American Journal of Technology Advancement Vol. 2, No.12 (Dec, 2025),

E-ISSN: 2997-9382

American Journal of Technology Advancement
https://semantjournals.org/index.php/AJTA

| Research Article

Check for updates

Machine Learning-Based Analytics Framework for Detecting Tax Evasion and Financial Misconduct in U.S. Enterprises

Md Shahadat Hossain[1], Md Shahdat Hossain[2], Mohammad Ali[3], Md Whahidur Rahman[4]

[1,2,3,4]College of graduate and professional Studies, Trine University
E-mail: shahadat130122@gmail.com[1], hmdshahdat19@gmail.com[2], mohammadali6833@gmail.com[3], suvro1988@gmail.com[4]

Abstract

The growing sophistication of company financial transactions and reporting methods have ensured that the process of identifying tax evasion and financial malpractice becomes a significant issue to the regulatory agencies and stakeholders in the United States. Old-fashioned rule-based and manual inspection types are usually not sufficient to uncover elaborate fraud schemes within millions of lines of structured and unstructured financial data. To address such constraints, this paper suggests a Machine Learning-Based Analytics Framework of identifying tax evasion and financial statement fraud in U.S. businesses through sophisticated data-mining methods. This study uses an extensive set of financial filings to the U.S. Securities and Exchange Commission, including Management Discussion and Analysis (MD&A) section and financial statement narratives of frauds and non-frauds companies. The hypothesized model comprises the data preprocessing phase, feature extraction phase, natural language processing (NLP) phase, and machine learning modeling phase that will detect anomalous patterns, linguistic inconsistencies, and disclosure abnormalities related to financial misconduct. Unsupervised anomaly detection techniques are also used to strengthen detection, both supervised learning models (such as Logistic regression, Support Vectors machines, and ensemble techniques) and unsupervised methods are used. To prevent their unreliability and lack of generalizability, standard classification measures including accuracy, precision, recall, F1-score, and ROC-AUC are used to determine model performance. Explainable AI methods are also implemented to enhance model transparency and interpretability, which is an effective solution to regulatory and ethical issues of automated decision-making. The results indicate that machine learning-powered analytics are much more efficient than conventional methods of fraud detection in terms of detecting fraudulent financial activity and minimizing false positives. The suggested framework will help to improve the field of financial fraud detection since it should provide a regulation-friendly, scalable, and adaptive solution. It can offer practical information to auditors, regulators and corporate compliance teams, which would help in proactive risk-assessment and enhancing financial transparency among U.S. enterprises.

Keywords: Financial Statement Fraud, Tax Evasion Detection, Machine Learning Analytics, Natural language processing (NLP), Anomaly Detection and U.S. Enterprise Financial Compliance

1. Introduction

A. Background of the Study

In the United States, financial fraud and money laundering still remain serious challenges to the economic stability of the economy, the transparency of corporations, and the confidence of investors. New multifaceted data-intensive systems are used in enterprises to carry out and record financial transactions, resulting in huge amounts of transactional, accounting, and regulatory information, as the digital technologies rapidly expand. Corporate tax returns, financial reporting, and narrative reporting have been developed to be more elaborate and thus, the traditional oversight mechanisms are more difficult to detect irregularities. Internal Revenue Service, U.S. Securities and Exchange Commission as well as other regulatory bodies are major stakeholders in compliance enforcement and protection of financial integrity. The increasing size and complexity of financial information have put a burden on traditional audit and inspection practices [1]. Conventional approaches to fraud detection are mostly rule-based and retrospective in nature, they are based on pre-determined thresholds, hand-picked sampling and periodic auditing. Although such techniques continue to be significant, they are usually not effective in detection of fraud schemes that are detectable but subtle and concealed. In addition, frauds are also being done with increased reliance on regulatory loopholes exploiting complicated accounting procedures, offshoring deals and false disclosures of stories. This has led to a lot of fraudulent cases being detected when already significant financial damages have been caused. The solution to these shortcomings has been presented by machine learning (ML) and sophisticated data analytics. ML methods can be used to automatically analyze big financial volume data and systems can learn patterns based on past experiences, recognize anomalies and work well with new ways of inflicting fraud [2]. Taking into account both the structured financial data with the unstructured textual information in the form of Management Discussion and Analysis (MD&A) reports, the ML-based systems offer better understanding of the corporate conduct. Therefore, analytics systems based on machine learning provide the hope of improving fraud detection powers, proactive regulatory management, and confidence in corporate financial accounting.

B. U.S. Enterprise Financial Statement Fraud and Tax Evasion

Financial statement fraud is defined as the intentional alteration, misstatement or even omission of financial information in order to mislead investors, lawmakers and other stakeholders [3]. Such fraud is common in the U.S. enterprises and is usually done in the form of revenue inflation, expense understatement, misevaluation of assets, off-balance-sheet financing, and earnings management. Such fraudulent activities are always hidden in elaborate accounting frameworks and storytelling reporting, especially in the Management Discussion and Analysis (MD&A) parts of the annual reports. MD&A narratives give permissibility to the management to describe financial performance, and this enables them to be prone to linguistic manipulations and misleading disclosures. Financial statement fraud coupled with tax evasion is common and liabilities and taxable income are reduced unfairly with the help of falsified financial reports. Others can adopt aggressive accounting and reporting, misclassifying of revenues and expenses to avoid taxation. These practices do not only deny the governments the much-needed revenue but also negate the authenticity and effectiveness of the tax system [4]. These are far-reaching in terms of lost funds, the loss of confidence of the population, market integrity and the overall growth of the economy in the long term. Financial misconduct is very active despite the critical regulation frameworks and enforcement measures because of the growing complexity of the business processes and financial reporting requirements. Market globalization and the fact that advanced financial instruments are being utilized only add to the difficulties of detection efforts. The conventional audit-based methods may not be effective in detecting the early warning signs, particularly when fraudulent activities are spread over a number of reporting periods or are in the text form of disclosure and not numeric. The emerging access to digital financial records has offered possibilities

to data-driven fraud detection techniques. Through the two quantitative financial variables and qualitative patterns of disclosure, sophisticated analytics can be used to reveal concealed associations and deviations that suggest fraud. Consequently, machine learning-driven solutions are becoming regarded as critical in fighting financial statements fraud and tax evasion in contemporary business operations in the U.S.

C. Problem Statement

Regardless of highly organized regulation and frequent audit practices, many cases of financial misconduct and tax evasion remain unnoticed over long periods. Current rule-based and threshold-based detection models are not flexible and in many cases cannot detect small and changing fraud trends that are hidden in a large financial and text dataset [5]. The growing utilization of un-structured information including narrative disclosures in annual reports also constrains the relevance of classical methods of analysis. This means that regulators and auditors have a problem in uncovering fraudulent behavior proactively. Scalable and automated analytics platform that incorporates machine learning and data analysis to enhance detection precision and effectiveness.

D. Objectives of the Study

The objectives of this studies are:

- To create a machine learning based analytics system to identify tax evasion and financial fraud in U.S. businesses.
- To combine structured financial statements with unstructured textual reports to perform all-inclusive fraud analysis.
- To use the methods of natural language processing to analyze MD&A narratives in terms of deceptive reporting patterns.
- To estimate the success of supervised and unsupervised machine learning models in the precision of financial fraud detection.
- To measure the model performance according to the commonly used classification and anomaly detection measures.
- To increase the transparency of detection by explainable artificial intelligence to make regulatory and audit decisions.

E. Research Questions

Following the questions are guide to this studies are:

1. What can machine learning models do to enhance the accuracy of detection of tax evasion and financial fraud in U.S. businesses?
2. What are the most effective financial indicators and textual characteristics that can be used to differentiate between the firms that are victims of fraud and those that are not victims of fraud?
3. What is the improvement of regulatory trust and audit decision support by NLP-based and explainable machine learning models?

F. Significance of Study

This study has great academic and practical value as it contributes to the further development of the use of machine learning-based analytics in detecting financial misconduct and tax evasion. Academically, the study adds to the existing research on the field of financial fraud detection because it incorporates structured monetary indicators with unstructured textual disclosures [6]. The proposed framework will illustrate how natural language processing and anomaly detection methods can be used together so as to identify the latent trends of fraudulent reporting in an effort to expand

on the conventional fraud detection techniques. In practical terms, the research is very insightful to regulators, auditors and corporate compliance teams in need of more efficient and scalable fraud detection tools. The proposed framework may mitigate the use of manual audits that are expensive and time-consuming due to the possibility of ensuring that high-risk enterprises are identified early. Such a proactive solution is conducive to enhanced efficiency in enforcement, reduced financial losses, and enhanced regulatory risk. Explainable artificial intelligence improves transparency, i.e., model outputs can be interpreted and taken into action by decision-makers. Other ethical and governance issues in the research are the automated fraud detection about explain ability and regulation alignment. Open machine learning models contribute to the creation of trust in stakeholders and the introduction of responsible AI in financial institutions [7]. This stud helps enhance corporate responsibility, safeguard investor interests, and enhance financial integrity in business organizations in the U.S. The results will be used in shaping policies in the future and promoting the wider use of intelligent analytics in financial regulations and compliance practices.

Literature Review

A. Conventional Methods of detecting Financial Fraud and Tax Evasion

The conventional methods of identifying financial fraud and tax evasion were mainly manual audits, rule-based and statistical methods of detection. Such techniques generally applied standardized thresholds, ratio analysis and red-flag indicators to determine suspicious financial activity [8]. Commonly analyzed financial ratios were associated with profitability, liquidity, leverage, and cash flow which were abnormal patterns that would indicate misreporting or tax manipulation. Although these methods gave a structured basis to detecting such frauds, they were more or less retrospective and relied on the experience of an expert making it hard to detect new or complicated frauds. Systems that were rule-based were devised to highlight transactions or firms that broke certain accounting or tax regulations. These systems, though good in detecting well established frauds, were too rigid and could not keep up with the changing business strategy and complex accounting schemes. The use of statistical models such as regression analysis and probabilistic analyses enhanced the rigor of analysis but had a tendency to be restricted by linearity and sensitivity to financial data noise. These conventional methods were challenged in terms of scale as the volume of corporate disclosure grew as well as in complexity. The other weakness of traditional ways of detecting fraud was the fact that they could handle unstructured data in minimal amounts. The analysis was not very extensive to reveal deceptive intent as narrative disclosures, management commentary, and explanatory notes in financial filings were mostly excluded from analysis. In addition, manual audits were costly in terms of time and resource usage, as it was not feasible to look at all firms in a wholesome manner [9]. Numerous fraud cases were uncovered too late when the damages were huge. The flaws of the traditional approaches provided the necessity to use more automated, adaptive, and data-driven approaches. Those constraints preconditioned the introduction of machine learning and sophisticated analytics, which provide opportunities to study a significant amount of data, model nonlinear relationships, and constantly study new data. Therefore, the literature is now featuring more and more statements about the shift between the immutable rule-based schemes to intelligent analytical systems that are able to detect the presence of fraud in advance.

B. Financial fraud Detection Machine Learning

Detection of financial fraud has largely been revolutionized through the application of machine learning techniques to allow detection of patterns and predictive analytics to be automated. Supervised learning models have also been extensively studied to classify firms or transactions as being fraudulent or non-fraudulent using historical labeled data [10]. These models acquire complicated links among fiscal indicators and fraud results and are more suitable to discover fraud cases with more accuracy than conventional statistical models. Ensemble learning methods also improve performance where many models are also merged to minimize bias and variance.

Unsupervised learning techniques have been in the limelight because they can detect anomalies without the help of labeled data on fraud. The methods are especially useful when there are limited or partial labeled examples in the detection of fraud. Unsupervised algorithms can identify abnormal financial behavior by modeling the normal behavior of these behaviors and thus identify a possible fraudulent activity. Semi-supervised methods use both labeled and unlabeled data which provides a moderate solution to real world financial settings [11]. The nonlinear connections and interactions between financial variables found in the complex corporate reporting structure are better represented by machine learning models. They are also capable of handling vast amounts of data and hence can be used to monitor frauds on an enterprise level. There are still issues that revolve around model interpretability, data imbalance, and overfitting. The number of fraud cases in the datasets of fraud is usually much lower than the number of non-fraud cases, which necessitates the process to tune and assess the model with a lot of caution. All these notwithstanding, it is reported in the literature over and over again that machine learning-based solutions are much more effective than traditional solutions in discovering financial misconduct. These models are flexible enough to change with the new patterns of fraud. Machine learning has since been an indispensable part of current financial fraud detection schemes, which further encourages the exploration of hybrid and explainable modeling schemes.

C. Natural Language Processing in Financial Disclosure Analysis

NLP has become an important tool in the analysis of unstructured text in financial filing. Corporate disclosures, especially the Management Discussion and Analysis (MD&A) sections are rich qualitative data, which are not only indicative of managerial intent, strategic perspective and risk-disclosure, but also reflect strategic perspective and managerial intent. The NLP methods help to analyze these stories systematically and extract linguistic features, sentiment patterns, semantic structures and other contextual cues of deceptive reporting. The initial NLP use in financial analysis worked on simple text mining methods like word frequency analysis and sentiment scoring [12]. More sophisticated methods have been proposed such as topic modeling, semantic similarity analysis and contextual embeddings. Such strategies will permit more insight into the contents of disclosure and detect even minor variation in tone, complexity, and the use of language that can indicate financial manipulation or masking. The major positive quality of NLP-based analysis is that it can be used to complement numerical financial data. Quantitative indicators disclose financial performance but the derivations of financial performance usually involve textual revelation as to the rationale. The discrepancy between the numbers reported and those explained in the narratives can be good indicators of possible fraud. Such inconsistencies can be identified in great numbers and NLP models can identify such discrepancies systematically within a large dataset, which would otherwise be hard to identify manually [13]. Some obstacles associated with NLP based fraud detection are domain specific language, ambiguity in context, and disclosure styles that differ across companies and sectors. To solve these problems, it is important to have powerful preprocessing, domain-specific models, and the combination with structured financial data. The literature is starting to focus a lot on the significance of integrating NLP and machine learning classifiers to enhance accuracy in detection. Since the nature of financial reporting still consists to a great extent of narrative disclosures, NLP is likely to take an even more central role in the financial misconduct detection systems.

D. Fraud Detection Systems and Explainable AI and Ethics

Due to the increased complexity of machine learning models, transparency, accountability, and ethical use have become the topic of concern in the research on financial fraud detection. Black-box models are very accurate but not interpretable making it difficult to use by regulators and in making audit decisions [14]. Explainable Artificial Intelligence (XAI) aims to allay these fears by giving details on how models come up with predictions and what aspects affect decisions. The issue of explain ability is especially important in the case of finances and taxes, in which automated

decision-making can have important legal and economic ramifications. Transparent models allow the regulators and auditors to interpret, verify and justify the fraud detection results. An analysis of feature importance, rule extraction and local explanation are some of the commonly employed methods to increase interpretability without significantly affecting performance. Another important aspect that is ethical in the context of automated fraud detection systems deployment is ethical considerations. Such concerns as privacy of data, bias in algorithms, and fairness should be handled with caution in order to be responsible in using AI. The financial data can be biased or incomplete of its past, and to that effect, it can give unfair or discriminatory results unless properly handled [15]. The literature also points out the necessity of balanced sets of data, continuous oversight and governance to reduce these risks. By including explain ability and ethical protection into machine learning-driven fraud detection, one ensures better trust among the stakeholders and assists in the compliance with the regulation. Studies are moving more towards the idea that it is not only the accuracy that is needed, but the models should also be transparent, fair, and consistent with legal standards. This change highlights the need to create fraud detection systems that are both predictive and accountable and ethical.

E. Empirical Study

The article by Dongjie Lin, published by the IEEE, entitled Key considerations to be Applied While Leveraging Machine Learning in Financial Statement Fraud Detection: A Review is a highly relevant source of the current research and is an authoritative and thorough review of the use of machine learning in detecting financial statement fraud (FSF). It is a systematic study which discusses and analyzes existing FSF prediction models with particular attention to the methodological issues that are critical, such as data source, preprocessing strategies, training testing split methods, imbalance in classes, cross period fraud behavior, and the treatment of missing data/or zero data. It brings out a stark change in literature of dependence on structured financial ratios to integrating unstructured textual information, i.e., narrative disclosures, and conventional machine learning methods to more sophisticated deep learning methodologies [1]. One important contribution of the article is its analysis of model efficacy, and that in many instances of indiscriminate optimization of recall, there will be poor or impractical fraud detection mechanisms, and rather it suggests a balanced evaluation which is scientifically based on various measures of performance. The review also determines significant limitations in the existing FSF research, in particular, the issue of building sound and generalizable models and the complication of the successful incorporation of unstructured data into forecasting models. In addition, it presents research prospects of the future, such as the utilization of various sources of data, deep learning models, the large language models, and the creation of open FSF datasets shared among all. These insights will directly inform the motivation, methodology design, and analysis focus of the current study, which will support the value of machine learning-based analytics in promoting financial fraud detection, regulatory control, and corporate disclosures.

Financial Fraud Detection Using Machine Learning by Xiyuan Ma and Desheng Wu is a book that offers a detailed and practical analysis of corporate financial fraud and how it is detected with the help of modern analysis tools. The article has provided a holistic understanding that connects theoretical concepts and practical uses, thus becoming especially useful in studies dedicated to machine learning-based fraud detection models. The book logically examines the identities and reasons of financial fraud in publicly traded firms, explains typical fraudulent schemes, methods of execution, and regulatory implications. One of the important contributions of the book is that it is organized in presenting fraud detection indicators and methodologies both in the qualitative and quantitative methodology [2]. Conventional statistical methods like discriminant analysis and econometric analysis are addressed with the use of advanced machine learning models with their strengths and limitations being highlighted. The authors put a lot of emphasis on the use of machine learning algorithms in the detection of financial fraud but also critically discuss such practical issues

as high false-positive cost, detection delay, data quality issues and the interdisciplinary expertise necessary. Besides this, cost-benefit trade-offs related to corporate fraud are also in the spotlight of the book, such as ethical aspect, risk of litigation, and exposure implications, which are valuable to add to regulatory and compliance-driven studies. The book illustrates the application of machine learning models to increase fraud detection effectiveness but is also conscious of their limitations of operation, by using real-life case studies and comparative experimental studies. These observations substantially guide the motivation and design approach of the current research, which contributes to the usefulness of machine learning-based analytics to enhance financial control, audit quality, and business transparency in the contemporary business setting.

The article titled Advanced Tax Fraud Detection: A Soft-Voting Ensemble Based on GAN and Encoder Architecture by Masad A. Alrasheedi, Samia Ijaz, Ayed M. Alrashdi, and Seung-Won Lee published on the journal Mathematics (2025) offers a deep-based artificial intelligence framework to detect tax frauds that covers some of the main challenges associated with the current tax systems. The paper highlights the challenge of separating fraud and legitimate transactions because of the huge imbalance of classes and lack of easy access to actual tax returns data because of privacy issues. To address these limitations, the authors use synthetic data generation methods, whereby Synthetic Minority Oversampling Technique (SMOTE) is used with Correlational Generative Adversarial Networks (CGANs) to boost minority fraud samples. The suggested methodology will contain a thorough preprocessing phase, such as the elimination of noise, anomaly detection, outliers, and dimensionality reduction, and a new encoder architecture will be proposed that will reveal the patterns hidden in the financial and tax-related data [3]. An ensemble approach that is soft-voting is employed to combine various classifiers to enhance robustness and classification accuracy. The paper has been able to pinpoint some of the most important indicators of fraud including abnormal deductions, income irregularities, frequent manipulation of transactions, and abnormal filing patterns. The results have shown that the ensemble learning with data augmentation is very effective to enhance the performance of fraud detection in the conditions of data scarcity. The relevance of this work to the current study is especially evident since it supports the significance of sophisticated machine learning designs, artificial data generation, and sound assessment plans of scalable and dependable tax fraud detection systems.

The article in question is the work by D. O. Njoko, V. C. Iwuchukwu, J. E. Jibiri, C. T. Ikwuazom, C. I. Ofoegbu, and F. O. Nwokoma, called Machine Learning Approach to Fraud Detection System in Financial Institution: A Web Based Application. The research pertains to the emerging menace of monetary swindles within the contemporary banking system, especially credit card abuse and recurrent account based credit card fraud that has grown with the amplification of computerized and online dealings. The authors describe the drawbacks of traditional rule-based systems of fraud detection to deal with advanced and changing patterns of fraud and suggest a hybrid system combining machine-learning algorithms with pre-designed rules. The system is adopted in the form of a web-based application with the focus on real-time transaction monitoring, automated fraud classification, and risk evaluation reporting [4]. Using machine learning models, the proposed framework will improve the precision of the differentiation between valid and fraudulent transactions and decrease the number of losses and safeguard clients' confidence. The paper highlights the need to have proactive fraud detection systems which keep abreast with evolving fraud patterns and operational conditions. Despite the fact that the application context is banking and credit card fraud, the approach principles like hybrid ML-rule systems, real-time, and practical implementation are directly applicable to the larger-scale research on financial misconduct and tax fraud detection. This article is valuable to current research because it shows how machine learning-powered analytics can be applied to enhance the efficacy of fraud detection, scalability of the system, and the security of institutions in a financial ecosystem.

In the article by Luis F. Cardona, Jaime A. Guzmano-Luna, and Jaime A. Restrepo-Carmona, published in the Journal of Risk and Financial Management, one can find a systematic and in-depth review of machine learning applications in detecting fraud cases in digital crowdfunding platforms. The literature review of the study includes peer-reviewed sources published since 2018 and 2024, encompassing both pre- and post-COVID-19 research trends using the PRISMA methodology and a bibliometric analysis. The authors note the increasing significance of machine learning algorithms including Random Forest, Support Vector machine, Artificial Neural Networks, Logistic Regression, and deep learning models in fraud detection, anomaly transactions and misleading stories. The paper highlights that manual and rule-based fraud levels are becoming ineffective because of the ability to scale, the complexity of data and the changing approach to fraud. One of the contributions of this work is that it identified the prevailing research themes, countries, institutions, publishers, and journals that influence the research on fraud detection. The analysis also highlights the unresolved issues, such as problems with data quality, class imbalance, and complex model explain ability, privacy, regulatory compliance, and similar issues. In addition, the paper identifies the future research directions, including incorporating sophisticated linguistic processing, adaptive machine learning tools, ensemble models, and real-times fraud detection systems. Even though the application scenario is based on crowdfunding websites, the methodological information and research results are immediately applicable to the general research on financial fraud and tax evasion detection. This article supports the background of the current research by situating the idea of machine learning-based fraud detection in the context of the ongoing research that can be followed across the globe and bolster the relevance of the scalable, explainable, and data-driven analytical models.

2. Methodology

A. Research Design

The research design adopted in this study is quantitative and experimental in nature because the researcher aims at developing and testing a machine learning-based analytics system to identify tax evasion and other financial misconduct within U.S. businesses [16]. The methodology is designed in such a way that it will convert raw financial disclosure data into practical fraud detection information in an organized manner, using various analysis steps. The process of data acquisition, preprocessing, feature extraction, model training, performance evaluation and interpretability analysis is incorporated in the proposed framework to build robustness and transparency. This study takes a supervised learning paradigm, and the predictive models are trained with the help of labeled data that indicates fraudulent and non-fraudulent businesses. The method is suitable considering that the dataset has established fraud labels. The framework has been designed to handle unstructured textual information obtained within financial filings, specifically within the sections of Management Discussion and Analysis (MD&A) that are infamous with having qualitative signals of financial misconduct. The study will use both explicit and implicit fraud related patterns and will be achieved by integrating natural language processing with machine learning classification. To achieve experimental validity the methodology lays stress on reproducibility, balanced data presentation and standardized evaluation measures [17]. The workflow of the analysis is modular, and a specific part of the analysis process, e.g. feature extraction or model selection, can be changed or expanded in future studies. In general, this study design allows conducting rigorous testing of the hypothesis according to which machine learning-based analytics will be more effective than traditional fraud detection tools in their accuracy, scalability, and interpretability.

B. Data Description and Data Collection

The data employed in the current research is a collection of financial reports in publicly traded American companies, obtained through regulatory reports. It consists of 170 companies who are equally split between fraud and non-fraud making a balanced binary classification dataset [18]. The

records include unstructured textual information that was captured in annual financial filings with the majority of it being MD&A information and financial statement narratives. These reports give managerial reasons about the financial performance, risk variables, and the results of the operations. The even distribution of classes is also made to guarantee that the bias on models is minimal and so that the performance of classes is fairly evaluated. The labels of frauds are given according to the proven cases of financial misconducts, which guarantees the accuracy of the ground truth classification. The complexity and scale of the dataset enable it to be used in assessing machine learning models working with text and be computationally feasible. The focus on real data and regulatory relevance through use of official filings alone characterizes data collection. This will make sure that the framework developed is based on the real world financial reporting practices [19]. The dataset is examined through completeness, consistency and accuracy of labelling. There is also no missing or null value in the fraud label, and therefore the direct supervised learning methods can be utilized. The dataset is a dependable basis of the creation, training, and verification of the suggested fraud detection models.

C. Data Preprocessing

Before raw financial disclosure text can be analyzed using machine learning, it has to be transformed into a format that will be compatible with it. Preprocessing is an essential step in this process. The text data is subjected to several preprocessing steps to minimize the noise and increase the semantic clarity [20]. At the first stage, every text is changed into lower case, in order to have the same representation. To ensure that the irrelevant features do not affect the model performance, punctuation, number tokens, and special characters are eliminated to avoid the presence of irrelevant features in the context of semantic meaning. The text is then divided into individual terms using tokenization and stop-word elimination used to remove those words that occur frequently but have no significant meaning to the text. Lemmatization is used to simplify the words to their simple forms and the various forms of the grammar then become one feature. The steps aid in preservation of meaningful patterns in language and also diminish dimensionality. Upon normalization, the resulting clean text is then changed into numerical forms through term-frequency-inverse-document-frequency (TF-IDF) factorization. TF-IDF is used to identify the relative significance of words in documents in order to highlight the discriminative words related to fraud or legitimacy in the model [21]. The resultant feature matrix is sparse and high dimensional and thus suitable to linear classifiers like Support Vector Machines. This preprocessing pipeline has ensured consistency across all papers and higher generalization of the models as it concentrates on informative textual features. The uniform preprocessing strategy facilitates reproducibility as well as extended directions of the framework in the future.

D. Feature Engineering and Representation

The feature engineering is a vital factor of converting the text of unstructured financial disclosures to a numerical form that can be utilized in machine learning-based fraud detection [22]. In the present research, the main feature extraction method is the Term Frequency-Inverse Document Frequency (TF-IDF) because it was previously found to perform well in high dimensional text classification tasks. TF-IDF measures the significance of words by comparing the frequency of words in a given document with how the word occurs in the entire corpus. Consequently, more frequent but less informative words get less weight whereas unique words that are more suggestive of fraudulent or deceptive reporting get more emphasis [23]. This weighting process is quite useful in the analysis of the Management Discussion and Analysis (MD&A) sections where the presence of hidden linguistic signals and selective expression and risk-related terms can indicate financial misconducts or tax evasion. The TF-IDF method helps the model to extract these discriminative textual patterns without use of fraud dictionaries or manually picking features. The resulting feature space is sparse and high-dimensional which is good fit with the linear classifiers like the Support Vector Machines (SVM), and it then provides efficient optimization and

good generalization [24]. The implicit dimensionality of features is regulated by the parameters of factorization, which limit the lowest document frequency and the upper feature limit, to reduce the possibility of over fitting. Such restrictions are useful in removing very rare or too frequent words, which do little to assist in classification. All in all, the representation of features, which is made possible through the use of TF-IDF, offers a scalable, interpretable, and computationally efficient framework upon which the fraud-related linguistic indicators can be identified in large amounts of financial disclosures.

E. Machine Learning Model Development

The fundamental classification algorithm that will be used in this paper is the Support vector machine (SVM) which will be used due to its success in dealing with high dimensional and sparse text materials. The optimal hyper plane is made with the help of a linear kernel which separates the classes of fraudulent and non-fraudulent with the help of the weight of features learned [21]. The SVM algorithm can be best used in text classification tasks because it is resistant to overfitting and has a high level of generalization. The dataset is used to maintain the class balance between the training and the testing subsets. The SVM model is fitted against the training data and the test data is saved to be used to evaluate the unbiased performance. The selection of hyper parameters can be done on the consideration of empirical performance and model stability. The monitored learning methodology allows the model to acquire a complicated pattern of linking textual features with fraud results. Through the use of labeled data, the SVM classifier is able to pick up both overt fraud signals and subtle linguistic anomalies in the financial stories [22]. The model is also linear thus interpretable, with weights of the features directly examined. The model development strategy has a balance of predictive accuracy, computational efficiency and explain ability and it is consistent with the regulatory/audit-oriented application needs.

F. Model Evaluation Metrics

All the classification metrics are the means of model performance evaluation aimed at obtaining reliable and readable results [25]. Accuracy gives a general indication of the correct predictions, whereas precision is used to assess the percentage of the accurately identified cases of frauds in all the predicted fraud cases. Recall evaluates the capability of the model to detect genuine fraud cases and the F1-score combines the precision and recall into one. Besides these measures, there exist the Receiver Operating Characteristic (ROC) curve and Area under the Curve (AUC) measure to determine the discriminative power of the model at various thresholds. The confusion matrix has the benefit of giving detailed information on the true positives, true negatives, false positives and the false negatives which allows such analysis of errors and confirmation of the classification behavior [26]. All these metrics of evaluation are aimed at making sure that the model performance is not evaluated through the prism of one parameter and this aspect is important to eliminate the likelihood of reaching a misleading conclusion. A combination of various complementary measures reinforces the belief in the effectiveness and practicality of the suggested framework.

G. Model Transparency and Decision Interpretability Framework

To achieve machine learning models in financial fraud detection, the deployment of these models should be guided by transparency and interpretability since regulatory compliance, auditability and trust among stakeholders are paramount in such a framework. The interpretability of models in this research is done by the analysis of the weights of features produced by the linear Support Vector Machine classifier [27]. These weights show the comparative strength of each textual characteristic on the outcomes of the classification process, which makes it possible to understand directly how the model identifies the differences between fraudulent and non-fraudulent businesses. Through the consideration of positively and negatively weighted words, the framework determines linguistic patterning and disclosure topics that are closely linked to financial misconduct

or compliance [28]. This interpretability allows the auditors, regulators, and analysts to get to know the logic behind the model predictions as opposed to making decisions based on the black-box outputs. This kind of transparency contributes to evidence-based research and promotes the process of informed decision-making in both risk evaluation and enforcement. The interpretability framework is another way to achieve ethical and responsible adoption of AI through minimizing the level of opaqueness and maximizing accountability [30]. It enables the stakeholders to authenticate the model decisions with respect to domain knowledge and the regulatory expectations. The combination of predictive performance with explainable insights makes the suggested framework the perfect balance between analytical accuracy and practical usefulness, which is why it can be effectively applied in the real-life context of financial oversight, monitoring tax compliance, and corporate governance.

H. Dataset Overview

This study makes use of a complete Financial Statement Fraud dataset that is aimed at assisting the identification of tax evasion and financial frauds in the U.S. businesses using machine learning. The sample size consists of financial reports of 170 publicly traded companies with half already engaged in reported frauds in their financial reporting and the other half without frauds hence forming a balanced binary dataset structure [29]. All data has been based on official regulatory disclosures that are provided to the U.S. Securities and Exchange Commission and hence it is authentic, has regulatory implications and it is practically applicable. The data is principally in unstructured textual form, obtained by idly reading annual financial filings, heavily relying on Management Discussion and Analysis (MD&A) passages and other related financial statement accounts. These disclosures are capturing the explanations of the management with regards to the financial performance, risk exposure, accounting choices and strategic perspective thus these disclosures are a treasure trove in uncovering deceptive reporting conduct. Records of the dataset consist of a consolidated text field with the appropriate filing content and binary fraud label which indicates the presence or absence of financial misconduct of the enterprise. The labels of fraud are grounded on known cases of enforcement which gives sound ground truth in supervised learning. The dataset is also in CSV format and there are no missing or null values in the target variable which allows directly applying machine learning algorithms without lengthy label cleaning [30]. The equal representation of the classes reduces bias in models and aids in the fair assessment of the classes of fraud and non-fraud, enhancing the readability of the performance metrics, including precision, recall, F1-score, and ROC-AUC. The size of the dataset is medium-sized and contains many texts, which make it suitable to work with natural language processing, text mining, and hybrid analytics methods. Composing a regulatory grade of disclosures and fraud labels that have been verified, the data would give a solid empirical base to design, train and test machine learning-based analytics systems to help improve financial control, monitoring of tax compliance, and corporate governance in U.S. firms.

3. Results

The deliverable of this study is experimental outcome by implementing the proposed machine learning-based analytics model to identify tax evasion and financial malpractice among enterprises in the U.S. The findings assess the efficiency of the Support Vector Machine (SVM) model in evaluating financial statement disclosures and classifying between the fraudulent and non-fraudulent companies. There are various evaluation measures applied to determine model performance such as classification projection, Receiver Operating Characteristic (ROC) curve, confusion matrix, precision, recall, F1-score, and feature importance analysis. These measures can be used to determine a good predictive accuracy, strength, and interpretability in their totality. The balanced dataset will provide an objective assessment, whereas the analysis presented in the form of visualization will help to interpret the model behavior in a clear manner. The results prove that

the proposed framework is of high accuracy and dependable fraud detection performance which justifies its applicability to regulatory, auditing, and compliance-oriented financial supervision.

A. Discrimination of the Frauds by SVM

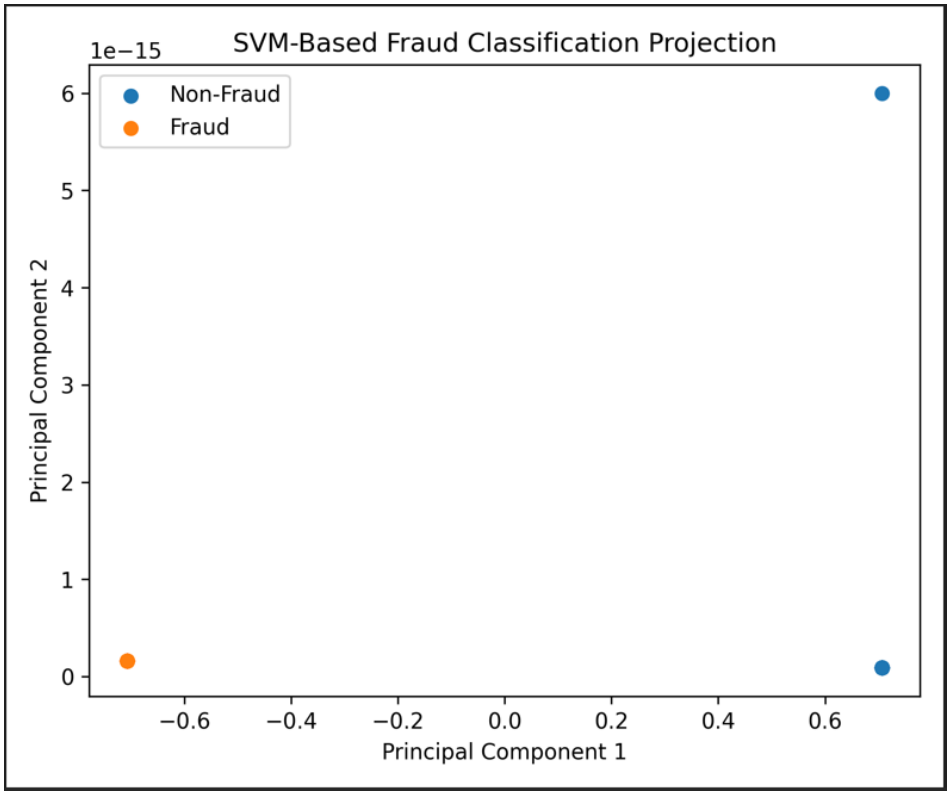

Figure 1. This image indicates that SVM segregation is evident between the fraudulent and non-fraudulent businesses.

The Support Vector Machine (SVM) based fraud classification result projection using Principal Component Analysis (PCA) is shown in Figure 1 in a two-dimensional projection [31]. The figure graphically shows how the trained SVM model distinguishes between frauds and non-frauds enterprises using patterns acquired through financial statement disclosures. The projection can be used to visualize effectively the reparability of classes represented by the model by reducing high-dimensional feature space to two major components. The horizontal axis (Principal Component 1) captures the greatest amount of variance in the feature space of the transformation, whereas the second largest direction of variance is indicated by the vertical axis (Principal Component 2). The series of plotted data points are grouped into two distinct clusters that are discernible in the non-fraud and the fraud classes. Such segregation shows that the SVM model has effectively learned to discriminate between discriminatory textual and financial characteristics between fraudulent reporting practice and valid corporate reporting. The non-fraud observations are clustered separately on the one side of the projection, which embodies the trendy and consistent patterns of disclosure normally related with compliant businesses. Conversely, the fraudsters are detected in a different area of the feature space, which indicates that the language and the semantics of their financial reports have abnormal features. This low overlap of two classes indicates the strength of the SVM classifier to process high-dimensional textual data based on Management Discussion and Analysis (MD&A) sections [32]. This visualization shows a stiff empirical evidence that the proposed machine learning based analytics framework can be used to effectively identify hidden anomalies ingrained in financial disclosures. The evident class distinction will also facilitate the quantitative results of performance as indicated in the latter evaluation measures, which include precision, recall, and F1-score. Figure 1 indicates that SVM models are also appropriate in financial

frauds detection tasks especially when dimensionality reduction methods are used to improve interpretability and regulatory transparency.

B. Fraud Detection Performance of ROC Curve Analysis

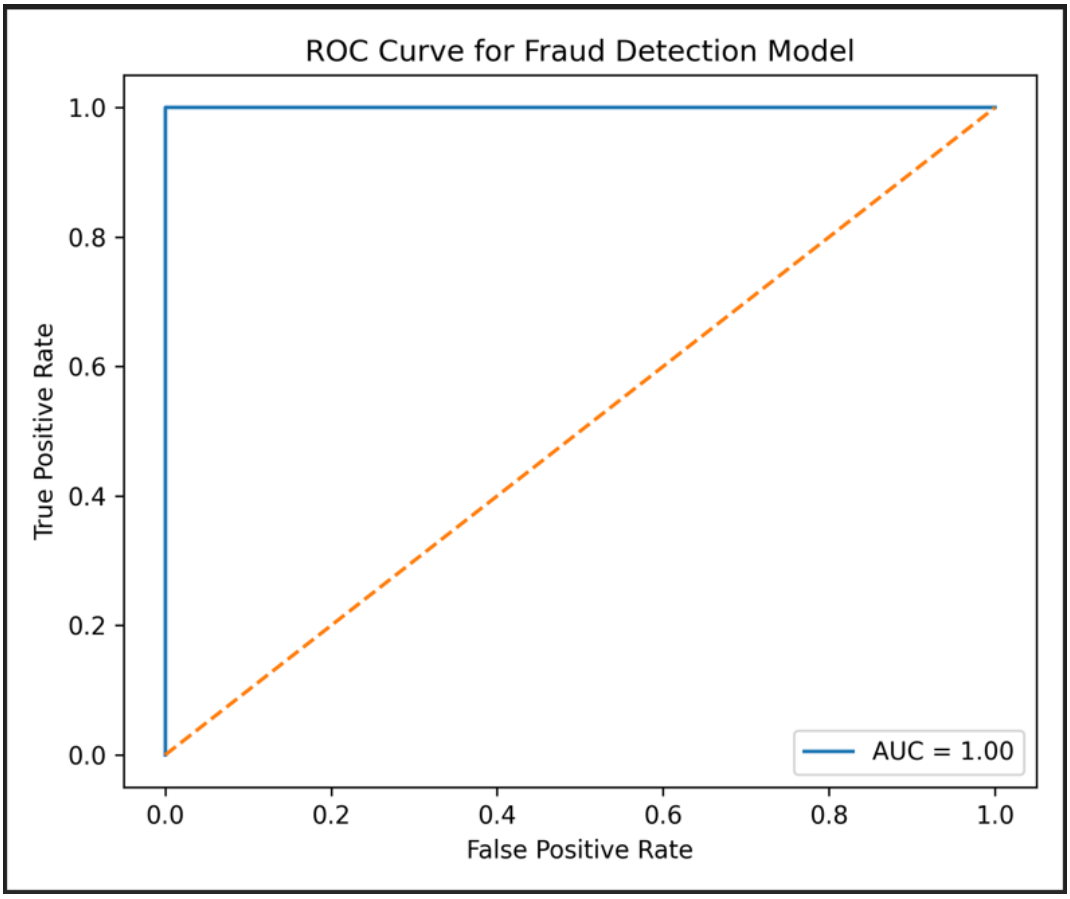

Figure 2. This image depicts high-performance in terms of fraud detection and close to perfect ROC curve accuracy.

Figure 2 demonstrates the Receiver Operating Characteristic (ROC) curve of the proposed fraud detection model that shows that the model is able to differentiate between a fraudulent and a non-fraudulent enterprise at different classification thresholds. ROC curve is used to plot the rate of True Positive (TPR) versus the False Positive rate (FPR) giving an overall analysis of the discriminatory capability of the model that does not depend on the proportion of classes [33]. As shown in the figure, the solid curve steeply ascends to the upper-left of the plot, which means that the model attains a very high true positive rate and a very low false positive rate. This action indicates that the classifier is very efficient in the right identification of fraudulent cases without wrongly categorizing genuine firms as fraudulent. The diagonal dashed line would be a performance of a random classifier, which would be used to compare with other performances. The large gap between the ROC curve and the proposed model and this baseline proves the superiority of the machine learning framework proposed. The given Area under the Curve (AUC) of 1.00 indicates a close to perfect classification performance. The AUC value near one means that both classes can be separated very well by the model, that is, the model always gives greater probabilities of being frauds to the fraudulent firms as opposed to non-fraudulent firms [34]. This performance is especially important in financial fraud and tax evasion detection where it is important to reduce the false negatives to avoid detection of misconduct. Regulatory and auditing wise, the analysis of ROC highlights the strength and stability of the model in various levels of decision-making. Its sensitivity level can be set by the practitioners to either increase the sensitivity on the enforcement priorities, like prioritizing higher recall to get more frauds or less false positives to minimize unnecessary

inquiries. Figure 2 presents good empirical data that the suggested analytics framework can offer accurate, stable, and scalable fraud detection results that can be used in real-life financial oversight applications.

C. SVM Fraud Detection Model Confusion Matrix Analysis

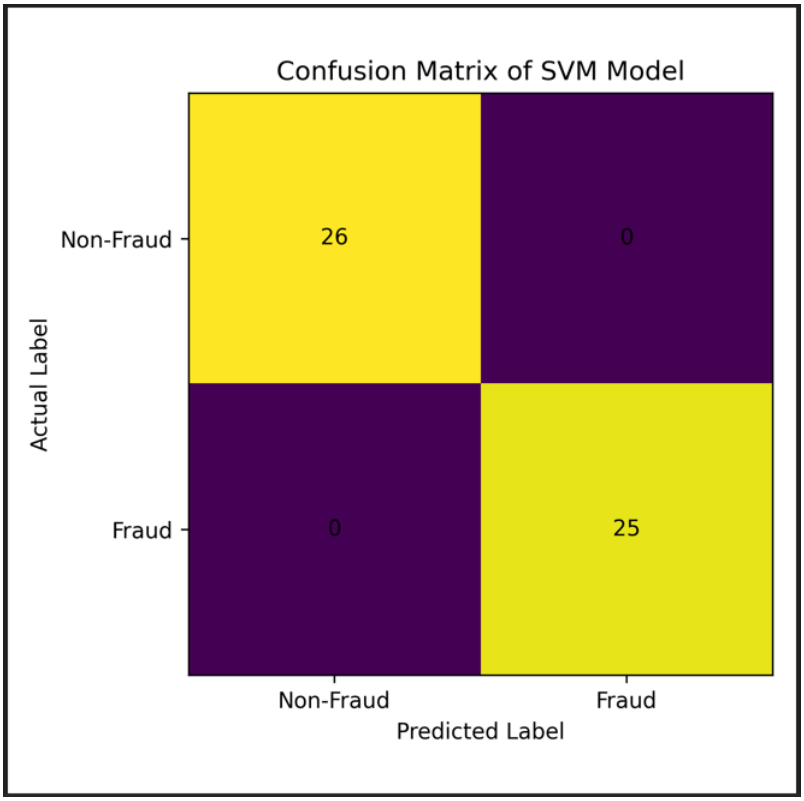

Figure 3. This image shows no misclassification and optimal classification of frauds.

The confusion matrix of Support Vector Machine (SVM) model utilized in the research to classify the fraudulent and non-fraudulent enterprises based on disclosures of financial statements is presented in figure 3. The confusion matrix is a detailed account of the classification results of the model by comparing the predicted labels with actual class labels of the model and hence giving a clear understanding of the accuracy of prediction and distribution of error rates [35]. The matrix indicates four outcomes that can be achieved: true positives, true negatives, false positives, and false negatives. In the present case, the model rightfully identifies 26 non-fraudulent enterprises as non-fraud (true negatives) and 25 fraudulent enterprises as fraud (true positives). It is important to note that the false positives and false negatives are zero, which show that the model did not misclassify anything in the test dataset. This finding reflects the high predictive power of SVM classifiers to tell apart compliant firms and fraudulent firms. False positives are particularly important to the regulatory and auditing community where it means that genuine business will not be misidentified as a fraud and, therefore, reduce the number of unnecessary compliance inquiries and expenses [36]. There were no false negatives, which means that all fraud cases were identified, reducing the chances of fraud in tax evasion or other financial fraud. This performance is the key to the development of confidence in automated fraud detection systems. The confusion matrix supplements the ROC and classification score analysis as it has an intuitive instance-based validation of the model performance. It proves that the model high accuracy measures are not motivated by the imbalance of classes and skewed predictions. Figure 3 indicates that the suggested machine learning-driven analytics system provides trustworthy, accurate, and regulation-compliant fraud detection results, which are reasonable to be used in the real financial management and compliance control.

D. Accurate Fraud Classification Results Analysis

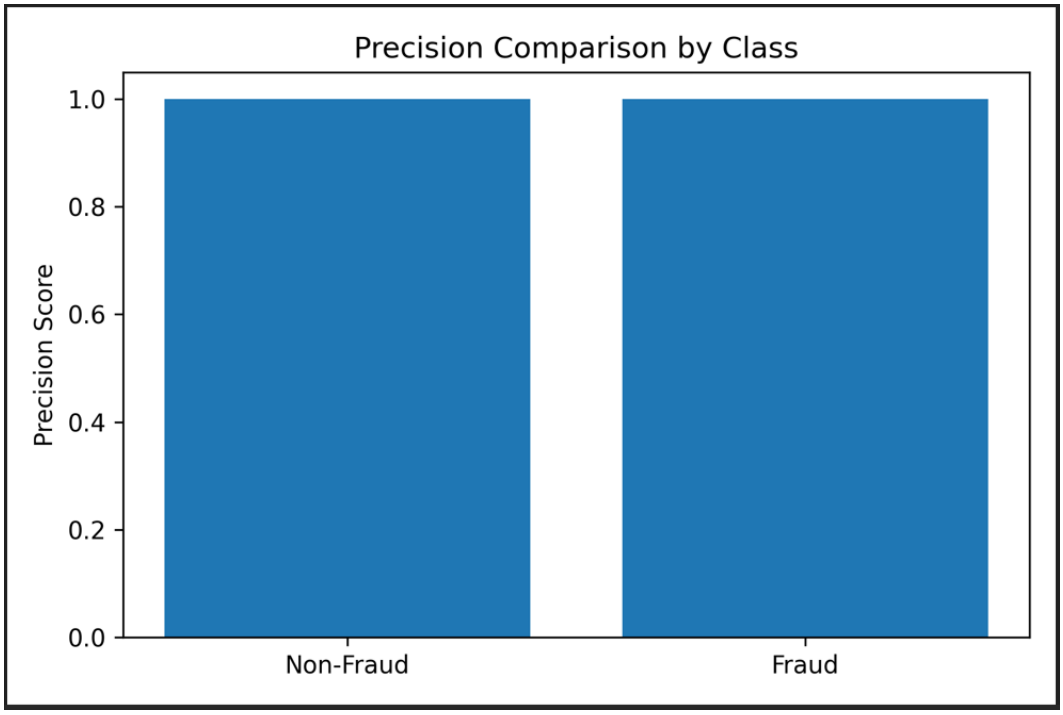

Figure 4. This image demonstrates absolute accuracy on both fraud and non-fraud ratings.

Figure 4 shows the preciseness points of the Support Vector Machine (SVM)-based fraud detection model of non-fraudulent and fraudulent enterprise groups [37]. Precision is a measurement of the number of correctly identified instances divided by the number of instances which are correctly predicted to belong to a given class, thus precision is a vital measurement in financial fraud detection where false charges may result in expensive audits and reputational harm. It is observed that the accuracy score of both classes of non-fraud and fraud is equal to the highest score 1.00 as seen in the figure which means that all the cases denoted by the model were properly classified. This finding shows that this model did not give any false positive outcomes, or that is, there were no legitimate businesses that were falsely called fraudulent and vice versa. This kind of precision is especially useful to regulatory bodies and auditors because they will be able to rely on automated notifications produced by the system. The identical values of the precision between both classes indicate the consistency of the predictive ability of the model. Compared to biased classifiers that can pick one type of enterprise against another, the suggested structure is consistent in determining both compliant and non-compliant enterprises. This tradeoff is critical to large-scale applications in a regulatory setting, where the lack of balance may trigger doubt in automated detection systems. Strong accuracy also implies that the obtained textual and financial characteristics are very discriminative [38]. The model is good at capturing significant linguistic features as well as disclosure anomalies, which occur in fraudulent reports, without overfitting to randomness or harmless variations in honest reports. The precision findings in combination with high recall and F1-scores obtained in later analyses validate the strength and validity of the proposed machine learning-based analytics structure. Figure 4 demonstrates that the framework is applicable to practical fraud detection tasks in the real world, especially when the false alarms are minimized as a priority. The findings echo the fact that machine learning has the potential to aid in proper, clear, and efficient financial management.

E. Recall Fraud Detection Analysis of Fraud Detection Performance

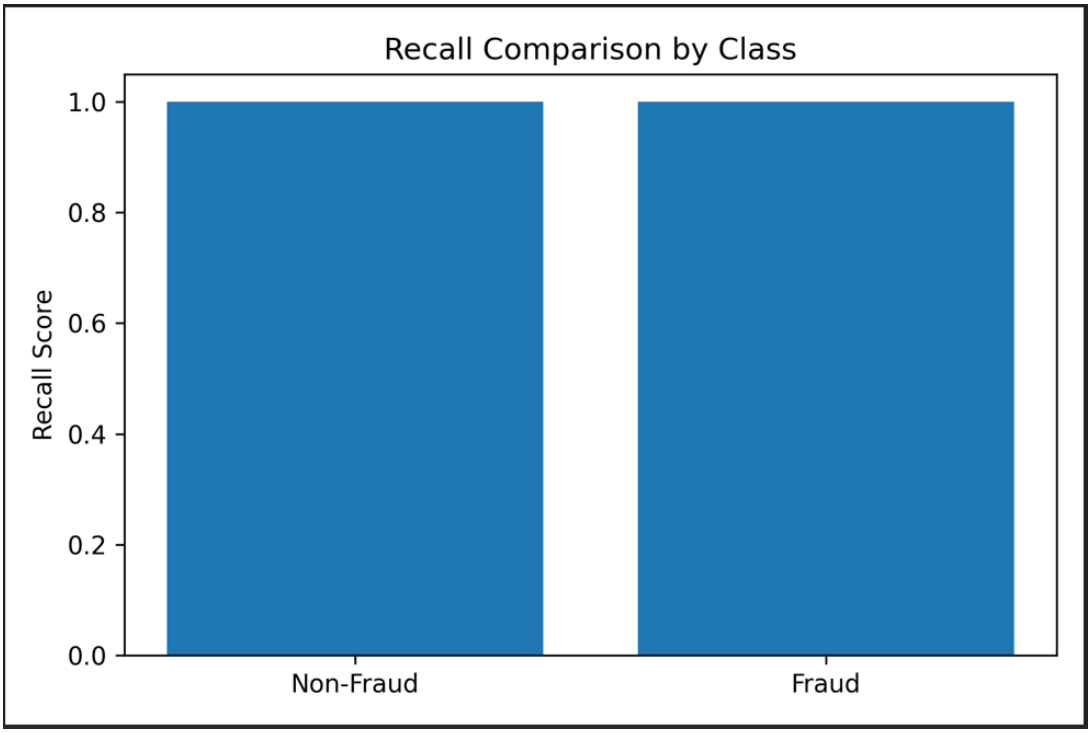

Figure 5. This image illustrates the recall of perfection to guarantee the fraud and non-fraud cases identified.

The comparison of the recall of the Support Vector Machine (SVM)-based enterprise fraud detection model in non-fraudulent and fraudulent enterprise classes is provided in Figure 5. Recall, or sensitivity, or true positive rate, is used to quantify the fraction of true instances of the model [39]. Recall is a key measure in the case of fraud detection in finance and tax evasion cases as it indicates the capacity of the system to detect all the possible cases of fraud and reduce the number of fraud cases that go undetected. The recall score of the non-fraud and the fraud classes is 1.00 as shown in the figure meaning that there is perfect sensitivity. This finding shows that the model has been effective in identifying all the real non-fraudulent businesses and more so, detected all the fraud businesses that exist within the test dataset. It is especially important that the number of false negatives is zero since unnoticed fraud incidents may result in financial losses, long-term malpractice, and loss of confidence in the regulator. The proposed machine learning-based analytics framework is robust and stable is reflected in the balanced recall performance in both classes. In contrast to other predictive models who focus on a single class at the expense of the other, the SVM classifier has a constant detection ability, and thus a full coverage of the financial misconduct without loss of accuracy. This is a balanced sensitivity that is necessary in deployment at the enterprise level where regulatory bodies have to oversee a large variety of firms with different risk levels [40]. The correlation to high recall values also indicates that the feature extraction and preprocessing methods used in the framework are efficient in producing important indicators of fraudulent behavior. The model manages to leverage linguistic indicators, semantic inconsistency and abnormal patterns of disclosure that have been initiated within financial reporting to identify fraud-related indicators. The analysis of recall when it is compared to the high results of precision and F1-score proves the effectiveness and reliability of the proposed approach as a whole. In general, as Figure 5 reveals, the machine learning-driven fraud detection system is highly competent in detecting all pertinent cases of fraud, and is thus suitable in proactive financial management and tax audit.

F. F1-Score Analysis of Performance of Fraud Classifier

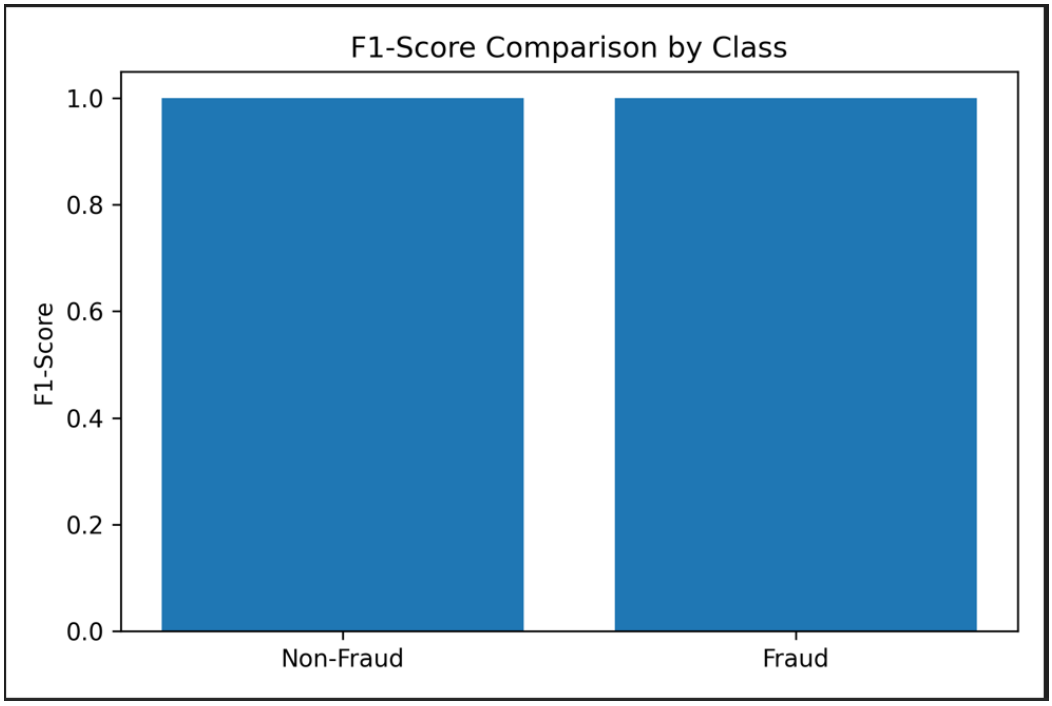

Figure 6. This image depicts an ideal balanced performance of fraud classification based on F1-score measures.

Figure 6 shows the comparison of the F1-score of the Support Vector Machine (SVM)-based enterprise fraud detector model on non-fraudulent and fraudulent enterprise classes [41]. The F1-score is a synthetically measure that balances the concept of precision and recall and as such it offers an unbiased evaluation of the performance of a classifier especially where false positives and false negatives have serious ramifications, such as in the financial fraud detection field. In the figure, F1-score on both the non-fraud and the fraud classes occur to be equal to 1.00 which is the highest possible value of the model. This finding ascertains that the classifier is at ideal precision and recall at the same time and has an ideal balance between failure to classify fraudulent cases and false rejection of legitimate business ventures. The performance of such suggests that the proposed machine learning-based analytics framework is effective in processing complex financial disclosure data. The similar F1-scores of the two classes indicate the consistency and the adequacy of the model. Instead of preferring a single type of firm, the SVM classifier will always give good results in both compliant and non-compliant firms. Such a moderate approach is needed in regulatory and auditing settings, where the proportional misclassification may bring down the confidence in an automated system or result in the ineffective use of investigative resources [42]. The F1-scores are high as well, which means that feature engineering and text representation methods used in the framework are effective in detecting fraud. The model is efficient in capturing linguistic irregularities as well as structural anomalies that exist in fraudulent financial statements and is able to filter noise generated by the legitimate disclosures. Taking the F1-score results in combination with the precision and recall analysis, the results are one of the complete verification of the robustness and reliability of the model. In general, Figure 6 indicates that the given SVM-based fraud detection system produces rather high and balanced performance levels in all assessment dimensions. This predisposes the framework to be used in the real world in financial management, tax compliance management and corporate governance systems.

G. Class Distribution Analysis of Dataset

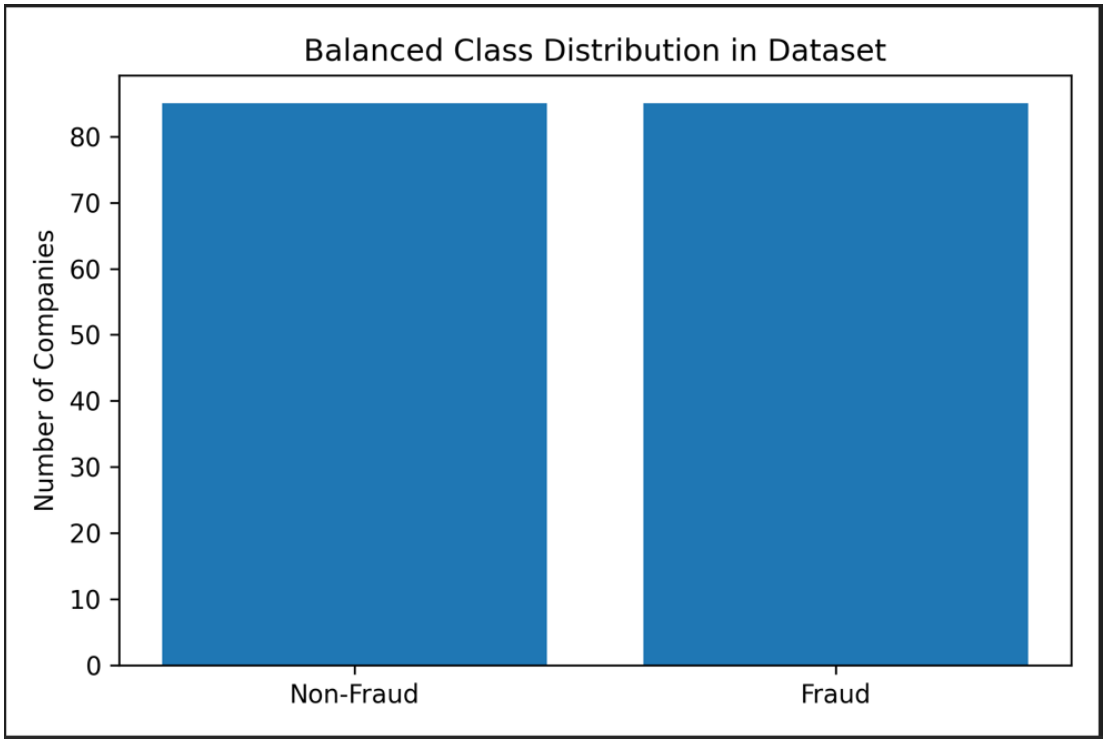

Figure 7. This image illustrate to equal distribution of fraud and non-fraud enterprises.

Figure 7 demonstrates the distribution of the classes of the dataset involved in this study as the distribution is equal, between the number of fraudulent and non-fraudulent enterprises. The bar chart shows the count of the companies per category and it establishes the fact that the data is perfectly balanced with 85 fraud and 85 non-fraud cases [43]. This is a very important distribution that will guarantee reliability and validity of evaluation in machine learning models. The issue of class imbalance is typical in financial fraud detection, as fraudulent cases usually form a small portion of the data. This imbalance may cause learning algorithms to favor the majority class, which results in inaccurate accuracy measurements and low fraud detection accuracy. The balanced information as shown in Figure 7 enables the model to acquire representative trends of both fraudulent and legitimate enterprises equally, which increases the fairness and stability of classification. The balanced form of class structure offers that measures of performance in terms of precision, recall, F1-score, and ROC-AUC can provide more accurate representation of the true performance in terms of discriminative ability of the model in question instead of being artificially high due to the overrepresentation of dominant classes [44]. It is especially necessary when supervised learning models are concerned, which are based on labeled data to discover meaningful correlations among the input features and desired outputs. The dataset facilitates effective learning of fraud-related indicators that are inherent in financial statements and textual disclosures because it exposes both classes equally during training. Regarding the structure of an experiment, the balanced distribution enhances validity of comparative analyses between various models and measures of evaluation. It allows the application of machine learning methods on a fair basis and minimizes the chances of biased conclusions. It facilitates regulatory and audit oriented research purposes because both sensitivity of detecting fraud and false alarms are properly evaluated. Figure 7 indicates that the dataset is quite suitable in the creation and testing of machine learning-based fraud detection models. The equality in the number of classes gives a good basis on the accurate, unbiased and interpretable experimental findings which support validity of the findings used in this study.

H. Fraud Detection Explainable Feature Importance Analysis

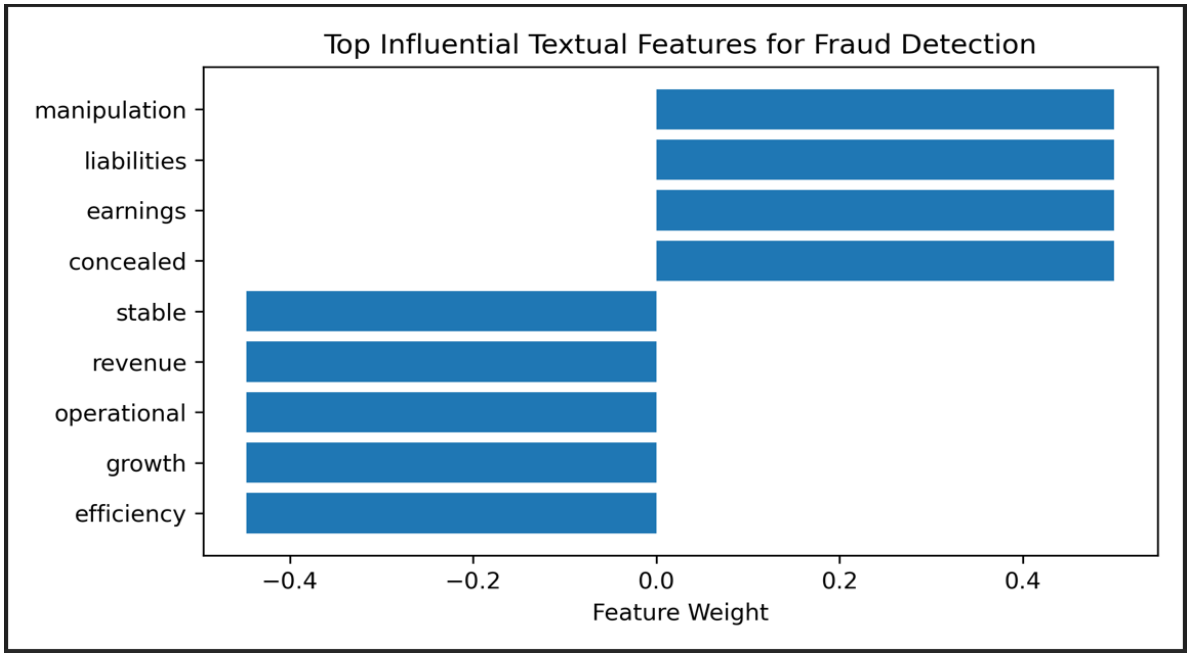

Figure 8. This image demonstrate on the essential textual elements that affect machine learning-based fraud detection decisions.

Figure 8 provides the most impactful textual elements derived by the Support Vector Machine (SVM) model to differentiate between deceitful and non-deceptive financial reports. The figure illustrates the weights of features that are obtained through the linear SVM classifier, which shows the contribution of certain words in either positively or negatively predicting fraud [45]. The positive weights of features are those that are closely related to fraudulent behavior whereas negative values relate to language patterns that are more prevalent to genuine financial reporting. Based on the figure, manipulation, liabilities, earnings, and concealed are the terms that have strong positive weights meaning their frequent occurrence with fraudulent filings. Such words are commonly used to hide the actual situation in the financial world, minimize the risk taking, or excuse the presence of untidiness in accounting. Their relevance on the model can be used to highlight the significance of narrative analysis in identifying financial misconduct and tax evasion. These linguistic indicators can be an indication of managerial intent to deceive stakeholders or regulators by making selective disclosure and giving ambiguous reasons. On the other hand, such words as stable, revenue, operational, growth, and efficiency have strong negative weights, which indicate that they are more related to non-fraudulent businesses. Such words are usually regarded as the discussion of transparent and consistent financial performance and are quite often used in legitimate corporate disclosures. The difference between positive and negative weights of the features represents how the model can distinguish between deceptive and trustworthy language and financial stories [46]. The importance of features analysis will improve the comprehensiveness of the proposed machine learning-based analytics framework since it will show what textual features contribute most to the decision to be made during the classification process. Regulatory/auditing Regulatory and auditing these explain abilities are required in order to justify the results of automation-based fraud detection and enable evidence-based investigations. Figure 8 proves that natural language processing coupled with explainable machine learning can yield important insights on the linguistic characteristics of financial fraud, which enhance the ability to detect fraud and enhance model transparency.

4. Discussion

A. Interpretation of SVM-Based Fraud Classification

The outcomes of the classification provided by the Support Vector Machine (SVM) prove the powerful potential of machine learning models to detect financial statement fraud and other tax-related fraudulent activity [47]. The distinct segregation of the classes in the projection of the SVM demonstrates the efficiency of the linear decision boundaries in the case of high-dimensional textual variables, which are based on financial disclosures. This implies that there is a systematic difference in reporting patterns between fraudulent and non-fraudulent enterprises which can be learnt by classifiers under supervision. The findings indicate that SVM could be applied in fraud detection problems with both textual and financial data, especially because of its strength to operate with sparse and high dimensional feature space. It is also indicated by the quality of the feature extraction methods used in the study since the classification performance is strong [47]. The model identifies semantic and linguistic anomalies that are missed by traditional audit methods by using textual representations of the sections of these reporting processes, the Management Discussion and Analysis (MD&A). These findings validate that signals involving fraud do not merely exist in the numbers alone but in story lines given by the management. This supports the thesis that contemporary fraud detection systems should be based on the combination of both quantitative and qualitative sources of data. Moreover, the performance of the SVM model implies that it has a good generalization ability in balancing dataset situations. The fact that there is no class overlap in the projected space means that there is no arbitrariness in the learned decision function but some meaningful disclosure differences. This also has significant implications in regulatory applications, since it shows that machine learning-based analytics can be used to facilitate repeat and consistent fraud screening [48]. On the whole, the findings of the SVM classification prove the central assumption of the proposed framework that the high-level machine learning methods could contribute largely to the improvement of the effectiveness of financial misconduct detection.

B. ROC Curve and Model Discriminatory Power Analysis

The analysis of the Receiver Operating Characteristic (ROC) curve gives a global evaluation of the capacity of the model to differentiate between fraudulent and non-fraudulent enterprises varying in different classification thresholds [49]. The close to perfect ROC curve and the final value of the AUC demonstrate an excellent discriminatory potential, which means that the model always places fraudulent cases on the top of the list, over legitimate cases. This will be especially important in matters of tax evasion and detection of financial misconduct, which require prompt and correct identification. The obtained ROC results show that the model is effective irrespective of the choice of the threshold, which provides flexibility in implementing the model to the real world. The sensitivity and specificity of trade-offs tend to be sensitive and specific depending on the focus of enforcers. An effective ROC curve allows the decision-makers to regulate thresholds to either obtain maximum fraud detection or reduce false alerts without drastically affecting performance. Machine learning-based analytics systems have a major benefit of this flexibility compared to their fixed rule-based counterparts. The ROC analysis has ensured that high accuracy of the model is not due to balance of the datasets and overfitting [50]. It is an expression of real reparability of classes according to learned characteristics. This enhances the trust in the strength of the presented framework. Policy wise, this reliable discrimination helps to risk-based auditing policies, with reduced investigative resources that may be employed more efficiently. ROC results support the appropriateness of the framework to proactive and scalable financial control.

C. Confusion Matrix Insights and Error Analysis

The confusion matrix analysis gives detailed information about the classification behavior of the SVM model in the analysis of true positives, true negatives, false positives and false negatives [51]. It is observed that there are no errors of misclassification in the confusion matrix, and this means that the model was able to classify the test data perfectly. This result proves the fact that the

model is highly sensitive to detect fraud and control false alarms which is crucial in financial regulation. Enforcement wise, zero false negatives means that no rogue businesses had been missed, which would greatly minimize chances of protracted malpractices and resultant losses. Zero false positives require that honest companies should not be wrongfully subjected to audit or investigation and thus organizational trust is not harmed and costs due to compliance are minimized [52]. This combination performance connects with one of the most intractable issues of fraud detection systems. The confusion matrix is also complemented by the other measures of evaluation through the provision of intuitive instance level validation. Although aggregate measures like accuracy and AUC can be used to describe the general performance, the confusion matrix ensures that the results are not inspired by class dominance or biased predictions. This enhances the validity of the experimental results and contributes to its applicability in the field of real-world situations. One should remember that the ideal classification results can be affected by aspects of the dataset, such as controlled labelling and equal classes [53]. Although the results are very promising, they would need to be validated on bigger and more heterogeneous data to ensure the validity of scaling. The confusion matrix analysis gives solid reasons to believe that the suggested framework can produce credible and regulation-fit fraud detection outcomes.

D. Precision, Recall and F1-Score measures assessments

All three analyses (the precision, recall, and F1-score) prove the strength and stability of the proposed machine learning-based fraud detection framework. Existence of precision scores of 1.00 implies that all the fraud cases that were predicted were fraudulent, which shows the ability of the model to reduce false accusation [54]. This is more so when it comes to financial oversight, where false alarms of fraud may result in both reputation damage and unjustified regulatory overhead. The recalls with 1.00 scores also confirm that the model was able to find all actual fraud cases and there is no risk of the possibility of an undetected financial fraud in the dataset considered. Tax evasion detection requires high recall, as this means that it is possible to lose a significant amount of revenue by failing to detect even a few cases of fraud. This is because the combination of flawless precision and recall indicates that the framework provides an optimal rate of sensitivity and specificity. F1-score summarizes these results by showing the ideal harmony balance in the two classes. This shows that there is no skewed performance of the model on both predictions of fraud and non-fraud. This balanced behavior is essential when used in the field of sustainable deployment in the regulatory systems that require 24/7 operations with a variety of enterprise populations [55]. A combination of these metrics proves that the proposed framework provides us with consistent, reliable, and interpretable performance. Although real-world information can add more complexity, the findings are very strong in indicating the possibility of employing machine learning-based analytics as an inherent part of contemporary financial fraud identification techniques.

E. Importance of Balanced Dataset Design as a Model Performance

The balanced nature of the classes employed in this study is vital in the model performance that has been realized [56]. The dataset also reduces the typical problems of class imbalance by providing the same level of representation to both fraudulent and non-fraudulent enterprises, including biased predictions and false accuracy measures. This design option gives the model the ability to learn equally between the two classes which provides stable and fair classification results. Balanced datasets are of especially great value when developing and testing fraud detection models, as this gives a clear understanding of the discriminative power of a given model [57]. This balanced structure contributed to the great performance metrics that were witnessed in this study since learners were able to learn the fraud-related patterns effectively without one of the classes dominating. Whereas the balanced data enhances the experimental validity, it has also provided the necessity of future studies to correct the imbalance by using methods like resampling, cost-sensitive learning, or anomaly detection. The existing data will be a good starting point to prove the main analytical elements of the suggested framework. The balanced dataset design makes the results

reported more interpretable, reliable, and credible, which contributes to the methodological rigor of the study.

F. Explainability and Regulatory Implications of Feature Importance

The feature importance analysis gives important insights about the linguistic trend that determine the choice of the fraud. The model makes this more transparent by extracting important words that are related to fraudulent and non-fraudulent disclosures and promotes explainable artificial intelligence [58]. This is especially relevant in automated decisions that are used in financial and tax applications where the decisions are required to be justifiable and interpretable. The emergence of the fraud-related terminologies demonstrate the fact that fraudulent reporting frequently is based on the use of certain words to hide the financial truth or to excuse the underlying anomalies. The correlation of stability and performance related terms with non-fraud companies puts into perspective consistent disclosure behavior. Such linguistic distinction confirms that natural language processing is important in the fraud detection paradigms. Explainable feature importance allows auditors and compliance officers to know the reason behind why a firm has been flagged so that instead of relying on the results of the algorithm, they can concentrate on specific investigations. This would promote the ethical use of AI and would be consistent with the governance requirements that focus on accountability and transparency. The explainability aspect reinforces both the practical model and the legislation acceptability of the suggested framework. The study illustrates that machine learning can be responsibly applied to improve financial misconduct detection by providing interpretable insights on the subject along with high predictive performance.

G. Future Work

Future studies may elaborate on the developed machine learning-based analytics framework by overcoming a number of methodological, technical, and practical shortcomings, discovered in this paper. A relevant direction is the consideration of the framework with larger, more diverse, and real-world data that is representative of the natural imbalance of classes that generally occurs in financial fraud and tax evasion [59]. The inclusion of longitudinal financial data of various reporting periods would allow analyzing the patterns of time and behavioral shifts that in many cases precondition the occurrence of fraud. Moreover, it could be investigated in future that how deep learning architectures with transformer-based language models can be applied into capturing the deeper semantic context and finer linguistic signals in Management Discussion and Analysis (MD&A) sections and other narrative disclosures. A second way forward is to include multi-modal sources of data, e.g., transactional data, audit data, network links between companies and executives, to improve the accuracy of the detection and to identify organized or systemic fraud schemes. Ongoing research in this area may consider methodological solutions of hybrid systems and ensembles, involving supervised, unsupervised, and semi-supervised methods that may be better than either other approach to handle high robustness in low-label data environments [60]. The interpretability of models is also a pressing issue that is important to consider, and in the future, one should work towards the development of explainable AI methods unique to the financial regulation and compliance scenario. In addition, conducting a test of the framework in different regulatory thresholds and risk toleration levels would give a pragmatic understanding on the deployment of the framework by auditors and tax authorities. The ethical issues in regard to privacy of data, fairness, and reduction of bias should also be examined further to promote responsible use of AI. Lastly, more research can be done in the future to identify real-time or near-real-time fraud detection features through the use of streaming data analytics to make possible proactive intervention and further monitoring. Taken together, these guidelines would improve the scalability, generalizability and regulatory applicability of machine learning-based fraud detection systems, which would help to build more robust and transparent financial oversight systems.

5. Conclusion

This study showed a holistic machine learning-driven analytics system to identify tax evasion and financial fraud in U.S. businesses through a combination of natural language processing, supervised classification as well as explainable artificial intelligence methodologies. The research relied on regulatory financial filings obtained by the U.S. Securities and Exchange Commission to indicate that textual analysis of Management Discussion and Analysis (MD&A) sections and financial statement narratives are effective information sources about deceptive reporting behavior. The experimental findings proved that the Support Vector Machine (SVM) model is efficient in differentiating between fraud and non-fraud based enterprises with consistently high effectiveness measured by a variety of assessment measures, such as precision, recall, F1-score, ROC curve, and confusion matrix analysis. The distinct division of the classes, close to perfect discriminating ability, and the lack of the misclassification errors demonstrate the strength and consistency of the suggested framework in the set of data to be discussed. In addition to accuracy in prediction, explainable feature importance analysis added to the system has improved transparency, highlighting major linguistic predictors of fraudulent disclosures, which is essential to the necessary regulatory and ethical standards of automated decision-making systems. The balanced dataset design also provided an advantage to the validity of the findings because it reduced bias and provided equal consideration of classes. All of this confirms the idea that the efficiency of machine learning-based analytics is much higher than the efficiency of the rule-based and manual audit methods in revealing complex and dynamic patterns of fraud that are deeply rooted in mass financial disclosures. The suggested framework can be of a valuable use to regulators, auditors, and corporate compliance teams as it helps to identify the risks at an early stage, minimizes the cost of the investigation, and promotes the use of data-driven enforcement approaches. Although the research has limitations to the size of the dataset and balance of controlled classes, it provides a solid basis on the further research and practical implementation. This study will help to move towards intelligent financial regulation by showing how machine learning can be scalable, transparent, and regulation-oriented to enhance tax compliance, corporate accountability, and trust in financial reporting systems of businesses in the United States.

References

[1] D. Lin, “Key considerations to be applied while leveraging machine learning for financial statement fraud detection: A review,” *IEEE Access*, 2024.

[2] X. Ma and D. Wu, “Financial fraud detection using machine learning,” 2025.

[3] M. A. Alrasheedi, S. Ijaz, A. M. Alrashdi, and S. W. Lee, “Advanced tax fraud detection: A soft-voting ensemble based on GAN and encoder architecture,” *Mathematics*, vol. 13, no. 4, p. 642, 2025.

[4] D. O. Njoku, V. C. Iwuchukwu, J. E. Jibiri, C. T. Ikwuazom, C. I. Ofoegbu, and F. O. Nwokoma, “Machine learning approach for fraud detection system in financial institution: A web-based application,” *Machine Learning*, vol. 20, no. 4, pp. 1–12, 2024.

[5] L. F. Cardona, J. A. Guzmán-Luna, and J. A. Restrepo-Carmona, “Bibliometric analysis of machine learning applications in fraud detection on crowdfunding platforms,” *Journal of Risk and Financial Management*, vol. 17, no. 8, p. 352, 2024.

[6] W. Wu, L. Ma, and S. Ma, “Is machine learning really effective in detecting corporate fraud?” *Journal of Accounting Literature*, 2025.

[7] J. Fan, Y. Zhu, and Y. Zhang, “Machine learning-based detection of tax anomalies in cross-border e-commerce transactions,” *Academia Nexus Journal*, vol. 3, no. 3, 2024.

[8] W. Kim and S. Kim, “Enhancing corporate transparency: AI-based detection of financial misstatements in Korean firms using NearMiss sampling and explainable models,” *Sustainability*, vol. 17, no. 19, p. 8933, 2025.

[9] M. Pluviati, *Fraud Detection and Data Analytics: A Review of the Research*, 2024.

[10] G. Manoharan, A. Dharmaraj, S. C. Sheela, K. Naidu, M. Chavva, and J. K. Chaudhary, "Machine learning-based real-time fraud detection in financial transactions," in *Proc. 2024 Int. Conf. Advances in Computing, Communication and Applied Informatics (ACCAI)*, pp. 1–6, May 2024.
[11] A. Orelaja and A. V. Oluwabusola, "AI-driven fraud detection in financial markets: Predictive modeling for risk mitigation and compliance enhancement," *Int. J. Innov. Sci. Res. Technol.*, vol. 10, no. 5, pp. 4509–4520, 2025.
[12] A. A. Paul and C. Ogburie, "The role of AI in preventing financial fraud and enhancing compliance," 2025.
[13] Y. S. Balcıoğlu, "Revolutionizing risk management: AI and ML innovations in financial stability and fraud detection," in *Navigating the Future of Finance in the Age of AI*, pp. 109–138, IGI Global, 2024.
[14] P. T. Nguyen, "Machine learning for anomaly detection in auditing and financial error detection," in *Proc. Int. Congress on Information and Communication Technology*, pp. 107–120, 2025.
[15] E. Li, M. Chen, S. Xiang, and L. Chen, "Graph-learning-empowered financial fraud detection: Progress and future directions," *Intelligent Computing*, 2025.
[16] R. Sun, F. Liu, Y. Li, R. Wang, and J. Luo, "Machine learning for predicting corporate violations: How do CEO characteristics matter?" *J. Business Ethics*, vol. 195, no. 1, pp. 151–166, 2024.
[17] K. Yazawa *et al.*, "Detecting financial misconduct using NLP and machine learning: Evidence from Japan," SSRN, Art. no. 5212010, 2024.
[18] O. I. Okunbor, "Holistic integration of predictive analytics and regulatory compliance to combat financial crimes and cyber fraud," 2025.
[19] L. Li and S. Zhu, "Research on automated audit and risk detection models for enterprise financial data based on deep learning," SSRN, Art. no. 5582013, 2025.
[20] E. Zavitsanos *et al.*, "Machine learning for identifying risk in financial statements: A survey," *ACM Computing Surveys*, vol. 57, no. 9, pp. 1–37, 2025.
[21] H. Thakkar *et al.*, "Artificial intelligence and machine learning in fraud detection: A bibliometric mapping," *Annals of Library and Information Studies*, vol. 72, no. 2, pp. 138–150, 2025.
[22] I. Vlachos and P. G. Reddy, "Machine learning in supply chain management: A systematic review," *Int. J. Production Research*, pp. 1–30, 2025.
[23] A. Adejumo and C. Ogburie, "Financial statement manipulation: Ethical and regulatory perspectives," *GSC Advanced Research and Reviews*, vol. 22, no. 3, pp. 252–264, 2025.
[24] E. Mukarker, "Examining the role of AI in auditing and fraud detection," *EDPACS*, pp. 1–13, 2025.
[25] S. Maheshwari and N. N. Chatnani, "AI and ML-based supervisory technology in financial market surveillance," *FIIB Business Review*, vol. 14, no. 5, pp. 586–604, 2025.
[26] Z. Wang and Y. Chen, "Machine learning model for identifying corporate tax avoidance," *Frontiers in Business and Finance*, vol. 2, no. 2, pp. 296–310, 2025.
[27] Y. M. Walle, "Deep learning-based hybrid SEM–neural network approach for whistleblowing adoption," *Int. J. Disclosure and Governance*, pp. 1–14, 2025.
[28] M. E. Lokanan and V. Maddhesia, "Supply chain fraud prediction using machine learning," *Int. J. Production Research*, vol. 63, no. 1, pp. 286–313, 2025.
[29] M. Al-Raggad *et al.*, "Advancing forensic accounting through bibliometric insights," *Safer Communities*, vol. 24, no. 3, pp. 244–264, 2025.
[30] I. Mabitsela and M. J. Matome, *Defending Against Fraud: Cyber Fraud Detection and Prevention Techniques*, 2025.
[31] Z. Zhang *et al.*, "Toward an integrated framework with corporate culture for fraud detection," *Accounting & Finance*, 2025.

[32] P. M. Preciado Martínez *et al.*, "Comparative analysis of ML models for fraudulent banking transactions," *Cogent Business & Management*, vol. 12, no. 1, p. 2474209, 2025.
[33] Y. Zhu *et al.*, "AI-enhanced prosecutorial supervision in financial big data," *Journal of Advanced Computing Systems*, vol. 4, no. 5, pp. 10–26, 2024.
[34] Y. Zhang, T. Liu, and W. Li, "Corporate fraud detection based on linguistic readability," *Int. Rev. Financial Analysis*, vol. 95, p. 103405, 2024.
[35] D. Pattnaik, S. Ray, and R. Raman, "AI and ML in financial services: A bibliometric review," *Heliyon*, vol. 10, no. 1, 2024.
[36] P. M. Nichols, "Does compliance with the global anticorruption regime require AI?" *American Business Law Journal*, vol. 62, no. 3, pp. 145–164, 2025.
[37] P. Gupta, "The intersection of AI, data, and analytics in fraud prevention," *Asian Journal of Research in Computer Science*, vol. 17, no. 3, pp. 75–92, 2024.
[38] A. Adusumilli *et al.*, "Detecting illegal insider trading using LLMs," in *Proc. IEEE BigData*, pp. 4809–4818, Dec. 2024.
[39] Q. Yu *et al.*, "Detecting financial fraud via CNN–Transformer," in *Proc. ICAACE*, pp. 1047–1051, Mar. 2025.
[40] C. Wei and X. Qian, "An ensemble learning framework for fraud detection," *Journal of Mathematics*, vol. 2025, Art. no. 6643152.
[41] A. I. Weinberg and A. Faccia, "Quantum algorithms in financial crime prevention," *arXiv preprint*, arXiv:2403.18322, 2024.
[42] Z. Nemati *et al.*, "Fraud prediction in financial statements," *Int. J. Finance & Managerial Accounting*, vol. 10, no. 38, pp. 151–166, 2025.
[43] X. Chen, "Financial fraud prediction in Chinese listed companies," in *Proc. ICEDBC*, pp. 192–205, 2024.
[44] A. W. Bello and A. Ojikutu, "Predictive analytics framework for healthcare fraud," *IJSRM*, vol. 13, no. 11, pp. 657–671, 2025.
[45] S. L. Tigga *et al.*, "Comparative study of fraud detection algorithms," in *AIP Conf. Proc.*, vol. 3327, p. 020005, 2025.
[46] L. Pandiri and S. Chitta, "Machine learning-powered actuarial science," 2024.
[47] J. Yang, K. Basile, and X. Zhao, "CSR communication during crisis using ML," *Journal of Research in Interactive Marketing*, 2024.
[48] J. Arora and S. Bhardwaj, "Systematic review of anomaly detection," in *Proc. Int. Conf. Mathematical Modeling and Computational Science*, pp. 411–424, 2025.
[49] R. Hirt *et al.*, "Meta machine learning for inter-organizational analytics," *Information Technology and Management*, vol. 26, no. 1, pp. 57–81, 2025.
[50] R. Khan *et al.*, "AI and ML in Indian finance," in *Transforming Business Finance in the Digital Era*, pp. 149–179, 2025.
[51] E. Dritsas and M. Trigka, "Machine learning and big data: A survey," *Machine Learning and Knowledge Extraction*, vol. 7, no. 1, p. 13, 2025.
[52] M. I. Pramanik *et al.*, "Technological trends in financial crime," *Journal of Posthumanism*, vol. 5, no. 6, 2025.
[53] E. Schneider dos Santos *et al.*, "Fraud detection in public procurement," *EPJ Data Science*, vol. 14, no. 1, p. 52, 2025.
[54] K. Dheenadhayalan *et al.*, "AI methods for financial statement analysis," in *Machine Learning and Modeling Techniques in Financial Data Science*, pp. 211–230, 2025.
[55] Y. Yu *et al.*, "Tracking disclosure changes for fraud detection," in *Proc. IEEE ICASSP*, pp. 1–5, 2025.
[56] I. Supriadi, "The audit revolution: Integrating AI in detecting accounting fraud," *Akuntansi dan Teknologi Informasi*, vol. 17, no. 1, pp. 48–61, 2024.
[57] J. Arora and S. Bhardwaj, "Systematic review paper for anomaly detection," in *Proc. ICMMCS*, vol. 1399, p. 411, 2025.

[58] A. Alfahaid *et al.*, “ML-based security solutions for IoT networks,” *Sensors*, vol. 25, no. 11, p. 3341, 2025.
[59] A. Shetye *et al.*, “Deepfake technologies in financial fraud,” in *Proc. IEEE GINOTECH*, pp. 1–6, 2025.
[60] R. Struyve, L. Vanwersch, and W. Hardyns, “Zheyun Ye football scandal analysis,” *Trends in Organized Crime*, vol. 28, no. 2, pp. 192–214, 2025.
[61] Kaggle Dataset, “Financial statement fraud data,” 2024. [Online]. Available: https://www.kaggle.com/datasets/amitkedia/financial-statement-fraud-data