ISBN: 978-628-96613-1-6. ISSN: 2414-6390. Digital Object Identifier: https://dx.doi.org/10.18687/LACCEI2025.1.1.1302

Financial Anomaly Detection Model Using Deep Neural Networks for Financial Statements in a Peruvian Organization

Delia Vasquez
Universidad Peruana de Ciencias Aplicadas
Lima, Peru
U20211C807@upc.edu.pe

Camilla Navinta
Universidad Peruana de Ciencias Aplicadas
Lima, Peru
U202123464@upc.edu.pe

Juan Mansilla Lopez
Universidad Peruana de Ciencias Aplicadas
Lima, Peru
pcsijman@upc.edu.pe

Gabriel Chicoma
Universidad Peruana de Ciencias Aplicadas
Lima, Peru
pcsigchi@upc.edu.pe

Ricardo Loza
Sheridan College
Oakville, Canada
ricardo.loza@sheridancollege.ca

Abstract– Financial fraud remains a critical challenge globally, with estimated annual losses reaching $5.38 trillion. In Peru, the absence of advanced technological solutions has intensified this issue, leading to significant economic losses of up to $100,000 per incident. Methods like internal and external audits are considered traditional for fraud detection and have proven to be insufficient, identifying only 15% and 4% of fraud cases. To address these shortcomings, this study proposes a deep neural network (DNN) model to detect anomalies in financial statements, leveraging machine learning techniques to improve fraud detection capabilities and provide security in finances. The model analyzes structured financial data, detecting irregularities through feature engineering and anomaly detection techniques. A dataset of 356 financial records from a Peruvian company in the hydrocarbons sector for the years 2021 and 2022 is utilized. The model's architecture consists of multiple densely connected layers optimized to capture nonlinear relationships within financial data. Furthermore, to compensate for the imbalance of classes, the Synthetic Minority Over-sampling Technique was used, enhancing the model's ability to identify fraudulent patterns with greater accuracy. The proposed model demonstrates a substantial improvement over conventional machine learning techniques, achieving an accuracy of 80%, a recall of 80%, and an AUC of 86%, great performance. Additionally, the model efficiently processes financial data in a faster manner, making it suitable for real-time fraud detection applications in high-risk environments. This study underlines the prospect of deep learning to improve anomaly detection, strengthen financial transparency, and enhance risk management in Peruvian organizations and beyond.

Keywords-- Financial anomaly detection, deep neural networks, financial statements, machine learning.

I. INTRODUCTION

Global organizations across various industries face a persistent threat of financial fraud, resulting in an annual loss of 5% of their revenue. Fraudulent practices, such as non-recognition of obsolete inventories and premature recording of revenue, compromise both market stability and investor and stakeholder confidence. by t Globally, the cost of financial fraud is estimated to amount to $5.38 trillion as shown the Association of Certified Fraud Examiners [1], highlighting the growing sophistication and prevalence of fraud in organizations' financial statements and posing significant challenges to financial risk management, transparency, and regulatory compliance in the marketplace.

In Peru, the situation is equally worrying. The prevalence of fraud is remarkable, and the lack of advanced technologies for its prevention has reduced investment in effective solutions. Liu et al. [2] mentions that both large corporations and small and medium-sized businesses in the country have experienced an increase in cases of financial fraud, with significant economic losses that can reach up to US$100,000 per incident.

Despite the efforts of regulators and auditors to mitigate fraud, they face inherent limitations in their capabilities, in along with costs associated with collecting and processing the information necessary for fraud detection. In addition, it has been identified that both internal and external auditors are only able to detect a small percentage of fraud cases, with efficacies of 15% and 4%, respectively shown by Liu et al. [2]. For this reason, there is a growing demand for automated systems that identify fraudulent financial statements.

Previously, models such as the M-Score have been proposed to address the detection of financial fraud, but they present problems related to scalability, adaptability and interpretability. There is a need for further study in this field and to improve the performance of existing methods. Exploration of fraud detection through accounting measures, financial ratios, and non-financial factors related to management decisions or corporate governance practices is on the rise. The use of automated processes based on computational intelligence has become a recent trend in the detection of fraud in financial statements.

In this context, it is critical to assess the underlying risks with fraud in financial statements, as this not only provides a competitive advantage, but also allows addressing the challenges faced by current fraud detection models. Therefore, the development of a model based on deep neural networks to detect possible anomalies in the financial statements of Peruvian organizations is proposed.

Research question: To what extent does a model based on deep neural networks improve the detection of anomalies in the financial statements of Peruvian companies, ensuring a balance between accuracy and efficiency in processing?

II. RELATED WORKS

A. Traditional Fraud Detection Strategies

As mentioned [3], [4], [5], [6], the continued use of traditional strategies such as internal and external audits, along with preventive methods based on specific laws, has been a common practice to detect financial fraud and mitigate risks. However, Song et al. [7] mentioned these strategies face significant limitations in terms of effectiveness and adaptability in an increasingly complex and digitized financial environment. According to Vilvanathan and Aamir [8] and Xu et al. [9], recent studies have shown that both internal and external audits are able to detect only a small proportion of fraud, with estimated effectiveness of 15% and 4%, respectively. This inefficiency has led to a growing demand for automated systems that can identify fraudulent patterns more efficiently. For example, Baesens et al. [3] and Kanapickienė and Grundienė [5] mentioned that traditional rule-based systems rely on predefined sets of rules to identify fraudulent behaviour, which limits their ability to adapt to new forms of fraud. As shown by Papík and Papíkov [6], these systems frequently produce a high rate of false positives due to their rigidity, resulting in the identification of legitimate activities as fraudulent and an additional burden for companies to investigate such alerts. In addition Vilvanathan and Aamir [8] and Baesens et al. [3] mentioned that statistical techniques, such as logistic regression and linear discriminant analysis, which seek to identify suspicious transactions based on historical patterns of data, present difficulties in capturing nonlinear relationships and complex patterns of fraud, especially in a digitized and constantly evolving financial environment. According to Butt et al. [4], these techniques are also sensitive to data quality and tend to fail when faced with unbalanced datasets, where frauds account for only a small fraction of total transactions.

Furthermore, Kanapickienė and Grundienė [5] and Papík and Papíkov [6] propose a robust framework for bank fraud detection using machine learning (ML) and neural networks (NN), which shows a significant improvement compared to traditional audit-based techniques. In addition, Song et al. [7] introduce the use of machine learning dynamic assembly models to detect fraud in financial statements, allowing for greater accuracy and adaptability when detecting anomalies in unbalanced financial data. As shown earlier by [8], [9], [3], these approaches built on artificial intelligence and machine learning have been key to overcoming the limitations of traditional methods, such as the inability to detect complex patterns and the over-reliance on historical data and predetermined rules.

Despite these technological advances, some authors such as Kanapickienė and Grundienė [5] highlight the importance of continuing to develop data engineering for fraud detection, addressing the challenges related to the quality and consistency of financial data. Additionally, Vilvanathan and Aamir [8] and Butt et al. [4] explore methods of detecting fraud in digital payments and find that models based on random forests and logistic regression are more effective in these environments compared to traditional financial audits. This evolution towards more dynamic and automated approaches represents a major shift in the way organizations should approach fraud prevention.

Moreover, some studies have incorporated the use of graph-based neural networks (GNN) and hierarchical attention techniques to improve the detection of fraud in financial transactions, such as the one proposed by Papík and Papíkov [6] and Song et al. [7], which highlights the potential of these technologies to handle multi-relational relationships and unbalanced data, common problems in financial fraud. According to Baesens et al. [3], these developments contrast with the limitations of traditional audit-based strategies, which often lack the capacity to adapt to the speed and complexity of contemporary fraud.

In conclusion, while traditional strategies continue to play an important role in controlling financial fraud in the Peruvian market, the increasing adoption of automated and computationally intelligence-based systems has proven to be more effective in identifying complex fraudulent patterns. These new approaches not only address the limitations of internal and external audits, but also offer greater scalability, adaptability, and accuracy in financial fraud detection.

B. Deep Learning applied to other types of financial fraud

Financial fraud, especially in credit card transactions and electronic payments, has increased in complexity and volume in recent years, leading to a growing need for advanced methods for its detection. As shown by Song et al. [7] and Vilvanathan and Aamir [8], traditional rule-based approaches and statistical analysis have become insufficient to adapt to the evolving and complex nature of modern fraud tactics. In this context Xu et al. [9] and Chavez and Ramos [10] mentioned that machine learning and deep learning models have proven to be promising tools. Techniques such as deep neural networks (DNNs), convolutional neural networks (CNNs), graph neural networks (GNNs) and generative adversarial networks (GANs) with self-service mechanisms are being applied to detect fraud patterns in large volumes of transactional data, overcoming the limitations of classical approaches by better adapting to data variability and complexity as shown by Qiao [11]. These studies focus on addressing issues such as class imbalance, data privacy, and real-time fraud detection, offering more accurate and scalable solutions to combat different types of financial fraud.

Additionally, Chang et al. [37] provide a comprehensive review of fraud detection methods in digital payment systems, highlighting the growing need for intelligent systems that can handle the evolving nature of online transactions. Similarly, Roseline et al. [38] present an autonomous approach for credit card fraud detection using machine learning techniques, demonstrating high efficiency and reduced human intervention in identifying fraudulent activities.

Furthermore, Cherif et al. [14] developed a hybrid deep neural network (DDNN) model for credit card fraud detection. This approach addresses the privacy risks and high costs associated with centralized approaches, enabling local training without the need to share sensitive data. They used distributed deep neural networks, resulting in improvements in accuracy and key metrics, overcoming the limitations of traditional centralized models. On the other hand, [14], [16], [17], [18] proposed Graph Neural Network (GNN) models to address the class imbalance in financial transactions. Their models, based on deep learning and graph structures, utilize subgraphs, neighborhood samplers, and message-passing techniques to enhance fraud detection performance in unbalanced datasets. Specifically, the systematic review by Motie & Raahemi highlights the effectiveness of GNNs in detecting complex fraud patterns in domains such as credit card fraud, cryptocurrency, online payments, and insurance claims.

On the other hand, Sulaiman et al. [19] and Tong and Shen [20] presented an approach based on fuzzy logic to handle uncertainty in real-time transactions. As shown by Trapero et al. [21] and Qayoom et al. [22], these models stand out for its ability to reduce false positives and adapt to changing fraud patterns which differentiates it from traditional models that do not adjust as well to transactional variability. Meanwhile Yajing et al. [23] and Zhao et al. [24] proposed hybrid models that combine oversampling algorithms and deep learning to balance data classes and improve sensitivity without compromising overall accuracy.

Chatterjee et al. [25] and Chatterjee et al. [11] examined the use of digital twins to enhance fraud detection, identifying opportunities and challenges. Bhowmik et al. [12] and Qiao [13] proposed solutions with the use of Digital Twins to simulate and analyse transactions, supported by adaptive fraud detection techniques. Cherif and Mahfoudhi [14] compared data balancing techniques to address imbalance in credit card fraud datasets, while Huang [15] and Puggaard et al. [16] proposed autonomous approaches using machine learning techniques.

Finally, Karthikeyan et al. [17] and Lei et al. [26] developed an encoder-decoder model based on graph networks to detect hidden fraud, capturing complex relationships between transactions and improving the accuracy and robustness of the model against more difficult-to-identify fraud.

Several studies demonstrate that traditional fraud detection methods fall short against complex patterns. Deep learning models like DNNs, GNNs, and GANs show superior performance in identifying anomalies [7], [8], [11], [12], [13]. Others improve accuracy and privacy using distributed architectures [15], [26] or handle class imbalance with encoder-decoder frameworks [14], [16], [17]. Hybrid and fuzzy logic approach further reduce false positives [19], [20]. Digital Twins also emerge as promising tools for simulating transactions and spotting fraud [11], [25], as also demonstrated by Chang et al. [37] and Roseline et al. [38], who showcase the effectiveness of intelligent and autonomous systems in the detection of digital and credit card fraud.

C. *Financial Statement Fraud Detection Algorithms*

In the field of financial fraud detection, several studies have explored alternative techniques to deep neural networks for managing common difficulties like data imbalance and the need for interpretability. Liu et al. [2] and Lui [28] proposed a weighted tree model based on XGBoost that improve the accuracy and interpretability of financial distress prediction by penalizing incorrect classifications of critical cases. Meanwhile Rahman and Zhu [29] and Xie and Yang [33] introduce the Deep Boosting Decision Trees (DBDT) model, which combines traditional boosting with deep learning, achieving improvements in the representation of unbalanced data without the need for resampling.

Wang et al. [30] and Wei [31] employed advanced clustering algorithms combined with smart city technologies, achieving more adaptive and precise accounting fraud analysis. These innovations have led to significant improvements in fraud detection capabilities, particularly when dealing with complex financial datasets. While Wu et al. [32] and Qayoom et al. [22] explore various machine learning models for detecting fraud in digital payments, finding that random forests and logistic regression are most effective in classifying fraudulent transactions, especially when enhanced with reinforcement learning techniques.

Xu et al. [9] developed knowledge graph-based analyses utilizing audit data to identify fraudulent companies. Their approach incorporates advanced graph theory concepts and machine learning algorithms to detect subtle patterns indicative of financial manipulation. Similarly, Cai and Xie [39] proposed a two-layer knowledge graph framework that enhances fraud detection explainability by integrating structured financial data and rule-based insights into a graph neural network (GNN) architecture. This method improves the interpretability of machine learning models by visualizing relationships between entities such as firms, transactions, and auditors. Li et al. [40], on the other hand, explored the fusion of textual analysis with graph-based representations, embedding abnormal managerial tones as node features within GNNs to improve fraud detection accuracy in financial disclosures. This work underscores the potential of GNNs to capture both semantic and relational cues across corporate networks. Furthermore, Zejun et al. [33] and Zhao et al. [24] introduced Self-Attention Generative Adversarial Networks (SAGANs), which enhance credit card fraud detection by generating realistic, complex simulated data that optimize model performance. This advancement has proven particularly effective in addressing the persistent issue of data imbalance in fraud detection.

The latest developments in this field, as demonstrated by Zheng et al. [34], utilize sophisticated analytical methods deliberately crafted to detect fraudulent accounting activities by leveraging the capabilities of smart city technologies. This integration of urban technological infrastructure with fraud detection systems represents a significant step forward in the field. Additionally, researchers like Ashta et al. [35] and Atif et al. [36] have contributed to developing enterprise-scale fraud prediction frameworks that combine multiple analytical approaches, improving the overall effectiveness of fraud detection in corporate environments.

Collectively, these studies underscore the ongoing evolution of financial statement fraud detection algorithms. While early approaches relied heavily on statistical methods, contemporary solutions increasingly incorporate hybrid models that combine traditional analytical techniques with advanced machine learning algorithms. The imperative for advanced detection strategies lies in their ability to adapt to emerging threats while preserving both interpretability and operational efficiency.

III. MODEL ARCHITECTURE

A. *Fundamentals of Deep Neural Networks*

Deep neural networks are invaluable in machine learning, primarily for their capability to distill intricate representations from otherwise unstructured information. Its architecture includes multiple layers of interconnected neurons that process inputs using activation functions such as ReLU or softmax. During training, they optimize weights through backpropagation and optimizers such as Adam or RMSprop to minimize the loss function. As information progresses, more abstract representations are generated, making it easier to spot hidden patterns. Their ability to adjust the relevance of variables makes them superior in classification, although they require high computational resources as show by Yajing et al. [23].

B. *Design and Configuration of the Model for Anomaly Detection*

Fig. 1, shows the architecture of this model, which is composed of three layers: the input layer, responsible for receiving and processing the pre-processed features; dense hidden layers, which facilitate the learning of nonlinear relationships through a hierarchical structure of neural connections; and the output layer, which generates a binary classification to distinguish normal from anomalous records.

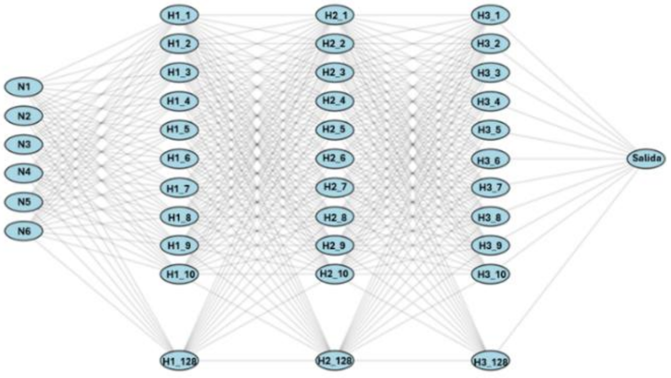


Fig. 1 Neural Network Design

The network input layer (N1 to N6) accepts a vector of high-dimensional features derived from financial data. This input vector includes both numerical and categorical variables, which have been preprocessed using techniques such as the OneHotEncoder encoder.

The network consists of three densely connected hidden layers (H1_128 to H3_128), each with 128 neurons and Rectified Linear Unit (ReLU) activation. The ReLU function is selected for its ability to model nonlinear relationships between variables. This multi-layered design allows the network to learn in a hierarchical manner, where each layer progressively transforms data and facilitates the discovery of complex patterns and latent features in financial records. Thus, the neural network can accurately differentiate between normal and abnormal transactions.

The output layer of the model consists of a single neuron with a sigmoid firing function, which produces a probability value between 0 and 1. This value represents the probability that a record is anomalous (1) or normal (0), based on a set probability threshold. The sigmoid function is suitable for this binary classification task, as it allows the model to assign an interpretive probability to each record, facilitating subsequent decision-making in the evaluation of the results. This aspect will be addressed in greater detail in the next section.

C. *Analysis and detection of anomalies*

Fig. 2 illustrates the proposed approach for the detection of anomalies in financial statements, composed of four phases: data preprocessing, calculus of financial ratios, use of the deep neural network and generation of the report. Each phase is methodologically justified below.

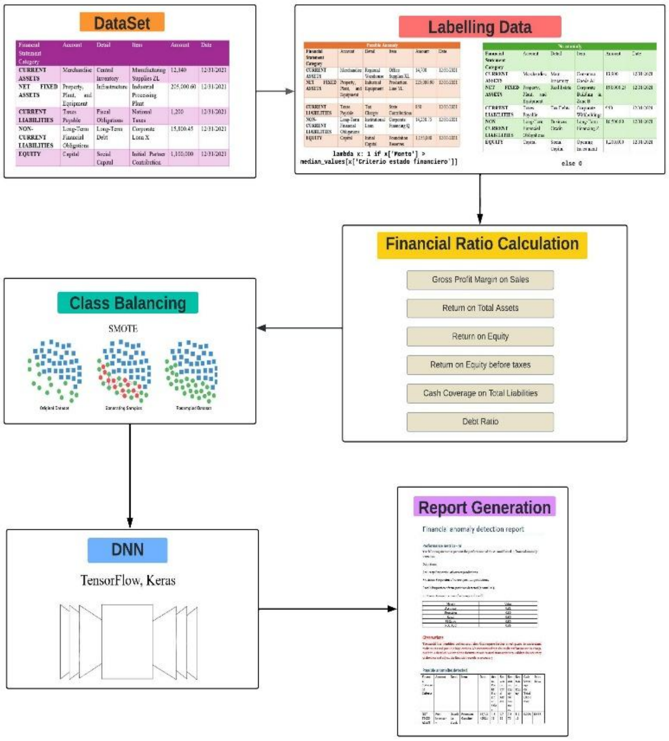

DataSet
Labelling Data
lambda x: 1 if x['Monto'] > median_values[x['Criterio estado financiero']]
else 0
Financial Ratio Calculation
Gross Profit Margin on Sales
Return on Total Assets
Return on Equity
Return on Equity before taxes
Cash Coverage on Total Liabilities
Debt Ratio
Class Balancing
SMOTE
DNN
TensorFlow, Keras
Report Generation
Financial anomaly detection report

Fig. 2 Financial statement anomaly detection process

In the first stage of the process, data preprocessing, the financial information that comes from an Excel file is cleaned. This step ensures the quality and consistency of the dataset, filling in missing values with zeros where there is empty data. In addition, a transformation of categorical columns is carried out – such as Financial Statement Criterion, Account, Detail, Item, Amount and Date – to the format of a text string (string). This standardizes the format of the data, facilitating compatibility and processing of information in the next stages of analysis.

After completing the cleaning and standardization, possible anomalies in the financial data are identified. To do this, a function is implemented that compares each value of the Amount variable with the median corresponding to its category in the Financial Statement Criterion column. For example, if a security belongs to the 'Current Assets' category, it is compared to the median of that category. The median is selected as the reference threshold due to its ability to represent the central value robustly, especially in the presence of outliers or extremes that could distort the mean. In this study, the median is determined by the data of the financial statement under analysis. For example, if the financial statement corresponds to the year 2021, the median of Current Assets is obtained by considering only the values of Amount belonging to Current Assets of that year. If an Amount value significantly exceeds the median of its group, it is assigned a Possible Anomaly label. This method allows detecting considerable deviations from typical behaviour, providing an efficient and reliable mechanism to identify outliers that may require additional attention in the analysis.

The second phase involves the calculation of financial ratios. According to Kanapickienė and Grundienė [5], critical financial ratios are identified and computed to provide a quantitative view of the financial stability and health of the entity under analysis, facilitating the preliminary identification of potentially anomalous patterns that may be indicative of fraudulent behaviour. For this analysis, financial ratios are calculated based on the company's annual financial statement, for example, for the year 2021. These ratios are compared with industry-standard benchmarks, allowing the entity's financial position to be assessed in its sectoral context. Importantly, using a single period can limit the ability to identify long-term trends or seasonal patterns. The ratios considered are as follows:

- Gross Profit Margin on Sales: Evaluates the company's profitability before discounting other operating costs, allowing to detect excessively low margins or inconsistencies that could indicate financial problems.

$$Gross\ Profit\ Margin = \frac{Gross\ Profit}{Total\ Sales}$$

- Return on Total Assets: Measures how efficiently a company uses its assets to generate profits. Unusually low values can indicate operational problems or mismanagement of resources.

$$Return\ on\ Total\ Assets = \frac{Net\ Profit}{Total\ Assets}$$

- Return on Equity before Tax: Indicates the return obtained before taxes. Significant deviations may suggest irregular accounting practices.

$$Profitability\ before\ taxes = \frac{Earnings\ before\ income\ taxes}{Equity}$$

- Return on Equity: Reflects the company's ability to generate value for its shareholders. Drastic changes can signal profit manipulation or financial imbalances.

$$Return\ on\ Equity = \frac{Net\ Profit}{Equity}$$

- Cash Coverage on Total Liabilities: Indicates the company's ability to cover its debts with available cash. Low coverage could signal liquidity issues or elevated financial risks.

$$\text{Cash Coverage on Total Liabilities} = \frac{Cash\ and\ cash\ equivalents}{Total\ Liabilities}$$

- Debt Ratio: Evaluates the level of financial leverage of the company. Excessive indebtedness can indicate elevated risks of insolvency.

$$\text{Debt Ratio} = \frac{Total\ Liabilities}{Equity}$$

The third phase of the anomaly detection process is implemented by the Deep Neural Network (DNN) model, designed to classify transactions as fraudulent or legitimate. This is configured using the build_model function, which defines an internal structure optimized for the analysis of financial data, including multiple densely connected layers that allow the capture of complex nonlinear relationships in the data set.

Since fraudulent transactions are significantly less frequent compared to legitimate ones, the model faces the challenge of a class imbalance, which could lead it to skew its predictions towards classifying transactions as legitimate. To mitigate this problem, the SMOTE (Synthetic Minority Over-sampling Technique) technique is implemented, an oversampling methodology that creates new synthetic samples of the minority class (fraud) by interpolating points in the characteristic space of existing fraudulent transactions.

$$X_{new} = X + (rand() \times (X_{nearest} - X))$$

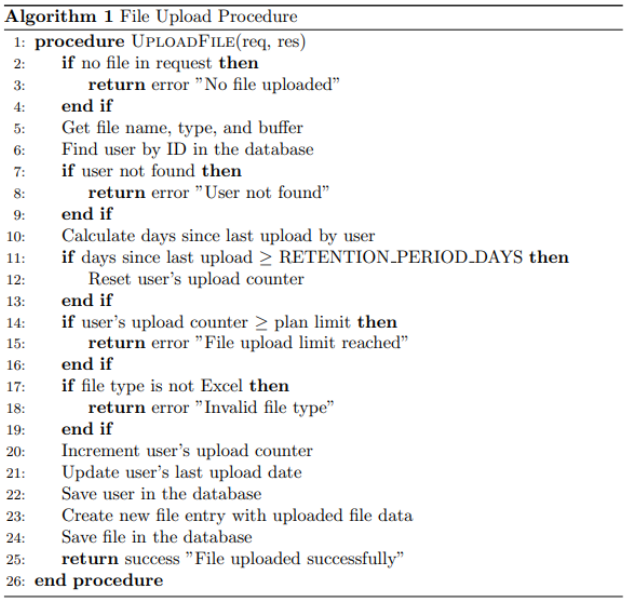
Unlike simple data duplication, SMOTE introduces meaningful variability by generating new data points between existing minority class samples and their nearest neighbors in the feature space. This results in synthetic observations with slightly altered combinations of input features, which helps the model to better generalize and learn the boundaries that distinguish fraud from normal behavior. In this implementation, SMOTE is embedded within a pipeline that first preprocesses the data using a ColumnTransformer to convert categorical variables into numerical format via one-hot encoding. Once the data is transformed, SMOTE is applied to the training subset, identifying fraudulent cases and synthetically expanding them based on their relationships in the multidimensional space of financial attributes. The resampled, balanced dataset is then used to train the deep neural network, ensuring that the model is exposed to a more representative set of fraudulent patterns during learning.

After this step, the model is trained through a network of densely connected layers. These layers allow the model to extract complex patterns in financial data by iteratively adjusting their internal weights. During training, the model learns to associate specific combinations of financial characteristics with both legitimate and fraudulent patterns of behaviour. For example, certain relationships between financial ratios may be indicative of financial risks or specific operating situations. However, not all unusual combinations are necessarily anomalies or frauds. A low gross profit margin combined with a high debt ratio, by itself, does not imply an irregularity, since this behaviour can be common in companies that resort to financing for their operations. As such, the model not only identifies these combinations, but compares them to industry-standard benchmarks to determine if they represent a significant deviation that warrants further evaluation.

Finally, the last layer of the model uses a sigmoid activation function, which transforms the numerical output of the model into a probability with values between 0 and 1. This function allows you to interpret the ranking as follows: if the probability assigned to a transaction is low (close to 0), it is considered legitimate, while if the probability is high (close to 1), the transaction is classified as potentially fraudulent. To define the decision point, a classification threshold is set at 0.90, which implies that only those transactions with a probability of fraud above this threshold will be classified as possible anomalies (1), while the others will be considered normal (0). Once the model has identified transactions that are highly likely to be anomalous, a detailed report is generated that includes not only the transactions classified as suspicious, but also the underlying reasons that may have influenced the classification, such as key financial ratio values. In addition, model evaluation metrics, such as accuracy, recall, F1 Score, and area under the ROC curve (AUC), are presented, providing an objective measure of its performance. The report suggests further review by the auditor or responsible person within the company to properly address the case.

D. Integration of the Fraud Detection Model to the system

Fig. 3, shows the logic of the file upload function. First, the user uploads an Excel file with the financial statements through the interface in React. The Node.js and Express backend validates the file format and, if valid, stores it in the MongoDB input files collection, associated with the user's ID. This ensures that the file is properly saved and ready for processing by the fraud detection model.

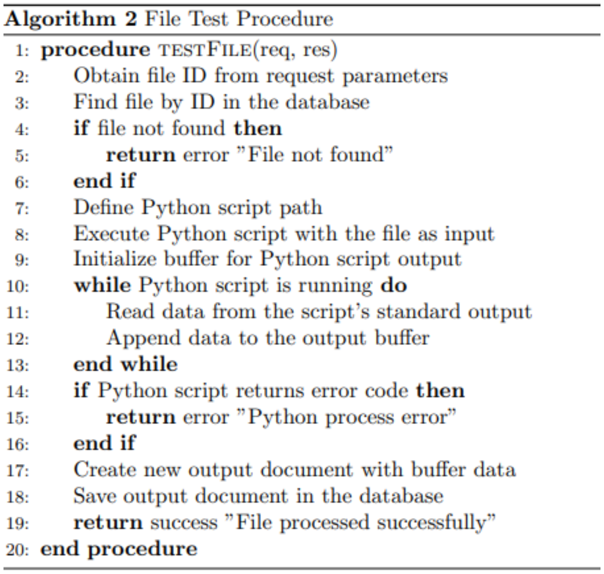


Algorithm 1 File Upload Procedure

```
procedure UPLOADFILE(req, res)
    if no file in request then
        return error "No file uploaded"
    end if
    Get file name, type, and buffer
    Find user by ID in the database
    if user not found then
        return error "User not found"
    end if
    Calculate days since last upload by user
    if days since last upload $\geq$ RETENTION_PERIOD_DAYS then
        Reset user's upload counter
    end if
    if user's upload counter $\geq$ plan limit then
        return error "File upload limit reached"
    end if
    if file type is not Excel then
        return error "Invalid file type"
    end if
    Increment user's upload counter
    Update user's last upload date
    Save user in the database
    Create new file entry with uploaded file data
    Save file in the database
    return success "File uploaded successfully"
end procedure
```

Fig. 3 Logic of the file upload function.

Fig. 4, shows the testing process of the model. When the user clicks the "Test" button, the financial statement file stored in MongoDB is retrieved and sent to the Python-developed deep neural network (DNN) model. The backend in Node.js executes the Python script using child_process (spawn), passing the Excel file through stdin for processing. The model analyzes the data for patterns indicative of potential financial fraud.

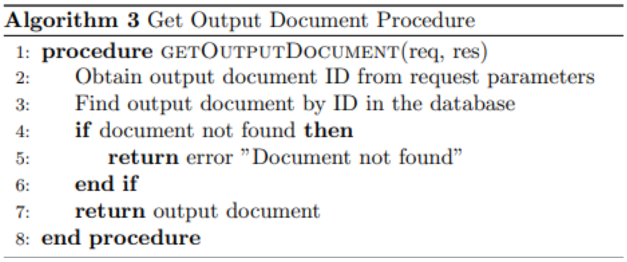


Algorithm 2 File Test Procedure

```
procedure TESTFILE(req, res)
    Obtain file ID from request parameters
    Find file by ID in the database
    if file not found then
        return error "File not found"
    end if
    Define Python script path
    Execute Python script with the file as input
    Initialize buffer for Python script output
    while Python script is running do
        Read data from the script's standard output
        Append data to the output buffer
    end while
    if Python script returns error code then
        return error "Python process error"
    end if
    Create new output document with buffer data
    Save output document in the database
    return success "File processed successfully"
end procedure
```

Fig. 4 Logic of the file test function.

Fig. 5, shows the process of generating and storing the report. After completing the analysis, the DNN model generates a report in Word format with the accounts where possible fraud was detected and metrics such as Recall and F1-score. This report is sent to the backend in Node.js, where it is stored in the MongoDB outputdocuments collection, linked to the same ID of the user who uploaded the original file. Thus, both the input Excel file and the fraud report are associated with the user, making it easier to manage in the application.



```
Algorithm 3 Get Output Document Procedure
procedure GETOUTPUTDOCUMENT(req, res)
    Obtain output document ID from request parameters
    Find output document by ID in the database
    if document not found then
        return error "Document not found"
    end if
    return output document
end procedure
```

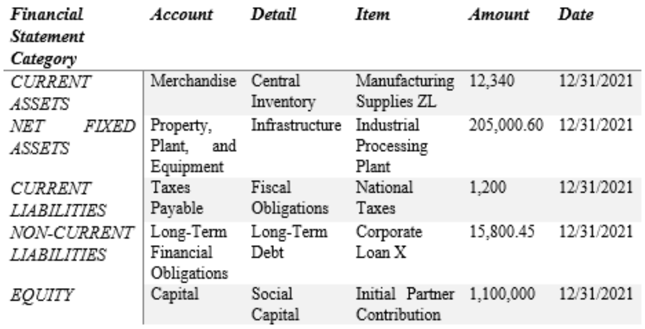
Fig. 5 Logic of generating the output document function.

In the web application dashboard, the input (financial statements) and output (fraud report) files are displayed in two separate tables. The user can view or download the fraud report from the outputdocuments table. By selecting download, the backend queries MongoDB to retrieve the Word file and allows the user to download it directly from the browser.

IV. METHOD

A. Dataset

Fig. 6 shows the structure of the data used in this study, which come from the financial statements of a company in the hydrocarbons sector. This company was strategically selected to build a representative and suitable dataset for the detection analysis of financial anomalies. 356 records corresponding to the years 2021 and 2022 were compiled, organized in a tabular format to facilitate their processing and analysis.



Financial Statement Category	*Account*	*Detail*	*Item*	*Amount*	*Date*
CURRENT ASSETS	Merchandise	Central Inventory	Manufacturing Supplies ZL	12,340	12/31/2021
NET FIXED ASSETS	Property, Plant, and Equipment	Infrastructure	Industrial Processing Plant	205,000.60	12/31/2021
CURRENT LIABILITIES	Taxes Payable	Fiscal Obligations	National Taxes	1,200	12/31/2021
NON-CURRENT LIABILITIES	Long-Term Financial Obligations	Long-Term Debt	Corporate Loan X	15,800.45	12/31/2021
EQUITY	Capital	Social Capital	Initial Partner Contribution	1,100,000	12/31/2021

Fig. 6 Data Set Category

Each record is structured into key variables. The financial statement criterion classifies records into categories such as net fixed assets, equity, current liabilities, and non-current liabilities, among others. The account variable identifies the specific accounting account to which each amount belongs, according to the corresponding financial statement criteria; Examples of these accounts include cash, investments, and accounts payable. The detail variable provides additional information relevant to each account criterion, while the item represents the specific variable associated with both the criterion and the account, covering specific subaccounts or items. The amount reflects the financial value attached to each record, and the date indicates the time period corresponding to the record.

B. Model Training and Fitting

The model uses the RMSprop optimizer with a learning rate of 0.001 and a decay of 0.9, allowing progressive adjustments in weights and stable convergence without the need for frequent manual adjustments. For the loss function, binary crossentropy is used, which is ideal for binary classification problems, as it measures the discrepancy between the model's predictions and the actual labels. Training takes place over 250 epochs, ensuring deep optimization without excessive risk of oversetting. A batch size of 32 samples is used, balancing accuracy and computational efficiency, facilitating regular updates of weights and improving the model's ability to generalize to unobserved data.

C. Model Evaluation Metrics and Validation

The validation process of the anomaly detection model aims to measure its performance and is structured in three key areas: comparison of metrics, analysis of processing times and evaluation using the confusion matrix.

First, the model was compared with other anomaly detection techniques using the same dataset. This analysis included approaches such as eXtreme Gradient Boosting (XGBoost), multilayer perceptron (MLP), convolutional neural networks (CNNs), and recurrent neural networks (RNNs). Key metrics such as accuracy, precision, recall, F1-score and area under the ROC curve (AUC) were used to evaluate performance.

Secondly, the processing times of the neural network were compared with those of the aforementioned techniques. This analysis is critical in financial applications, where speed in detecting anomalies is a critical factor.

Finally, the validation incorporated the confounding matrix to evaluate the accuracy of the DNN model in the classification of anomalies. This matrix represents true and false positives, as well as true and false negatives, allowing the balance between correct and incorrect predictions to be analysed. In addition, it facilitates a detailed examination of classification errors and the rate of false positives, key factors in financial applications.

V. RESULTS

This section presents the results obtained by comparing various machine learning models in terms of accuracy, anomaly classification capability, and efficiency. The key findings of each analysis are detailed below and the effectiveness of DNN is assessed.

A. Model Performance in Classification Metrics

Fig. 7 presents the analysis of the performance metrics of the machine learning models, evaluated in terms of Accuracy, Precision, Recall, F1-Score, and AUC. The goal is to determine which offers an optimal balance between effective anomaly detection and classification error reduction. These metrics are critical in the financial realm, where accurate detection is crucial to mitigate risks and prevent significant losses.

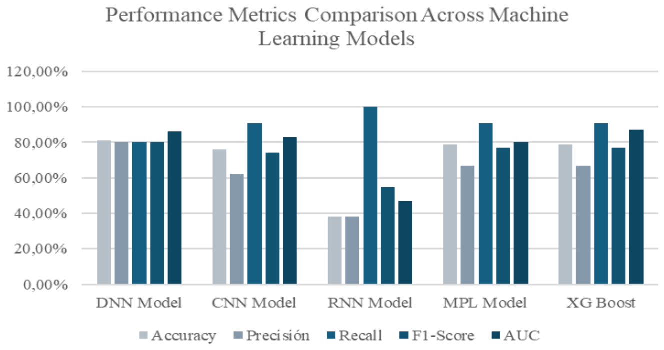


Fig. 7 Performance metrics comparison

The deep neural network model showed outstanding performance in this analysis, reaching 80% in accuracy and recall. High accuracy indicates that most of the anomalies detected by the model are actually anomalies, reducing the number of false alarms (false positives). On the other hand, high recall demonstrates the model's ability to identify most existing anomalies, minimizing false negatives and ensuring that few anomalies go unnoticed.

The DNN's F1-score, with a value of 80%, confirms its ability to balance accuracy and recall, consolidating its performance in terms of overall performance. Additionally, the AUC of 86% suggests a high discriminative capacity between classes, allowing a clear and reliable differentiation between anomalous and non-anomalous instances.

In comparison, the RNN model achieved a 100% recall, detecting all anomalies, but at the cost of a lower accuracy of 68%, which implies a higher rate of false positives. This represents a disadvantage in financial applications, where too many erroneous alerts can lead to unnecessary investigations and additional costs.

On the other hand, XGBoost obtained a higher AUC, but presented lower values in accuracy and F1-score, suggesting a less balanced performance in the anomaly classification. As for the other models, their performance varied between 60% and 85%, but they showed a lack of stability in the metrics, which indicates less consistency in their performance.

B. *Accuracy in Anomaly Classification*

Fig. 8, presents the confusion matrix, which provides a detailed view of the DNN's ability to correctly classify instances as "Anomaly" or "Non-Anomaly". This representation is especially useful in contexts where classification errors can have significant repercussions, such as in financial analysis.

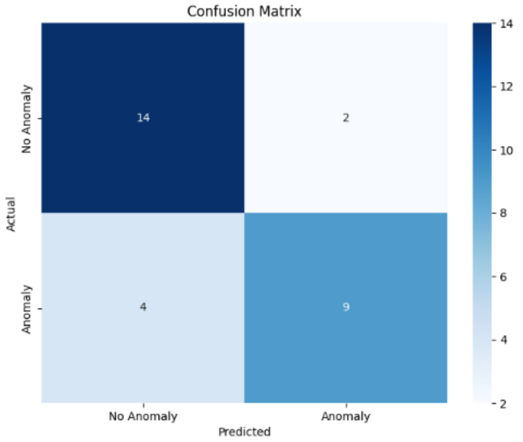


Fig. 8 Performance of Confusion Matrix

The DNN model showed solid accuracy in classification, with 14 true negatives and 9 true positives, indicating that in most cases it managed to correctly distinguish between anomalous and normal instances. However, 4 false negatives and 2 false positives were identified. False negatives represent anomalies not detected by the model, which could be problematic in financial analysis by allowing potential irregularities to go unnoticed. This result suggests that while DNN is highly accurate, further optimization could reduce the rate of false negatives, improving its sensitivity in detecting critical anomalies.

In contrast, the reduced false positive rate strengthens the DNN's ability to minimize unnecessary alerts, which is crucial in financial applications, where each anomaly investigation bears substantial costs. The balance observed in the confounding matrix demonstrates that the model delivers strong performance for environments where accuracy and reduction of misclassification are a priority.

C. *Processing Time Efficiency*

Fig. 9 shows the average processing time of each model over 250 epochs, a key factor in applications that require real-time execution. In addition to accuracy and classification capability, processing speed is crucial to assess the viability of each model in high-demand environments.

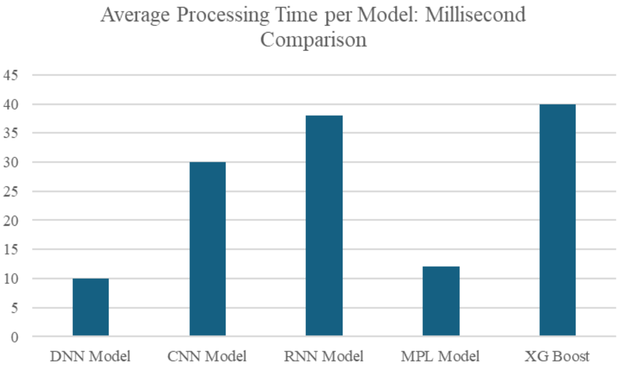


Fig. 9 Average processing DNN Model

The DNN demonstrated outstanding efficiency with an average processing time of approximately 10 ms per epoch, positioning itself as one of the fastest and most efficient models compared to other approaches such as RNN and XGBoost, which recorded average times of 37 ms and 40 ms, respectively. This speed advantage ensures an agile response in anomaly detection, a crucial aspect in financial applications.

In addition, the DNN showed remarkable stability in its runtime, maintaining consistent performance without major fluctuations throughout the 250 epochs. In contrast, other models exhibited variability in their processing times, which could limit their applicability in environments that demand speed and operational stability. The low processing time of the DNN also implies a lower consumption of computational resources, optimizing the required infrastructure and allowing scalable analysis without compromising the accuracy of the model.

VI. CONCLUSIONS

Evidence from this study shows that the DNN model greatly improves the detection of anomalies in financial statements, with a careful balance between accuracy and processing efficiency. By leveraging advanced learning capabilities, the model shows a marked improvement over other technical approaches, offering a more precise and dynamic method for identifying irregular financial patterns. This advancement is particularly valuable in environments where reliability and timely anomaly detection are critical to financial integrity and regulatory compliance.

One significant advantage of the DNN model is its aptitude for identifying shifts in financial behavior, this reduces the risk of both false positives, which can trigger unnecessary investigations, and false negatives, which may allow financial misstatements to go unnoticed. Such improvements contribute to more efficient financial review processes, optimizing resource allocation and ensuring that critical cases receive prompt attention.

When tested on a dataset of 356 financial records from a peruvian company efficiently the model shows that organizations can maintain rigorous financial controls without the delays typically associated with manual review processes. The integration of deep learning techniques allows for continuous adaptation, enhancing the model’s effectiveness in identifying evolving financial anomalies over time.

Beyond improving anomaly detection, the model also represents a progress towards greater automation in financial oversight. By streamlining financial audits and minimizing the need for manual intervention, it enables financial teams to focus on higher-level strategic analysis rather than routine anomaly detection. The results highlight the promising role of deep neural networks in financial analysis, particularly in its ability to enhance both accuracy and efficiency in anomaly detection. As financial landscapes continue to evolve, models like this will play an increasingly vital role in reinforcing trust, transparency, and resilience in financial reporting.

VII. RECOMMENDATIONS

To maximize the impact of the Deep Neural Network (DNN) model in the detection of financial anomalies and ensure its effective application in production environments, the following recommendations are proposed. These suggestions are focused on the scalability of the model and its adaptation to new areas of analysis of financial anomalies.

Although the DNN was validated in a set of 356 registries, its performance should be evaluated in significantly larger volumes, typical of large-scale financial institutions. To achieve efficient scalability without compromising accuracy, it is recommended to explore advanced partitioning techniques that optimize data distribution, as well as the implementation of distributed processing architectures that allow information to be divided and analyzed in parallel.

The DNN’s success in detecting financial anomalies highlights its potential for broader risk analysis. It could enhance fraud detection, transaction monitoring, and compliance auditing by identifying distinct patterns. Expanding its application would validate its adaptability across financial sectors, reinforcing its value in detecting irregularities.

REFERENCES

[1] Association of Certified Fraud Examiners. (2020). Report to the Nations: 2020 Global Study on Occupational Fraud and Abuse. Association of Certified Fraud Examiners. Retrieved from https://legacy.acfe.com/report-to-the-nations/2020/

[2] Liu, W., Fan, H., Xia, M., & Pang, C. (2022). Predicting and interpreting financial distress using a weighted boosted tree-based tree. Engineering Applications of Artificial, 116 Intelligence. https://doi.org/10.1016/j.engappai.2022.105466

[3] Baesens, B., Hoppner, S. & Verdonck, T. (2021). Data Engineering for Fraud Detection. Decision Support Systems. https://doi.org/10.1016/j.dss.2021.113492

[4] Butt, I., Fares, O.H., & Lee, S.H.M. Utilization of artificial intelligence in the banking sector: a systematic literature review. J Financ Serv Mark 28, 835–852 (2023). https://doi.org/10.1057/s41264-022-00176-7

[5] Kanapickienė, R., & Grundienė, Ž. (2015). The Model of Fraud Detection in Financial Statements by Means of Financial Ratios. https://doi.org/10.1016/j.sbspro.2015.11.545

[6] Papík, R., & Papíkov, M. (2022). Detecting accounting fraud in companies reporting under US GAAP through data mining. International Journal of Financial Analysis, 49, 87-98. https://doi.org/10.1016/j.accinf.2022.100559

[7] Song, Y., Sung, W., Jang, Y., & Jung, W. (2020). Application of an artificial neural network in predicting the effectiveness of Co2 sequestration in saline aquifers, 98. International Journal of Greenhouse Gas, 98. https://doi.org/10.1016/j.ijggc.2020.103042

[8] Vilvanathan, L., & Aamir, R. B. (2023). State of the art in financial statement fraud detection: A systematic review. Technological Forecasting & Social Change, 192, 40-1625. https://doi.org/10.1016/j.techfore.2023.122527

[9] Xu, X., Xiong, F. & An, Z. Using Machine Learning to Predict Corporate Fraud: Evidence Based on the GONE Framework. J Bus Ethics 186, 137–158 (2023). https://doi.org/10.1007/s10551-022-05120-2

[10] Charizanos, G., Demirhan, H., & İçen, D. (2024). An online fuzzy fraud detection framework for credit card transactions. Expert Systems With Applications, 252, 957-4174. https://doi.org/10.1016/j.eswa.2024.124127

[11] Mohammed, E. (2024). Leveraging Digital Twins in Adaptive Fraud Detection Systems. Computers & Security, 132, 102675. https://doi.org/10.3390/electronics13193941

[12] Bhowmik, M & Siri, T. & Rudra, B. (2022). A Comparative Study of Machine Learning Algorithms for Financial Fraud Detection in Cryptocurrency Markets. Journal of Financial Markets, 52, 89-105. https://doi.org/10.1109/ICCMC51019.2021.9418470

[13] Qiao, Y. (2024). Effectiveness of long-short term memory network in financial fraud detection. https://doi.org/10.21203/rs.3.rs-4608608/v1

[14] Cherif, A., Ammar, H., Kalkatawi, M., Alshehri, S., & Imine, A. (2024). Encoder–decoder graph neural network for credit card fraud detection. *Journal of King Saud University - Computer and Information Sciences, 36*(3), 102003. https://doi.org/10.1016/j.jksuci.2024.102003

[15] Huang, H., Liu, B., Xue, X., Cao, J., & Chen, X. (2024). Imbalanced credit card fraud detection data: A solution based on hybrid neural network and clustering-based undersampling technique. *Applied Soft Computing, 154*, 111368. https://doi.org/10.1016/j.asoc.2024.111368

[16] Puggaard de Oliveira Hansen, J., Ribeiro da Silva, E., & Bilberg, A. (2024). Agile digital machine development. Computers in Industry, 155, 104061. https://doi.org/10.1016/j.compind.2023.104061

[17] Karthikeyan, T., Govindarajan, M., & Vijayakumar, V (2023). An effective fraud detection using competitive swarm optimization based deep neural network. Measurement: Sensors, 27. https://doi.org/10.1016/j.measen.2023.100793

[18] Motie, S., & Raahemi, B. (2024). Financial fraud detection using graph neural networks: A systematic review. *Expert Systems with Applications, 240*, 122156. https://doi.org/10.1016/j.eswa.2023.122156

[19] Sulaiman, S. S., Nadher, I., & Hameed, S. M. (2024). Credit card fraud detection using improved deep learning models. *Computers, Materials & Continua, 78*(1), 1049-1069. https://doi.org/10.32604/cmc.2023.046051

[20] Tong, G., & Shen, J. (2023). Financial transaction fraud detector based on imbalance learning and graph. Applied Soft Computing, 149, 1568-4946. https://doi.org/10.1016/j.asoc.2023.110984

[21] Trapero, J. R., Holgado de Frutos, E., & Pedregal, D. J. (2024). Demand forecasting under lost sales stock policies. International Journal of Forecasting, 40(1), 111-122. https://doi.org/10.1016/j.ijforecast.2023.09.004

[22] Qayoom, A., Yadong, W., Song, W., Abbas, S., & Ghafoor, N. (2024). A Deep Reinforcement Learning Framework for Detecting Fraudulent Bank Account Openings. Sir Syed University Research Journal of Engineering & Technology, 14(2), 85–92. https://doi.org/10.33317/ssurj.653

[23] Yajing, L., Zhengya, S., & Wensheng, Z. (2023). Improving fraud detection via hierarchical attention-based Graph Neural Network. Journal of Information Security and Applications, 72. https://doi.org/10.1016/j.jisa.2022.103399

[24] Zhao, C., Sun, X., Wu, M., & Kang, L. (2024). Advancing financial fraud detection: Self-attention generative. Finance Research Letters. https://doi.org/10.1016/j.frl.2023.104843

[25] Chatterjee, P. & Das, D. & Rawat, D. (2024). Digital twin for credit card fraud detection: opportunities, challenges, and fraud detection advancements. Future Generation Computer Systems, 158. https://doi.org/10.1016/j.future.2024.04.057

[26] Lei, H., Zhang, T., & Liu, Y. (2024). A Distributed Deep Neural Network Approach for Credit Card Fraud Detection with Privacy-Preserving Capabilities. Finance Research Letters, 19(1), 1-12. https://doi.org/10.1016/j.frl.2023.104547

[27] Lui, Y. (2023). Design of XGBoost prediction model for financial operation fraud of listed companies. International Journal of System Assurance Engineering and Management, 14, 2354-2364. https://doi.org/10.1007/s13198-023-02083-z

[28] Rahman, M., & Zhu, H. (2024). Detecting accounting fraud in family firms: Evidence from machine learning approaches, 64, 100722. Advances in Accounting. https://doi.org/10.1016/j.adiac.2023.100722

[29] Rahman, M., & Zhu, H. (2024). Predicting financial distress using machine learning approaches: Evidence China. Journal of Contemporary Accounting & Economics, 20(1). https://doi.org/10.1016/j.jcae.2024.100403

[30] Wang, X., Liu, J., Liu, Z., & Liu, J. (2023). Fraud detection on multi-relation graphs via imbalanced and. Information Sciences. Information Sciences, 642, 119153. https://doi.org/10.1016/j.ins.2023.119153

[31] Wei, D., Nan, H., & Fujing, X. (2024). The information content of financial statement fraud risk: An ensemble. Decision Support, Systems 182 (2024) 114231 https://doi.org/10.1016/j.dss.2024.114231

[32] Wu, H., Chang, Y., Li, J., & Zhu, X. (2022). Financial fraud risk n audit information knowledge graph. The 8th International Conference o ogy and Quantitative Management. https://doi.org/10.1016/j.procs.2022.01.097

[33] Zejun, Z., Zhao W., & Lixin Cai (2025). Predicting financial fraud in Chinese listed companies: An enterprise portrait and machine learning approach. Pacific-Basin Finance Journal, 90. https://doi.org/10.1016/j.pacfin.2025.102665

[34] Zheng, X., Ali, M., Hamid, A., & Hou, Y. (2024). Data mining algorithm in the identification of accounting fraud by smart city information technology. Heliyon. https://doi.org/10.1016/j.heliyon.2024.e30048

[35] Astha, V., et al. (2024). A Robust Framework for Fraud Detection in Banking using ML and NN. Proceedings of the National Academy of Sciences, India Section A: Physical Sciences. https://doi.org/10.1007/s40010-024-00871-1

[36] Atif Khan, M., & Juan, P. (2023). Detecting financial statement fraud using dynamic ensemble machine learning. International Review of Financial Analysis, 89. https://doi.org/10.1016/j.irfa.2023.102827

[37] Chang, V., Minh, L., Doan, T., Di Stefano, A., Sun, Z., & Fortino, G. (2022). Digital payment fraud detection methods in digital ages and. Computers and Electrical Engineering. https://doi.org/10.1016/j.compeleceng.2022.107734

[38] Roseline, B. & Naidu, G & Samuthira, V. & Rajasree, A. & Mageswari, N (2022). Autonomous credit card fraud detection using machine learning approach. Computers and Electrical Engineering, 102. https://doi.org/10.1016/j.compeleceng.2022.108132

[39] Cai, S., & Xie, Z. (2024). Explainable fraud detection of financial statement data driven by two-layer knowledge graph. *Expert Systems with Applications, 246*, 123126. https://doi.org/10.1016/j.eswa.2023.123126

[40] Li, J., Guo, C., Lv, S., Xie, Q., & Zheng, X. (2024). Financial fraud detection for Chinese listed firms: Does managers' abnormal tone matter? *Emerging Markets Review, 62*, 101170. https://doi.org/10.1016/j.ememar.2024.101170