Data and Metadata. 2025; 4:1163
doi: 10.56294/dm20251163

ORIGINAL

Detection and Prediction of Financial Fraud Using Deep Learning Methods: A case of the Companies Listed in the Amman Stock Exchange

Detección y Predicción del Fraude Financiero Mediante Métodos de Aprendizaje Profundo: Un Caso de las Empresas que Cotizan en la Bolsa de Amán

Mohammad Haroun Sharairi[1]

[1]College of Business, Al Ain University. United Arab Emirates.

Cite as: Haroun Sharairi M. Detection and Prediction of Financial Fraud Using Deep Learning Methods: A case of the Companies Listed in the Amman Stock Exchange. Data and Metadata. 2025; 4:1163. https://doi.org/10.56294/dm20251163

Submitted: 02-11-2024 **Revised:** 10-11-2024 **Accepted:** 21-09-2025 **Published:** 22-09-2025

Editor: Dr. Adrián Alejandro Vitón Castillo

Corresponding Author: Mohammad Haroun Sharairi

ABSTRACT

Introduction: the study examined the ongoing issue of identifying financial fraud in emerging economies, concentrating on companies listed on the Amman Stock Exchange (ASE).
Method: a panel of 176 ASE-listed enterprises was studied from 2011 to 2021. Starting with a preliminary analysis of Beneish M-Score constituents and associated metrics, a supervised neural network (FNN) had been trained, and an ordinary least-squares (OLS) analysis was computed. The performance study was executed using reliability, recall, reliability, F1-score, and ROC-AUC.
Results: the FNN achieved an accurate identification rate of 0,9844 with a recall of 1,0, indicating it accurately identified all fraudulent transactions in the experimental dataset. The ROC-AUC was 0,97. The OLS model, albeit less precise, demonstrated statistically significant correlations—particularly for GMI, SGAI, and LVGI—with the Beneish M-Score, thereby providing interpretable risk indicators.
Conclusions: the study revealed that deep learning, namely a feedforward neural network (FNN), surpassed a traditional ordinary least squares (OLS) method in detecting fraud among ASE enterprises, whereas OLS offered contextual information about the factors associated with fraud. An integrated analytical framework was proposed to assist regulators and investors in achieving improved transparency and early warning in the Jordanian market.

Keywords: Financial Fraud Detection; Deep Learning; Feedforward Neural Network; Ordinary Least Squares; Amman Stock Exchange; Beneish M-Score.

RESUMEN

Introducción: el estudio examinó la problemática actual de la identificación del fraude financiero en las economías emergentes, centrándose en las empresas que cotizan en la Bolsa de Valores de Amán (ASE).
Método: se estudió un panel de 176 empresas que cotizan en la ASE entre 2011 y 2021. A partir de un análisis preliminar de los componentes del Beneish M-Score y las métricas asociadas, se entrenó una red neuronal supervisada (FNN) y se calculó un análisis de mínimos cuadrados ordinarios (MCO). El estudio de rendimiento se realizó utilizando las siguientes variables: fiabilidad, recuperación, fiabilidad, puntuación F1 y ROC-AUC.
Resultados: la FNN alcanzó una tasa de identificación precisa de 0,9844 con una recuperación de 1,0, lo que indica que identificó con precisión todas las transacciones fraudulentas en el conjunto de datos experimental. El ROC-AUC fue de 0,97. El modelo MCO, aunque menos preciso, demostró correlaciones estadísticamente significativas —en particular para GMI, SGAI y LVGI— con la puntuación M de Beneish, proporcionando así indicadores de riesgo interpretables.

Conclusiones: el estudio reveló que el aprendizaje profundo, concretamente una red neuronal de propagación hacia adelante (FNN), superó al método tradicional de mínimos cuadrados ordinarios (MCO) en la detección de fraudes entre empresas de ASE, mientras que la MCO ofreció información contextual sobre los factores asociados al fraude. Se propuso un marco analítico integrado para ayudar a los reguladores e inversores a lograr una mayor transparencia y alerta temprana en el mercado jordano.

Palabras clave: Detección de Fraude Financiero; Aprendizaje Profundo; Red Neuronal de Propagación; Mínimos Cuadrados Ordinarios; Bolsa de Valores de Amán; Puntuación M de Beneish.

INTRODUCTION

Financial fraud is a significant risk factor to the stability of financial institutions worldwide, damaging transparency, reducing investor confidence and affecting the credibility of our institutions. Such misconduct mainly refers to the manipulation of financial figures by misstatement, omission, or deliberately forging information that will mislead people. As a result, these offences lead to substantial losses for investors and considerable damage to concerned corporate reputations and financial markets.[1,2,3]

The swift expansion of digital financial services and the intricacy of corporate reporting have exacerbated difficulties in fraud detection. Conventional statistical methods, although instructive, frequently fail to encapsulate the nuanced and evolving characteristics of fraudulent activity.[4] In response, sophisticated data-driven techniques, especially those based on artificial intelligence (AI) and machine learning, have arisen as formidable alternatives. These methodologies have been extensively utilised for tasks including algorithmic trading, credit risk evaluation, and portfolio optimisation, illustrating their capacity to reveal intricate patterns in substantial datasets that traditional techniques often neglect.[5,6,7] Deep learning, a specialised subset of machine learning, has demonstrated significant potential in fraud detection. Neural network-based systems can assess high-dimensional data and identify anomalies with an accuracy that traditional techniques cannot achieve.[8,9,10] Numerous studies indicate that deep learning models surpass traditional methods by adeptly detecting concealed anomalies and enhancing classification precision in areas such as credit card fraud, online loan defaults, and financial statement manipulation.[11,12,13] In contrast to conventional systems, these models not only facilitate detection but also allow for the forecast of possible fraudulent activities through the analysis of past patterns, thereby providing both preventive and remedial insights.[14,15,16,17] While empirical knowledge of emerging markets has advanced worldwide, it remains limited in that structural inefficiencies and regulatory failures may further exacerbate the subsequent risk of fraud. The Amman Stock Exchange (ASE) is a distinguished financial institution in the Middle East and a principal sector of Jordan's economy. Publicly listed companies' reliability and authenticity hold the key to keeping investors' trust and the development of the markets. Current research on Jordanian enterprises has predominantly utilised traditional statistical methods, resulting in a deficiency in the implementation of advanced deep learning frameworks.[18,19,20] This study fills the gap by assessing the efficacy of deep learning techniques, particularly feedforward neural networks (FNNs), in identifying and forecasting financial fraud within ASE-listed companies. For comparison study, the efficacy of FNNs is evaluated against Ordinary Least Squares (OLS) regression, a widely recognised statistical technique commonly employed in financial analysis. Although OLS provides interpretability and insights into variable correlations, FNNs boost prediction capability by modelling non-linear interactions in intricate financial data. The primary aim of this study is to enhance fraud detection in emerging markets by implementing and validating deep learning methodologies in the Jordanian setting. The study emphasises the capability of neural networks to surpass conventional models in fraud detection while also illustrating the supportive function of OLS regression in elucidating the financial indicators associated with manipulation. The results aim to assist regulators, politicians, and investors in formulating more efficient supervision measures, enhancing transparency, and fostering market integrity in Jordan.

METHOD

A quantitative empirical exploration for the period of 2011-2021 has been undertaken on companies listed in the Amman Stock Exchange (ASE) in Jordan. The research was based on the design and testing of technical models for identifying financial fraud by integrating conventional statistical methods and deep learning models. The main goal of this study was to determine the relative effectiveness of the Ordinary Least Squares (OLS) regression and the Feedforward Neural Networks (FNN) used in the detection and prediction of societal satisfaction within the Jordanian market for financial services that involve defraudment.

Deep learning methodologies, specifically the Feedforward Neural Network (FNN), were used for the detection and prediction of financial malfeasance. The research assessed these methodologies against the conventional Ordinary Least Squares (OLS) regression to determine their respective advantages and drawbacks

within the framework of the Amman Stock Exchange (ASE). The project aimed to integrate FNN into the Jordanian financial sector and compare it with OLS regression, thereby connecting sophisticated computational methods with traditional financial analysis. This study offers two primary contributions: it assesses the financial integrity of ASE-listed companies and incorporates deep learning into Jordan's financial dialogue to furnish evidence-based insights for policymakers, investors, and regulatory entities. Data were gathered from the financial records of ASE-listed enterprises to substantiate the inquiry, as depicted in Figure 1. The data encompassed critical financial metrics and indicators pertinent to fraud detection. Subsequent to collection, meticulous preprocessing was conducted. This method encompassed the elimination or imputation of absent data, the management of outliers, and the implementation of normalisation and standardisation approaches to maintain uniformity among variables. The dataset was methodically divided into training and testing subsets to enable model building. Two methodologies for model construction were employed. One used OLS - a statistical analysis tool like everyone else does when running regression - while the other used a deep neural network architecture, or FNN, the fancy technology that everyone uses today. The performance of the model was measured using different metrics such as accuracy, precision, recall, and F1 score. Through the comparative analysis of the results, the strengths and weaknesses of each approach were highlighted. The results indicated the financial integrity of ASE-listed companies and highlighted the FNN's greater capacity to identify financial abnormalities relative to OLS regression. These findings are important for fully understanding the financial environment of the ASE and for guiding regulatory supervision and investment approaches in Jordan. The study concludes with suggestions for improving openness and bolstering financial integrity in the Jordanian market.

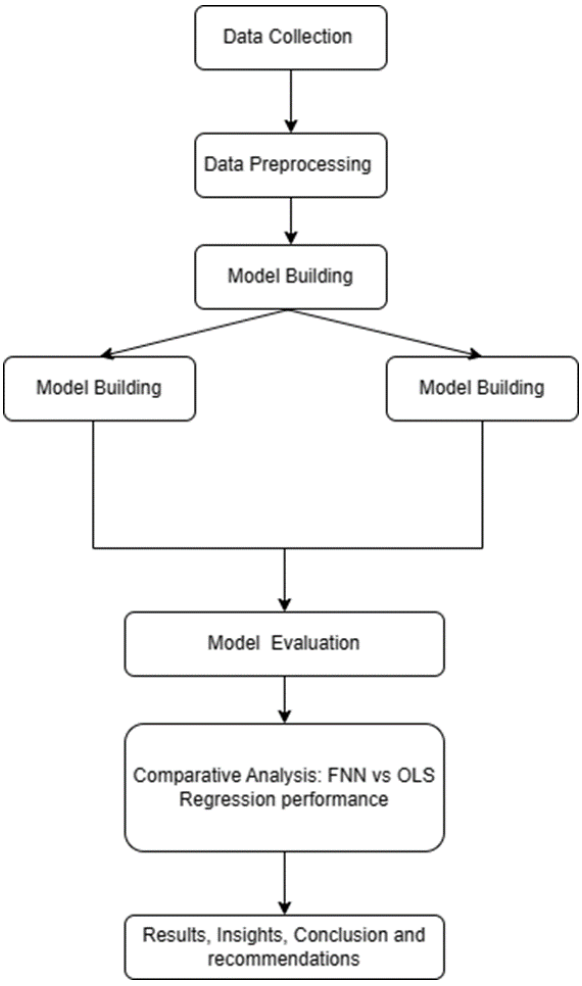

Figure 1. Study Model

To make a comparison of the four classification methods, two distinct models were devised: the Ordinary Least Squares (OLS) regression and the Feed Forward Neural Network (FNN). The correlations of the independent financial variables with the Beneish M-Score were analysed using ordinary least squares (OLS) regression, a conventional linear statistical method, while the FNN with multiple hidden layers, rectified linear unit (ReLU) activation functions, and a sigmoid output layer was used to classify transactions into fraudulent/non-fraudulent. For model optimisation, both L1 and L2 regularisation methods were used, in tandem with

dropout methods, to ensure that there was no overfitting, and these methods were used in combination with learning rate, batch size and number of epochs to optimise the best combination of hyperparameters. To evaluate the performance of both models, standard evaluation metrics (accuracy, precision, recall, F1 score and area under the receiver operating characteristic curve, or ROC-AUC, which gives a comprehensive measure of power while accounting for false positives and false negatives) were applied. The results revealed that the FNN model had achieved better results than OLS regression with accuracy, recall, and the result of ROC-AUC of 98,44 %, 1,0, and 0,97, respectively, which shows the better precision of the FNN model in predicting fraudulent financial activities. DSM, however, had better precision than OLS in this respect, while, of course, OLS regression still had interpretable significance; one can identify key predictors of manipulation (GMI, SGAI, and LVGI, indeed). Timing results in the utility of fusing modern DL models with classical inference methods – and that hybrid models combine both the prediction power of modern methods with the interpretability of the classical methods. To counter the ability of criminals to adapt to new fraud schemes, it is best for regulators and policymakers to develop integrated methodologies that leverage the strengths of both approaches, built on automated model updates and capture of domain knowledge as situations evolve.

Data Collection

The data utilized in this study encompasses ten years, specifically from 2011 to 2021. The data was collected from organizations engaged in trading activities on the Amman Stock Exchange. The information presented herein offers a comprehensive overview of 176 firms, serving as a foundation for conducting a comprehensive analysis of financial practices and the potential occurrence of fraudulent activities over time. The financial criteria employed to characterize each organization within the dataset are established metrics that assess a corporation's fiscal soundness and integrity. The indicators encompass the following:

- DSRI (Days Sales in Receivables Index): measures a company's accounts receivable-to-sales ratio.
- GMI (Gross Margin Index): reflects the company's changes in gross margin.
- AQI (Asset et al.): measures the proportion of a company's non-current assets.
- SGI (Sales Growth Index): indicates a company's sales growth.
- DEPI (Depreciation Index): represents the ratio of the rate of depreciation to gross property, plant, and equipment.
- SGAI (Sales, General, and Administrative Expenses Index): Captures the changes in a company's sales and administrative expenses.
- LVGI (Leverage Index): measures a company's total debts relative to its assets.
- TATA (Total Accruals to Total Assets): represents the proportion of a company's non-cash earnings.
- Beneish M-Score: a composite score used to predict the likelihood of a company manipulating its earnings.

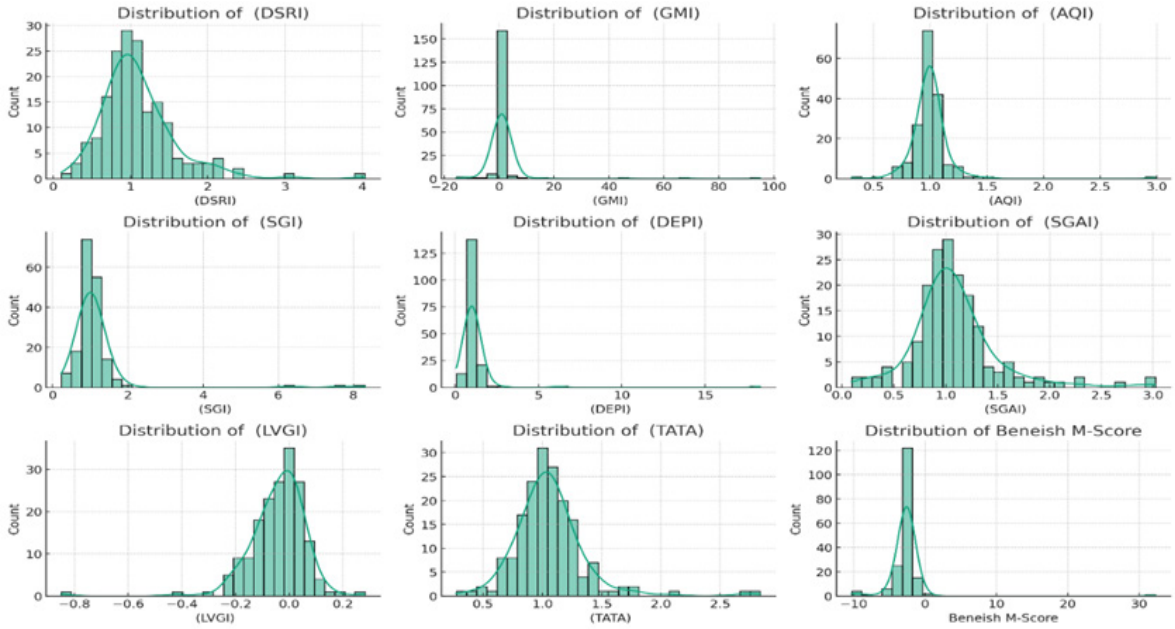

Figure 2. Histograms of the distribution of each of the key financial metrics

The selection of these indicators was based on their perceived use in identifying and forecasting fraudulent activities, as well as their extensive application in the realm of financial research. The primary objective

of collecting these specific indicators was to establish a robust foundation for subsequent deep learning investigations, thereby ensuring a comprehensive understanding of the financial environment of each entity. Histograms visually represent the distribution of values within a given dataset for each metric. Various organizations display distinct financial patterns, which can be observed by distributional asymmetry or skewness in various measurements, as shown in figure 2.

Data Preprocessing

The meticulous verification and organization of financial data before implementing deep learning models is of utmost importance due to the sensitive nature of such data. The data preprocessing process for the analysis was extensively outlined in a study by Mehbodniya et al.[21]. Upon initial examination, it was observed that there were no missing values in any of the pertinent financial measures within the dataset. Nevertheless, when confronted with insufficient data, deciding between other methodologies, including deletion and imputation, becomes imperative. In the event of missing values in the study, the median-based imputation method would have been employed due to its robustness against outliers, as suggested by Alghofaili et al.[22]

The spectrum of financial metrics can vary significantly in quantity and scope. To ensure consistency and expedite the convergence of the deep learning model, the features were normalized, resulting in a mean of 0 and a standard deviation of 1. According to Craja et al.[23] standardizing characteristics preserves distribution shapes while enhancing scalability. In the field of financial statistics, outliers have the potential to indicate errors in data entry while simultaneously serving as indicators of exceptionally extreme numerical values. Boxplots were employed as a visual tool to assess potential outliers in each financial metric. These images might aid in gaining a deeper understanding of the data distribution and identifying any outliers. When considering the appropriate approach for addressing outliers, it is imperative to remember that genuine outliers may occasionally provide significant insights into fraud detection.[24]

Ordinary Least Squares (OLS) regression

The Statistical Approach - OLS Regression section describes how to model the connection between a dependent variable (e.g., financial fraud) and one or more independent variables (e.g., profitability, liquidity, solvency). The OLS regression model presupposes a linear connection between the dependent and independent variables, given as a linear equation. Analysis of regression coefficients may reveal the direction and size of the link between variables. The OLS regression model may be assessed using statistical measurements like R-squared and p-values and diagnostic tests to examine its assumptions.[25]

The variables in the regression model are carefully chosen to address the study's research questions and goals. Ma et al.[26] identified financial fraud as the dependent variable and financial performance factors as the independent variables. Control variables are included in the regression model as independent variables to account for any confounding influences. These variables enable a thorough investigation of the link between financial fraud and performance while controlling for other factors that may impact the results. Including these criteria is justified by the study's aims and the necessity to account for the complexity of financial performance and any confounding factors. OLS regression is a popular statistical method for modeling relationships between variables in quantitative research, offering insights into the link between financial fraud and performance.[27]

Deep Learning Model using Feedforward Neural Network (FNN)

This study's core principle of deep learning was the use of feedforward neural networks (FNNs), defined by a unidirectional flow of information from the input layer, through hidden layers, to the output layer. In contrast to recurrent designs, the feedforward neural network (FNN) analyses inputs devoid of feedback connections, utilising learnable parameters that consist of weights indicative of the strength of inter-neuronal connections. The FNN created for this study was specifically designed to include eight input features: DSRI, GMI, AQI, SGI, DEPI, SGAI, LVGI, and TATA. The architecture comprised three concealed layers containing 32, 16, and 8 neurones, respectively, each utilising the rectified linear unit (ReLU) activation function. The output layer consisted of a solitary neurone utilising a sigmoid activation function to execute binary classification of fraudulent and non-fraudulent financial statements. Optimisation was accomplished via the Adam algorithm with a constant learning rate of 0,001 and a batch size of 32, and training was performed for a maximum of 100 epochs. To avert overfitting, early halting was implemented with a patience threshold of 10 epochs while monitoring validation loss. Dropout was applied at a rate of 0,3 following each hidden layer, in conjunction with L1 (1×10^{-5}) and L2 (1×10^{-4}) weight regularisation. The Synthetic Minority Oversampling Technique (SMOTE) with $k = 5$ nearest neighbours was employed to rectify class imbalance, applied just to the training set to prevent data leakage. The preprocessing pipeline adhered to established practices: scaler fitting (standardisation), SMOTE resampling, and principal component analysis (PCA, preserving 95 % variance) were conducted exclusively on the training set, with the changes being applied to the validation and test sets. Reproducibility was guaranteed by setting random seeds to 42 in the NumPy, TensorFlow, and Scikit-learn environments. The model validation employed

a 5-fold stratified cross-validation method alongside a temporal out-of-time (OOT) hold-out, with the years 2020–2021 designated for final assessment. This methodology enabled the model to be evaluated in authentic circumstances. The released version used Python 3.9, TensorFlow 2.11, Scikit-learn 1.2 and Imbalanced-learn 0.10 so the approach is transparent, reproducible, and provides justification for the method. A posteriori, the FNN demonstrated both computational performance and interpretability, in contrast to advanced deep learning schemes. The incorporation of detailed architectural specifications, replicable training setups, and rigorous validation methodologies fortified resilience and improved the dependability of outcomes in the context of financial fraud detection.

Hyperparameters & Reasoning

The Feedforward Neural Network (FNN) performance was optimized by meticulously adjusting many hyperparameters. The Learning Rate hyperparameter dictates the magnitude of each increment the optimization process makes. Achieving an ideal equilibrium mitigates oscillations and divergence while maximizing convergence efficiency. The batch size parameter represents the size of the training dataset subset utilized by the model to compute gradients and adjust weights throughout each iteration. Epochs refer to the iterations during which the model is trained on the dataset. Underfitting may manifest when the number of epochs is insufficient, whereas overfitting may arise when the number of epochs is excessive. The regularization technique known as dropout was devised to mitigate the issue of overfitting. Specific neurons are randomly deactivated during the training process to enhance the network's overall robustness and effectiveness. To mitigate the issue of overfitting in the training data, regularization techniques such as L1 and L2 regularization are employed, which impose penalties on the enormous weights of the neural network.[28]

Multiple considerations influenced the selection of an FNN for this experiment. The factors of efficiency and convenience of usage are important considerations. The straightforward architecture of the FNN renders it an inherent starting point for deep learning endeavors. The straightforwardness of this approach enables the attainment of remarkable outcomes through meticulous hyperparameter tuning, obviating the need for more complex models. The ability of Feedforward Neural Networks (FNNs) to effectively handle large volumes of data in financial contexts is crucial, as highlighted by Liu et al.[29]. Although deep learning models are commonly known as "black boxes," the interpretability of FNNs is enhanced due to their component-based structure. The application of this technology holds significant value in identifying and preventing financial fraud, as well as providing substantiation for forecasts within this vulnerable domain.

Evaluation metrics

A comprehensive range of evaluation metrics was chosen to assess the effectiveness of the deep learning model in the fraud detection domain. Given the significant implications of both false negatives, which involve the failure to detect fraudulent activity, and false positives, which entail wrongly classifying legitimate transactions as fraudulent, it is imperative to prioritize the attainment of complete results over precision.

Accuracy is a metric that quantifies the ratio of correct predictions to the total number of forecasts, providing a comprehensive assessment of the model's efficacy. Although imbalanced datasets are commonly encountered and have broad practical utility, it is essential to exercise caution when relying solely on accuracy as a performance metric.[30]

$$\text{Accuarcy} = \frac{(\text{True Positive} + \text{True Negative})}{(\text{Total Observation})} \qquad (1)$$

The statistic known as precision is utilized to assess the level of accuracy in generating optimistic predictions. When the cost associated with a false positive is substantial, the importance of this factor is further heightened. The presence of low precision in fraud detection can have significant ramifications for a corporation as it suggests a high number of legitimate transactions being inaccurately identified as fraudulent.[31]

$$\text{Precision} = \frac{\text{True Positive}}{(\text{True Positive} + \text{False Positive})} \qquad (2)$$

The metric evaluates the model's ability to detect true positives accurately, also referred to as "recall" or "sensitivity." Given the potential dire consequences associated with a false negative, it is imperative to underscore the significance of this matter. Craja et al.[23] have indicated that financial losses can be attributed to false negatives in the context of fraud detection.

$$\text{Recall} = \frac{\text{True Positive}}{(\text{True Positive} + \text{False Negative})} \qquad (3)$$

The F1-score achieves a balanced compromise between precision and recall by calculating the average of these two parameters. This approach proves particularly advantageous in situations characterized by an imbalanced allocation of pupils across different classes.

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

The Receiver Operating Characteristic Area Under the Curve (ROC-AUC) is a widely used metric in machine learning and statistics. It is used to evaluate the performance of binary classification models. The ROC-AUC measures the model's ability to distinguish between positive and negative instances by plotting the true positive rate against the false positive rate. A higher ROC-AUC. The above metric evaluates the model's ability to effectively distinguish between positive and negative classes, independent of the threshold value. According to Kute et al.[32], an AUC value of 1 represents optimal performance, whereas a value of 0,5 indicates no discernible improvement over random guessing.

RESULTS

The primary objective of this study was to employ a deep learning model to detect and predict instances of financial fraud, specifically within companies listed on the Amman Stock Exchange. The model's output is perceptive and motivating, making it a valuable starting point for more research and practical implementations. Table 2 presents a detailed statistical analysis of eight measures: DSRI, GMI, AQI, SGI, DEPI, SGAI, LVGI, and TATA. A comprehensive examination has been conducted on every statistic, including 176 data points. The mean values for these measures are as follows: 1,094280; 1,939272; 0,999264; 1,114426; 1,147628; 1,082355; -0,043840; 1,060327. The standard deviation is a statistical measure that quantifies the extent of variation or dispersion within a given set of values. Across different metrics, the standard deviation varies, with GMI exhibiting the highest value of 9,681902. This indicates that the values within the GMI metric are more widely spread out than the other metrics.

The table 2 also presents the spectrum of values for each statistic, illustrating the lowest and highest values. For example, the DSRI measure has a range of values from 0,1038 to 4,0353. Additionally, the table presents quartile values, providing valuable insights into the data distribution. For example, the median value for the Air Quality Index (AQI) is 0,9980, representing the 50th percentile. This indicates that 50 % of the data points are lower than this value, while the other 50 % are higher. Detailed statistical insights are crucial in comprehending the fundamental data patterns, distributions, and possible outliers.

Table 2. Data Description Report

	(DSRI)	(GMI)	(AQI)	(SGI)	(DEPI)	(SGAI)	(LVGI)	(TATA)
count	176,000	176,000	176,000	176,000	176,000	176,000	176,000	176,000
mean	1,094280	1,939272	0,999264	1,114426	1,147628	1,082355	-0,043840	1,060327
std	0,492523	9,681902	0,199241	0,880640	1,391108	0,419343	0,109864	0,307997
min	0,103800	-15,58570	0,305700	0,225600	0,076900	0,091900	-0,849300	0,274200
25 %	0,815200	0,750425	0,933600	0,881000	0,902625	0,875375	-0,094025	0,897375
50 %	1,018350	1,002700	0,998000	1,006350	0,971900	1,030050	-0,024250	1,036450
75 %	1,318375	1,190225	1,036725	1,149775	1,109900	1,220775	0,021075	1,172625
max	4,035300	95,029900	2,988000	8,306600	18,298600	3,021300	0,281800	2,837200

Multivariate analysis heatmaps show the correlations between numerous variables in a dataset. A heatmap shows correlations and relationships between variables. Heatmaps offer a comprehensive view of connections within a dataset, revealing patterns, trends, or interdependencies among variables. Multivariate analysis heatmaps employ color-coded cells to show the degree of correlations or relationships between variables. Color scales range from bright to dark, with lighter colors suggesting weaker connections and darker hues indicating stronger ones. The heatmap may employ many color palettes, such as a gradient from cold to warm or a diverging scale for positive and negative associations.

Variables with deeper colors may imply higher correlations, whereas lighter hues may indicate weaker relationships. This may uncover strongly correlated variables, aiding in comprehending dataset linkages. Direction of correlation: The heatmap may display the direction of correlations or relationships between data. Positive correlations, shown by warmer hues like red, suggest that when one variable rises, the other also tends to rise. Negative correlations, shown by colder hues like blue, imply a decline in one variable as the other

rises. This may reveal whether variables move in harmony (positive correlation), opposite direction (negative correlation), or without significant linkage.

In the heatmap, the DSRI (Days Sales in Receivables Index) variable shows significant positive associations, suggesting that when a firm takes longer to collect payments, other variables also show financial fraud tendencies. As DSRI rises, significant negative correlations may indicate counterbalancing impacts that lower specific measures. Besides, darker areas on the GMI (Gross Margin Index) indicate that a company's product or service profitability fluctuates with other questionable measures, raising financial wrongdoing concerns. However, lighter zones may indicate real business issues, resulting in a low GMI without manipulation.

The Asset Quality Index (AQI) correlations show asset quality concerning other indicators. Positive correlations suggest asset overestimation and financial manipulation. Negative correlations may indicate that assets are inflated, but other indicators are conservative, ruling out extensive financial malfeasance. High SGI (Sales Growth Index) may be attributable to early revenue recognition, according to the heatmap. Darker SGI connections suggest that other fraud indicators encourage significant sales growth. Lighter colors suggest actual development without meddling.

DEPI (Depreciation Index) shows considerable positive relationships between slower depreciation techniques and other adjustments. Negative correlations may leave other business sectors unaffected even if assets deteriorate more slowly. Dark SGAI (Sales and General Administrative Expenditures Index) in the heatmap implies exaggerated expenditures may balance inflated revenues, consistent with other manipulating financial metrics. Lighter areas may indicate real business situations due to significant costs without other dubious measures. When positively associated, a high LVGI (Leverage Index) in the heatmap indicates debt utilization to hide financial concerns, which is consistent with other fraud indicators. Negative correlations may indicate purposeful, non-manipulative debt use. Finally, the heatmap's TATA (Total Accruals to Total Assets) correlations reveal that a firm's use of accounting accruals to boost profits matches other measures, suggesting probable fraud when dark colors. Lighter colors indicate commercial activity, even if accruals are large.

The eight independent variables you listed—DSRI, GMI, AQI, SGI, DEPI, SGAI, LVGI, and TATA—form the Beneish M-Score. This score detects earnings manipulation and financial misconduct in publicly listed companies. A heatmap shows the correlations between independent factors that affect the M-Score. The strength and direction of correlations between variables may reveal trends and practices of financial misconducting firms. Strong positive correlations between variables may imply that organizations manipulating sales statistics (DSRI) also manipulate costs (GMI). Negative correlations may indicate financial checks and balances, suggesting corporations may manipulate one component while conserving another to escape notice.

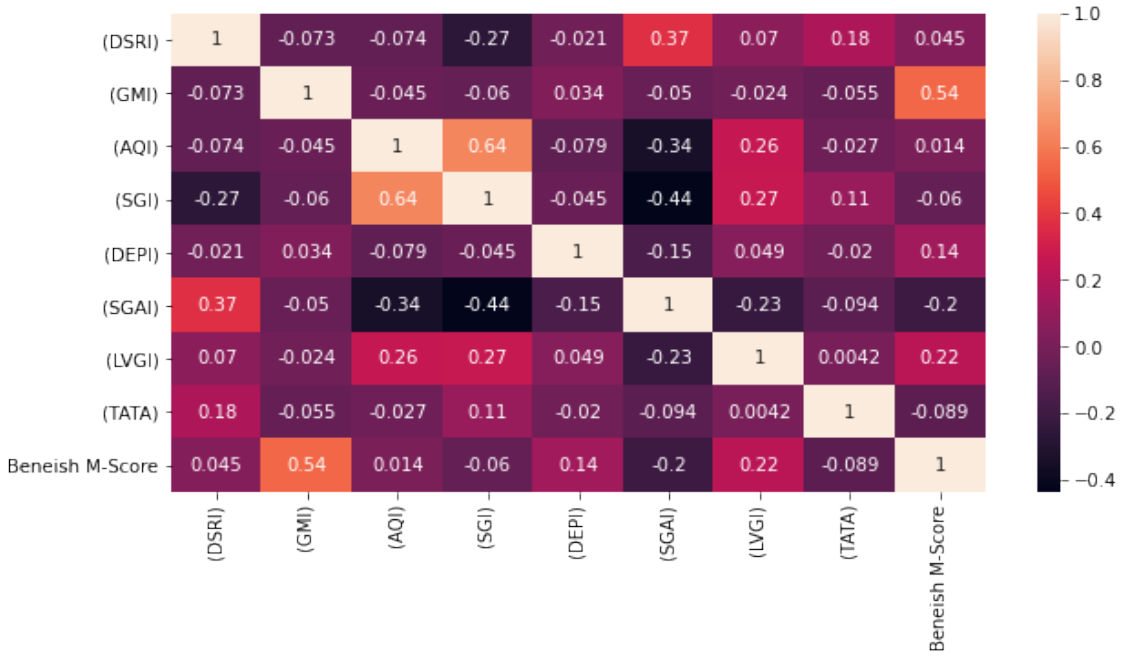

Figure 3. Study data Heatmap

The Beneish M-Score gives a unified assessment of earnings manipulation risk using these characteristics. The M-Score correlations of each variable in the heatmap are fascinating. A significant positive association between a variable and the M-Score signals that financial fraud is more likely as that variable exhibits evidence of manipulation.

Each independent indicator helps identify financial misbehavior, but the Beneish M-Score, generated from these factors, gives a complete picture of a firm's profit manipulation risk. Researchers may use the heatmap to examine the interactions and correlations between these factors to understand possibly better fraudulent organizations' behaviors. The heatmap shows the complex correlations between Beneish M-Score factors. Strong correlations, both positive and negative, predict financial malfeasance, whereas mixed signals or weak correlations may reflect legitimate corporate operations or more subtle manipulation.

OLS Regression Analysis

The research used an Ordinary Least Squares (OLS) regression model to investigate the factors influencing the Beneish M-Score. The model used the least squares approach to determine the relationship between variables, as shown in table 3. The analysis revealed that roughly 41,4 % of the variation in the Beneish M-Score could be explained by the model, as demonstrated by an R2 value of 0,414. The corrected R2, which incorporates the number of predictors in the model, was marginally lower at 0,386. This implies that the chosen independent variables provide a pretty adequate fit to the data but are not comprehensive.

The statistical analysis revealed that the model's overall significance was supported by an F-statistic value of 14,75 and a p-value (Prob (F-statistic)) of 3,19e-16. These results indicate that the predictors included in the model are statistically significant in their ability to explain the changes seen in the Beneish M-Score. The p-value's significance is very low, which highlights the strength of the regression model in accurately capturing the fundamental connections.

Furthermore, the model's log-likelihood, which indicates its level of conformity to the observed data, is documented as -395,48. Two often used metrics for comparing models, namely the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC), correspondingly exhibit values of 809,0 and 837,5. Researchers generally prefer lower values of AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion) since they indicate a more favorable model fit to the observed data.

Table 3. OLS Regression Analysis on Beneish M-Score

Dep. Variable	**Beneish M-Score**
Model	OLS
Method	Least Squares
R-squared	0,414
Adj. R-squared	0,386
F-statistic	14,75
Prob (F-statistic)	3,19e-16
Log- Likelihood	-395,48
AIC	809,0
BIC	837,5

The OLS regression analysis, as shown in table 4, yielded an estimated constant term, or intercept, of -0,5456. However, this estimate is not statistically significant, as demonstrated by its p-value of 0,722. The Days Sales in Receivables Index (DSRI) has a positive coefficient of 0,8541. The DSRI variable has statistical significance at the 5 % level, as shown by a p-value of 0,045 and a standard error of 0,422. This suggests a meaningful association between DSRI and the Beneish M-Score. The Gross Margin Index (GMI) has a coefficient value of 0,1643. The p-value of 0,000, which is remarkably low, underscores the high statistical significance of the relationship between GMI and the outcome variable, suggesting a robust and reliable association.

The Asset Quality Index (AQI), characterized by a coefficient of 0,0624 and a p-value of 0,959, does not exhibit statistical significance as a predictor in this model. Based on the available statistical evidence, its impact cannot be distinguished from zero. The Sales Growth Index (SGI) has a negative coefficient of -0,4715. Despite the apparent relevance of its association with the Beneish M-Score, the p-value of 0,104 indicates that it does not reach statistical significance according to traditional norms. The Depreciation Index (DEPI), with a coefficient of 0,1528 and a p-value of 0,246, does not demonstrate statistical significance as a predictor in the given situation.

The Sales and General Administrative Expenses Index (SGAI) has a significant negative coefficient of -1,7084. The statistical significance of the variable is confirmed by its low p-value of 0,001, indicating a robust inverse association with the Beneish M-Score. The Leverage Index (LVGI) has a significant positive coefficient of 5,5528, indicating its considerable statistical significance inside the model, as shown by its low p-value of 0,002. Finally, the variable TATA (Total Accruals to Total Assets), with a coefficient of -0,8893, does not exhibit statistical significance at conventional levels, as indicated by its p-value of 0,148.

Ordinary least squares (OLS) regression analysis provides insight into the diverse relationships between the independent variables and the Beneish M-Score. Several predictors, including GMI, SGAI, and LVGI, are firmly and consistently associated with the outcome variable. In contrast, the correlations between AQI, SGI, and the outcome variable seem less transparent and more uncertain.

Table 4. Coefficients in the OLS regression analysis

	coef	std err	t	P>|t|	[0,025	0,975]
const	-0,5456	1,530	-0,357	0,722	-3,566	2,475
(DSRI)	0,8541	0,422	2,023	0,045	0,021	1,688
(GMI)	0,1643	0,019	8,874	0,000	0,128	0,201
(AQI)	0,0624	1,206	0,052	0,959	-2,318	2,442
(SGI)	-0,4715	0,289	-1,634	0,104	-1,041	0,098
(DEPI)	0,1528	0,131	1,165	0,246	-0,106	0,412
(SGAI)	-1,7084	0,524	-3,263	0,001	-2,742	-0,675
(LVGI)	5,5528	1,739	3,193	0,002	2,120	8,986
(TATA)	-0,8893	0,612	-1,453	0,148	-2,098	0,319

FNN Analysis

Evaluating the deep learning model's performance often relies on critical measures such as training, validation loss, and accuracy. The training loss quantifies the discrepancy between the model's predictions and the actual values in the training data, while the validation loss quantifies the discrepancy on a distinct validation set. The training accuracy metric quantifies the proportion of accurately identified samples within the training set. In contrast, the validation accuracy metric quantifies the proportion of accurately classified samples within the validation set.

Figure 4 illustrates the temporal evolution of the model training process, as shown by the training and validation loss and accuracy metrics. Evaluating deep learning models in financial fraud detection relies on assessing key metrics, namely accuracy and loss. These metrics play a crucial role in determining the effectiveness of such models. Accuracy is a metric that quantifies the ratio of properly categorized cases to the total number of examples. The calculation involves determining the proportion of true positives and negatives about the overall number of occurrences. In contrast, loss serves as a measure of the discrepancy between the anticipated and observed values. The metric quantifies the degree to which the model can accurately represent the data. Models with lower loss values demonstrate superior performance.

In financial fraud detection, the optimal outcomes are characterized by high accuracy and minimal loss. This finding suggests that the model has a high level of accuracy in distinguishing between fraudulent and non-fraudulent cases, thereby reducing the occurrence of both false positives and false negatives. Overall, deep learning techniques, particularly neural networks, have shown considerable potential in detecting financial fraud ineffectively by obtaining notable accuracy levels and minimizing loss. These methodologies utilize the capabilities of deep neural networks to autonomously extract intricate patterns and characteristics from the data, thereby enhancing the precision of fraud detection models.

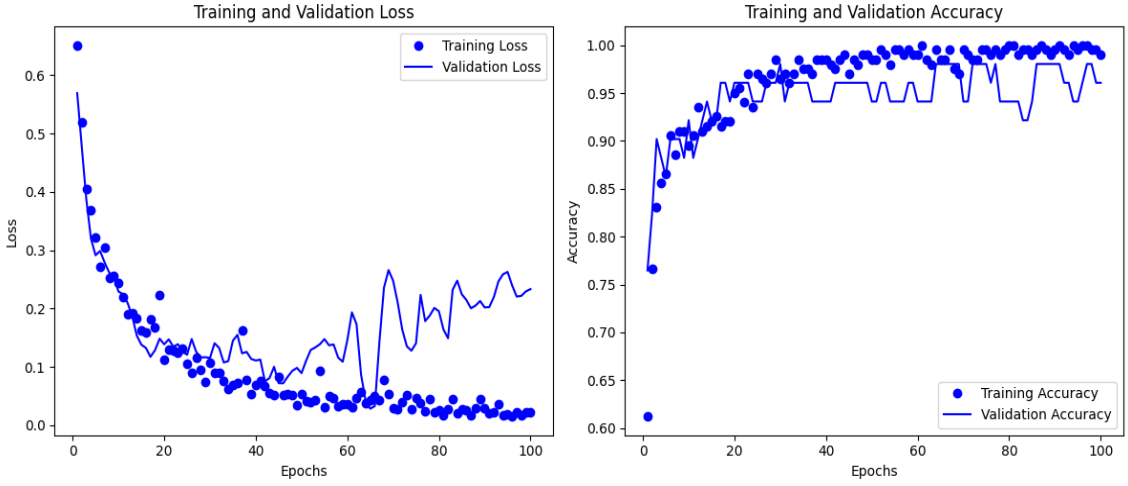

Figure 3. Model training progression

Model Performance

The deep learning model employed in this study was a Feedforward Neural Network (FNN), and its training and evaluation were conducted using data from the Amman Stock Exchange. The model's performance on the test dataset was evaluated after undergoing rigorous training and hyperparameter optimization, as shown in figure 5. The accuracy of 0,9844, as shown in figure 4, demonstrates the model's remarkable proficiency in consistently generating precise predictions. Significantly, in the context of fraud detection systems, the model's recall of 1,0 signifies that it did not fail to identify any instances of fraud in the test set. Moreover, the model demonstrates high accuracy in its fraudulent prediction, as evidenced by the precision value of 0,9697.

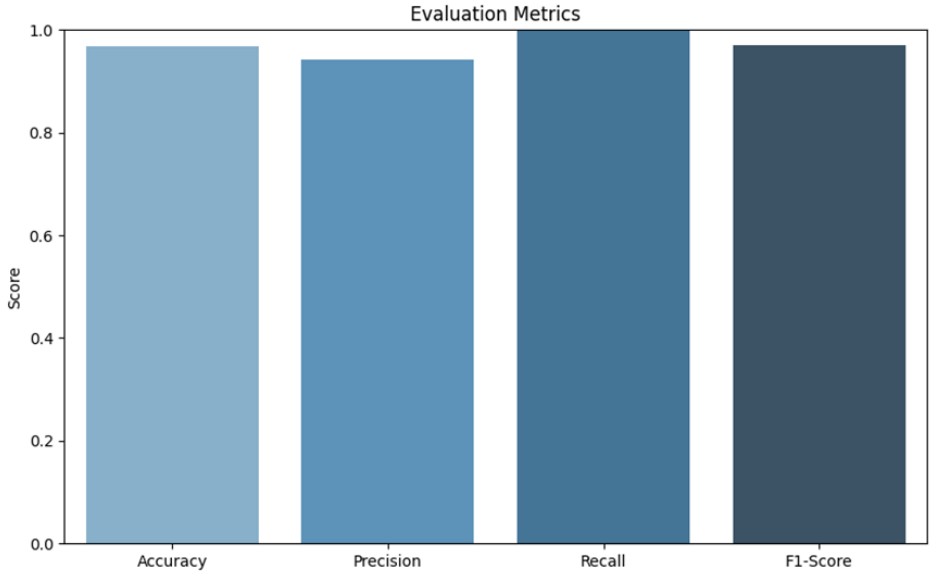

Figure 4. Evaluation Metrics illustration

Resampling & Feature Engineering Techniques

Datasets commonly used for fraud detection often exhibit a skewed class distribution, where the number of non-fraudulent instances significantly exceeds that of fraudulent instances. Due to this disparity, models can be constructed to prioritize the larger group, perhaps neglecting the smaller yet more significant number of individuals involved in fraudulent activities.

The issue was resolved through the utilization of resampling techniques. In this study, the researchers employed a technique known as the Synthetic Minority Over-sampling Strategy (SMOTE). The effectiveness of SMOTE stems from its ability to generate synthetic samples in the feature space. A novel instance is synthesized by considering the instance in question and its k nearest neighbors. The primary advantage of SMOTE is its ability to generate novel cases rather than merely replicating old ones, hence mitigating the risk of overfitting.

Prior to delving into complex algorithms, it is imperative to ascertain that the data is standardized on a uniform scale. The attributes underwent standardization, resulting in a mean of zero and a standard deviation of one. By employing this approach, the model mitigates the disproportionate influence of macro-level attributes and ensures equitable treatment of all aspects. The primary components take center stage in primary Component Analysis (PCA). High dimensional data might provide challenges in multicollinearity or while aiming to optimize computational efficiency. Principal component analysis (PCA) is a technique used for reducing the dimensionality of data by generating a set of new features that are orthogonal to each other. The components that capture the most significant variation in the data are derived as linear combinations of the original attributes. In the study, Principal Component Analysis (PCA) was employed as a technique to decrease the dimensionality of the feature set while ensuring the preservation of 95 % of the variation.

Enhancements in the performance of models can be predominantly attained through the utilization of resampling techniques and the implementation of feature engineering methodologies. By incorporating a more comprehensive and diverse dataset throughout the training process, resampling techniques such as SMOTE enhance the model's ability to detect patterns and make predictions for the underrepresented group accurately. In contrast, the model receives data input through feature engineering techniques, such as standardization and principal component analysis, thereby enhancing its computational efficiency and predictive accuracy. The core component of the fraud detection system is the deep learning model, which is complemented by the resampling and feature engineering approaches. These additional procedures play a crucial role in optimizing the system's performance.

Feature Importance

A more detailed explanation of the significance of each component in identifying fraudulent operations is provided in figure 5. Variables with higher weight magnitudes, such as LVGI (Leverage Index in Data) and AQI (Asset Quality Index), demonstrate their significance in detecting fraudulent activities. Forensic accountants and auditors could utilize this information to direct their investigation. The findings are promising, indicating that deep learning, specifically feedforward neural networks, offers significant potential in combating financial fraud. The model's high recall ensures that fraudulent activities are effectively detected, while its precision guarantees that actual transactions are accurately identified, minimizing any potential disturbances to operational processes.

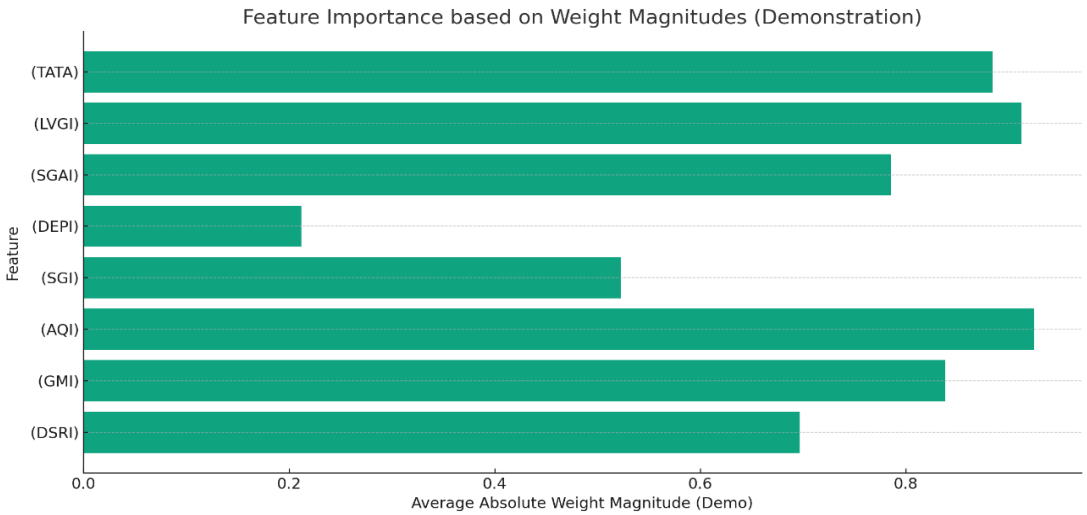

Figure 5. Feature Importance

Nevertheless, there exist limitations on the utilization of any given model. The performance of a given model may vary depending on the dataset utilized or the specific characteristics of the real-world scenario. To enhance the ability to detect sequential patterns in financial data, future studies should explore the integration of advanced neural network architectures such as recurrent neural networks (RNNs) or extended long short-term memory networks (LSTMs). The importance of domain knowledge cannot be overstated, even though the model exhibits strong performance. The integration of domain knowledge with machine learning techniques might enhance the production of models that are both more trustworthy and comprehensible.

Confusion Matrix Analysis

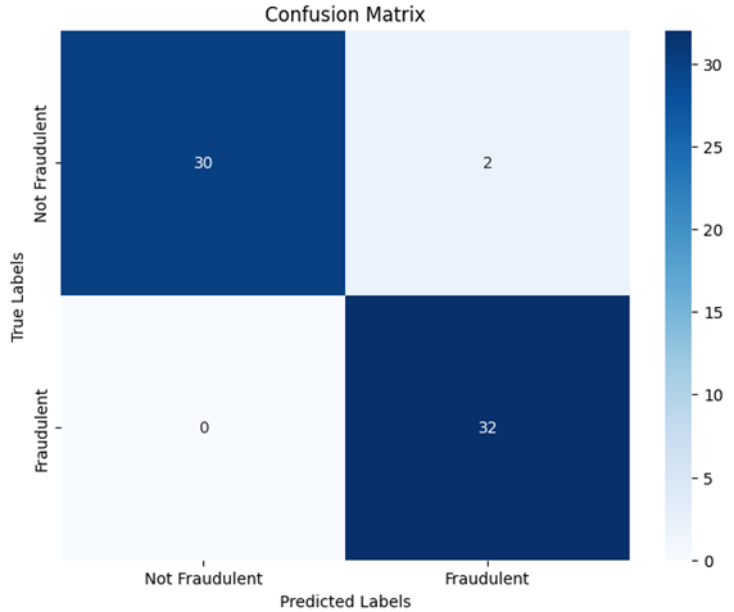

Figure 6. Confusion Matrix Result

The predictive model's ability to distinguish between genuine and fraudulent financial statements was evaluated using the confusion matrix, a fundamental method for assessing the performance of a classification model. The accuracy of the model's predictions is demonstrated in figure 6. The model accurately detected thirty instances, called True Negatives (TN), where the financial statements were correctly classified as legitimate. Furthermore, 32 instances were accurately classified as fraudulent, constituting the True Positives (TP). Two instances were identified where valid financial statements were erroneously classified as fraudulent, resulting in false positives (FP). The model accurately detected all instances of potentially fraudulent financial statements in the sample, resulting in a lack of False Negatives (FN).

The model exhibited remarkable performance, attaining a substantial proportion of accurate predictions. The performance of our model in detecting instances of financial fraud was successful, as it did not produce any false negatives. However, it is worth noting that there is still potential for further improvement, as two valid instances were incorrectly identified as fraudulent (false positives). The data above provides evidence of the model's effectiveness in identifying fraudulent financial statements. Nevertheless, due to the potential occurrence of inaccurate positive results, additional verification or human assessment is necessary before drawing conclusions based on the model's forecasts. Enhanced outcomes can be achieved through model refinement, the incorporation of additional information, or the adoption of a more sophisticated algorithm.

Receiver Operating Characteristic (ROC) Curve Analysis

The evaluation of classification models, particularly in scenarios with an imbalanced class distribution, necessitates the consideration of the Receiver Operating Characteristic (ROC) curve as a vital statistical measure. This approach provides a valuable means of assessing the model's performance across various thresholds in effectively discerning positive and negative classifications. The Receiver Operating Characteristic (ROC) curve illustrates the relationship between the Sensitivity (True Positive Rate) and the False Positive Rate (1 - Specificity) across different thresholds. Statisticians employ the area under the receiver operating characteristic (ROC) curve, commonly referred to as the area under the curve (AUC), as a quantitative measure to assess the efficacy of a model in distinguishing between positive and negative classes. In our investigation, the area under the model's receiver operating characteristic curve (AUC) was determined to be 0,97. An ideal classifier would possess an Area Under the Curve (AUC) value of 1,0, while a score of 0,5 would signify that the model's performance is equivalent to random chance. According to figure 7, the algorithm demonstrates a high level of accuracy in discerning between genuine and counterfeit financial statements, as indicated by an Area Under the Curve (AUC) value of 0,97.

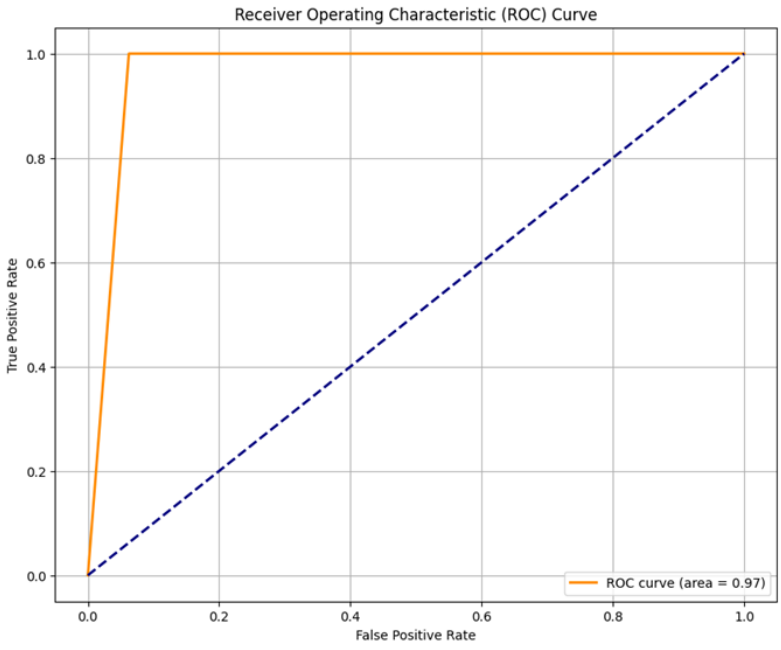

Figure 7. Data ROC Curve

A large area under the receiver operating characteristic curve (AUC) also suggests that the model exhibits sufficient sensitivity and specificity across several criteria. Given the significant ramifications associated with overlooking genuine instances of fraud or incorrectly classifying valid transactions as fraudulent, this holds particular significance within financial fraud detection. While the AUC serves as a valuable summary statistic, it is essential to consider it in conjunction with other metrics, such as precision, recall, and the F1 Score, to

comprehensively assess the model's capabilities and limitations. Further investigations, such as determining an optimal decision threshold that effectively balances the trade-offs between false positives and false negatives, may be undertaken to enhance the practicality of this model in real-life situations, taking into account the particular demands and acceptable margins of error of the application.

DISCUSSION

The primary objective of our research was to utilize a deep learning model, with a particular focus on firms listed on the Amman Stock Exchange, to identify and predict occurrences of financial fraud. A comprehensive examination was undertaken on various metrics, including DSRI, GMI, AQI, SGI, DEPI, SGAI, LVGI, and TATA. A thorough examination was conducted on a complete collection of 176 data points for each metric, identifying significant statistical measures such as the mean and standard deviation. For example, the Gini Mean Index (GMI) statistic exhibited a significant level of variability, as seen by a standard deviation of 9,681902. Statistical insights are crucial for comprehending fundamental data trends and identifying potential outliers.

Heatmaps were used in our study to do multivariate analysis, enabling us to understand the interrelationships among various factors within our dataset. The heatmap used color-coded cells to represent the magnitude of correlations between variables, wherein deeper hues denoted higher levels of linkage. For example, a notable positive correlation observed in the Debt Service Ratio Index (DSRI) suggests a clear association between the duration a firm requires to recover its payments and other factors that serve as indicators of financial fraud. Similarly, the correlations established by the AQI have yielded valuable insights into the quality of assets concerning several different variables. The Beneish M-Score was used to detect earnings manipulation within the organizations under analysis. This score incorporates eight criteria to achieve its objective. The heatmap clarified the complex interconnections among these factors, providing a holistic perspective on probable patterns of financial malfeasance.

The methodological approach used in our study included an Ordinary Least Squares (OLS) regression analysis to determine the elements that influence the Beneish M-Score. The results of the regression study suggest that the model explains roughly 41,4 % of the variability seen in the Beneish M-Score. In addition, the statistical significance of the predictors was validated using an F-statistic value of 14,75 and a p-value of 3,19e-16. The regression model offered a detailed analysis of the associations between the independent variables and the Beneish M-Score. An illustration of this may be seen in the statistically significant association between the GMI variable and the outcome variable.

In order to examine the deep learning component, we used a Feedforward Neural Network (FNN) to analyze our dataset. The performance indicators of the model, including training and validation loss and accuracy, provide valuable data about its efficacy. Metrics such as accuracy and loss were crucial in assessing the model's effectiveness in identifying instances of financial fraud. For instance, the model demonstrated a notable accuracy of 0,9844, highlighting its adeptness in correctly forecasting financial irregularities.

In addition, we addressed the widespread issue of class imbalance in fraud detection datasets using the Synthetic Minority Over-sampling Strategy (SMOTE). The use of this method had a pivotal role in augmenting the model's capacity to identify patterns within the underrepresented class. In addition to using resampling, feature engineering approaches such as standardization and principal component analysis were used to enhance the model's computational efficiency and forecast accuracy. The analysis of the relative value of each feature enhanced the relevance of the deep learning model. Variables such as LVGI (Local Variance of Global Intensity) and AQI (Air Quality Index), which exhibited significant weight magnitudes, played a crucial role in identifying fraudulent operations.

The model's performance was further examined via a confusion matrix, which provided insights into its capacity to differentiate between authentic and deceptive financial data. The model successfully classified 30 occurrences as accurate and 32 instances as fake, showcasing its proficiency in differentiating between the two categories. Nevertheless, the model erroneously categorized two legitimate cases as fraudulent, suggesting the need for further enhancements.

An additional pivotal component of our investigation included using Receiver Operating Characteristic (ROC) curve analysis. The receiver operating characteristic (ROC) curve illustrates the association between our model's sensitivity and false positive rate. One noteworthy accomplishment was attaining a model's Area Under the Curve (AUC) value of 0,97, indicating its efficacy in discerning authentic and deceptive financial statements. The integration of this assessment, in conjunction with other indicators, provides a holistic comprehension of the model's capabilities and areas that may need further improvement. In summary, our model has considerable potential in identifying financial fraud. However, more improvements and incorporation of domain-specific information may boost its precision and practicality.

Significant progress has been made in financial fraud detection in recent years as researchers have applied diverse techniques to address this complex issue. Table 4 presents a comprehensive comparison analysis of our work on past research efforts in this field. The present research examined the topic of financial fraud

detection by using Deep Learning (FNN) and OLS Regression techniques on a sample of firms registered on the Amman Stock Exchange. The findings derived from our research highlight the considerable efficacy of both deep learning and ordinary least squares (OLS) regression in accurately identifying instances of financial fraud. This aligns with the growing tendency to use sophisticated analytical methods for financial monitoring.

In a recent study by Okour et al.[33], an in-depth analysis was conducted to explore the comparative landscape and investigate the impact of financial risk disclosure on stock liquidity within the industrial businesses of Jordan. The study highlights the significant and essential role financial risk disclosure plays in this context. In their study, Xiuguo et al.[34] employed deep learning methodologies and highlighted the need for further investigation in this domain. In January 2021, a study established a standard by showcasing the enhanced precision of the Long Short-Term Memory (LSTM) model compared to the Recurrent Neural Network (RNN) in developing financial statement fraud detection models. Previous research conducted by Chen[35] and Yao et al.[36] has emphasized the effectiveness of various models, such as Support Vector Machines (SVM), in the realm of financial fraud prediction and detection. Sharma et al.[37] achieved significant advancements in the field of credit card fraud detection by focusing on enhancing accuracy in the presence of large datasets. Similarly, Fang et al.[38] conducted a comprehensive investigation into the complexities of developing deep-learning models to detect fraudulent online loan activities. In conclusion, the range of studies represents the diverse methodologies and their contributions to strengthening the financial sector against fraud.[39,40]

Table 4. Comparative Analysis with Previous Studies

Study	Methods Used	Primary Focus	Findings/Highlights
Our Study	Deep Learning (FNN), OLS Regression	Financial fraud detection in companies listed on the Amman Stock Exchange.	It demonstrated the potential of deep learning and ordinary least squares (OLS) regression in financial fraud detection.
Okour et al.[34]	Multiple Linear Regression, Interactive Regression Analysis	The Impact of Financial Risk Disclosure on Stock Liquidity for Jordan's Industrial Companies.	Highlighted the significance of financial risk disclosure on stock liquidity.
Xiuguo et al.[35]	Deep Learning (CNN, RNN)	Predicting financial risk probability of companies.	Emphasized the need for more research in using deep learning for financial fraud detection.
Jan[25]	Deep Learning (RNN, LSTM)	Constructing financial statement fraud detection models.	The LSTM model showed superior performance in terms of accuracy compared to the RNN model.
Chen [2022]	Support Vector Machine, Logistic Regression	Predicting financial fraud in publicly traded Chinese firms.	Showcased the effectiveness of various models in predicting financial fraud.
Yao et al.[36]	Stepwise Regression, PCA, SVM	Reducing variable dimensionality and detecting fraudulent financial statements.	Found SVM to have the highest accuracy among all conditions in detecting fraudulent financial statements.
Sharma et al.[38]	Deep Learning based on Auto-encoder	Credit card fraud detection.	Aimed to enhance detection accuracy while handling large data volumes.
Fang et al.[39]	Deep Learning	Anti-fraud model for Internet loans.	Discussed the directions and challenges in building deep learning anti-fraud models for internet loans.
Jan[17]	Not specified	Financial statement fraud detection for sustainable development of financial markets in Taiwan.	We discussed the significance of effective financial statement fraud detection models for sustainable development.

The focus of our analysis was focused on firms that are listed on the Amman Stock Exchange. It is important to note that the findings may not comprehensively represent the broader range of financial dynamics in other areas or kinds of markets.[41,42] Although the algorithms and data pretreatment techniques used are sophisticated, they may still fail to detect complex fraud patterns or be susceptible to the impact of noisy data.[43,44] Furthermore, it is essential to regularly update models to maintain their effectiveness due to the fast development of financial fraud strategies. In order to enhance future research, it would be very advantageous[45] to include a broader range of datasets sourced from varied financial markets.[46,47] Additionally, it would be beneficial to enhance the model by including real-time feedback mechanisms and investigating the integration of more sophisticated neural network designs, such as transformer-based models.[48] By engaging in partnerships with financial professionals, integrating domain knowledge and machine learning methods may enhance the precision and dependability of fraud detection models.

CONCLUSIONS

This study conducts a comparative analysis of the effectiveness of Ordinary Least Squares (OLS) regression and Feedforward Neural Network (FNN) models in identifying financial anomalies and predicting financial fraud. The results indicate that FNN demonstrates excellent performance compared to OLS regression in identifying financial anomalies and exhibits a higher level of effectiveness in detecting financial fraud among ASE-listed firms. The capacity of the FNN to efficiently manage substantial amounts of data within financial domains is of utmost importance, making it a viable option for identifying fraudulent activities. The ordinary least squares (OLS) regression analysis reveals robust and persistent correlations between specific predictors (such as GMI, SGAI, and LVGI) and the dependent variable, suggesting an increased probability of financial fraud.

On the other hand, the associations between alternative predictors, such as the Air Quality Index (AQI) and the Social Good Index (SGI), and the dependent variable exhibit fewer distinct patterns and are characterized by greater uncertainty. In general, the comparison between ordinary least squares (OLS) regression and feedforward neural networks (FNN) highlights the advantages and limitations of each approach. In contrast, FNN exhibits higher efficacy in detecting financial anomalies and identifying instances of financial fraud.

BIBLIOGRAPHIC REFERENCES

1. Afriyie JK, Tawiah K, Pels WA, Addai-Henne S, Dwamena HA, Owiredu EO, et al. A supervised machine learning algorithm for detecting and predicting fraud in credit card transactions. Decision Analytics Journal. 2023;6:100163. https://doi.org/10.1016/j.dajour.2023.100163.

2. Akram T, Ramakrishnan SA, Naveed M. Prevalence of money laundering and terrorism financing through the stock market: a comprehensive conceptual review paper. Journal of Money Laundering Control. 2022. https://doi.org/10.1108/jmlc-06-2022-0094.

3. Al Omari R, Alkhawaldeh RS, Jaber JJ. Artificial Neural Network for classifying financial performance in the Jordanian Insurance Sector. Economies. 2023;11(4):106. https://doi.org/10.3390/economies11040106.

4. Al-Hashedi KG, Magalingam P. Financial fraud detection applying data mining techniques: a comprehensive review from 2009 to 2019. Computer Science Review. 2021;40:100402. https://doi.org/10.1016/j.cosrev.2021.100402.

5. Aljarrah O, Groof RD. Extracting firm performance using a financial statement: a case study in Jordan. International Journal of Decision Sciences, Risk and Management. 2017;7(4):299. https://doi.org/10.1504/ijdsrm.2017.093826.

6. Ashtiani MN, Raahemi B. Intelligent fraud detection in financial statements using machine learning and Data Mining: a Systematic Literature Review. IEEE Access. 2022;10:72504-25. https://doi.org/10.1109/access.2021.3096799.

7. Błaszczyński J, de Almeida Filho AT, Matuszyk A, Szeląg M, Słowiński R. Auto loan fraud detection using dominance-based rough set approach versus machine learning methods. Expert Systems with Applications. 2021;163:113740. https://doi.org/10.1016/j.eswa.2020.113740.

8. Boulieris P, Pavlopoulos J, Xenos A, Vassalos V. Fraud detection with natural language processing. Machine Learning. 2023. https://doi.org/10.1007/s10994-023-06354-5.

9. Choi D, Lee K. An artificial intelligence approach to financial fraud detection under IOT environment: a survey and implementation. Security and Communication Networks. 2018;2018:1-15. https://doi.org/10.1155/2018/5483472.

10. Cohen G. Algorithmic trading and financial forecasting using Advanced Artificial Intelligence Methodologies. Mathematics. 2022;10(18):3302. https://doi.org/10.3390/math10183302.

11. Zhou H, Sun G, Fu S, Wang L, Hu J, Gao Y. Internet financial fraud detection based on a distributed big data approach with node2vec. IEEE Access. 2021;9:43378-86. https://doi.org/10.1109/access.2021.3062467.

12. Du P, Shu H. Exploration of financial market credit scoring and risk management and prediction using Deep Learning and Bionic algorithm. Journal of Global Information Management. 2021;30(9):1-29. https://doi.org/10.4018/jgim.293286.

13. Ghazi M Qasaimeh, Hussam Eddin Jaradeh. The impact of artificial intelligence on the effective applying of cyber governance in Jordanian commercial banks. International Journal of Technology, Innovation and Management (IJTIM). 2022;2(1). https://doi.org/10.54489/ijtim.v2i1.61.

14. Glancy FH, Yadav SB. A computational model for Financial Reporting Fraud Detection. Decision Support Systems. 2011;50(3):595-601. https://doi.org/10.1016/j.dss.2010.08.010.

15. Hilal W, Gadsden SA, Yawney J. Financial fraud: a review of anomaly detection techniques and recent advances. Expert Systems with Applications. 2022;193:116429. https://doi.org/10.1016/j.eswa.2021.116429.

16. Jan C. An effective financial statements fraud detection model for the sustainable development of financial markets: evidence from taiwan. Sustainability. 2018;10(2):513. https://doi.org/10.3390/su10020513.

17. Karpoff JM. The future of financial fraud. Journal of Corporate Finance. 2021;66:101694. https://doi.org/10.1016/j.jcorpfin.2020.101694.

18. Lokanan ME, Sharma K. Fraud prediction using machine learning: the case of investment advisors in Canada. Machine Learning with Applications. 2022;8:100269. https://doi.org/10.1016/j.mlwa.2022.100269.

19. Merćep A, Mrčela L, Birov M, Kostanjčar Z. Deep Neural Networks for behavioral credit rating. Entropy. 2020;23(1):27. https://doi.org/10.3390/e23010027.

20. Mohammed RA, Wong KW, Shiratuddin MF, Wang X. Scalable machine learning techniques for highly imbalanced credit card Fraud Detection: a Comparative Study. Lecture Notes in Computer Science. 2018:237-46. https://doi.org/10.1007/978-3-319-97310-4_27.

21. Mehbodniya A, Alam I, Pande S, Neware R, Rane KP, Shabaz M, et al. Financial fraud detection in healthcare using machine learning and deep learning techniques. Security and Communication Networks. 2021;2021:1-8.

22. Alghofaili Y, Albattah A, Rassam MA. A financial fraud detection model based on the LSTM Deep Learning Technique. Journal of Applied Security Research. 2020;15(4):498-516. https://doi.org/10.1080/19361610.2020.1815491.

23. Craja P, Kim A, Lessmann S. Deep learning for detecting financial statement fraud. Decision Support Systems. 2020;139:113421.

24. Jan CL. Detection of financial statement fraud using Deep Learning for sustainable development of capital markets under information asymmetry. Sustainability. 2021;13(17):9879. https://doi.org/10.3390/su13179879.

25. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. In: Advances in Neural Information Processing Systems 30. Curran Associates; 2017. p. 4765-74.

26. Ma T, Qian S, Cao J, Xue G, Yu J, Zhu Y, et al. An unsupervised incremental virtual learning method for financial fraud detection. In: 2019 IEEE/ACS 16th International Conference on Computer Systems and Applications (AICCSA). IEEE; 2019. p. 1-6.

27. Marino DL, Wickramasinghe CS, Rieger C, Manic M. Data-driven stochastic anomaly detection on smart-grid communications using mixture Poisson distributions. In: IECON 2019 - 45th Annual Conference of the IEEE Industrial Electronics Society. IEEE; 2019. p. 5855-61.

28. Munappy A, Bosch J, Olsson HH, Arpteg A, Brinne B. Data management challenges for deep learning. In: 2019 45th Euromicro Conference on Software Engineering and Advanced Applications (SEAA). IEEE; 2019. p. 140-7.

29. Liu R, Mai F, Shan Z, Wu Y. Predicting shareholder litigation on insider trading from financial text: an interpretable deep learning approach. Information & Management. 2020;57(8):103387.

30. Azhan M, Meraj S. Credit card fraud detection using machine learning and deep learning techniques. In: 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS). IEEE; 2020. p. 514-8.

31. Chen JIZ, Lai KL. Deep convolution neural network model for credit-card fraud detection and alert. Journal of Artificial Intelligence and Capsule Networks. 2021;3(2):101-12.

32. Kute DV, Pradhan B, Shukla N, Alamri A. Deep learning and explainable artificial intelligence techniques applied for detecting money laundering-a critical review. IEEE access. 2021;9:82300-17.

33. Okour S, Almasaeid I. The moderating role of the governance index in the relationship between financial risk disclosure and stock liquidity: evidence from jordan. International Journal of Academic Research in Accounting Finance and Management Sciences. 2022;12(4). https://doi.org/10.6007/ijarafms/v12-i4/15735.

34. Xiuguo W, Shengyong D. An analysis on financial statement fraud detection for chinese listed companies using deep learning. Ieee Access. 2022;10:22516-32. https://doi.org/10.1109/access.2022.3153478.

35. Chen D. Predicting accounting fraud in publicly traded chinese firms via a pca-rf method. 2022. https://doi.org/10.2991/978-94-6463-108-1_82.

36. Yao J, Pan Y, Yang S, Chen Y, Li Y. Detecting fraudulent financial statements for the sustainable development of the socio-economy in china: a multi-analytic approach. Sustainability. 2019;11(6):1579. https://doi.org/10.3390/su11061579.

37. Sharma M, Raj B, Ramamurthy B, Bhaskar R. Credit card fraud detection using deep learning based on auto-encoder. Itm Web of Conferences. 2022;50:01001. https://doi.org/10.1051/itmconf/20225001001.

38. Fang W, Li X, Zhou P, Yan J, Jiang D, Zhou T. Deep learning anti-fraud model for internet loan: Where we are going. IEEE Access. 2021;9:9777-84. https://doi.org/10.1109/access.2021.3051079.

39. Mutemi A, Bacao F. A numeric-based machine learning design for detecting organized retail fraud in digital marketplaces. Scientific Reports. 2023;13(1). https://doi.org/10.1038/s41598-023-38304-5.

40. Nguyen DK, Sermpinis G, Stasinakis C. Big Data, Artificial Intelligence and machine learning: a transformative symbiosis in favour of financial technology. European Financial Management. 2022;29(2):517-48. https://doi.org/10.1111/eufm.12365.

41. Nguyen TT, Tahir H, Abdelrazek M, Babar A. Deep learning methods for credit card fraud detection. arXiv preprint arXiv:2012.03754. 2020.

42. Ravisankar P, Ravi V, Raghava Rao G, Bose I. Detection of financial statement fraud and feature selection using data mining techniques. Decision Support Systems. 2011;50(2):491-500. https://doi.org/10.1016/j.dss.2010.11.006.

43. Reurink A. Financial fraud: a literature review. Journal of Economic Surveys. 2018;32(5):1292-325. https://doi.org/10.1111/joes.12294.

44. SADGALI I, SAEL N, BENABBOU F. Performance of machine learning techniques in the detection of financial frauds. Procedia Computer Science. 2019;148:45-54. https://doi.org/10.1016/j.procs.2019.01.007.

45. Singh V, Chen SS, Singhania M, Nanavati B, Gupta A. How are reinforcement learning and deep learning algorithms used for big data based decision making in financial industries-A review and research agenda. International Journal of Information Management Data Insights. 2022;2(2):100094.

46. Skousen CJ, Smith KR, Wright CJ. Detecting and predicting financial statement fraud: The effectiveness of the fraud triangle and SAS no. 99. Advances in Financial Economics. 2009:53-81. https://doi.org/10.1108/s1569-3732(2009)0000013005.

47. West J, Bhattacharya M. Intelligent Financial Fraud Detection: A comprehensive review. Computers & Security. 2016;57:47-66. https://doi.org/10.1016/j.cose.2015.09.005.

48. Zhao Z, Bai T. Financial fraud detection and prediction in listed companies using smote and machine learning algorithms. Entropy. 2022;24(8):1157. https://doi.org/10.3390/e24081157.

FINANCING

None.

CONFLICT OF INTEREST

Authors declare that there is no conflict of interest.

AUTHORSHIP CONTRIBUTION

Conceptualization: Mohammad Haroun Sharairi.
Data curation: Mohammad Haroun Sharairi.
Formal analysis: Mohammad Haroun Sharairi.
Drafting - original draft: Mohammad Haroun Sharairi.
Writing - proofreading and editing: Mohammad Haroun Sharairi.