyearsonOtherQuotedBo 0.008845829926433915
Attrib AveragetimeinroleforNEDs -0.026743964016109605
Attrib zcrrntliabtot 0.02202695116810304
Attrib CurrentLiabilitiesTotal 0.8789962033215836
Attrib Goodwill -6.539039417564979
Attrib zgoodwill -5.26864886390566E-4
Attrib LiquidWealthEDAverage -33.31034984527476
Attrib LiabilitiesTotal -10.236471316404812
Attrib zliabtot 0.04234443770858716
Attrib zintastot -0.023517073489716694
Attrib IntangibleAssetsTotal -7.5939568257713335
Attrib zaccpaytra -0.00329018675171349
Attrib accountspayabletrade -1.9600566396135581
Attrib AverageTotalDirectCompensatio -1.4515658565511778
Attrib Fiscalyear -1.2413444050210538
Attrib zdivtot -0.019406773339655976
Attrib dividendstotal -0.4318909588340421
Attrib AssetTurnoverRatio 0.04026683559218542
Attrib zassetturnrat 0.008937175847293288
Attrib AverageSalaryNEDs -0.2831471893255082
Attrib CEOandChairmanRolesarecombi=Yes 0.020867774858992445
Sigmoid Node 11
Inputs Weights
Threshold 0.33607658147597863
Attrib Pillar10 0.1550784128994592
Attrib Pillar1 0.1848952031010029
Attrib NationalityMixED 0.013385702435259083
Attrib GenderMaleEDs 3.2944181588639747
Attrib NationalityMixNED 0.024661296818211803
Attrib Pillar11 0.1910192248940649
Attrib TotalNumberofBoardMembers 0.6017678536101665
Attrib Pillar5 0.14845619739579669
Attrib AccountingStandard=DS -0.044988024198658515
Attrib AccountingStandard=DI 0.032701540371905743
Attrib AccountingStandard=US 0.044338544914295105
Attrib AccountingStandard=DU 0.01759340353505618

Attrib AccountingStandard=ND 0.014034440127083755
Attrib Pillar7 0.15486282559648573
Attrib TotalNumberofEDs 0.13502163196667297
Attrib Pillar3 0.1738252779959582
Attrib TotalNumberofNEDs 0.38540977104131197
Attrib Pillar6 0.19722996586957722
Attrib Pillar9 0.14203287106098256
Attrib Pillar8 0.11756911411238448
Attrib AverageAgeEDs 1.978916225980282
Attrib ChiefFinancialOfficerSOXCert=N -0.025566936177958525
Attrib ChiefFinancialOfficerSOXCert=Y 0.017846853712367074
Attrib ChiefFinancialOfficerSOXCert=E -0.029214706206707023
Attrib ChiefExecutiveOfficerSOXCert=N 0.007684642253608789
Attrib ChiefExecutiveOfficerSOXCert=Y -0.0012757647158077323
Attrib ChiefExecutiveOfficerSOXCert=E 0.03072385519007052
Attrib AveragetimeinroleforEDs 0.11404596617928288
Attrib AverageNumberofEducationEDs 0.08340826684222959
Attrib InventoriesOther 0.04991252115420925
Attrib zinvoth 0.015264307366418364
Attrib GenderMaleNED 3.8827610696919024
Attrib WorkingCapitalBalanceSheet -0.7783886425356108
Attrib zworkcap -0.02327135869699081
Attrib AverageyearsonOtherQuotedBo 0.045687638413710076
Attrib AveragetimeinroleforNEDs 0.18588513918169566
Attrib zcrrntliabtot -0.04680377477521298
Attrib CurrentLiabilitiesTotal -0.16862152760681554
Attrib Goodwill 0.023787833540055193
Attrib zgoodwill 0.01555878659586409
Attrib LiquidWealthEDAverage 9.175518839810328
Attrib LiabilitiesTotal 0.471981833741865
Attrib zliabtot 0.01956055830903266
Attrib zintastot 0.028717079996359047
Attrib IntangibleAssetsTotal -0.4727495815901476
Attrib zaccpaytra -0.04010570178764506
Attrib accountspayabletrade -0.36063246190521797
Attrib AverageTotalDirectCompensatio 27.012283436135256
Attrib Fiscalyear 31.706540942488846
Attrib zdivtot -0.027026182108441058
Attrib dividendstotal -0.03465472004154929
Attrib AssetTurnoverRatio 0.005411992401506249
Attrib zassetturnrat -0.04623597036563807
Attrib AverageSalaryNEDs 1.2511525387386973
Attrib CEOandChairmanRolesarecombi=Yes 0.04501735126071838
Sigmoid Node 12
Inputs Weights
Threshold -0.9187752359535939
Node 2 0.06833603804867323
Node 3 1.203187596364194

Node 4 6.552254150798696
Node 5 2.1103909853228413
Node 6 -1.8964057092283046
Node 7 -3.022049154100978
Node 8 -0.06361993271472847
Node 9 0.8879719596181005
Node 10 1.5073614837320557
Node 11 -1.1176646609959218
Sigmoid Node 13
Inputs Weights
Threshold -1.591222652808006
Node 2 -0.6122635182906935
Node 3 0.5386964943623569
Node 4 4.823479489507344
Node 5 1.5608628689960087
Node 6 -2.547649292890014
Node 7 -2.889986025291388
Node 8 -0.7330349840576
Node 9 1.4716418664978415
Node 10 1.595402234110429
Node 11 -1.6083334385508787
Sigmoid Node 14
Inputs Weights
Threshold -0.6100989976406945
Node 2 0.5995643461535892
Node 3 2.3059123507583967
Node 4 7.475567999177689
Node 5 4.602028007645106
Node 6 -2.0312821656014255
Node 7 -0.9196231404360361
Node 8 0.48264450412197757
Node 9 -6.709381590102851
Node 10 1.526299602308262
Node 11 -0.7034251040896767
Sigmoid Node 15
Inputs Weights
Threshold -0.6753776307739172
Node 2 -0.15335510251480403
Node 3 0.27448642921617683
Node 4 -0.24593503578390574
Node 5 0.4623058913338887
Node 6 -1.8755169668607754
Node 7 0.11946115922554049
Node 8 -0.22558389757484396
Node 9 -1.015013565545301
Node 10 0.37786670290089314
Node 11 -1.5770060139145963
Sigmoid Node 16

Inputs Weights
Threshold 0.14423596773897926
Node 2 -0.25963629353609946
Node 3 -0.7569389706633741
Node 4 -1.6756356426444026
Node 5 -4.398650481312167
Node 6 0.611432213821398
Node 7 0.44348698054869684
Node 8 -0.22479446621587
Node 9 -0.3871213429214805
Node 10 -0.14034903191574938
Node 11 -0.0014708878598727148
Sigmoid Node 17
Inputs Weights
Threshold -1.6350741138733957
Node 2 -0.8719455223460933
Node 3 0.08835245308724567
Node 4 2.9552821476585693
Node 5 0.8980782789216257
Node 6 -2.376702123813909
Node 7 -1.3918370101690274
Node 8 -0.9978398120868028
Node 9 0.3493546319581283
Node 10 1.1717035132941787
Node 11 -2.0121006208527947
Sigmoid Node 18
Inputs Weights
Threshold -0.10479129719002594
Node 2 0.0556498108430522
Node 3 0.08526915493395801
Node 4 -0.5756961080644106
Node 5 -1.0750280867442732
Node 6 -1.412724237926555
Node 7 0.4493125045567374
Node 8 0.021974469803656253
Node 9 -0.9376885784432086
Node 10 0.07105287273965241
Node 11 -1.6030567844055688
Sigmoid Node 19
Inputs Weights
Threshold -0.5632532010086324
Node 2 -0.04720964829233953
Node 3 0.23193333864975238
Node 4 -0.35524335270505963
Node 5 0.1411135002298517
Node 6 -1.8287465698171497
Node 7 0.16375835248436013
Node 8 -0.13082321378292022

Node 9 -0.9604617884172918
Node 10 0.3158124117285389
Node 11 -1.5973912001542507
Sigmoid Node 20
Inputs Weights
Threshold -0.6401997440840291
Node 2 -0.07224250111363438
Node 3 0.3821325314027302
Node 4 -0.067593886284139
Node 5 0.736827797314425
Node 6 -2.271039129094986
Node 7 0.08099971892389957
Node 8 -0.1359101626331764
Node 9 -1.0942443590786557
Node 10 0.523236303559416
Node 11 -1.7998434384681858
Sigmoid Node 21
Inputs Weights
Threshold 0.41236019636923077
Node 2 -0.29851225155641503
Node 3 -1.0724951211152747
Node 4 -2.040221325237673
Node 5 -5.484169002227882
Node 6 1.1767781873087684
Node 7 0.5189803718073845
Node 8 -0.2575910592816257
Node 9 -0.15922918591094076
Node 10 -0.16114651282458564
Node 11 0.25606288903596575
Sigmoid Node 22
Inputs Weights
Threshold -0.13635098367975584
Node 2 0.06360123110381954
Node 3 0.08146216291028834
Node 4 -0.5926651659408982
Node 5 -1.0331136302221093
Node 6 -1.4601585244061261
Node 7 0.45190131074125334
Node 8 0.02238824767843508
Node 9 -0.91399039146548
Node 10 0.06077258976287376
Node 11 -1.5780482927635837
Sigmoid Node 23
Inputs Weights
Threshold -1.2584757383921896
Node 2 -0.5524570958827417
Node 3 0.2900724367155166
Node 4 1.277918191407262

Node 5 0.9783104454254199
Node 6 -2.0773132764636273
Node 7 -0.7722404129663273
Node 8 -0.6615860252286306
Node 9 -0.21381035916655544
Node 10 0.9869123575819506
Node 11 -1.87975242194931
Sigmoid Node 24
Inputs Weights
Threshold -0.5222015686083747
Node 2 -0.0522442558158236
Node 3 0.1823094581454472
Node 4 -0.40384002093035354
Node 5 -0.011185183078204462
Node 6 -1.7274382280613505
Node 7 0.20771337956531755
Node 8 -0.10946457879109263
Node 9 -0.9365851738188554
Node 10 0.28904914569444723
Node 11 -1.5837879321653336
Sigmoid Node 25
Inputs Weights
Threshold -0.3105074730488689
Node 2 -0.0071136245481539075
Node 3 0.13475146883360248
Node 4 -0.43044530039900025
Node 5 -0.5074120619042826
Node 6 -1.602358772604541
Node 7 0.3574292447164527
Node 8 0.01061560784204945
Node 9 -0.9829467013370636
Node 10 0.11887006507767996
Node 11 -1.626219062329602
Sigmoid Node 26
Inputs Weights
Threshold -0.4187926851559866
Node 2 -0.023149645260898022
Node 3 0.1911990400219948
Node 4 -0.34171558826889215
Node 5 -0.19973272043312293
Node 6 -1.630517735117883
Node 7 0.2712201184616594
Node 8 -0.1250317908510869
Node 9 -0.973761893347796
Node 10 0.150579905432877
Node 11 -1.6390315718688733
Class 0
Input

Node 0
Class 1
Input
Node 1

Appendix B. Odds Ratios of the Logistic Function

Variable	Class 0
Pillar10	21.1129
Pillar1	1.0683
NationalityMixED	0.0138
GenderMaleEDs	1.0329
NationalityMixNED	0.0394
Pillar11	0.0088
TotalNumberofBoardMembers	1.2851
Pillar5	399.6621
AccountingStandard=DS	0.7994
AccountingStandard=DI	3.237
AccountingStandard=US	1.5791
AccountingStandard=DU	0.082
AccountingStandard=ND	1
Pillar7	74.2888
TotalNumberofEDs	0.5335
Pillar3	0.2857
TotalNumberofNEDs	0.6379
Pillar6	21.6108
Pillar9	0.3949
Pillar8	1.1474
AverageAgeEDs	1.0057
ChiefFinancialOfficerSOXCert=N	0.674
ChiefFinancialOfficerSOXCert=Y	1.1379
ChiefFinancialOfficerSOXCert=E	0.9751
ChiefExecutiveOfficerSOXCert=N	0.674
ChiefExecutiveOfficerSOXCert=Y	1.1379
ChiefExecutiveOfficerSOXCert=E	0.9751
AveragetimeinroleforEDs	1.0289
AverageNumberofEducationEDs	1.1228
InventoriesOther	1.0088
zinvoth	2.0281
GenderMaleNED	0.9949
WorkingCapitalBalanceSheet	1
zworkcap	1.037
AverageyearsonOtherQuotedBo	0.9474
AveragetimeinroleforNEDs	1.0568
zcrrntliabtot	1.0207
CurrentLiabilitiesTotal	1
Goodwill	1
zgoodwill	1.0293
LiquidWealthEDAverage	1
LiabilitiesTotal	1

Appendix C. Employed pillars and subpillars of the Global Competitiveness Index of World Economic Forum

1st pillar: Institutions	**7th pillar: Labor market efficiency**
1.01 Property rights	7.01 Cooperation in labor-employer relations
1.02 Intellectual property protection	7.02 Flexibility of wage determination
1.03 Diversion of public funds	7.03 Hiring and firing practices
1.04 Public trust in politicians	7.04 Redundancy costs weeks of salary
1.05 Irregular payments and bribes	7.05 Effect of taxation on incentives to work
1.06 Judicial independence	7.06 Pay and productivity
1.07 Favoritism in decisions of government officials	7.07 Reliance on professional management
1.08 Efficiency of government spending	7.08 Country capacity to retain talent
1.09 Burden of government regulation	7.09 Country capacity to attract talent
1.10 Efficiency of legal framework in settling disputes	7.10 Female participation in the labor force ratio to men
1.11 Efficiency of legal framework in challenging regulations	
1.12 Transparency of government policymaking	**8th pillar: Financial market development**
1.13 Business costs of terrorism	8.01 Availability of financial services
1.14 Business costs of crime and violence	8.02 Affordability of financial services
1.15 Organized crime	8.03 Financing through local equity market
1.16 Reliability of police services	8.04 Ease of access to loans
1.17 Ethical behavior of firms	8.05 Venture capital availability
1.18 Strength of auditing and reporting standards	8.06 Soundness of banks
1.19 Efficacy of corporate boards	8.07 Regulation of securities exchanges
1.20 Protection of minority shareholders' interests	8.08 Legal rights index
1.21 Strength of investor protection	
	9th pillar: Technological readiness
3rd pillar: Macroeconomic environment	9.01 Availability of latest technologies
3.01 Government budget balance % GDP	9.02 Firm-level technology absorption
3.02 Gross national savings % GDP	9.03 FDI and technology transfer
3.03 Inflation annual % change	9.04 Internet users % pop.
3.04 Government debt % GDP	9.05 Fixed-broadband Internet subscriptions /100 pop.
3.05 Country credit rating	9.06 Internet bandwidth kb/s/user
	9.07 Mobile-broadband subscriptions /100 pop.
5th pillar: Higher education and training	
5.01 Secondary education enrollment rate	**10th pillar: Market size**
5.02 Tertiary education enrollment rate	10.01 Domestic market size index
5.03 Quality of the education system	10.02 Foreign market size index
5.04 Quality of math and science education	10.03 GDP (PPP) PPP $ billions
5.05 Quality of management schools	10.04 Exports % GDP

5.06 Internet access in schools	
5.07 Local availability of specialized training services	**11th pillar: Business sophistication**
5.08 Extent of staff training	11.01 Local supplier quantity
	11.02 Local supplier quality
6th pillar: Goods market efficiency	11.03 State of cluster development
6.01 Intensity of local competition	11.04 Nature of competitive advantage
6.02 Extent of market dominance	11.05 Value chain breadth
6.03 Effectiveness of anti-monopoly policy	11.06 Control of international distribution
6.04 Effect of taxation on incentives to invest	11.07 Production process sophistication
6.05 Total tax rate % profits	11.08 Extent of marketing
6.06 No. of procedures to start a business	11.09 Willingness to delegate authority
6.07 Time to start a business days	
6.08 Agricultural policy costs	
6.09 Prevalence of non-tariff barriers	
6.10 Trade tariffs % duty	
6.11 Prevalence of foreign ownership	
6.12 Business impact of rules on FDI	
6.13 Burden of customs procedures	
6.14 Imports % GDP	
6.15 Degree of customer orientation	
6.16 Buyer sophistication	

Appendix D. Visualized Decision Tree

Number of Leaves : 65 Size of the tree : 117