ISBN: 978-628-96613-3-0. ISSN: 2414-6390. Digital Object Identifier: https://dx.doi.org/10.18687/LACCEI2026.1.1.2042

Machine learning to detect fraud-indicating anomalies in corporate financial statements

Fernando Gutierrez Portela[1]; Luisa Ximena Bustamante Molano[2]; Ludivia Hernandez Aros[3]
[1,3]*Universidad Cooperativa de Colombia, Colombia, fernando.gutierrez@campusucc.edu.co, ludivia.hernandez@campusucc.edu.co*
[2]*Universidad Da Vinci, México, luisa.bustamante@campusucc.edu.co*

Abstract*–Financial statement fraud poses a threat to the transparency of financial information reported in the markets and to investor confidence. Given this scenario, the development and implementation of AI-based systems that incorporate ML techniques allow for the analysis of large volumes of financial data and the early detection of anomaly and patterns associated with fraudulent practices in financial reports. The CRISP-ML(Q) methodology is used, along with a database of 146,045 records from the SEC and COMPUSTAT, with a marked imbalance of fraud cases (0.0066%). Supervised (random forest, RF; logistic regression, LR) and unsupervised (isolation forest, IF; one-class SVM; local outlier factor, LOF) models were used, along with feature selection, normalization, and oversampling with SMOTE. The results show that the RF model achieved a high overall accuracy of 98.18%, standing out for its ability to correctly classify most observations. Unsupervised models achieved a high level of recall in fraud detection (80%), demonstrating their effectiveness in identifying a greater number of fraudulent cases. It is concluded that effective fraud detection requires more robust approaches that combine traditional techniques with cost-sensitive algorithms, adaptive thresholds, and supervised models, strengthening anomaly identification and supporting financial risk auditing and management.

***Keywords*—anomaly, machine learning, fraud detection, financial statements, unsupervised models, supervised models**

I. Introduction

Fraud in financial statements (FSF) is one of the most complex, damaging, and difficult to detect forms of economic crime affecting public and private organizations [1], [2]. Investigations that address this problem affecting the organization and its finances, as well as the collateral damage to different stakeholders such as investors, creditors, regulators, and other interested parties, are required. From an academic perspective, it is defined as the deliberate action of manipulating, altering, or omitting material information contained in the FSE to deceive internal or external users, present a distorted financial picture, and obtain undue benefits. This fraudulent practice includes overestimating assets/income, underestimating liabilities and expenses, and any other accounting manipulation that affects the entity's economic, financial, and equity situation's true presentation [2], [3]. Additionally, FSF undermines accounting integrity, corporate financial reporting transparency, and information reliability that supports decision-making [1], [2].

Various investigations and forensic studies, such as reports by the Association of Certified Fraud Examiners (ACFE), highlight that FSF is the least common category of occupational fraud, accounting for approximately 5% of detected schemes [4]. However, despite its relatively low frequency, the economic impact is disproportionately high: the median losses associated with this form of fraud amount to USD 766,000 per case, a figure that significantly exceeds the losses caused by other categories of organizational fraud [4]. In contrast, asset misappropriation, which includes the misappropriation or unauthorized use of entity resources such as cash, accounts for approximately 89% of cases, although with median losses of less than 120,000 USD (ACFE, 2024). Corruption cases, which include bribery, conflicts of interest, and extortion, generate median losses of $200,000 per incident [4].

Given this scenario, although asset misappropriation and corruption are more common, FSF has the most severe economic and reputational impact, as it affects the accounting structure and distorts key financial performance indicators. Therefore, detecting this type of fraud is one of the most complex challenges, as those responsible often hold high-level positions with access to sensitive information and the ability to circumvent internal controls [5], [6], [7].

Despite regulatory and technological advances, traditional practices of document review and manual verification of accounting records are insufficient to identify sophisticated and well-concealed irregularities [5], [6], [7], [8], [9]. New forms of fraud emerge as technologies advance, and variables such as the fraud triangle, pentagon, and now hexagon point to a need to adjust the audit engagement in its planning, execution, reporting, and monitoring stages. Multiple studies have highlighted that conventional internal control systems tend to be reactive and rely excessively on the auditor's expertise to detect atypical patterns, limiting the ability to prevent and respond to complex and structured fraud schemes [5], [6], [7], [8], [9].

In this context, the integration of advanced analysis tools using artificial intelligence (AI) and machine learning (ML) techniques to strengthen financial information auditing, monitoring, and supervision processes becomes relevant [10], [11]. Techniques based on anomaly detection (AD) allow for the analysis of large volumes of accounting and transactional data, the identification of statistically significant deviations, and the generation of preventive alerts, offering a more proactive and adaptive detection capability in the face of increasingly dynamic fraud schemes [10], [11].

With all of the above, the study contributes to the development of intelligent systems that use ML models, incorporating AD techniques aimed at the early identification of FSF, in response to auditors' new challenges in financial audit engagements. In line with the provisions of ISA 240, updated by the International Auditing and Assurance Standards

Board (IAASB), which requires a more critical, professionally skeptical, and active attitude throughout the audit process [12], the proposed approach considers the use of technology as an enabling element to support the detection of irregularities, especially in contexts of large volumes of data and highly unbalanced data.

From this perspective, the use of ML models applied to intelligent systems contributes to strengthening the quality of corporate financial reporting, optimizing internal control systems, and improving transparency and accountability in organizations through technological solutions with potential application in real-world audit scenarios. In this context, the research question guiding the study is as follows: To what extent do ML models effectively detect FSF in highly unbalanced data contexts, and what techniques can improve their performance in real-world audit scenarios?

II. METODOLOGHY

The research adopted a quantitative descriptive methodological approach, applying ML techniques to analyze fraud detection models' performance/validation metrics in financial statements [13]. The main purpose of this study was to design an intelligent system that would measure the behavior of different supervised and unsupervised models in a real scenario characterized by a marked imbalance of fraud classes in relation to normal data. The target population corresponds to companies that report financial information to capital market control and supervision agencies, specifically audited and consolidated financial statements. The sample analyzed comprises 146,045 accounting records corresponding to companies registered with the Securities and Exchange Commission (SEC) and the COMPUSTAT database [14]. The sample includes both fraud cases and non-fraudulent reports, contrasting statistical behaviors, and accounting patterns. The selection is based on the availability, reliability, and comprehensiveness of these sources for empirical studies on fraud.

The study period for the dataset, covering 1990 to 2014, is justified for two reasons: (1) it is a sufficiently long period to capture structural changes in financial and regulatory reporting systems, including corporate scandals (Enron, WorldCom), regulatory reforms, and technological improvements; and (2) it ensures temporal coverage that observes fraud patterns in different economic contexts and market cycles. Additionally, a systematic review of the literature from 2019 to date has been conducted, incorporating recent findings in the field of AD and FSF using ML [15], [16], strengthening the research's theoretical and methodological framework.

The methodological strategy followed the standard cross-industry process for the development of ML applications with the quality control methodology (CRISP-ML(Q)), widely recognized in the literature for its effectiveness in ML projects [17].

The process was structured in six iterative phases as follows:

1. Understanding the business and the data: The business/model objectives and success criteria are defined, and the data's quality/viability is evaluated.
2. Data engineering: Data are prepared through selection, cleaning, transformation, and balancing processes.
3. ML model engineering: Design, train, and document ML models (performance metrics).
4. ML model evaluation: The model's performance, robustness, and explainability are validated before implementation.
5. Model deployment: The model is integrated into the production environment and existing software systems.
6. Monitoring and maintenance: Model performance is continuously monitored, and adjustments or retraining are performed when necessary.

Each phase was developed cyclically and incrementally, allowing for successive iterations that progressively refine and optimize the architecture of the AD system applied to FSF, in line with previous studies' best practices.

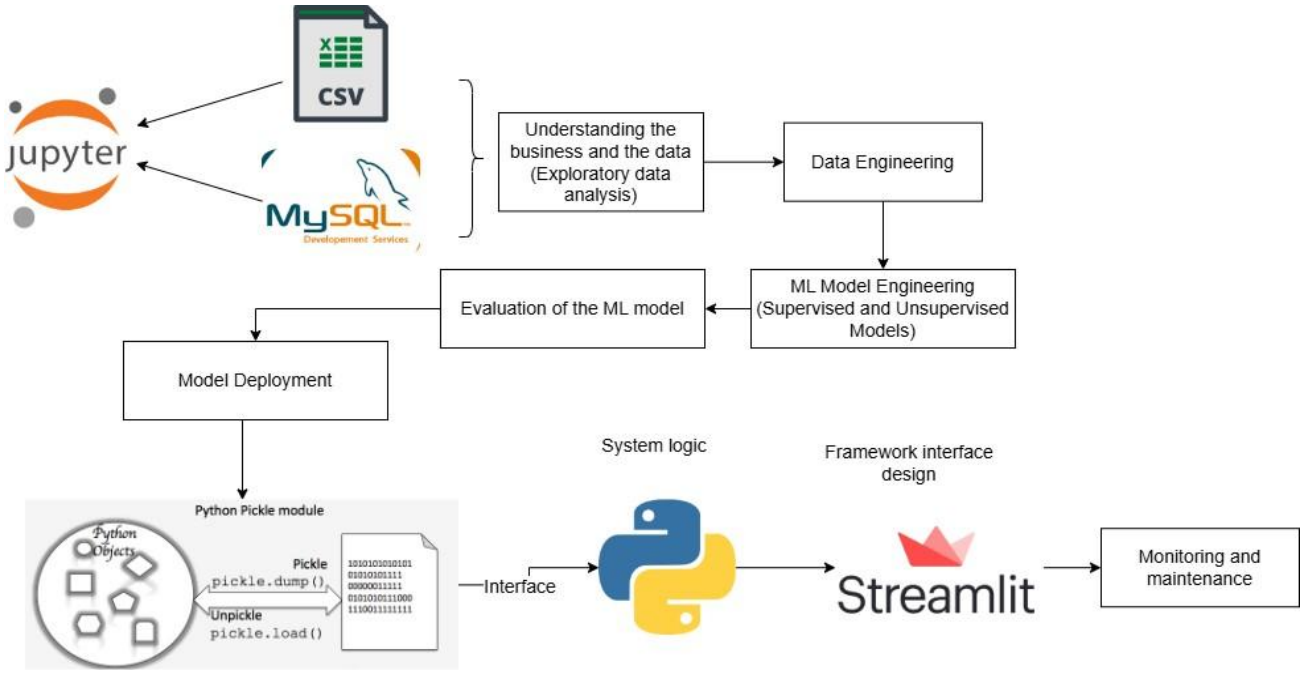


Fig 1 System architecture
Source: Authors

III. RESULTS

This section presents the findings obtained using the CRISP-ML(Q) methodology for detecting FSF using ML techniques.

Understanding the business and data

The business problem was expressed as the need to proactively detect anomalous patterns in key accounting indicators, reducing the exclusive reliance on ex post facto audits. In line with auditing principles, an early detection system was assumed to prioritize minimizing false negatives, given that overlooking fraud causes greater reputational and economic damage than generating false alerts, with implications that affect the financial auditor if each of the Information Assurance Standards (NAI) has not been followed.

The main source of the dataset is [14], which contains data from the AAER issued by the Securities and Exchange Commission (SEC) between May 17, 1982, and September 30, 2016, as well as additional data manually extracted from the SEC through December 31, 2018. The COMPUSTAT dataset is used to obtain financial accounting information, covering the period from 1991 to 2014. The total volume of transactions analyzed amounts to 146,045 records, of which 145,081 correspond to normal records and 964 to cases of fraud, representing 0.0066% of the data. Table 1 presents the details of the dataset.

TABLE 1.
DESCRIPTION OF THE STUDY DATASET

N	Column	Type	Column description
1	Fyear	Int64	Year
2	Gvkey	Int64	A unique four-digit number assigned to each company.
3	P_aaer	Float64	Consecutive reporting periods.
4	Misstate	Int64	Fraud label (1 fraud and 0 no fraud)
The 28 gross accounting variables are described as follows:			
5	Act	Float64	Current assets, total
6	Ap	Float64	Accounts payable, Trade
7	At	Float64	Assets, Total
8	Ceq	Float64	Common/ordinary equity, total
9	Che	Float64	Cash and Short-Term Investments
10	Cogs	Float64	Cost of the goods sold
11	Csho	Float64	Outstanding common shares
12	Dlc	Float64	Current liabilities, total
13	Dltis	Float64	Long-Term Debt Issued
14	Dltt	Float64	Long-term debt, total
15	Dp	Float64	Depreciation and amortization
16	Ib	Float64	Income before the extraordinary items
17	Invt	Float64	Inventories, total
18	Ivao	Float64	Investments, advances, and other
19	Ivst	Float64	Short-term investments, total
20	Lct	Float64	Current liabilities, total
21	Lt	Float64	Liabilities, total
22	Ni	Float64	Net income (loss)
23	Ppegt	Float64	Total property, plant, and equipment
24	Pstk	Float64	Preferred/preference stock (capital), total
25	Re	Float64	Retained earnings
26	Rect	Float64	Total accounts receivable
27	Sale	Float64	Sales/revenue (net)
28	Sstk	Float64	Sale of common and preferred stocks
29	Txp	Float64	Income taxes payable
30	Txt	Float64	Total income taxes
31	Xint	Float64	Total interest and related expenses
32	prcc_f	Float64	Closing price (annual and fiscal)
The 14 variables of the financial ratio are described as follows:			
33	dch_wc	Float64	WC accumulations

34	ch_rsst	Float64	RSST accumulations
35	dch_rec	Float64	Changes in accounts receivable
36	dch_inv	Float64	Change in the inventory
37	soft_assets	Float64	% Soft assets
38	ch_cs	Float64	Change in cash sales
39	ch_cm	Float64	Change in the cash margin
40	ch_roa	Float64	Change in the return on assets
41	Issue	Int64	Effective issuance
42	Bm	Float64	Market reserve
43	Dpi	Float64	Depreciation index
44	Reoa	Float64	Retained earnings on total assets
45	EBIT	Float64	Earnings before interest and taxes on the total assets
46	ch_fcf	Float64	Change in FCFs

Source: [14], [15]

Following on from the above, the characteristics of the classes have been labeled in binary form as "misstate," where a value of 1 indicates fraud and 0 indicates no fraud. The dataset includes four identification and label columns (fyear, gvkey, p_aaer, and misstate), 28 characteristics: assets, liabilities, sales, inventories, cash flows, debt, and equity), and 14 financial ratios (changes in accounts receivable, changes in inventories, return on assets, depreciation index, and free cash flow). The fraud cases identified by the AD models include misleading statements or omissions in the offering/sale of securities, inaccurate financial reports filed with the SEC, failure to file reports, failure to maintain books and internal accounting controls, and accounting errors in provisions for uncollectible, reflecting a realistic scenario where multiple types of fraud can occur simultaneously.

Data Engineering

Several key steps were taken to prepare the data. During cleaning, the key and p_aaer columns were removed. The gray variable corresponds to a unique identifier for each company, which is considered irrelevant for prediction, whereas p_aaer was discarded to control for serial fraud. In the feature selection and reduction stage, the random forest algorithm was applied to estimate the importance of the variables, StandardScaler was used to normalize the data, and SMOTE was implemented to balance the minority class corresponding to fraud cases. Finally, the feature set was reduced from 46 to the 10 most significant variables: sstk, dltis, soft_assets, prcc_f, ap, sale, rect, txp, dlc, and csho.

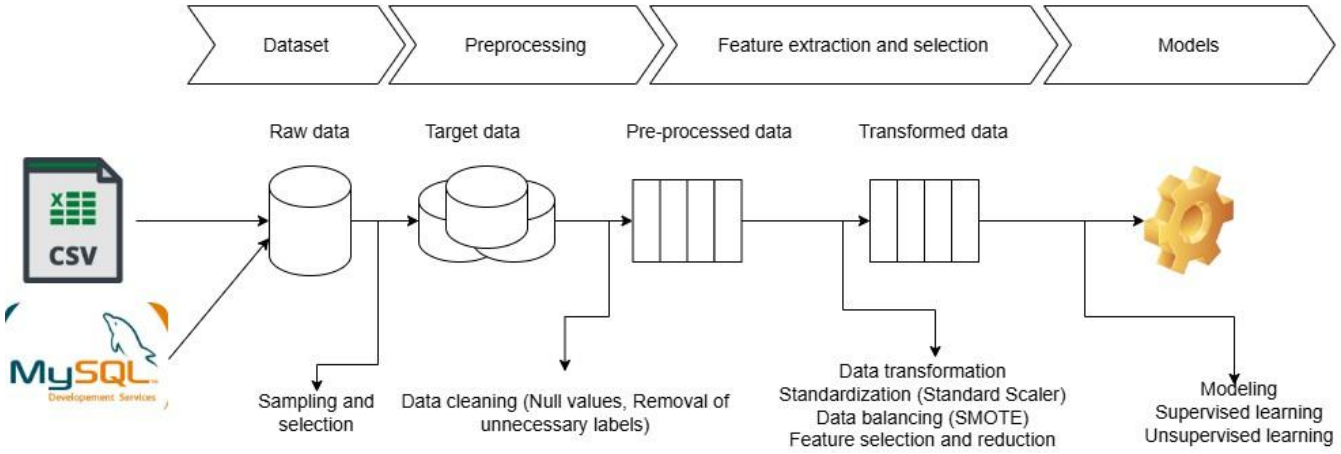


Fig. 2 Preprocessing methods
Source: Authors

ML model engineering

A robust approach was implemented during the modeling stage to detect FSF by combining a preprocessing pipeline, feature selection, and sampling techniques. To this end, five different models were trained, including supervised and unsupervised models, to compare their performance and ability to identify fraudulent observations in a highly unbalanced data context. Table 2 showing the models used, along with a brief description of each and the main parameters used during their training, is provided below.

TABLE 2.
ML MODELS USED IN THE STUDY

Type	Model	Description	Parameters
Supervised	Logistic regression (LR)	A supervised linear model for binary classification was developed. Estimates the fraud probability using a linear combination of selected variables.	Max_iter = 1000 (maximum convergence iterations)
	Random Forest Classifier	Decision tree ensemble. Supervised classifier that combines multiple trees to improve accuracy and reduce overfitting.	Random_state = 42 (reproducibility)

Unsupervised	Isolation forest (IF)	Isolates anomalous observations by generating random data partitions.	n_jobs = -1 (use all cores)
	One-Class SVM	Learns the boundary that encloses most of the normal data.	n_estimators = 50 (number of trees)
	Local Outlier Factor	Measures local density to identify abnormal observations based on their velocity.	Random_state = 42

Source: Authors

All models were trained using a unified preprocessing pipeline to ensure comparability and robustness of the results. The process included variable normalization using StandardScaler, selection of the most relevant features to reduce dimensionality and noise, and the application of SMOTE as an oversampling technique to increase the representativeness of the minority class (fraud). This combination helped mitigate bias toward the majority class and improved the model's ability to detect accounting fraud patterns, which is critical in real-world auditing and financial assurance contexts.

The dataset was split into training (70%) and testing (30%) sets using the train_test_split function from sklearn.model_selection (random_state = 42). This split was performed without stratification; therefore, class proportions were not explicitly preserved across the subsets. While this approach introduces potential variability in class distribution, the subsequent use of SMOTE on the training data helps compensate for class imbalance during model learning.

The methodological design has implications for professional practice, as it proposes an intelligent system based on AI and ML techniques to support fraud prevention and detection tasks in financial contexts. On the one hand, the inclusion of supervised models provides advantages in terms of accuracy, traceability, and interpretability, allowing the results generated by the system to be backed up with quantifiable and verifiable evidence, useful for analysis and incorporation into control, review, or subsequent evaluation processes. However, the incorporation of unsupervised models expands the system's ability to explore data and detect previously undefined atypical behaviors, strengthening the early identification of financial risks. The hybrid approach positions the intelligent system as a technological complement to traditional analysis, capable of generating alerts, patterns, and risk signals that support professional judgment without replacing it, thus contributing to a more robust and proactive assessment of possible signs of fraud.

ML model evaluation

The evaluation was performed using metrics adjusted to the asymmetric nature of the problem (Table 3). The ROC-AUC was used as a robust discrimination measure, reflecting the model's ability to distinguish between fraud and non-fraud, regardless of the classification threshold. Similarly, weighted precision (w) and sensitivity (recall) were analyzed to minimize false negatives, which represent a significant risk in a financial/forensic audit engagement.

TABLE 3.
RESULTS OF THE MODELS USED

Modelo	Accuracy	Precision (w)	Recall (w)	F1-score (w)	ROC AUC	Recall (Fraude - Clase 1)
LR	61.92%	98.86%	61.92%	75.82%	69.28%	62.00%
RF	98.18%	98.70%	98.18%	98.43%	75.84%	11.00%
IF	8.41%	97.71%	8.41%	14.50%	45.05%	82.00%
One-Class SVM	17.55%	98.01%	17.55%	29.08%	42.85%	69.00%
LOF	14.74%	98.50%	14.74%	24.74%	49.16%	84.00%

Source: Authors

The empirical results reveal marked differences in the behavior of supervised and unsupervised models under conditions of extreme class imbalance. RF achieved a high overall accuracy (98.18%) and weighted F1-score (98.43%). However, these aggregate metrics are not informative in this setting, as the fraud class represents only 0.0066% of the observations. Under such imbalance, accuracy and weighted metrics are dominated by the majority class and may systematically overestimate model performance. Indeed, a trivial classifier predicting exclusively the non-fraud class would yield similarly high accuracy.

A more appropriate evaluation based on class-conditional metrics shows that RF attains a recall of approximately 11% for the minority (fraud) class, indicating limited sensitivity to fraudulent instances. The discrepancy between the high weighted recall (98.18%) and the low fraud recall arises from the aggregation scheme, in which class-wise contributions are weighted by support, thereby masking poor performance on rare events.

In contrast, LR exhibits substantially lower overall accuracy (61.92%) but achieves a fraud recall of 61.92%, reflecting a markedly higher true positive rate for the minority class. This result underscores a fundamental trade-off between global accuracy and minority class sensitivity, and suggests that, in applications such as fraud detection, models should be evaluated and selected based on their ability to identify rare but high-impact events rather than on aggregate predictive performance.

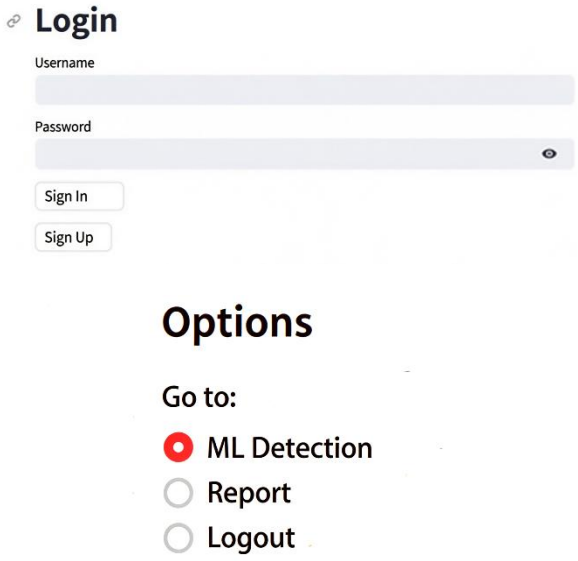
Unsupervised approaches, including IF, One-Class SVM, and LOF, demonstrate limited effectiveness. While some models achieve relatively high recall for the fraud class, this is accompanied by extremely low specificity, resulting in a large number of false positives and, consequently, poor F1-scores and overall accuracy. Furthermore, their ROC-AUC values, which are close to or below 0.5, indicate an inability to reliably discriminate between fraudulent and non-fraudulent observations. This outcome can be attributed to the underlying assumption of AD methods that fraudulent instances constitute statistically separable outlier. In financial reporting date, however, fraudulent behavior is often structurally similar to legitimate activity, leading to significant overlap in feature space and reducing the separability required for unsupervised detection.

Model deployment

A prototype fraud detection system from the study by [15], [16] was used, developed under a frontend–backend architecture, which separates the user interaction layer from the analytical and data processing logic. The frontend was implemented using Streamlit, supported by visualization libraries such as Altair, Matplotlib, Seaborn, and PyDeck, for the interactive graphical interface that supported data loading in Excel and CSV formats, connection to relational databases, configuration of analysis parameters, and visualization of performance metrics/statistical results. In this way, components were incorporated for input validation and navigation between views, facilitating the auditor or financial analyst's usability of the system.

The backend was developed in Python, integrating Pandas, NumPy, Scikit-learn, and SciPy for data processing, transformation, and analysis. The ML models (supervised and unsupervised) were designed, trained, and adjusted in Jupyter Notebook, allowing for controlled experimentation, metric evaluation, and process documentation. SQLAlchemy and MySQL were used for data management to facilitate the persistence, querying, and administration of information.

In addition, the system incorporates frameworks such as Flask and authentication and access control libraries, ensuring proper communication between the system layers. The separation between frontend and backend favors the reproducibility of experiments, software quality control, and future integration of the prototype into broader financial analysis and audit environments.



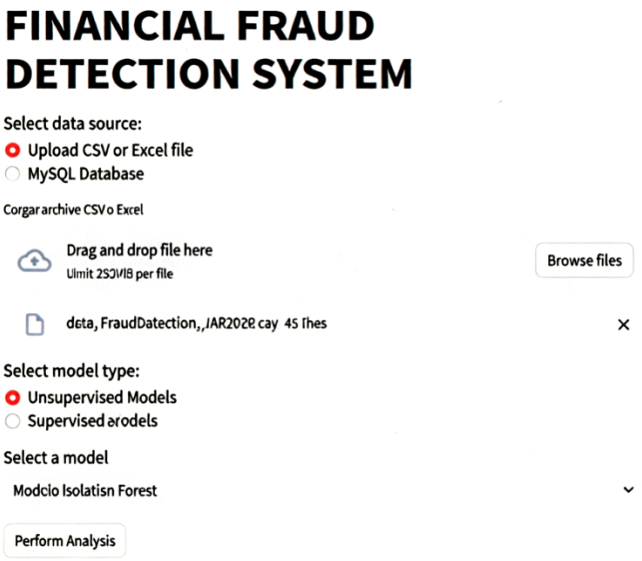


Fig. 3 Designed interface
Source: Authors

Monitoring and maintenance

The intelligent system is continuously monitored to improve the results/performance of the models and their metrics, which leads to adjustments or retraining when necessary.

IV. CONCLUSION AND DISCUSSION

This study provides evidence of the methodological and practical challenges associated with FSF under extreme class imbalance. The results underscore that model evaluation in such context is highly dependent on the choice of performance metrics, and that reliance on aggregate measures alone may lead to misleading conclusions regarding model effectiveness [18].

A central implication of the analysis is that commonly reported global metrics, when used in isolation, are not sufficient to assess predictive utility in rare-event detection problems. In highly skewed distributions, performance indicators are dominated by the majority class, which can obscure deficiencies in identifying minority-class observations. This issue has been widely recognized in the literature, which emphasizes the importance of class-sensitive evaluation frameworks in imbalanced learning scenarios [19], [20].

The finding also reinforces prior research indicating that supervised learning approaches tend to outperform unsupervised alternatives when labeled data are available, even under severe imbalance conditions [18]. Nevertheless, their effectiveness is fundamentally constrained by the scarcity and heterogeneity of fraudulent instances, which limits the ability of standard learning algorithms to construct robust decision boundaries that generalize to unseen fraud patterns.

Although resampling strategies such as SMOTE are commonly used to mitigate class imbalance, their effectiveness in extreme imbalance scenarios is inherently limited. In particular, when the minority class represents a fraction as small

as that observed in this study, synthetic sampling may introduce artificial structures that do not fully reflect the underlying data-generating process. This limitation has been documented in prior studies, which highlight that oversampling techniques may improve class balance in training but do not necessarily enhance true minority-class generalization [15], [21].

From a methodological perspective, the absence of stratified or temporally structured validation constitutes an important limitation. Financial statement data typically exhibit temporal dependencies and non-stationary behavior, which are not fully captured by random train-test splits. As noted in prior research, the lack of time-aware validation can lead to overly optimistic or unstable estimated of model performance in financial applications [1], [19], [22]. This suggests that future work should incorporate cross-validation strategies that preserve temporal ordering or reflect realistic deployment conditions.

The analysis of unsupervised models further indicates that AD techniques such as IF, One-Class SVM, and LOF are sensitive to distributional assumptions that are often violated in financial fraud contexts. Although these methods are theoretically well-suited for identifying rare events, their performance deteriorates when fraudulent behavior does not exhibit clear statistical separation from legitimate transactions. This aligns with previous findings showing that unsupervised methods tend to suffer from high false-positive rates and limited discriminative stability in complex financial environments [19], [20].

Overall, the results suggest that improving fraud detection in financial reporting requires moving beyond algorithmic selection toward more comprehensive learning strategies. These include cost-sensitive learning, decision threshold optimization, and hybrid frameworks that integrate supervised and unsupervised components. Additionally, enhancing dataset representativeness remains a fundamental challenge, as the performance ceiling of predictive models is strongly constrained by the availability and quality of labeled fraud cases.

REFERENCES

[1] T. Shahana, V. Lavanya, and A. R. Bhat, "State of the art in financial statement fraud detection: A systematic review," Technol. Forecast. Soc. Change, vol. 192, p. 122527, Jul. 2023, doi: 10.1016/j.techfore.2023.122527.

[2] C. Onwubiko, "Fraud matrix: A morphological and analysis-based classification and taxonomy of fraud," Comput. Secur., vol. 96, p. 101900, Sep. 2020, doi: 10.1016/j.cose.2020.101900.

[3] L. Chen, B. Xiu, and Z. Ding, "Finding Misstatement Accounts in Financial Statements through Ontology Reasoning," IEEE Access, pp. 1–1, 2024, doi: 10.1109/ACCESS.2020.3014620.

[4] ACFE, "Occupational fraud 2024: A report to the Nations," 2024. [Online]. Available: https://www.acfe.com/-/media/files/acfe/pdfs/rttn/2024/2024-report-to-the-nations.pdf

[5] L. Sabauri, "Internal Audit's Role in Supporting Sustainability Reporting," Int. J. Sustain. Dev. Plan., vol. 19, no. 5, pp. 1981–1988, May 2024, doi: 10.18280/ijsdp.190537.

[6] V. R. Encalada, "Auditoría forense: riesgo de auditoría, fraude y materialidad," Suma Negocios, vol. 14, no. 31, pp. 122–135, Dec. 2023, doi: 10.14349/sumneg/2023.V14.N31.A4.

[7] C. A. Ocampo, "Detectando el Fraude con Inteligencia Artificial: Una Perspectiva Avanzada en Auditoría Forense," Rev. la Junta, vol. 6, no. 2, pp. 13–40, Dec. 2023, doi: 10.53641/junta.v6i2.116.

[8] L. Hernandez, J. J. Sarmiento, and J. J. Moreno, "Structured Framework for Dealing with Types of Financial Statement Fraud, Integrating Common Modalities, Variables, Strategies, and Patterns," J. Risk Financ. Manag., vol. 19, no. 2, p. 157, Feb. 2026, doi: 10.3390/jrfm19020157.

[9] L. Hernandez, J. Juliao-Rossi, and F. Gutierrez Portela, "A bibliometric and content analysis of financial statement fraud: focus on the use of Industry 4.0 technologies," Int. Rev. Econ. Financ., vol. 106, p. 105002, Mar. 2026, doi: 10.1016/j.iref.2026.105002.

[10] M. Lokanan and S. Sharma, "The use of machine learning algorithms to predict financial statement fraud," Br. Account. Rev., vol. 56, no. 6, p. 101441, Nov. 2024, doi: 10.1016/j.bar.2024.101441.

[11] M. N. Ashtiani and B. Raahemi, "Intelligent Fraud Detection in Financial Statements Using Machine Learning and Data Mining: A Systematic Literature Review," IEEE Access, vol. 10, pp. 72504–72525, 2022, doi: 10.1109/ACCESS.2021.3096799.

[12] IFAC, "Basis for conclusions: International Standard on Auditing (ISA) 240 (Revised) ," 2025. [Online]. Available: https://ifacweb.blob.core.windows.net/publicfiles/2025-07/IAASB-ISA-240-Revised-Fraud-Basis-Conclusions.pdf

[13] R. Hernández-Sampieri and C. P. Mendoza-Torres, Metodología de la investigación: Las rutas cuantitativa, cualitativa y mixta, McGRAW-HILL. Ciudad de México, 2018.

[14] Y. Bao, B. Ke, B. Li, J. Yu, and J. Zhang, "Detecting Accounting Frauds in Publicly Traded U.S. Firms: New Perspective and New Method," SSRN Electron. J., 2015, doi: 10.2139/ssrn.2670703.

[15] L. X. Bustamante, L. Hernández Aros, and F. Gutiérrez Portela, "Financial Fraud Detection Through the Application of Machine Learning Techniques with an Anomaly-Based Approach," 2025, pp. 159–172. doi: 10.1007/978-3-031-91328-0_13.

[16] L. Hernandez, L. X. Bustamante Molano, F. Gutierrez-Portela, J. J. Moreno Hernandez, and M. S. Rodríguez Barrero, "Financial fraud detection through the application of machine learning techniques: a literature review," Humanit. Soc. Sci. Commun., vol. 11, no. 1, p. 1130, Sep. 2024, doi: 10.1057/s41599-024-03606-0.

[17] M. Kretschmer, T. Margoni, and P. Oruç, "Copyright Law and the Lifecycle of Machine Learning Models," IIC - Int. Rev. Intellect. Prop. Compet. Law, vol. 55, no. 1, pp. 110–138, Jan. 2024, doi: 10.1007/s40319-023-01419-3.

[18] J. West and M. Bhattacharya, "Intelligent financial fraud detection: A comprehensive review," Comput. Secur., vol. 57, pp. 47–66, 2016, doi: 10.1016/j.cose.2015.09.005.

[19] M. Abbas, D. Almulla, A. Y. Alghasra, and M. Al-Shammari, "Applying Machine Learning to Detect Fraud of Financial Statements: A Systematic Literature Review," in 2024 International Conference on Decision Aid Sciences and Applications (DASA), IEEE, Dec. 2024, pp. 1–7. doi: 10.1109/DASA63652.2024.10836535.

[20] M. Suljić, A. Jusufović, S. Filipović, E. Suljić, and A. Gadžo, "Data mining approach in detecting inaccurate financial statements in government-owned enterprises," Croat. Oper. Res. Rev., vol. 16, no. 1, pp. 1–15, 2025, doi: 10.17535/crorr.2025.0001.

[21] M. A. K. Achakzai and J. Peng, "Using machine learning Meta-Classifiers to detect financial frauds," Financ. Res. Lett., vol. 48, p. 102915, Aug. 2022, doi: 10.1016/j.frl.2022.102915.

[22] S. Gaffaroglu and S. Alp, "Detecting frauds in financial statements: a comprehensive literature review between 2019 and 2023 (June).," Pressacademia, Feb. 2024, doi: 10.17261/Pressacademia.2023.1849.