

Table of Contents

Introduction	7
 Chapter 1	
Review of the Results of Previous Studies on Detecting Financial Manipulation in Enterprises	9
1.1. Popular Fraud Detection Models	9
1.1.1. M-Score Model[1]	9
1.1.2. F-Score Model by Dechow, Ge, Larson and Sloan	12
1.1.3. Christensen, Paik and Williams' Model	15
1.1.4. Armstrong, Jagolinzer and Larcker's Models[2]	16
1.1.5. Beasley's Models[3]	18
1.1.6. Bonner, Palmrose and Young's Models[4]	22
1.1.7. Brazel, Jones and Zimbelman's Models[5]	25
1.1.8. Richardson, Tuna and Wu's Models[6]	26
1.1.9. Erickson, Hanlon and Maydew's Models[7]	28
1.1.10. Feng, Ge, Luo and Shevlin's Models[8]	31
1.1.11. Hribar, Kravet and Wilson's 2014 Models[9]	33
1.1.12. The Study by Price, Sharp and Wood[15]	36
1.1.13. Persons' Model[20]	40
1.1.14. Summers and Sweeney's Model[21]	41
1.1.15. Models by Spathis et al.[22]	42
1.1.16. Models by Yuan et al.[23]	43
1.1.17. Bell and Carcello's Model[24]	44
1.1.18. Spathis' Model[22]	45
1.1.19. Model by Spathis et al.[25]	46
1.2. Results of Previous Studies Using Black Box Algorithms	47
 Chapter 2	
The Biggest Financial Scandals of the 20th Century	59
2.1. List of US-Listed Companies Responsible for the Biggest Financial Scandals of the 20th Century	59

2.2. Company Mechanisms Relevant to Studies of Financial Statement Manipulation	68
2.3. Manipulation Mechanisms Used by the Studied US-Listed Companies	74

Chapter 3

Creating Own Models for Detecting Financial Manipulation	97
3.1. Source of Data and Description of the Study Method	97
3.2. Description of Machine Learning Algorithms	103
3.2.1. Linear Regression	106
3.2.2. Logistic Regression	108
3.2.3. Classification and Regression Trees (CARTs)	109
3.2.4. Random Forest	111
3.2.5. Support-Vector Machines (SVMs)	112
3.2.6. Classification Using Naive (Simplified) Bayesian Inference	118
3.2.7. Linear Discriminant Analysis	122
3.2.8. Gradient Boosting	124
3.2.10. Regression Estimation by the Gaussian Process Method	128
3.2.11. Deep Neural Networks (DNNs)	129
3.3. Results of Training Algorithms Based on the Biggest Scandals of the 20th Century	135
3.3.1. Logistic Regression	137
3.3.2. Decision Tree	139
3.3.3. XGBoost (Extreme Gradient Boosting)	143
3.3.4. Random Forest	153
3.3.5. Deep Neural Network	164
3.3.6. Naive Bayesian Inference	167
3.3.7. Linear Discriminant Analysis	177
3.3.8. Support Vector Machine (SVM)	178
3.4. Summary and Final Conclusions	181
Appendices	188
Bibliography	197