Title: The Application of Machine Learning and Artificial Neural Networks Algorithms to Predict Financial Distress

Data availability statement: The dataset will be made available on Royal Roads University data management repository

Funding statement: Three is no funding associated with this paper

Conflict of interest disclosure: We confirm that there is no conflict of interest with this paper

Ethics approval statement: There are no ethical issues associated with this paper

Permission to reproduce material from other sources: We confirm that there is no materials from other sources in this paper

Author Contribution: Both authors contributed equally to this paper

Authors' Details

Mark Lokanan

Mark Lokanan is a data scientist and white-collar criminologist and an Associate Professor in the Faculty of Management at Royal Roads University. He is a graduate from the School of Criminology, Simon Fraser University, Canada.

Mark E. Lokanan Ph.D
Associate Professor
Faculty of Management | **Royal Roads University**
T 250.391.2600 ext. 4386#
2005 Sooke Road, Victoria, BC Canada V9B 5Y2 | royalroads.ca
Mark.Lokanan@royalroads.ca

Sana Ramzan

Sana Ramzan is a doctoral student at Royal Roads University. She specializes in fraud and predicting financial distress using machine learning.

Graduate Student

2005 Sooke Road
Victoria, BC Canada
V9B 5Y2
sana.ramzan@royalroads.ca

1. Introduction

In 1999, Livent, a Canadian live entertainment corporation, was charged with manipulating its books by Canadian and US authorities. Then in 2002, Nortel Networks, a Canadian technology champion, was charged with manipulating its earnings to inflate profits. There are other well-known international cases of companies charged with manipulating their financial statements (e.g., Enron, WorldCom, Tyco, Tesco, Parmalat, Satyam, and Toshiba). Each one of these companies experienced financial distress that led to their executives performing actions well beyond the generally accepted accounting principles (GAAP) and International Financial Reporting Standards guidelines. Although GAAP allows accountants to adopt different methods to reflect the reality of companies' financial position, earnings manipulation becomes fraud when the accountants intentionally produce false and misleading information to mislead shareholders and other stakeholders. The relationship between earnings management and fraud is well-established in the literature (Perols, 2011; Ramírez-Orellana et al., 2017; Trompeter et al., 2013).

Companies can manipulate their financials by overstating their revenues and understating their expenses (Achakzai & Peng, 2023). Usually, financial statement manipulation is perpetuated to increase earnings through a series of events that lead to the aggressive interpretation of accounting rules. When management needs to increase earnings, the possibilities include using legitimate aggressive accounting practices or issuing fraudulent financial statements. Financial data are susceptible to manipulation and can be misrepresented intentionally or inadvertently (Lokanan, 2015). Earnings management occurs over several years and can increase the likelihood of committing fraud. The typical case of earnings manipulation involves managing earnings using discretionary accruals or committing fraud by creating fictitious sales revenues and under-accrued expenses to meet analysts' expectations (Trompeter et al., 2013). When businesses experience financial distress, they are motivated to commit fraud by manipulating their financial statements (Achakzai & Peng, 2023; Bao et al., 2022; Eilifsen & Knivsflå, 2021; Perols & Lougee, 2011). Transactions that were once stationary practices to smooth out earnings periodically can lead to intentional manipulation of the financials, questionable ethical practices, and outright fraud.

The advent of machine learning (ML) and artificial intelligence (AI) presents innovative opportunities for auditors and regulators to proactively identify high-risk firms and prevent potentially fraudulent activities. Specifically, supervised learning within ML/AI holds promise in unveiling earnings management practices and addressing some of the limitations inherent in conventional audit methodologies, especially when dealing with extensive datasets. In this paper, we apply ML and AI algorithms to predict earnings manipulations using the Beneish (1999) M-score for companies listed on the Toronto Stock Exchange (TSX). Our primary aim is to construct a novel model for forecasting earnings manipulations, utilizing readily available financial statement data from TSX-listed companies. We operate under the assumption that firms are more inclined to manipulate their earnings when facing financial distress. Thus, our central research question centers on utilizing ML and AI algorithms in predicting financial distress and identifying earnings manipulation in firms by leveraging the Beneish M-score.

In this paper, we will use various ML algorithms ranging from logistic regression and support vector machines (SVM) to ensemble-based classifiers and feedforward artificial neural networks (ANN) to predict financial distress and detect earnings manipulation in firms listed on the TSX. We also employed the classification and regression trees (CART) algorithm with bootstrapping of samples to enhance the interpretability of our models. By virtue of being one of the most established algorithms, logistic regression will be used to build the base model to compare performance with the other ML and neural network classifiers..

This paper adds to the growing body of research that leverages ML algorithms for financial distress prediction (Achakzai & Peng, 2023; Bao et al., 2022; Mselmi et al., 2017; J. Sun et al., 2014; Zhao et al., 2023).While previous studies in accounting and finance primarily relied on single predictive modelling techniques, such as linear and ensemble algorithms, SVM, and ANN for financial distress prediction (Achakzai & Peng, 2023; Bao et al., 2022; Mselmi et al., 2017), this paper introduces a novel approach by conducting four distinct experiments: predictive modelling, recursive feature elimination with cross-validation (RFECV), an AI sequential model, and CART with bootstrapping samples. By expanding the spectrum of classifiers beyond individual predictive modelling and traditional AI models that have dominated financial distress research (see Bao et al., 2022; Mselmi et al., 2017; Zhao et al., 2023), this paper makes a significant contribution by addressing the call for advanced ML techniques to fill the gaps discussed by previous research (Bao et al., 2022; Campa & Camacho-Miñano, 2015; Habib et al., 2020; Kuizinienė et al., 2022; J. Sun et al., 2014).

This study makes a significant methodological contribution to the literature by employing ML to predict financial distress with different samples. Unlike most prior research that employed ML to predict financial distress primarily using samples of Asian firms, often relying on the China Stock Market & Accounting Research Database for Chinese listed firms (Achakzai & Peng, 2023; Jiang & Jones, 2018; Tang et al., 2020; Zhao et al., 2023), this study extends the scope by utilizing data from the TSX. While some studies have predicted financial distress in publicly listed Taiwanese firms (Y.-P. Huang & Yen, 2019), South Korean manufacturing industries (Bae, 2012), French small and medium-sized firms (Mselmi et al., 2017), and United

Arab Emirates firms (Sreedharan et al., 2020), our study offers insights by broadening the applicability of ML techniques to predict financial distress using samples from a different context and jurisdiction.

To our knowledge, this study is the first to apply RFECV and CART with Bootstrapping in financial distress research. There is no universally superior model applicable in all contexts (Bao et al., 2020; Fern´andez-Delgado et al., 2014). Consequently, it becomes an empirical inquiry to investigate whether RFECV or CART with bootstrapping can surpass the traditional predictive modelling algorithms commonly utilized for predicting financial distress in finance and accounting research (Achakzai & Peng, 2023; Z. Li et al., 2021; Zhao et al., 2023). The aim is to empirically evaluate whether these approaches can provide a better understanding of financial distress in the specific context of the TSX-listed firms. Instead of generalized insights, this recognition highlights the importance of considering the unique characteristics and intricacies of specific datasets and problem domains when developing predictive models and underscores the need for context-specific solutions rather than relying solely on predictive modelling for classification tasks.

The remainder of this paper is organized as follows: Section 2 reviews the existing literature, emphasizing contemporary audit methodologies and the use of ML in research on predicting financial distress. Section 3 details the experimental design, covering aspects like data collection, preprocessing, the algorithms used, and performance metrics. Section 4 presents the results and offers a thorough analysis of the findings. Lastly, Section 5 concludes the paper and highlights potential directions for future research.

2. Current Audit Methodology

Financial statement manipulation is becoming increasingly prevalent owing to the inability of auditors to identify red flags of fraud. Typically, fraudsters exaggerate positive financial positions and conceal negative ones, making it difficult to detect manipulation in financial statements (Blay et al., 2007; Firth et al., 2011; Hartwig et al., 2017; Hilal et al., 2022; Lokanan & Sharma, 2022). Companies manipulate their earnings either by inflating their revenues or deflating their expenses (Eulerich et al., 2018; Filip & Raffournier, 2014; Perols & Lougee, 2011). As such, stakeholders' access to accurate and reliable financial information is limited in situations of corporate distress (DeZoort & Harrison, 2018; Khaksar et al., 2022). Not only does misstatement of financial information undermine trust between businesses and their users, but it could also lead to disastrous consequences for investors and creditors. It is therefore imperative that steps be taken to ensure that users have access to accurate financial statements through independent audits (DeZoort & Harrison, 2018; Eulerich et al., 2018; Khaksar et al., 2022).

Auditors play a crucial role in ensuring the integrity of financial statements and should continuously enhance their audit procedures to detect any potential irregularities or manipulations (Akther & Xu, 2020; Oyerogba, 2021). However, they often face challenges in identifying transaction omissions due to audit procedures that may be inadequately designed (Hamilton & Smith, 2021). Given this context, both financial statement users and auditors should prioritize vigilant monitoring of companies to uncover discrepancies, anomalies, or overstated positives in their engagements (M. E. Lokanan & Sharma, 2022). Such rigorous monitoring can serve as early warning signs, enabling timely actions to mitigate the adverse effects of distorted financial statements (Aslam et al., 2022; Aviantara, 2021; Dechow et al., 1996).

Despite the use of the engagement approach to audit financial statements, many researchers have highlighted the outdated nature and inability of traditional auditing procedures to identify red flags that may signify potential fraud (Aubert et al., 2019; DeZoort & Harrison, 2018; Sun et al., 2014). The fraud triangle, a measure outlined in SAS No. 99, evaluates pressure, opportunity, and rationalization to detect fraudulent activities (Aubert et al., 2019; Davis & Pesch, 2013; DeZoort & Harrison, 2018; Morales et al., 2014). However, the fraud triangle has been criticized for its limited capabilities to assess patterns of financial distress that ultimately set the groundwork for pressures and opportunities to commit fraud (Davis & Pesch, 2013; Firth et al., 2011; Lokanan, 2018; Nasir et al., 2019). Consequently, researchers have concluded that more comprehensive investigative strategies are required to properly identify suspicious activities and prevent future instances of fraud (Lokanan, 2015; Morales et al., 2014). One promising approach involves the utilization of ML and AI-based technologies to identify predictors of financial distress (Campa & Camacho-Miñano, 2015; Mselmi et al., 2017; Zhao et al., 2023).

2.1 Machine Learning for Financial Distress Prediction

Although the application of computational technology is in its early stages, recent studies have shown that ML and AI hold a significant advantage over traditional audit approaches in detecting red flags of fraud in financial statements. More specifically, these researchers found that ML- and AI-based algorithms yield high precision, sensitivity, and accuracy scores in detecting fraud (Achakzai & Peng, 2023; Bao et al., 2020; Zhao et al., 2023). Essentially, ML- and AI-based tools are more adept at detecting anomalies and red flags of fraud in financial statements than the traditional sampling approach (Hilal et al., 2022; Mselmi et al., 2017). More specifically, researchers have found that ML and AI technologies can identify and report more accurately on true positives and negatives, lending greater support to the use of analytics in the fraud risk management processes (Blay et al., 2007; Kuizinienė et al., 2022; M. E. Lokanan & Sharma, 2022; Qiu et al., 2021; van der Heijden, 2022).

Researchers have employed various ML and AI techniques to predict fraud in financial statements. Techniques such as neural networks, SVM, and decision trees have been employed to predict anomalies and uncover fraudulent accounting practices in organizations (Achakzai & Peng, 2023; Hilal et al., 2022; Kim & Kogan, 2014; M. Lokanan et al., 2019; Mselmi et al., 2017). The application of ML- and AI-based technologies enable more effective processing of larger volumes of data while generating more accurate predictions to reduce instances of false positives (Achakzai & Juan, 2022; Hilal et al., 2022; Kim & Kogan, 2014). Others have found that ML models capture patterns and trends in the data and provide opportunities to gain deeper insight into potential warning signs of fraud within financial statements that may often be missed by human analysts (Cho et al., 2020; Hajek & Henriques, 2017).

Integrating ML- and AI-based analytics with traditional statistical methods has proven more efficient than manual analysis to detect anomalies in financial statements (Achakzai & Juan, 2022; Albizri et al., 2019; Alden et al., 2012; Gupta & Mehta, 2021; Kim & Kogan, 2014). In particular, research have found that combining ML with the existing Beneish M-score produces superior results compared to manual computation as a means of predicting financial distress (Papík & Papíková, 2019). Subsequent work by Papík and Papíková (2022) also showed that predictive models developed using data mining and ML effectively identify accounting fraud at a much higher rate than traditional methods.By combining and utilizing the best features of ANN, SVM, and ensemble models, these algorithms have proven to be highly effective in predicting financial distress and detecting fraud (Mselmi et al., 2017; M. Nasir et al., 2021; Zhao et al., 2023).

The use of ML- and AI-based technologies to predict financial distress and fraud is a well-established area of inquiry (Achakzai & Peng, 2023; Bao et al., 2020; M. E. Lokanan, 2022; Mselmi et al., 2017; Zhao et al., 2023). However, there is a need for more nuanced research focused on applying ML and AI classifiers to predict financial distress and detect earnings manipulation using financial ratios. This paper aims to bridge the knowledge gap by presenting an investigation in which ML and AI technology are used to predict financial distress and detect earnings manipulations using data from the TSX. The Beneish M-score serves as a proxy for manipulation. Based on the preceding literature review, we attempt to predict financial distress and detect earnings manipulations by asserting that:

Financial Distress = *fd* (profitability, liquidity, efficiency, solvency, and operational performance indicators)

3. Experimental Setting

This paper employs classification models to assess financial distress and detect earnings manipulation in TSX-listed companies. Our analysis aims to reveal critical factors for predicting financial distress and assess the likelihood of future distress occurrences. Our objective is to identify the most relevant features for predicting financial distress and build models that accurately classify firms at risk based on various financial ratios. This predictive approach has demonstrated effectiveness in the early detection of financial distress and warning of earnings manipulation (Mselmi et al., 2017; J. Sun et al., 2014; Zhao et al., 2023).

3.1. Data Collection

The data for this project were collected from the TSX. The TSX is the ninth largest stock exchange in the world and is a member of the S&P/TSX Composite Index, an index of the top five hundred companies in Canada. With a market capitalization of CAD$2.1 trillion, it is home to many large Canadian banks, insurance companies, and other financial institutions. The data were collected for the fiscal year ending 31 March 2021. After filtering out companies for which data were unavailable, the final sample consisted of 464 observations (i.e., companies). As shown in Appendix 1, the variables collected represent profitability, liquidity, efficiency, operational performance, and solvency ratios as predictors of financial distress. By investigating these ratios, we can gain insights into which companies are at risk of financial distress and be able to detect earnings manipulation.

3.2. Variables and Measurements

3.2.1. Independent Variables

The independent variables used to predict financial distress are shown in Appendix 1. The measures focus on important indicators of corporate financial health, such as profitability, liquidity, efficiency, solvency, and operating performance. These measures are further broken down into several distinct ratios that can be used to assess a company's current and future financial performance. Because financial health is an integral factor in any decision-making framework, readily available comprehensive metrics offer decision-makers more reliable insights into which variables would prove most effective in predicting financial distress.

3.2.2. Profitability Ratios

Profitability ratios are excellent determinants of financial performance (Demetriades & Owusu-Agyei, 2022; Fang et al., 2017). Previous research has found that return on asset (ROA) helps predict financial statement fraud. ROA assesses the company's performance and determines the effectiveness of the assets in generating revenues (Fang et al., 2017; Gupta & Mehta, 2021; Hajek & Henriques, 2017; J. Sun et al., 2014). Several other profitability ratios, such as gross profit margin (GPM), net profit margin (NPM), return on equity (ROE), and net profit to total assets (NPTA) have been used effectively to predict financial distress in companies (Craja et al., 2020; Fang et al., 2017; Gupta & Mehta, 2021; S. Y. Huang et al., 2017; Song et al., 2014). Profitability ratios are one of the most fraud-sensitive types of ratio in that their values differ significantly between financially distressed companies (true positives) and non-distressed companies (false positives), making them useful for predicting financial distress (Gupta & Mehta, 2021).

3.2.3. *Liquidity Ratios*

Liquidity ratios evaluate a firm's ability to meet short-term financial obligations (Gupta & Mehta, 2021; Hajek & Henriques, 2017; S. Y. Huang et al., 2017). Some liquidity ratios used in previous research to predict financial distress are the current ratio (CR), quick ratio (QR), days receivable turnover (AccTR), working capital to total assets (WCTA), cash-to-current liabilities (CCL), and cash-to-total assets (CTA) (Albizri et al., 2019; Fang et al., 2017; Gupta & Mehta, 2021; Hajek & Henriques, 2017; Song et al., 2014). These liquidity ratios have proven useful in predicting financial distress in companies. The use of liquidity ratios in this study is important because firms with perceived high liquidity can attract investors, which proves to be one of the motivations for manipulation by financially distressed firms (Gupta & Mehta, 2021).

3.2.4. *Efficiency Ratios*

Previous research has shown that efficiency ratios, such as accounts days receivable turnover (AR), days accounts payable turnover (AP), asset turnover (AT), and inventory turnover (InvT), have been widely recognized as valuable indicators of a company's financial health (Albizri et al., 2019; Dimitropoulos & Asteriou, 2009; Fang et al., 2017; Gupta & Mehta, 2021; Habib et al., 2020; Zainudin & Hashim, 2016). These ratios offer insights into a company's competitiveness and profitability, making them crucial in the prediction of financially distressed firms (Aviantara, 2021; Habib et al., 2020). Notably, higher efficiency ratios are associated with a greater likelihood of account manipulation, adding to their relevance in financial distress prediction (Gupta & Mehta, 2021; Song et al., 2014).

3.2.5. Solvency Ratios

Another valuable indicator in determining a company's financial health is identifying the level of debt in an organization's capital structure using safety or solvency ratios. Solvency ratios such as debt to equity (DEQ), total liabilities to total assets (TLTA), net profit to total liabilities (NPTL), cash to total liabilities (CTL), current assets to total assets (CATA), and current liabilities to total assets (CLTA) have been employed to assess organizations' probability of default (Albizri et al., 2019; Craja et al., 2020; Fang et al., 2017; Gupta & Mehta, 2021; S. Y. Huang et al., 2017). Lenders and investors desire low solvency ratios because they indicate protection in the event of a business collapse and ensure the return on the initial investment in liquidation (Gupta & Mehta, 2021).

3.2.6. Operating Performance

Others have employed earnings and cash flow ratios to assess the operational performance of companies in the context of financial distress prediction. Ratios such as Earnings Before Interest and Taxes (EBIT), Earnings Before Interest, Taxes, Depreciation, Amortization, and Rent (EBITDAR), and Earnings Before Interest, Taxes, Depreciation, and Amortization (EBITDA) have been widely used in financial distress research (Albizri et al., 2019; Campa & Camacho-Miñano, 2015; N. Li, 2016). These ratios offer valuable insights into a firm's operational performance and financial health, making them essential components of predictive modeling for financial distress.

3.2.7. Control variables

Several control variables were utilized to isolate and measure the specific effect of the independent variable on financial distress while holding other relevant factors constant (Bernerth & Aguinis, 2016; Farber, 2005). In this study, a comprehensive set of control variables was incorporated to evaluate the precise impact of profitability, liquidity, efficiency, solvency, and operational variables on predicting financial distress. These variables encompassed the firm's credit rating, R-Score, Enterprise Value (EV), Market Capitalization (MarketCap), revenue growth (Rev_G), net income growth rates (NI_G), debt growth (D_G), Corporate Governance Score (CG_Score), Audit and Risk Oversight (ARO), Board Structure (BoardS), Shareholder Relations (ShareR), and Compensation (Comp). By including these control variables in our analysis, we sought to account for potential confounding factors and variations in firm characteristics (Becker et al., 2016; Bernerth & Aguinis, 2016). This approach allowed us to assess more accurately the individual and combined contributions of the chosen financial and operational indicators in predicting financial distress, thereby enhancing the robustness of the models and reliability of our conclusions.

3.3. Dependent Variable

The dependent variable for this project is the Beneish M-score (Beneish, 1999). The Beneish M-score is one of the best accounting indicators for predicting financial distress and detecting earnings manipulation (Albizri et al., 2019; Aviantara, 2021; Maniatis, 2022; Repousis, 2016). The Beneish M-score is a probabilistic model that uses financial ratios, which are categorized into eight variables to detect earnings manipulation in a company's earnings. Companies with higher M-Scores tend to manipulation their earnings because they are experiencing financial distress. Others noted that companies could use the "M-score models and data mining for an early indication of financial distress or red flags for detecting financial fraud" (Tarjo & Herawati, 2015, p. 2). In a more recent study of the M-score model, Beneish and Vorst claimed that a cost-based assessment of models is preferable to traditional model comparison measures (2022). For this paper, the Beneish M-score was recoded as a binary classifier of 0 and

1. If the Beneish M-score is > -2.22, the company is coded as (0=no-manipulator); if the Beneish M-score is < -2.22, the company is coded as (1=manipulator) (Maniatis, 2022). The formula to represent financial distress in the earnings manipulation modeling is shown in equation 1:

$$y = \{1, \text{Financial distress} \\ 0, \text{No-financial distress}\} \qquad \textit{eq.1}$$

3.4. Machine Learning Workflow

Figure 1 illustrates the workflow employed in this project to construct a classification model capable of assessing whether a firm has engaged in financial manipulation. The process begins by gathering data suitable for building the classification model, followed by preprocessing of the features to identify outliers or noise in the data. Preprocessing also includes standardizing the data so that all observations are within the same range. Subsequently, the data are split into train and test sets, with the test set acting as an unseen sample from which to evaluate model performance. Finally, the evaluation of the model takes place using validation techniques on the test set before the final prediction.

3.4.1. Data cleaning and Preprocessing

Figure 1 displays the ML workflow. Data cleaning is a critical step in the data preprocessing pipeline and is essential for obtaining meaningful and reliable insights from the data. A check for duplicates indicates that there were none. The feature "Rating" represents the ratings from Moody's and ranges from "Aaa" to "C". These letters were replaced with integers, where "Aaa" = 1 and "C" = 9 the lowest rating. Outliers were identified using the Z-score. The most common threshold value for detecting outliers using the Z-score is ± 3 (Salgado et al., 2016). A Z-score threshold of ± 3 was applied to detect potential outliers in the dataset. Data points with Z-scores beyond this range are typically considered outliers (Chikodili et al., 2021). Missing values were imputed using K-nearest neighbors (K-NN). K-NN is the preferred choice for imputation when the missing data is ≤ 40% (Zhang, 2012). To test for multicollinearity, any feature with a Pearson correction of .70 and above was removed from the dataset.

The dataset exhibited a notable imbalance, with approximately 88% of the firms categorized as "not distressed" and the remaining 12% classified as "distressed." To rectify this imbalance and create a more equitable representation of both classes, we employed the SMOTE+ENN technique. SMOTE+ENN combines the Synthetic Minority Over-sampling Technique (SMOTE) with the Edited Nearest Neighbors (ENN) algorithm to effectively address the class imbalance issue (Puri & Kumar Gupta, 2022). SMOTE generates synthetic samples for the minority class by interpolating between existing instances, thus augmenting the distressed samples in the dataset (M. E. Lokanan, 2023). Subsequently, the ENN algorithm identifies and eliminates noisy or redundant data points to enhance the overall quality of the dataset (Sisodia et al., 2017). This process ensures that the models have a more balanced and representative dataset to learn from, ultimately improving their predictive performance in identifying financial distress. As seen in equation 2, the financial distress class consists of only 12% of the labeled data. After oversampling, the train and test dataset had equal samples (195 instances) of class 1 and class 0.

$$\textit{Financial distress}(Fd)_{Yes} = \text{(Total transactions)/(Distress transactions)} * 100$$
$$Fd_{yes} = (464/56)*100 = 12\% \qquad \textit{eq. 2}$$

The min-max scaling technique was applied to the dataset to normalize the features, ensuring that they all fall within the same range. Normalization is a crucial preprocessing step in ML as it compresses all the variables to a common scale, preventing certain features from dominating others during the modeling process (Islam, 2021; Singh & Singh, 2020). Min-max scaling specifically transforms the values of each feature to lie between a specified range, typically 0 and 1. By compressing the variables between 0 and 1, min-max normalization helps to maintain the integrity of the data's distribution while ensuring that all attributes contribute equally to the modeling process.

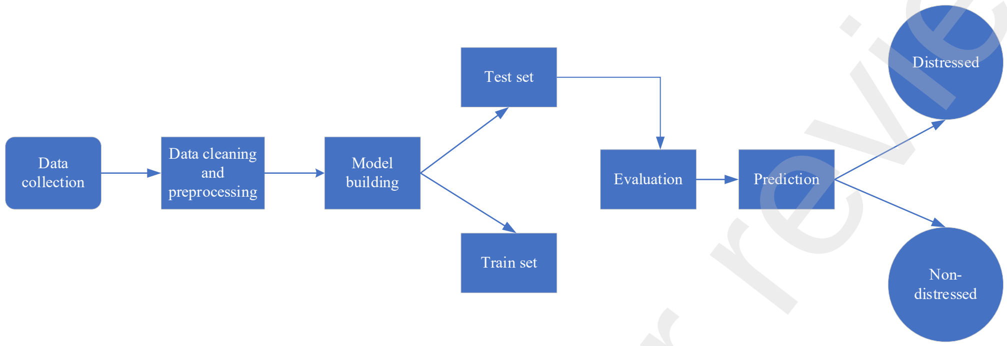

Figure 1: Overview of the Machine Learning Workflow

3.4.2. Types of Experiments

As can be seen in Table 1, the experiments conducted in this study encompass a diverse set of classifiers, algorithms, and techniques. Linear classifiers, such as logistic regression, are employed for predictive modeling. Ensemble classifiers, notably random forest, serve a dual purpose, facilitating both predictive modeling and feature selection. Tree-based classifiers, including decision trees and CART, are utilized for predictive modeling and RFECV. Finally, neural networks, specifically ANN, are harnessed for predictive modeling. These experiments provide a comprehensive evaluation of the effectiveness of these various methods in predicting financial distress within the specific context of TSX-listed firms.

To perform our experiments, we divided the dataset into two subsets: a training set comprising 75% of the data and a test, or holdout, set containing the remaining 25%. The model was then trained using the training data, enabling it to make predictions for the response variable on the test data, ensuring evaluation on unseen observations. By randomly selecting samples from the dataset, we prevent any inherent data ordering from biasing our results. We also employed GridSearchCV to optimize the predictive models (Botchey et al., 2020a; M. E. Lokanan, 2022).

Table 1: Types of Experiments

Classifiers	Algorithms	Predictive	RFECV	Bootstrapping
Linear	Logistic Regression	**x**	**x**	
Ensemble	Random Forest	**x**	**x**	**x**
Tree-Based	Decision Tree	**x**	**x**	
	CART	**x**	**x**	**x**
Neural Networks	ANN	**x**		

3.4.3. Hypermeters Used for Tuning

Table 2 shows the parameters used to tune the models. While running GridSearchCV helped to tune and optimize the predictive and AI Sequential models, they were computationally expensive and resource intensive. For the RFECV and CART models, a manual tuning approach was adopted, which proved to be effective in enhancing model performance compared to default settings. The manual calibration of these models allowed for a more tailored and efficient configuration, ultimately yielding improved predictive accuracy. By manually fine-tuning the hyperparameters, we were able to gain a better understanding of how each parameter influenced the model's predictions within the specific context of financial distress prediction. This level of control and customization is especially valuable when dealing with complex financial datasets or when specific constraints or requirements are imposed on the model for accurate distress prediction.

Table 2: Hyperparameter Tuning

Algorithms	Predictive Modellign with GridSearchCV
Logistic Regression	'C': 100, 'max_iter': 100, 'penalty': 'l2', 'solver': 'newton-cg
Decision Tree	max_depth=12, max_features='auto', min_samples_leaf= 4, splitter= 'best
Random Forest	'max_depth': 4, 'max_features': 'auto', 'min_samples_leaf': 2, 'n_estimators': 10
SVM	'C': 100, 'degree': 2, 'gamma': 1, 'kernel': 'rbf'

Algorithms	RFECV
Logistic Regression	solver='lbfgs', penalty='l2', C=1.0, max_iter=100, class_weight='balanced', random_state=42
Decision Tree	max_depth=3,min_samples_split=2,criterion='gini', random_state=42
Random Forest	n_estimators=100, max_depth=3, min_samples_split=2, criterion='gini', random_state=42
SVM	kernel='sigmoid', C=1.0, random_state=42

Algorithms	Bootstrapping of Samples
CART	n_estimators=100, max_depth=3, max_features='sqrt', random_state=42

Algorithms	Artificial Neural Network
AI Sequential Model	'dropout_rate': 0.0, 'learn_rate': 0.001, 'unit': 5

3.5. Algorithms and Parameters Tuning

3.6.1 Logistic Regression

The benchmarked algorithm used to evaluate model performance is logistic regression. Logistic regression is a supervised learning technique that estimates or predicts the likelihood of a binary event occurring on linear data (Lokanan & Sharma, 2022; Papík & Papíková, 2022). The mathematical formula of logistic regression is represented in equation 3:

$$y = e^{(b0 + b1*x)} / (1 + e^{(b0 + b1*x)}) \qquad \textit{eq. 3}$$

Where

y is the predicted output,
b0 is the bias or intercept term, and
b1 is the coefficient for the single input value (x).

Each column in the input data contains a *b* coefficient derived from the training data (Lokanan & Sharma, 2022; Papík & Papíková, 2022). Logistic regression is one of the most common ML classification techniques owing to its ease of implementation, functionality, and effectiveness in categorizing new information (Nasir et al., 2021; Wang & Song, 2011). To evaluate if a company is experiencing financial distress, logistic regression fits the data into a logistic function and compares their value against a predetermined threshold (Abbas et al., 2020; Papík & Papíková, 2022; Wang & Song, 2011).

3.6.2. Decision Tree Classifier

Decision trees are a type of ML algorithm used to help solve regression and classification problems. The algorithm works by creating a tree-like structure, with each branch representing a different decision (Botchey et al., 2020b; Sahin et al., 2013). The tree is created by splitting the dataset into smaller and smaller subsets until each subset only contains one data point. Once the tree is created, it can be used to predict new data points. Decision trees are generally highly accurate and are often used in predictive modelling tasks (Sahin et al., 2013; Tian et al., 2020). However, a single decision tree is often insufficient to produce effective results. For more accurate prediction, random forest classifiers (RFC) have proven to be significantly effective in fraud classification tasks (Aslam et al., 2022; Lokanan, 2022; Nami & Shajari, 2018).

3.6.3. Random Forest

The random forest algorithm is an extension of the decision tree algorithm that builds multiple trees and combines their predictions (Aria et al., 2021). The random forest classifier (RFC) is useful because it addresses nonlinearity in the data. Unlike the single decision tree, the RFC is a collection of decision-tree classifiers that generate a collection of decorrelated trees (random forest) based on multiple simulations of the actual training sample (Bhattacharyya et al., 2011; Breiman, 2001; Lokanan & Sharma, 2022; Papík & Papíková, 2022). As a set of decision trees, the RFC achieves decreased variance and decreased sensitivity to the training data. The benefits of RFC over a single decision tree are increased stability, efficiency for large and small datasets, increased accuracy, robustness to noise, reduction of overfitting, adaptivity in handling multiple data attributes, and computational speed that is faster than other ensemble methods (Bhattacharyya et al., 2011; Breiman, 2001; Papík & Papíková, 2022; Wang & Song, 2011). Entropy is used to split the trees for the random forest model. Mathematically, entropy is represented in equation 4:

Where
P is the probability of the +ve class,
(1-p) = The probability of the -ve class,
Hence, we are able to calculate H(y) for the dependent variable, financial distress.

$$H(y) = -p\log2(p) - (1-p)\log2(1-p) \qquad eq.\ 4$$

3.6.4. Support Vector Machines (SVM)

SVM is a powerful tool for ML models and has been successfully used in a variety of tasks such as classification, regression, and outlier detection (Ghosh et al., 2019; Rtayli & Enneya, 2020). The SVM algorithm works by finding a hyperplane that best separates the data into classes. To do this, the algorithm first computes a set of support vectors, which are points in the data that are closest to the hyperplane (Botchey et al., 2020b; Rtayli & Enneya, 2020). The distance between the support vectors and the hyperplane is called the margin (Lokanan & Liu, 2021, p. 11). The SVM algorithm then maximizes the margin to find the optimal decision boundary (p. 11). When the data are linearly separable from the hyperplanes, then the SVM is represented by the linear formula in equation 5:

Where

$B=(B_1,\ldots B_n)$, and
$X=(X1,\ldots Xn)$ are n-dimensional vectors

$$B_0+B_1X_1+\ldots+B_nX_n=0 \qquad eq.\ 5$$

In the event that the data are linearly inseparable, then the kernel tricks such as polynomial, Gaussian radial basis, sigmoid, and hyperbolic tangent are applied to map the data into a higher dimensional space (Botchey et al., 2020). SVM has proven to be a robust algorithm for fraud detection tasks (Botchey et al., 2020; Rtayli & Enneya, 2020). Researchers have used SVM to predict data patterns that may indicate fraud and aberrant patterns in financial transactions (Botchey et al., 2020; Lokanan & Liu, 2021). By studying historical transaction data, SVM can learn to recognize "typical" transaction characteristics from abnormal transactions (Liu et al., 2021; Rtayli & Enneya, 2020).

3.6.5. ANN

ANN has been previously used in fraud prediction and financial distress research with very good performance (Bao et al., 2020; Mselmi et al., 2017). Owing to the presence of multiple neurons, ANN is a reliable algorithm because it can analyze large datasets, efficiently handle and process nonlinear data, and easily solve complicated tasks (Johnson & Khoshgoftaar, 2019; M. Lokanan, 2022). Figure 2 presents an illustration of a feedforward ANN. The neural network takes in three input features, processes them through two hidden layers, and produces a binary prediction (either "Distress" or "No Distress") based on the activations of the neurons in the output layer. The feedforward method highlights factors influencing the output during every step in a neural network, thus allowing for improved predictive accuracy. The mathematical representation of the feedforward ANN is shown in equation 6:

Where

$\sum_{i=1}$ represents the summation of the weighted inputs from all connected neural networks.
W_i are the weights assigned to the connections between the input features (predictors) X_i and the neural network.
b is the bias or error term.
$\sum(x)$ is the function that computes the weighted sum of inputs and bias to predict the output y.

$$\sum i=1(Wi\cdot Xi)+b=y \qquad eq.\ 6$$

This equation represents the basic operation of a neural network where inputs are multiplied by their respective weights, summed up, and then passed through an activation function to produce the output.

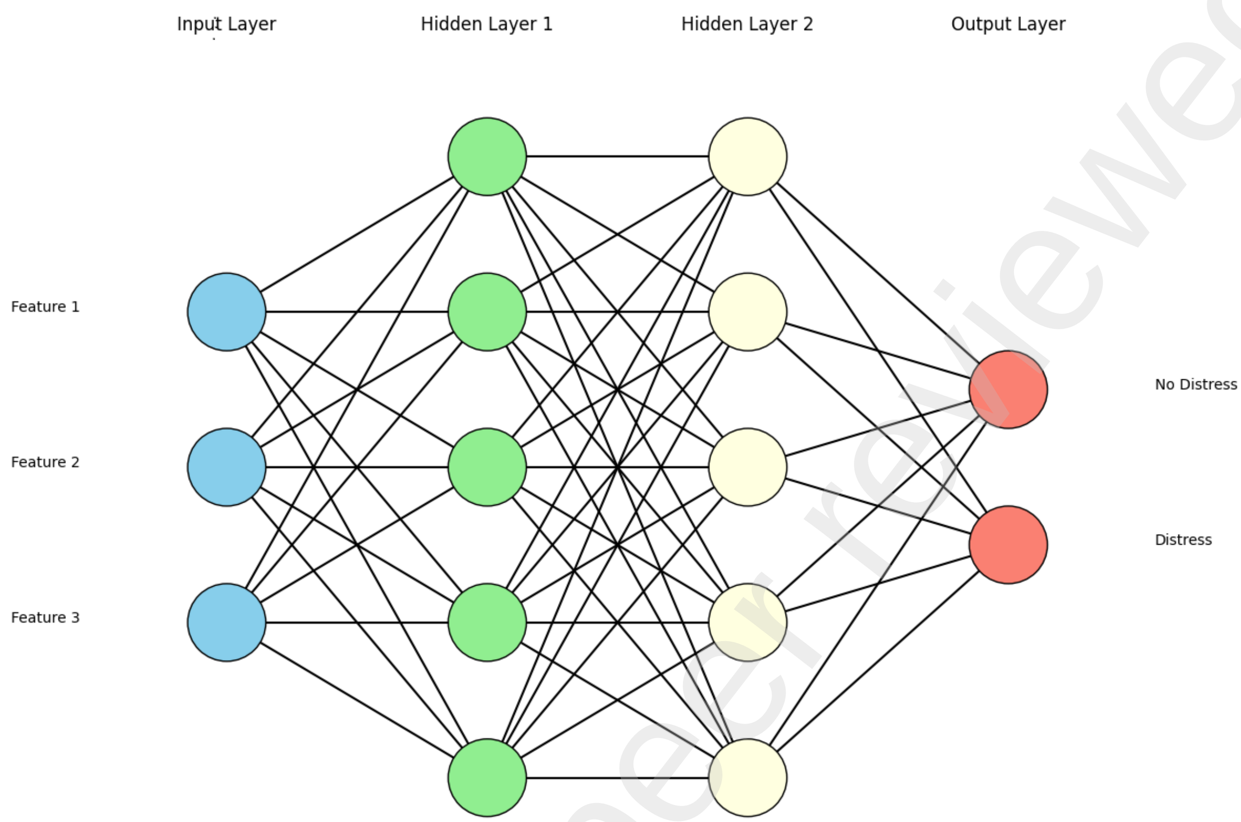

Figure 2: Neural Network Model

3.6.6. CART with Bootstrapped Samples

Bootstrapping involves resampling the dataset with replacement to create multiple subsets. These bootstrapped samples will be analyzed using CART to assess the robustness and stability of the model when applied to different variations of the data. The bootstrapped samples consist of various ratios ranging from 1:1 to 10:1 and represent different ratios of non-distress firms to financial distress firms in a dataset. Each sample is characterized by the ratio, the number of non-distress firms, and the number of financial distress firms. For instance, the first sample has a 1:1 ratio, meaning there are an equal number of non-distressed and financial distress firms. Subsequent samples exhibit varying levels of class imbalance, with the ratio increasing from 2:1 to 10:1, indicating a decreasing proportion of financial distress firms relative to non-distressed firms. These bootstrapped samples allow for the exploration of model performance under different class distribution of the data, aiding in the evaluation and optimization of predictive models for financial distress. Furthermore, bootstrapping helps estimate the variability of model performance and allows the model to generalize effectively across different data subsets.

3.7. Performance Measures

The performance of a classifier can be visualized through a confusion matrix. As can be seen in Figure 3, the matrix is made up of four quadrants, each of which represents the predicted and actual values for one class. The quadrants are labelled true positive (TP), false positive (FP), true negative (TN), and false negative (FN). The true positive rate (TPR) is the proportion of cases that are correctly classified as positive, while the false positive rate (FPR) is the proportion of cases that are incorrectly classified as positive. The true negative rate (TNR) is the proportion of cases that are correctly classified as negative, while the false negative rate (FNR) is the proportion of cases that are incorrectly classified as negative. In general, a classifier with a high TPR and FPR will be more accurate than one with a lower TPR and a higher FPR (Botchey et al., 2020; Lokanan, 2022).

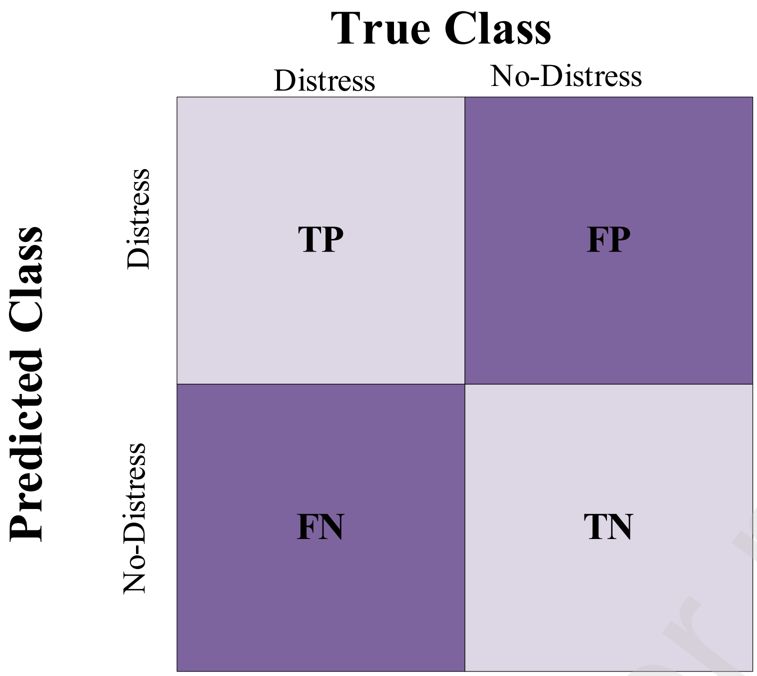

Figure 3: Confusion Matrix

The performance metrics used in this paper are shown in Table 3. Accuracy is the portion of total observations that the model predicted correctly. Precision is a measure of accurate prediction. In the context of this paper, precision measures how often the classifier is correct when it predicts that a firm is experiencing financial distress. Precision attempts to answer the following questions: *what proportion of the positive identifications was actually correct*? Recall is a measure of the completeness of the model. Recall measures how correct the model was in predicting actual financial distress and answers the following question: *what proportion of actual distressed firms were identified correctly*? The F-measure combines precision and recall into a single metric, while the receiver operating characteristics (ROC) curve visualizes the trade-off between sensitivity and specificity. Ultimately, the objective is to choose the model that strikes the best balance between accuracy and precision for deployment (see Lokanan, 2022; Lokanan & Sharma, 2022).

Table 3: Evaluation Metrics

Metrics	Formulae
Accuracy	TP+TN/(TP+TN+FN+FP
Error rate	ER=1-Accuracy
Sensitivity@k	TP/(TP+FN)
Specificity@k	TN/(TN+FP)
Precision	TP/(TP+FP)
F-measure	2*(Precision * Recall)/Precision+Recall
ROC Area Under the Curve (AUC)	Plot the TPR [TP/TP+FN] and FPR [FP/TN+FP] to a single score

4. Experimental Results

4.1. Summary Statistics

Table 4 shows the summary results of the explanatory features and their effects on financial distress. While ratios from profitability, solvency, and operational performance show potential signs of financial distress, there are no specific signs in the liquidity and efficiency ratios that strongly indicate financial distress. In terms of profitability, the mean values for ROA, GPM, NPM, ROE, and NPTA are generally lower than zero, indicating potential financial distress. These ratios typically reflect a firm's ability to generate profits in relation to its assets and equity. As for the solvency ratios, the mean values for DEQ, TLTA, and CLTA are relatively high, suggesting potential financial distress. Higher debt ratios and total liabilities ratios may indicate increased financial risks. The mean values of the operational performance ratios, namely, EBITDAR, EBITDA, and EBIT, are relatively low compared to their respective maximum values. These findings indicate potential financial distress, as these ratios represent the firm's operational profitability and ability to cover its operating expenses and debts.

Table 4: Summary Results

Profitability						
index	ROA	GPM	NPM	ROE	NPTA	
count	463	463	463	463	463	
mean	-0.02	38.29	-15.56	-0.77	-0.02	
std	0.28	35.5	177.52	18.75	0.28	
min	-3.5	-376.67	-2156.49	-400.22	-3.5	
max	0.85	127.05	818.22	32.99	0.85	
Liquidity						
index	CR	QR	AccPT	CCL	WCTA	CTA
count	463	463	463	463	463	463
mean	2.67	2.09	55.01	1.21	0.14	0.13
std	4.06	3.03	114.29	2.44	0.22	0.16
min	0.06	0.04	-271.39	0	-0.53	0
max	49.67	22.9	1819.35	18.4	0.86	0.8
Efficiency						
index	AT	InvT	AR	AP		
count	463	463	463	463		
mean	0.59	152.21	370684533.5	300977347.9		
std	0.52	963.67	1222854468	1537013650		
min	-0.43	0	0	0		
max	3.81	18971.07	16098000000	29136000000		
Solvency						
index	DEQ	TLTA	NPTL	CTL	CATA	CLTA
count	463	463	463	463	463	463
mean	2.95	0.52	-0.02	0.62	0.34	0.2
std	90.89	0.31	1.25	1.69	0.25	0.17
min	-666.29	0.01	-6.47	0	0	0
max	1835.89	3.7	16.4	13.98	0.99	1.25
Operational Performance						
index	EBITDAR	EBITDA	EBIT			
count	463	463	463			
mean	0.08	620425545.5	0.07			
std	1.66	1834219690	0.28			
min	-20.37	-1742816000	-3.26			
max	7.62	20331000000	0.87			

4.2. Control Variables

Table 5 displays the results of the control variables. The results provide valuable insights into various factors that can potentially influence a company's financial health and stability. Among these metrics, credit ratings and R_Score are key indicators of a firm's overall financial health, with an average Rating of 6.79 suggesting that many companies in the dataset are relatively stable. However, there is some variability, with scores ranging from 1 to 9, indicating the presence of both strong and distressed firms. When examining financial performance, Enterprise Value and Market Capitalization are critical. The substantial variation in these scores, with a mean Enterprise Value of approximately $6.1 billion and a mean Market Capitalization of around $4.3 billion, reflects differences in company sizes and valuations within the dataset. It is noteworthy that there are companies with negative Enterprise Value, which could indicate severe financial distress. Revenue and income growth are also vital for assessing a company's performance. The average Revenue Growth of 0.87% suggests relatively modest revenue expansion, while the average Net Income Growth of 1.51% shows a similar trend in profitability. However, the wide standard deviations in these figures suggest significant variations among firms, with some experiencing negative growth, possibly indicating financial distress. Debt Growth is another critical factor. The average Debt Growth of 0.58 suggests moderate debt accumulation, but the presence of negative values implies that some companies are reducing their debt load to reduce their financial leverage.

Corporate governance is fundamental to long-term stability (Farber, 2005; N. A. B. M. Nasir et al., 2019). The analysis of corporate governance measures (Corporate Governance Scores, Audit and Risk Oversight, Board Structure, Shareholder Relations, and Compensation scores) indicates moderate to high scores on average but with significant variability among companies. These findings indicate that, on average, companies exhibit moderate to high scores in various aspects of corporate governance, such as the strengths of audit committees, board structure, executive compensation, and overall corporate governance practices. However, the significant variability among companies in these measures suggests that financial distress is not solely determined by corporate governance scores. While stronger governance practices may exist on average, the presence of weaker governance in some companies implies that financial distress cannot be solely attributed to governance deficiencies. Other factors and circumstances likely play a substantial role in the occurrence of financial distress among firms.

Table 5: Summary Results Control Variables

index	count	mean	std	min	max
Rating	463	6.79	2.51	1	9
R_Score	463	4.16	0.67	2	5
EV	463	6085702971	19364571360	-431125203	2.67566E+11
MarketCap	463	4338120539	12669013099	0	1.33198E+11
Rev_G	463	0.87	7.54	-4.81	156.6
NI_G	463	1.51	19.66	-270.21	174.81
D_G	463	0.58	3.92	-1	65.26
CG_Score	463	5.46	2.77	1	10
ARO	463	5.52	2.78	1	10
BoardS	463	5.42	2.77	1	10
ShareR	463	4.9	2.91	1	10
Comp	463	5.47	2.86	1	10

Table 6 looks at the financial ratios indicating signs of financial distress in comparison to non-financial distress firms. As expected, financially distressed firms exhibit concerning signs. On average, these firms exhibit a negative ROA of -0.05%, indicating inefficiency in generating profits from their assets. While they manage a positive GPM of 33.28%, the strikingly negative NPM of -69.94% is alarming, pointing to significant losses. Interestingly, the ROE remains positive at 0.42%, indicating that these firms generate some returns for their equity holders. It's worth noting that some distressed firms maintain a negative DEQ ratio, implying a more equity-heavy capital structure. The negative EBITDAR of -0.42 suggests potential operational challenges. The results suggest that distressed firms generally struggle with profitability, incur substantial losses, and have a more conservative debt structure compared to their non-distressed counterparts. Non-distressed firms generally exhibit stronger profitability, more efficient operations, and a more stable financial position compared to their financially distressed counterparts. While these ratios show signs of potential financial distress, further statistical analysis is needed to confirm their predictive power.

Table 6: Summary Results of Distressed versus Non-Distressed Firms

	Financially Distressed Firms										
	ROA	**GPM**	**NPM**	**ROE**	**NPTA**	**DEQ**	**TLTA**	**CLTA**	**EBITDAR**	**EBITDA**	**EBIT**
mean	-0.05	33.28	-69.94	0.42	-0.05	-0.01	0.42	0.2	-0.42	5.92E+08	0.04
std	0.32	-63.11	252.01	4.5	0.32	3.84	0.23	0.19	2.21	1.51E+09	0.34
min	-1.3	-376.67	-1282.79	-3.3	-1.3	-27.71	0.05	0.02	-9.23	-3.81E+08	-0.92
max	0.63	99.88	89.45	32.99	0.63	4.06	1.01	0.74	1.5	8.25E+09	0.87
	Non-Financially Distressed Firms										
mean	-0.01	38.98	-8.08	-0.93	-0.01	3.36	0.54	0.2	0.15	6.24E+08	0.07
std	0.28	29.84	163.69	19.93	0.28	96.93	0.32	0.17	1.57	1.88E+09	0.27
min	-3.5	-193.02	-2156.49	-400.22	-3.5	-666.29	0.01	0	-20.37	-1.74E+09	-3.26
max	0.85	127.05	818.22	27.88	0.85	1835.89	3.7	1.25	7.62	2.03E+10	0.87

We use the t-test to assess whether there is a statistically significant difference between distressed and non-distressed firms. Table 7 displays the results from the t-tests for distressed and non-distressed firms. The results from the t-test indicate that several financial ratios were found to be statistically significant in distinguishing distressed firms from non-distressed firms. These statistically significant ratios include the QR, NPM, CCL, TLTA, CTL, EBITDAR, and CRA. To put it into context, distressed firms exhibit a significantly higher quick ratio and a higher cash conversion cycle, indicating potential liquidity challenges. They also tend to have significantly lower net profit margins, which imply operational difficulties. Additionally, distressed firms show a higher TLTA, suggesting a heavier debt burden. Distressed firms tend to have lower EBITDAR compared to non-distressed firms. On the other hand, non-distressed firms have a higher proportion of CTA and CTL, indicating a stronger cash position. These findings underscore the importance of considering these financial metrics when assessing the financial health and risk profiles of companies in both distressed and non-distressed scenarios.

Table 7: Results from t-tests

Variable	Mean (Distressed)	Mean (Non-Distressed)	T-Statistic	P-Value
ROA	-0.05	-0.01	-0.99	0.32
CR	3.59	2.55	1.81	0.07
QR	3.22	1.94	2.99	0.00
AccRT	92.81	80.56	0.56	0.58
AccPT	79.24	51.68	1.70	0.09
InvT	175.93	148.94	0.20	0.84
AT	0.58	0.59	-0.12	0.90
GPM	33.28	38.98	-1.13	0.26
NPM	-69.94	-8.08	-2.46	0.01
ROE	0.42	-0.93	0.50	0.61
CCL	2.00	1.11	2.59	0.01
NPTA	-0.05	-0.01	-0.99	0.32
NPTL	-0.12	-0.01	-0.63	0.53
DEQ	-0.01	3.36	-0.26	0.80
TLTA	0.42	0.54	-2.74	0.01
EBITDAR	-0.42	0.15	-2.40	0.02
EBIT	0.04	0.07	-0.78	0.44
WCTA	0.18	0.13	1.65	0.10
CTA	0.18	0.12	2.67	0.01
CTL	1.11	0.56	2.31	0.02
CATA	0.38	0.33	1.41	0.16
CLTA	0.20	0.20	-0.08	0.93
AR	231118971.91	389887657.39	-0.91	0.36
AP	160845429.63	320258398.04	-0.73	0.47
EBITDA	591939418.25	624345012.67	-0.12	0.90

4.3. Results from Machine Learning Classifiers

The results from the predictive modeling experiments reveal insightful patterns. As shown in Table 8, all of the predictive models demonstrate high overall accuracy, with logistic regression achieving 91% accuracy, decision tree, random forest, and SVM all achieving 96% accuracy. However, it's worth noting that these models have varying performance in terms of precision, recall, and F1-score. Decision tree and random forest outperform logistic regression and SVM in terms of recall and F1-score, suggesting that they may have better capabilities to correctly identify distressed firms.

The RFECV experiments indicate a slight decrease in the overall accuracy across all algorithms. Logistic Regression, decision tree, random forest, and SVM achieve accuracies of 81%, 91%, 91%, and 91%, respectively. Notably, the precision, recall, and F1-score for logistic regression have decreased compared to the predictive model, indicating a potential loss of predictive power. Decision tree and random forest maintain relatively stable performance, with strong precision and recall scores, suggesting their robustness in predicting financial distress.

These findings highlight the importance of balancing different evaluation metrics in predictive modeling for financial distress. Note that neither the predictive modeling nor the RFECV model exhibits signs of overfitting. While high accuracy is desirable, it is equally crucial to consider precision, recall, and the F1-score, as they provide insights into the model's ability to correctly identify distressed firms and minimize false positives. Decision tree and random forest, with their balanced precision and recall scores, appear to be well-suited for predicting financial distress, even when feature selection is applied. Logistic Regression, while initially accurate, may require careful feature selection to optimize its performance. Overall, these results emphasize the significance of algorithm selection and experimental engineering in developing effective predictive models for financial distress.

Table 8: Predictive Modelling and RFECV Results

Predictive Modelling					
Algorithm	**Train Accuracy**	**Test Accuracy**	**Precision**	**Recall**	**F1-Score**
Logistic Regression	0.91	0.91	1.00	0.23	0.38
Decision Tree	0.98	0.96	1.00	0.92	0.96
Random Forest	0.97	0.96	1.00	0.93	0.97
SVM	0.99	0.96	1.00	0.92	0.96
RFECV Model					
Algorithm	**Train Accuracy**	**Test Accuracy**	**Precision**	**Recall**	**F1-Score**
Logistic Regression	0.85	0.81	0.25	0.28	0.26
Decision Tree	0.94	0.91	0.75	0.43	0.59
Random Forest	0.94	0.91	0.75	0.43	0.59
SVM	0.92	0.91	0.76	0.44	0.60

Detecting financial distress among firms comes down to professional judgment and domain knowledge. As seen in Figure 4, a more aggressive approach that flags a huge number of firms as experiencing financial distress would have a high recall, given that auditors would catch many earnings manipulations that occur (Perols, 2011; Williams, 2021). However, this approach would also have low precision because it would flag a lot of healthy firms not involved in earnings manipulation as experiencing financial distress. In contrast, a highly conservative approach that only flags the most obvious cases of earnings management would probably have high precision (Perols & Lougee, 2011). However, it would miss the subtler cases of earnings manipulation and would thus have a lower recall. To strike the right balance between recall and precision, analysts need to carefully consider how they set up their systems and what criteria they use for flagging potential cases of earnings manipulation or firms experiencing financial distress (Baryannis et al., 2019; Lokanan, 2022; Williams, 2021).

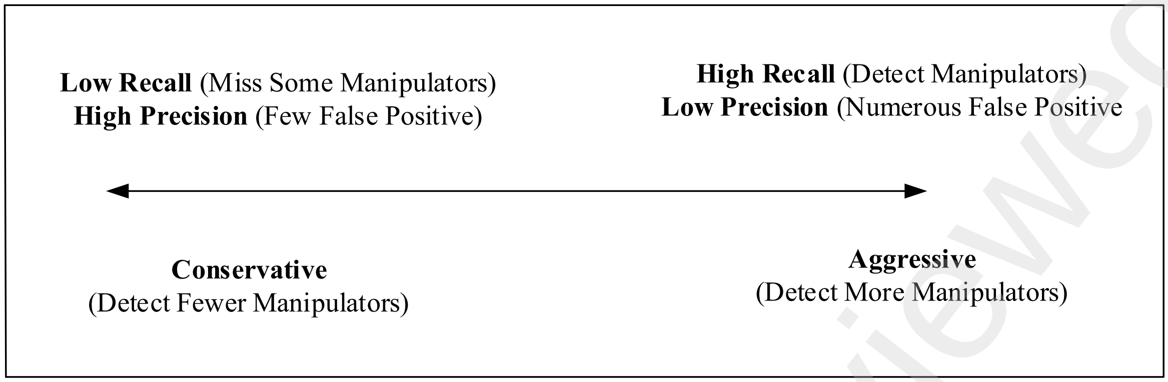

Figure 4: Approaches to Detect Distress in Firms

4.4. Results from ANN

In our modeling approach, we have employed both linear and nonlinear algorithms, with the nonlinear methods outperforming the linear logistic regression algorithm. ANN has proven to be particularly effective due to its ability to optimize complex cost functions, as it is inherently nonlinear (M. E. Lokanan, 2022; Modaresi et al., 2018) Figure 5 illustrates the use of a fully connected neural network, leveraging its feedforward architecture for efficient information propagation and adept modeling of complex datasets. As displayed in Figure XXX, the AI sequential model underwent meticulous optimization and tuning, encompassing batch size, epochs, dropout rate, learning rate, as well as activation and optimization functions to fine-tune the model architecture and optimize performance.

```
model = Sequential()
model.add(Dense(units=15, input_dim=len(features),kernel_regularizer=regularizers.l2(0.001), activation='hard_sigmoid'))
model.add(Dropout(0.2))
model.add(Dense(units=15, activation='hard_sigmoid'))
model.add(Dropout(0.2))
model.add(Dense(units=12, activation='hard_sigmoid'))
model.add(BatchNormalization())
model.add(Dense(units=11, activation='hard_sigmoid'))
model.add(Dense(units=11, activation='hard_sigmoid'))
model.add(Dense(units=1, activation='sigmoid'))
model.summary()
```

Figure 5: ANN-Optimized Model

As displayed in Table 9, the ANN model emerges as the most proficient algorithm for forecasting financial distress. The ANN model surpasses all other ML classifiers in performance, underscoring the potential for neural network models to enhance predictive accuracy in financial distress forecasting when contrasted with conventional classifiers. Notably, the recall score assumes particular significance in this context, as it effectively signifies the model's capability to identify instances of financial distress, a vital aspect for company stakeholders aiming to preempt bankruptcy risks (Almaskati et al., 2021; Aviantara, 2021; Y. Sun et al., 2021). The noticeable uptick in predictive accuracy offered by the ANN model, in comparison to its traditional ML classifiers, substantiates the rationale for corporations to persist in their investments in AI-based tools for financial distress prediction and detection.

Table 9: ANN Performance

ANN Sequential Model	Precision	Recall	F1-Score	Accuracy
Train Scores	1.00	1.0	1.0	1.0
Test Scores	.98	.96	.97	.98

4.5. CART with Bootstrapping

Table 10 presents the outcomes derived from the CART model employing bootstrapped samples. The results indicate varying performance levels across different sample ratios, with the 3:01 and 6:01 ratios consistently outperforming the other samples. These two ratios consistently achieve higher test accuracy, test F1-scores, and test AUC values. Several factors could contribute to their superior performance, including the potential presence of more pertinent features for distinguishing between distressed and non-distressed firms or the model's capacity to glean insights from the limited distressed samples. Additionally, dataset-specific characteristics or distinct financial behaviors exhibited by firms in the 3:01 and 6:01 ratios may favor the CART model's predictive abilities. The model appears to generalize well to unseen data and accurately identify instances of financial distress. As the class imbalance ratio decreases, transitioning from 3:01 to 6:01, the models maintain their proficiency in accurately classifying the majority class (normal transactions) during training. Nevertheless, there is a slight decline in overall performance, particularly concerning the classification of the minority class (financial distress transactions). This decline is expected, as classifying the minority class becomes more challenging in the presence of a higher-class imbalance.

Table 10: Bootstrapped Samples with CART

Samples	Ratio	Normal Transactions	Financial Distress Transactions	Training Accuracy	Training F1-Score	Training AUC	Test Accuracy	Test F1-Score	Test AUC
0	1:01	173	173	1.00	1.00	1.00	0.84	0.34	0.63
1	2:01	173	86	1.00	1.00	1.00	0.89	0.43	0.66
2	3:01	173	57	1.00	0.99	0.99	0.92	0.57	0.71
3	4:01	173	43	0.99	0.96	0.97	0.91	0.50	0.67
4	5:01	173	34	0.97	0.89	0.90	0.91	0.50	0.67
5	6:01	173	28	0.98	0.90	0.91	0.92	0.57	0.71
6	7:01	173	24	0.98	0.91	0.92	0.90	0.33	0.60
7	8:01	173	21	0.98	0.89	0.90	0.89	0.24	0.57
8	9:01	173	19	0.98	0.91	0.92	0.90	0.33	0.60
9	10:01	173	17	0.98	0.90	0.91	0.90	0.33	0.60

4.6. ROC Curve

Figure 6 shows the results of the ROC curve. The ROC curve is a graphical representation of the performance of a binary classifier as it consists of a varied discrimination threshold. The curve is created by plotting the TPR (sensitivity) against the FPR (1 - specificity) at various threshold settings (Lokanan, 2022). The area under the ROC curve (AUCROC) measures how well the classifier discriminates between positive and negative examples, with a perfect classifier having an AUC of 1.0. Classifiers with an AUC near 1 are said to have good discriminatory power, whereas those with an AUC near 0.5 have poor discriminatory power. As can be seen in Figure 7, the ANN model AUROC score is 1.0, while the ensemble classifiers have scores of .95, respectively. The SVM's AUROC score was .98, while that of the logistic regression was .50. As a measure of separability, the AUROC did an impressive job in comparing different binary classifiers. The results from the AUROC provide additional evidence to show that the ensemble algorithms and the SVM classifiers were consistent across all the measures of performance: accuracy, precision, and recall.

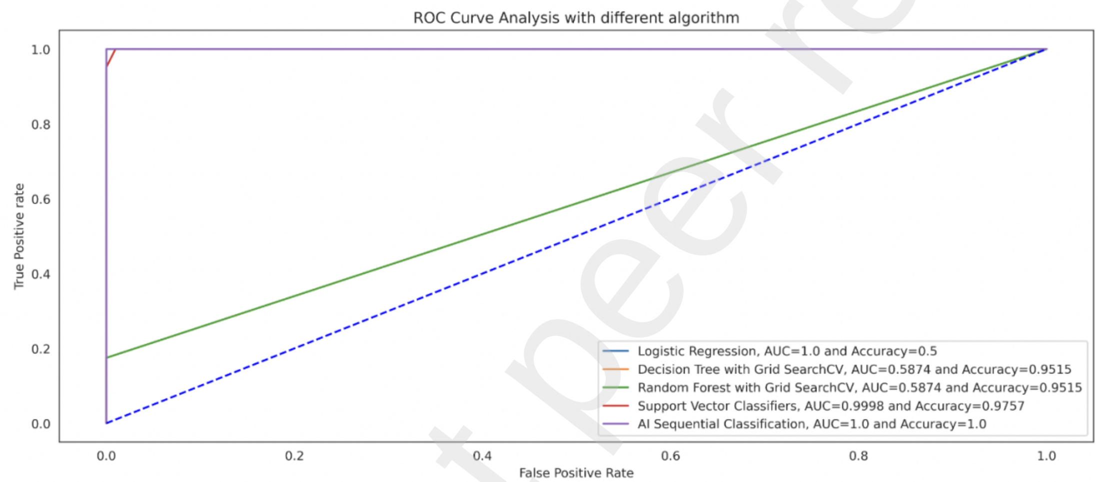

Figure 6: Results of the ROC Curves

4.7. Feature Relevance

Figure 7 displays the coefficient analysis of the top ten features for predicting financial distress. Revenue growth emerges as the most significant predictor, indicating that the rate of revenue growth strongly influences the likelihood of financial distress. Following closely is the dividend growth rate, which also plays a substantial role in predicting financial distress. A higher dividend growth rate can be a warning sign for fraud, as it may suggest that the company is overcommitting to its shareholders at the expense of its financial stability.

The cash conversion cycle is another noteworthy feature among the top predictors. A longer cash conversion cycle can signal liquidity challenges, making it difficult for the company to meet its short-term obligations. Additionally, the company's profit margins stand out as an essential predictor. Lower GPM values may indicate that the company is struggling to maintain profitability, potentially leading to financial distress. The amount of cash available to pay off liabilities is also significant in predicting financial distress. A higher CTL ratio suggests that the company has sufficient cash to cover its liabilities, which can be a protective factor against distress. EBIT and EBITDAR are also prominent features. Their presence in the top 10 predictors underscores the importance of operational profitability in financial stability.

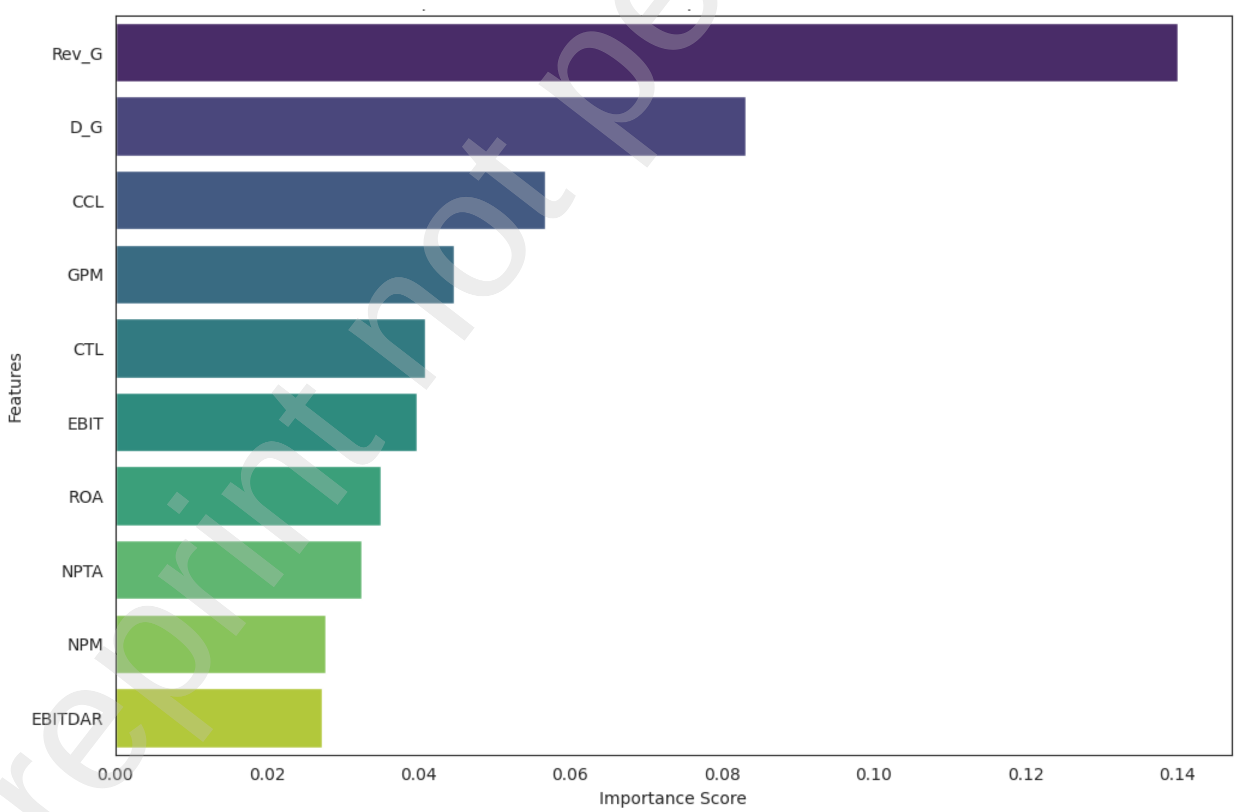

Figure 7: Top 10 Feature Relevance

Table 11 provides insights into key features for predicting financial distress by analyzing the mean and standard deviation. Distressed firms tend to have higher mean values in revenue growth but with significant variability. Their dividend growth rates are lower on average, with greater variations. The cash conversion cycle is higher for distressed firms, indicating potential liquidity challenges with relatively lower variability. Distressed firms have lower profitability, although with substantial variability. The CTL ratio suggests distressed firms tend to have a higher cash reserve relative to total liabilities, but with greater variability. EBIT and EBITDAR ratios show minimal variations and minor differences in means. NPM stands out with significantly lower mean values and wide standard deviation for distressed firms, highlighting their financial struggles and considerable profitability variation. These findings underscore the need for a nuanced financial distress prediction approach, considering central tendencies and variability.

Table 11: Top Features Mean and Standard Deviation

Index	Feature	Non-Financial Distress	Distress
0	Rev_G	0.28 ± 1.17	5.14 ± 21.14
1	D_G	0.39 ± 2.50	1.92 ± 8.98
2	CCL	1.11 ± 2.19	2.00 ± 3.72
3	GPM	38.98 ± 29.84	33.28 ± 63.11
4	CTL	0.56 ± 1.58	1.11 ± 2.30
5	EBIT	0.07 ± 0.27	0.04 ± 0.34
6	ROA	-0.01 ± 0.28	-0.05 ± 0.32
7	NPTA	-0.01 ± 0.28	-0.05 ± 0.32
8	NPM	-8.08 ± 163.69	-69.94 ± 252.01
9	EBITDAR	0.15 ± 1.57	-0.42 ± 2.21

5. Discussion and Conclusion

In this study, we explored the potential of ML and AI algorithms for predicting financial distress on the TSX using the Beneish M-score as a proxy for the target outcome. The TSX dataset contains a wide range of financial variables for listed companies which, when evaluated using ML- and AI-based algorithms, can construct efficient predictive models. The results reveal critical insights into the potential indicators of the financial health for companies listed on the TSX. Among these features, revenue and dividend growth, the cash conversion cycle, gross profit margins, and cash available to fund liabilities were the top features to predict financial distress. Our results corroborate previous findings that distressed firms have lower liquidity, profit, and solvency ratios in predicting financial distress (Hajek & Henriques, 2017; Mselmi et al., 2017; Zhao et al., 2023). Collectively, these findings emphasize that a combination of factors related to revenue, dividend management, liquidity, profitability, and operational performance holds the key to predicting financial distress, underlining the significance of monitoring these indicators for proactive financial risk mitigation and maintenance of financial health in businesses (Bao et al., 2020; Campa & Camacho-Miñano, 2015; Habib et al., 2020; Mselmi et al., 2017).

In keeping with the existing finance and accounting literature, we employed standalone ML models for financial distress prediction (Achakzai & Peng, 2023; Bao et al., 2020; Hajek & Henriques, 2017; Mselmi et al., 2017; Zhao et al., 2023). However, our study departs from the existing financial distress research in several important ways. First, we used RFECV to systematically identify and select the most relevant features, enhancing the interpretability and efficiency of the models. One important aspect of developing accurate predictive models for financial distress is feature selection. RFECV contributes to the literature by providing a structured methodology for feature selection, which is crucial for identifying the features to detect financially distressed firms (Rtayli & Enneya, 2020). Second, while CART is not a novel algorithm, its application to the specific context of Canadian firms listed on the TSX for financial distress prediction contributes to the literature by providing a practical and interpretable framework tailored to this specific domain. Many studies have focused on ML techniques, but CART's simplicity and transparency can complement existing research, offering a different perspective and valuable insights from the stand-alone ML models. In this regard, the application of CART to a specific dataset, such as TSX-listed firms, extends the body of literature by offering a practical and transparent approach to detect risks in firms. Third, our study diverges from the prevailing literature on financial distress prediction by extending the generalizability of research in accounting and finance. Our study employs data from the TSX, offering a different geographical context and thereby broadening the scope of financial distress prediction research.

Both the ANN and ML classifiers were highly effective in predicting financial distress when tested on unseen data. The ANN model had the highest performance accuracy (98%), followed by the ensemble methods (random forest and decision tree), and the SVM predictive models with 96%. The logistic regression model had the lowest performance accuracy (91%). When comparing accuracy, precision, recall, and the F1 score, the ANN model consistently outperformed the other classifiers in predicting financial distress. These results demonstrate that the ANN model is capable of leveraging intricate patterns in larger datasets and can be utilized to more accurately classify the data to predict financial distress.

These findings corroborate previous research on the effectiveness of ML algorithms, especially ensemble, SVM, and AI-based models in predictive financial distress (Achakzai & Peng, 2023; Mselmi et al., 2017; Zhao et al., 2023). ANNs' powerful ability to find correlations between data points enhances its ability to identify concealed relationships more efficiently than the traditional ML classifiers in predicting financial distress. The RFECV's models had the poorest performance across all performance measures. The CART model emphasizes the importance of dataset balance in financial distress prediction models and the potential benefits of ensuring an adequate representation of both normal and distress samples.

Using ML- and AI-based tools, the findings from this study can provide valuable insights and enhance the detection of manipulation in financial statements (Hammami & Hendijani Zadeh, 2022; Ramírez-Orellana et al., 2017; van der Heijden, 2022). The use of ML and AI predictive techniques is not only useful in predicting financial distress but can also be applied to detect earnings manipulation that resulted from insolvency (Kim & Kogan, 2014). The results from this study are also beneficial to auditors who need additional support while conducting their analyses and audit functions (DeZoort & Harrison, 2018; Dimitrijevic et al., 2021). Applying predictive techniques to audit engagements improves accuracy, providing greater confidence in the reliability of the underlying financial statements (DeZoort & Harrison, 2018; Hamilton & Smith, 2021; Khaksar et al., 2022). The application of ML- and AI-based algorithms to predict irregularities in financial statements opens up the possibility for companies and regulators to react swiftly when signs of manipulation are identified, thereby enabling auditors to protect the integrity of their engagements (Hilal et al., 2022; van der Heijden, 2022).

5.3. Limitations and Future Research

Despite the success of using ML- and AI-based tools to predict earnings manipulation, this study suffers from a few limitations. First and foremost, owing to the lack of available data, the variables used to predict earnings manipulation were mainly limited to financial/firm-level ratios. As such, future researchers should use corporate governance, auditors' characteristics, and enforcement-related variables to improve the predictive power of similar models. Furthermore, the benchmarks for forecasting accuracy should be established to ensure that future models improve upon the success of the models employed in this paper.

Another limitation in predicting financial distress is the challenge of dealing with evolving macro-economic and changing market conditions. Financial distress prediction models often rely on historical financial data and predefined features. However, these models may struggle to adapt to sudden economic shifts, regulatory changes, or unforeseen external events, such as economic crises or global pandemics. These unforeseen factors can significantly impact a company's financial health, and traditional models may not capture such dynamic changes effectively. An area of future research in financial distress prediction could involve the integration of real-time data and advanced technologies like natural language processing and sentiment analysis. By incorporating news articles, social media sentiment, and other real-time data sources, predictive models could gain a more comprehensive understanding of a company's financial standing and its reputation in the market. This would enable the development of more adaptive and robust models capable of responding to rapidly changing financial conditions, offering stakeholders more timely and accurate insights into potential distress situations.

References

Abbas, K., Afaq, M., Ahmed Khan, T., & Song, W.-C. (2020). A Blockchain and Machine Learning-Based Drug Supply Chain Management and Recommendation System for Smart Pharmaceutical Industry. *Electronics*, *9*(5), 852. https://doi.org/10.3390/electronics9050852

Achakzai, M. A. K., & Juan, P. (2022). Using machine learning Meta-Classifiers to detect financial frauds. *Finance Research Letters*, *48*, 102915. https://doi.org/10.1016/j.frl.2022.102915

Achakzai, M. A. K., & Peng, J. (2023). Detecting financial statement fraud using dynamic ensemble machine learning. *International Review of Financial Analysis*, *89*, 102827. https://doi.org/10.1016/j.irfa.2023.102827

Akther, T., & Xu, F. (2020). Existence of the Audit Expectation Gap and Its Impact on Stakeholders' Confidence: The Moderating Role of the Financial Reporting Council. *International Journal of Financial Studies*, *8*(1), 4. https://doi.org/10.3390/ijfs8010004

Albizri, A., Appelbaum, D., & Rizzotto, N. (2019). Evaluation of financial statements fraud detection research: A multi-disciplinary analysis. *International Journal of Disclosure and Governance*, *16*(4), 206–241. https://doi.org/10.1057/s41310-019-00067-9

Alden, M. E., Bryan, D. M., Lessley, B. J., & Tripathy, A. (2012). Detection of Financial Statement Fraud Using Evolutionary Algorithms. *Journal of Emerging Technologies in Accounting*, *9*(1), 71–94. https://doi.org/10.2308/jeta-50390

Almaskati, N., Bird, R., Yeung, D., & Lu, Y. (2021). A horse race of models and estimation methods for predicting bankruptcy. *Advances in Accounting*, *52*, 100513. https://doi.org/10.1016/j.adiac.2021.100513

Aria, M., Cuccurullo, C., & Gnasso, A. (2021). A comparison among interpretative proposals for Random Forests. *Machine Learning with Applications*, *6*, 100094. https://doi.org/10.1016/j.mlwa.2021.100094

Aslam, N., Khan, I. U., Alansari, A., Alrammah, M., Alghwairy, A., Alqahtani, R., Alqahtani, R., Almushikes, M., & Hashim, M. A. (2022). Anomaly Detection Using Explainable Random Forest for the Prediction of Undesirable Events in Oil Wells. *Applied Computational Intelligence and Soft Computing*, *2022*, 1–14. https://doi.org/10.1155/2022/1558381

Aubert, F., Wang, J. J., & Grudnitski, G. (2019). Convergence consensus analyst earnings estimates and option pricing in modeling material accounting misstatements. *Review of Accounting and Finance*, *18*(1), 134–156. https://doi.org/10.1108/RAF-05-2017-0101

Aviantara, R. (2021). Scoring the financial distress and the financial statement fraud of Garuda Indonesia with «DDCC» as the financial solutions. *Journal of Modelling in Management*. https://doi.org/10.1108/JM2-01-2020-0017

Bae, J. K. (2012). Predicting financial distress of the South Korean manufacturing industries. *Expert Systems with Applications*, *39*(10), 9159–9165. https://doi.org/10.1016/j.eswa.2012.02.058

Bao, Y., Hilary, G., & Ke, B. (2022). Artificial Intelligence and Fraud Detection. In V. Babich, J. R. Birge, & G. Hilary (Eds.), *Innovative Technology at the Interface of Finance and Operations* (Vol. 11, pp. 223–247). Springer International Publishing. https://doi.org/10.1007/978-3-030-75729-8_8

Bao, Y., Ke, B., Li, B., Yu, Y. J., & Zhang, J. (2020). Detecting Accounting Fraud in Publicly Traded U.S. Firms Using a Machine Learning Approach. *Journal of Accounting Research*, *58*(1), 199–235. https://doi.org/10.1111/1475-679X.12292

Baryannis, G., Dani, S., & Antoniou, G. (2019). Predicting supply chain risks using machine learning: The trade-off between performance and interpretability. *Future Generation Computer Systems*, *101*, 993–1004. https://doi.org/10.1016/j.future.2019.07.059

Becker, T. E., Atinc, G., Breaugh, J. A., Carlson, K. D., Edwards, J. R., & Spector, P. E. (2016). Statistical control in correlational studies: 10 essential recommendations for organizational researchers. *Journal of Organizational Behavior*, *37*(2), 157–167. https://doi.org/10.1002/job.2053

Beneish, M. D. (1999). The Detection of Earnings Manipulation. *Financial Analysts Journal*, *55*(5), 24–36. https://doi.org/10.2469/faj.v55.n5.2296

Beneish, M. D., & Vorst, P. (2022). The Cost of Fraud Prediction Errors. *The Accounting Review*, *97*(6), 91–121. https://doi.org/10.2308/TAR-2020-0068

Bernerth, J. B., & Aguinis, H. (2016). A Critical Review and Best-Practice Recommendations for Control Variable Usage. *Personnel Psychology*, *69*(1), 229–283. https://doi.org/10.1111/peps.12103

Bhattacharyya, S., Jha, S., Tharakunnel, K., & Westland, J. C. (2011). Data mining for credit card fraud: A comparative study. *Decision Support Systems*, *50*(3), 602–613. https://doi.org/10.1016/j.dss.2010.08.008

Blay, A. D., Sneathen, L. D., & Kizirian, T. (2007). The Effects of Fraud and Going-concern Risk on Auditors' Assessments of the Risk of Material Misstatement and Resulting Audit Procedures: Effects of Fraud and Going-concern Risk on Auditors' Assessments of RMM. *International Journal of Auditing*, *11*(3), 149–163. https://doi.org/10.1111/j.1099-1123.2007.00362.x

Botchey, F. E., Qin, Z., & Hughes-Lartey, K. (2020a). Mobile Money Fraud Prediction—A Cross-Case Analysis on the Efficiency of Support Vector Machines, Gradient Boosted Decision Trees, and Naïve Bayes Algorithms. *Information*, *11*(8), Article 8. https://doi.org/10.3390/info11080383

Botchey, F. E., Qin, Z., & Hughes-Lartey, K. (2020b). Mobile Money Fraud Prediction—A Cross-Case Analysis on the Efficiency of Support Vector Machines, Gradient Boosted Decision Trees, and Naïve Bayes Algorithms. *Information*, *11*(8), Article 8. https://doi.org/10.3390/info11080383

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Campa, D., & Camacho-Miñano, M.-M. (2015). The impact of SME's pre-bankruptcy financial distress on earnings management tools. *International Review of Financial Analysis*, *42*, 222–234. https://doi.org/10.1016/j.irfa.2015.07.004

Chikodili, N. B., Abdulmalik, M. D., Abisoye, O. A., & Bashir, S. A. (2021). Outlier Detection in Multivariate Time Series Data Using a Fusion of K-Medoid, Standardized Euclidean Distance and Z-Score. In S. Misra & B. Muhammad-Bello (Eds.), *Information and Communication Technology and Applications* (Vol. 1350, pp. 259–271). Springer International Publishing. https://doi.org/10.1007/978-3-030-69143-1_21

Cho, S., Vasarhelyi, M. A., Sun, T. (Sophia), & Zhang, C. (Abigail). (2020). Learning from Machine Learning in Accounting and Assurance. *Journal of Emerging Technologies in Accounting*, *17*(1), 1–10. https://doi.org/10.2308/jeta-10718

Craja, P., Kim, A., & Lessmann, S. (2020). Deep learning for detecting financial statement fraud. *Decision Support Systems*, *139*, 113421. https://doi.org/10.1016/j.dss.2020.113421

Davis, J. S., & Pesch, H. L. (2013). Fraud dynamics and controls in organizations. *Accounting, Organizations and Society*, *38*(6–7), 469–483. https://doi.org/10.1016/j.aos.2012.07.005

Dechow, P. M., Sloan, R. G., & Sweeney, A. P. (1996). Causes and Consequences of Earnings Manipulation: An Analysis of Firms Subject to Enforcement Actions by the SEC. *Contemporary Accounting Research*, *13*(1), 1–36. https://doi.org/10.1111/j.1911-3846.1996.tb00489.x

Demetriades, P., & Owusu-Agyei, S. (2022). Fraudulent financial reporting: An application of fraud diamond to Toshiba's accounting scandal. *Journal of Financial Crime*, *29*(2), 729–763. https://doi.org/10.1108/JFC-05-2021-0108

DeZoort, F. T., & Harrison, P. D. (2018). Understanding Auditors' Sense of Responsibility for Detecting Fraud Within Organizations. *Journal of Business Ethics*, *149*(4), 857–874. https://doi.org/10.1007/s10551-016-3064-3

Dimitrijevic, D., Jovkovic, B., & Milutinovic, S. (2021). The scope and limitations of external audit in detecting frauds in company's operations. *Journal of Financial Crime*, *28*(3), 632–646. https://doi.org/10.1108/JFC-11-2019-0155

Dimitropoulos, P. E., & Asteriou, D. (2009). The value relevance of financial statements and their impact on stock prices: Evidence from Greece. *Managerial Auditing Journal*, *24*(3), 248–265. https://doi.org/10.1108/02686900910941131

Eilifsen, A., & Knivsflå, K. H. (2021). Core earnings management: How do audit firms interact with classification shifting and accruals management? *International Journal of Auditing*, *25*(1), 142–165. https://doi.org/10.1111/ijau.12214

Eulerich, M., Theis, J. C., Lao, J., & Ramon, M. (2018). Do fine feathers make a fine bird? The influence of attractiveness on fraud-risk judgments by internal auditors. *International Journal of Auditing*, *22*(3), 332–344. https://doi.org/10.1111/ijau.12137

Fang, V. W., Huang, A. H., & Wang, W. (2017). Imperfect Accounting and Reporting Bias. *Journal of Accounting Research*, *55*(4), 919–962. https://doi.org/10.1111/1475-679X.12170

Farber, D. B. (2005). Restoring Trust after Fraud: Does Corporate Governance Matter? *The Accounting Review*, *80*(2), 539–561. https://doi.org/10.2308/accr.2005.80.2.539

Fern´andez-Delgado, M., Cernadas, E., & Sen´en, B. (2014). *Do we Need Hundreds of Classifiers to Solve Real World Classification Problems? 15*, 3133–3181.

Filip, A., & Raffournier, B. (2014). Financial Crisis And Earnings Management: The European Evidence. *The International Journal of Accounting*, *49*(4), 455–478. https://doi.org/10.1016/j.intacc.2014.10.004

Firth, M., Rui, O. M., & Wu, W. (2011). Cooking the books: Recipes and costs of falsified financial statements in China. *Journal of Corporate Finance*, *17*(2), 371–390. https://doi.org/10.1016/j.jcorpfin.2010.09.002

Ghosh, S., Dasgupta, A., & Swetapadma, A. (2019). A Study on Support Vector Machine based Linear and Non-Linear Pattern Classification. *2019 International Conference on Intelligent Sustainable Systems (ICISS)*, 24–28. https://doi.org/10.1109/ISS1.2019.8908018

Gupta, S., & Mehta, S. K. (2021). Data Mining-based Financial Statement Fraud Detection: Systematic Literature Review and Meta-analysis to Estimate Data Sample Mapping of Fraudulent Companies Against Non-fraudulent Companies. *Global Business Review*, 097215092098485. https://doi.org/10.1177/0972150920984857

Habib, A., Costa, M. D., Huang, H. J., Bhuiyan, Md. B. U., & Sun, L. (2020). Determinants and consequences of financial distress: Review of the empirical literature. *Accounting & Finance*, *60*(S1), 1023–1075. https://doi.org/10.1111/acfi.12400

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. https://doi.org/10.1016/j.knosys.2017.05.001

Hamilton, E. L., & Smith, J. L. (2021). Error or Fraud? The Effect of Omissions on Management's Fraud Strategies and Auditors' Evaluations of Identified Misstatements. *The Accounting Review*, *96*(1), 225–249. https://doi.org/10.2308/tar-2017-0355

Hammami, A., & Hendijani Zadeh, M. (2022). Predicting earnings management through machine learning ensemble classifiers. *Journal of Forecasting*, *41*(8), 1639–1660. https://doi.org/10.1002/for.2885

Hartwig, M., Voss, J. A., Brimbal, L., & Wallace, D. B. (2017). Investment Professionals' Ability to Detect Deception: Accuracy, Bias and Metacognitive Realism. *Journal of Behavioral Finance*, *18*(1), 1–13. https://doi.org/10.1080/15427560.2017.1276069

Hilal, W., Gadsden, S. A., & Yawney, J. (2022). Financial Fraud: A Review of Anomaly Detection Techniques and Recent Advances. *Expert Systems with Applications*, *193*, 116429. https://doi.org/10.1016/j.eswa.2021.116429

Huang, S. Y., Lin, C.-C., Chiu, A.-A., & Yen, D. C. (2017). Fraud detection using fraud triangle risk factors. *Information Systems Frontiers*, *19*(6), 1343–1356. https://doi.org/10.1007/s10796-016-9647-9

Huang, Y.-P., & Yen, M.-F. (2019). A new perspective of performance comparison among machine learning algorithms for financial distress prediction. *Applied Soft Computing*, *83*, 105663. https://doi.org/10.1016/j.asoc.2019.105663

Islam, T. U. (2021). Min-max approach for comparison of univariate normality tests. *PLOS ONE*, *16*(8), e0255024. https://doi.org/10.1371/journal.pone.0255024

Jiang, Y., & Jones, S. (2018). Corporate distress prediction in China: A machine learning approach. *Accounting & Finance*, *58*(4), 1063–1109. https://doi.org/10.1111/acfi.12432

Johnson, J. M., & Khoshgoftaar, T. M. (2019). Medicare fraud detection using neural networks. *Journal of Big Data*, *6*(1), 63. https://doi.org/10.1186/s40537-019-0225-0

Khaksar, J., Salehi, M., & Lari DashtBayaz, M. (2022). The relationship between auditor characteristics and fraud detection. *Journal of Facilities Management*, *20*(1), 79–101. https://doi.org/10.1108/JFM-02-2021-0024

Kim, Y., & Kogan, A. (2014). Development of an Anomaly Detection Model for a Bank's Transitory Account System. *Journal of Information Systems*, *28*(1), 145–165. https://doi.org/10.2308/isys-50699

Kuizinienė, D., Krilavičius, T., Damaševičius, R., & Maskeliūnas, R. (2022). Systematic Review of Financial Distress Identification using Artificial Intelligence Methods. *Applied Artificial Intelligence*, *36*(1), 2138124. https://doi.org/10.1080/08839514.2022.2138124

Li, N. (2016). Performance Measures in Earnings-Based Financial Covenants in Debt Contracts. *Journal of Accounting Research*, *54*(4), 1149–1186. https://doi.org/10.1111/1475-679X.12125

Li, Z., Huang, M., Liu, G., & Jiang, C. (2021). A hybrid method with dynamic weighted entropy for handling the problem of class imbalance with overlap in credit card fraud detection. *Expert Systems with Applications*, *175*. Scopus. https://doi.org/10.1016/j.eswa.2021.114750

Liu, Q., Wang, C., Zhang, P., & Zheng, K. (2021). Detecting stock market manipulation via machine learning: Evidence from China Securities Regulatory Commission punishment cases. *International Review of Financial Analysis*, *78*, 101887. https://doi.org/10.1016/j.irfa.2021.101887

Lokanan, M. (2017). Self-regulation and compliance enforcement practices by the Investment Dealers Association in Canada: 1984 to 2008. *Journal of Financial Regulation and Compliance*, *25*(1), 2–21. https://doi.org/10.1108/JFRC-04-2016-0038

Lokanan, M. (2018). Theorizing Financial Crimes as Moral Actions. *European Accounting Review*, *27*(5), 901–938. https://doi.org/10.1080/09638180.2017.1417144

Lokanan, M. (2022). The determinants of investment fraud: A machine learning and artificial intelligence approach. *Frontiers in Big Data*, *5*, 961039. https://doi.org/10.3389/fdata.2022.961039

Lokanan, M. E. (2015). Challenges to the fraud triangle: Questions on its usefulness. *Accounting Forum*, *39*(3), 201–224. https://doi.org/10.1016/j.accfor.2015.05.002

Lokanan, M. E. (2022). Predicting Money Laundering Using Machine Learning and Artificial Neural Networks Algorithms in Banks. *Journal of Applied Security Research*, *0*(0), 1–25. https://doi.org/10.1080/19361610.2022.2114744

Lokanan, M. E. (2023). Predicting mobile money transaction fraud using machine learning algorithms. *Applied AI Letters*, *4*(2), e85. https://doi.org/10.1002/ail2.85

Lokanan, M. E., & Sharma, K. (2022). Fraud prediction using machine learning: The case of investment advisors in Canada. *Machine Learning with Applications*, *8*, 100269. https://doi.org/10.1016/j.mlwa.2022.100269

Lokanan, M., & Liu, S. (2021a). Predicting Fraud Victimization Using Classical Machine Learning. *Entropy*, *23*(3), 300. https://doi.org/10.3390/e23030300

Lokanan, M., & Liu, S. (2021b). Predicting Fraud Victimization Using Classical Machine Learning. *Entropy*, *23*(3), Article 3. https://doi.org/10.3390/e23030300

Lokanan, M., Tran, V., & Vuong, N. H. (2019). Detecting anomalies in financial statements using machine learning algorithm: The case of Vietnamese listed firms. *Asian Journal of Accounting Research*, *4*(2), 181–201. https://doi.org/10.1108/AJAR-09-2018-0032

Maniatis, A. (2022). Detecting the probability of financial fraud due to earnings manipulation in companies listed in Athens Stock Exchange Market. *Journal of Financial Crime*, *29*(2), 603–619. https://doi.org/10.1108/JFC-04-2021-0083

Modaresi, F., Araghinejad, S., & Ebrahimi, K. (2018). A Comparative Assessment of Artificial Neural Network, Generalized Regression Neural Network, Least-Square Support Vector Regression, and K-Nearest Neighbor Regression for Monthly Streamflow Forecasting in Linear and Nonlinear Conditions. *Water Resources Management*, *32*(1), 243–258. https://doi.org/10.1007/s11269-017-1807-2

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. https://doi.org/10.1016/j.aos.2014.01.006

Mselmi, N., Lahiani, A., & Hamza, T. (2017). Financial distress prediction: The case of French small and medium-sized firms. *International Review of Financial Analysis*, *50*, 67–80. https://doi.org/10.1016/j.irfa.2017.02.004

Nami, S., & Shajari, M. (2018). Cost-sensitive payment card fraud detection based on dynamic random forest and k-nearest neighbors. *Expert Systems with Applications*, *110*, 381–392. https://doi.org/10.1016/j.eswa.2018.06.011

Nasir, M., Simsek, S., Cornelsen, E., Ragothaman, S., & Dag, A. (2021). Developing a decision support system to detect material weaknesses in internal control. *Decision Support Systems*, *151*, 113631. https://doi.org/10.1016/j.dss.2021.113631

Nasir, N. A. B. M., Ali, M. J., & Ahmed, K. (2019). Corporate governance, board ethnicity and financial statement fraud: Evidence from Malaysia. *Accounting Research Journal*, *32*(3), 514–531. https://doi.org/10.1108/ARJ-02-2018-0024

Oyerogba, E. O. (2021). Forensic auditing mechanism and fraud detection: The case of Nigerian public sector. *Journal of Accounting in Emerging Economies*, *11*(5), 752–775. https://doi.org/10.1108/JAEE-04-2020-0072

Papík, M., & Papíková, L. (2019). DETECTION MODELS FOR UNINTENTIONAL FINANCIAL RESTATEMENTS. *Journal of Business Economics and Management*, *21*(1), 64–86. https://doi.org/10.3846/jbem.2019.10179

Papík, M., & Papíková, L. (2022). Detecting accounting fraud in companies reporting under US GAAP through data mining. *International Journal of Accounting Information Systems*, *45*, 100559. https://doi.org/10.1016/j.accinf.2022.100559

Pech, R. (n.d.). *Fraud detection in mobile money transfer as binary classification problem*. 16.

Perols, J. (2011). Financial Statement Fraud Detection: An Analysis of Statistical and Machine Learning Algorithms. *AUDITING: A Journal of Practice & Theory*, *30*(2), 19–50. https://doi.org/10.2308/ajpt-50009

Perols, J., & Lougee, B. (2011). The relation between earnings management and financial statement fraud. *Advances in Accounting*, *27*(1), 39–53. https://doi.org/10.1016/j.adiac.2010.10.004

Puri, A., & Kumar Gupta, M. (2022). Improved Hybrid Bag-Boost Ensemble With K-Means-SMOTE–ENN Technique for Handling Noisy Class Imbalanced Data. *The Computer Journal*, *65*(1), 124–138. https://doi.org/10.1093/comjnl/bxab039

Qiu, S., Luo, Y., & Guo, H. (2021). Multisource evidence theory-based fraud risk assessment of China's listed companies. *Journal of Forecasting*, *40*(8), 1524–1539. https://doi.org/10.1002/for.2782

Ramírez-Orellana, A., Martínez-Romero, M. J., & Mariño-Garrido, T. (2017). Measuring fraud and earnings management by a case of study: Evidence from an international family business. *European Journal of Family Business*, *7*(1–2), 41–53. https://doi.org/10.1016/j.ejfb.2017.10.001

Repousis, S. (2016). Using Beneish model to detect corporate financial statement fraud in Greece. *Journal of Financial Crime*, *23*(4), 1063–1073. https://doi.org/10.1108/JFC-11-2014-0055

Rtayli, N., & Enneya, N. (2020). Enhanced credit card fraud detection based on SVM-recursive feature elimination and hyper-parameters optimization. *Journal of Information Security and Applications*, *55*, 102596. https://doi.org/10.1016/j.jisa.2020.102596

Sahin, Y., Bulkan, S., & Duman, E. (2013). A cost-sensitive decision tree approach for fraud detection. *Expert Systems with Applications*, *40*(15), 5916–5923. https://doi.org/10.1016/j.eswa.2013.05.021

Salgado, C. M., Azevedo, C., Proença, H., & Vieira, S. M. (2016). Noise Versus Outliers. In MIT Critical Data, *Secondary Analysis of Electronic Health Records* (pp. 163–183). Springer International Publishing. https://doi.org/10.1007/978-3-319-43742-2_14

Singh, D., & Singh, B. (2020). Investigating the impact of data normalization on classification performance. *Applied Soft Computing*, *97*, 105524. https://doi.org/10.1016/j.asoc.2019.105524

Sisodia, D. S., Reddy, N. K., & Bhandari, S. (2017). Performance evaluation of class balancing techniques for credit card fraud detection. *2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI)*, 2747–2752. https://doi.org/10.1109/ICPCSI.2017.8392219

Song, X.-P., Hu, Z.-H., Du, J.-G., & Sheng, Z.-H. (2014). Application of Machine Learning Methods to Risk Assessment of Financial Statement Fraud: Evidence from China: Risk Assessment of Financial Statement Fraud. *Journal of Forecasting*, *33*(8), 611–626. https://doi.org/10.1002/for.2294

Sreedharan, M., Khedr, A. M., & El Bannany, M. (2020). A Comparative Analysis of Machine Learning Classifiers and Ensemble Techniques in Financial Distress Prediction. *2020*

17th International Multi-Conference on Systems, Signals & Devices (SSD), 653–657. https://doi.org/10.1109/SSD49366.2020.9364178
Sun, J., Li, H., Huang, Q.-H., & He, K.-Y. (2014). Predicting financial distress and corporate failure: A review from the state-of-the-art definitions, modeling, sampling, and featuring approaches. *Knowledge-Based Systems*, *57*, 41–56. https://doi.org/10.1016/j.knosys.2013.12.006
Sun, Y., Fang, S., & Zhang, Z. (Justin). (2021). Impression management strategies on enterprise social media platforms: An affordance perspective. *International Journal of Information Management*, *60*, 102359. https://doi.org/10.1016/j.ijinfomgt.2021.102359
Tang, X., Li, S., Tan, M., & Shi, W. (2020). Incorporating textual and management factors into financial distress prediction: A comparative study of machine learning methods. *Journal of Forecasting*, *39*(5), 769–787. https://doi.org/10.1002/for.2661
Tarjo, & Herawati, N. (2015). Application of Beneish M-Score Models and Data Mining to Detect Financial Fraud. *Procedia - Social and Behavioral Sciences*, *211*, 924–930. https://doi.org/10.1016/j.sbspro.2015.11.122
Tian, Z., Xiao, J., Feng, H., & Wei, Y. (2020). Credit Risk Assessment based on Gradient Boosting Decision Tree. *Procedia Computer Science*, *174*, 150–160. https://doi.org/10.1016/j.procs.2020.06.070
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*(Supplement 1), 287–321. https://doi.org/10.2308/ajpt-50360
van der Heijden, H. (2022). Predicting industry sectors from financial statements: An illustration of machine learning in accounting research. *The British Accounting Review*, *54*(5), 101096. https://doi.org/10.1016/j.bar.2022.101096
Wang, Y. Q., & Song, Y. M. (2011). Appling Data Mining Technology to Analasis Clearance Efficiency in the Port Logistics. *Key Engineering Materials*, *480–481*, 1144–1149. https://doi.org/10.4028/www.scientific.net/KEM.480-481.1144
Williams, C. K. I. (2021). The Effect of Class Imbalance on Precision-Recall Curves. *Neural Computation*, *33*(4), 853–857. https://doi.org/10.1162/neco_a_01362
Zainudin, E. F., & Hashim, H. A. (2016). Detecting fraudulent financial reporting using financial ratio. *Journal of Financial Reporting and Accounting*, *14*(2), 266–278. https://doi.org/10.1108/JFRA-05-2015-0053
Zhang, S. (2012). Nearest neighbor selection for iteratively kNN imputation. *Journal of Systems and Software*, *85*(11), 2541–2552. https://doi.org/10.1016/j.jss.2012.05.073
Zhao, Q., Xu, W., & Ji, Y. (2023). Predicting financial distress of Chinese listed companies using machine learning: To what extent does textual disclosure matter? *International Review of Financial Analysis*, *89*, 102770. https://doi.org/10.1016/j.irfa.2023.102770

Appendix 1

Independent Variables

Variables	Description	Papers
	Independent Variables	
	Profitability Ratios	
ROA	Return on Assets	Demetriades & Owusu-Agyei, 2022; Fang et al., 2017; Gupta & Mehta, 2020; Hajek & Henriques, 2017; Song et al., 2014; Sun et al., 2014
ROE	Return on Equity	Albizri et al., 2019; Craja et al., 2020; Gupta & Mehta, 2020; Huang et al., 2017
NPM	Net Profit Margin	Albizri et al., 2019; Craja et al., 2020; Fang et al., 2017; Gupta & Mehta, 2020; Hajek & Henriques, 2017; Huang et al., 2017; Song et al., 2014
GPM	Net Profit Margin	
NPTA	Net Profit to Fixed Assets	Albizri et al., 2019; Gupta & Mehta, 2020
	Liquidity Ratios	
CR	Current Ratio	Albizri et al., 2019; Fang et al., 2017; Gupta & Mehta, 2020; Hajek & Henriques, 2017; Huang et al., 2017; Song et al., 2014
QR	Quick Ratio	Albizri et al., 2019; Fang et al., 2017; Gupta & Mehta, 2020; Hajek & Henriques, 2017; Huang et al., 2017; Song et al., 2014
AccRT	Accounts Receivable Turnover in Days	Fang et al., 2017; Gupta & Mehta, 2020; Hajek & Henriques, 2017; Huang et al., 2017; Song et al., 2014
WCTA	Working Capital to Total Assets	Albizri et al., 2019
CCL	Cash to Current Liabilities	Albizri et al., 2019; Gupta & Mehta, 2020
CTA	Cash to Total Assets	Albizri et al., 2019; Gupta & Mehta, 2020

Efficiency Ratios		
AT	Asset Turnover	Albizri et al., 2019; Fang et al., 2017; Gupta & Mehta, 2020; Huang et al., 2017
InvT	Inventory Turnover	Zainudin & Hashim, 2016; Dimitropoulos & Asteriou (2009)
AR	Accounts Receivable Turnover in Days	Shaked & Altman, 2016
AP	Accouns Payable Turnover in Days	Habib et al., 2018
Solvency Ratios		
DEQ	Debt to Equity	Albizri et al., 2019; Craja et al., 2020; Gupta & Mehta, 2020; Huang et al., 2017
TLTA	Liabilities to Assets	Albizri et al., 2019; Gupta & Mehta, 2020
NPTL	Net Profit to Total Liabilities	Craja et al. 2020
CTL	Cash to Total Liabilities	Albizri et al., 2019; Gupta & Mehta, 2020
CATA	Current Assets to Total Assets	Albizri et al., 2019; Fang et al., 2017; Gupta & Mehta, 2020
CLTA	Current Liabilities to Total Assets	Albizri et al., 2019; Gupta & Mehta, 2020
Operating Performance		
EBIT	Earnings Before Interest and Taxes	Albizri et al., 2019
EBITDAR	Earnings Before Interest, Taxes, Depreciation, Amortization, and Rent	Li, 2016
EBITDA	Earnings Before Interest, Taxes, Depreciation, and Amortization	Campa & Camacho-Miñano, 2015