International Journal for Research in Business, Management, and Accounting

Audit Risk Prediction and Fraudulent Firm Classification Using Business Analytics

Published By:
Green Publications Services Private LTD.

ISSN(Online)**: 2455-6114**
DOI: 10.53555/bma.v11i2.2513
Volume 11 Issue 02 June 2025

Dr. Rakesh Kumar Sharma[1]*
Dr. Ananya Verma[2]

[1]Department of Commerce, Meerut College, Meerut, Uttar Pradesh, India
Email: rakesh.sharma.mrt@gmail.com

[2]Department of Management Studies, Dayalbagh Educational Institute, Agra, Uttar Pradesh, India

Corresponding Author:
Dr. Rakesh Kumar Sharma
Email: rakesh.sharma.mrt@gmail.com

Article History:
Article Type: Research

Received Date: 23/03/2025 **Revised Date: 25/04/2025**
Acceptance Date: 20/05/2025 **Published Date: 28/06/2025**

ABSTRACT

Audit risk prediction has become increasingly important as auditors face growing volumes of complex financial and operational data. Business analytics offers a structured approach for identifying high-risk firms and supporting fraudulent firm classification through measurable audit-risk indicators. This study investigates the role of audit-risk indicators in distinguishing risky firms from non-risky firms. A quantitative research design was adopted, using descriptive statistics, comparative analysis, correlation analysis, and classification-based assessment. The analysis considered major audit-risk dimensions, including inherent risk, monetary exposure, total audit parameter exposure, component risk indicators, control risk, and overall audit risk. The results reveal that risky firms recorded substantially higher values for inherent risk, monetary exposure, total audit parameter exposure, and overall audit risk compared with non-risky firms. Correlation findings indicate that composite audit score, monetary exposure score, audit parameter scores, control risk, district-level loss, and inherent risk were meaningfully associated with final risk classification. The classification assessment further confirmed that audit-risk indicators can effectively separate risky firms from non-risky firms. The findings highlight the practical value of business analytics in strengthening audit planning, improving risk-based screening, and supporting data-driven decision-making among auditors, managers, regulators, and governance stakeholders. The study contributes to accounting and auditing literature by demonstrating how structured audit-risk indicators can support fraudulent firm classification and audit risk assessment.

KEYWORDS: Audit risk prediction; fraudulent firm classification; business analytics; audit data analytics; risk assessment; fraud detection

1. **Introduction**

Auditing has progressively evolved from manual, judgment-based procedures toward technology-supported methods for risk analysis. In the modern business setting, firms generate large volumes of financial and operational data related to their operations, making it increasingly difficult for auditors to use only traditional sampling and experience in performing audit work. This has seen audit risk prediction become an important area in accounting and business analysis that enables auditors to determine which firms are risky before carrying out the entire audit process. The introduction of artificial intelligence in auditing is also proving useful in ensuring audit efficiency, effectiveness, and detecting suspicious financial record activities (Fedyk et al., 2022).

In relation to business analytics and auditing, the issue is important because the occurrence of audit risk does not usually come from one reason alone. On the contrary, audit risk normally emerges because of a combination of financial risk, poor internal controls, past indicators of risk, industry-specific factors, and the collection of audit-related elements. There has been a growing emphasis on the idea that both artificial intelligence and analytical models could act as tools which could assist in changing auditing into an ongoing task rather than a reactionary one (Leocadio et al., 2024). When it comes to auditing from an external auditor's point of view, artificial intelligence would be able to improve audit quality because of better data processing abilities and detection of risk factors (Noordin et al., 2022).

In addition, the application of big data and data analytics technology has brought about changes to the expectations of auditors. The auditor is now expected to work with structured and unstructured data, spot anomalies, analyze trends, and make decisions based on evidence obtained. The effect of big data analytics technology on auditing as a profession is especially relevant for developing countries because auditing can become modernized and improve the proficiency and effectiveness of auditors (Abdelwahed et al., 2025). Similarly, internal audits would be positively affected by the use of analytics technology (Alvarez-Foronda et al., 2023).

Firm classification related to fraud is highly linked to the discovery of financial statement fraud. According to previous data mining researches, combining analysis techniques would assist in detecting financial statement fraud based on hidden patterns in financial measures (Chen, 2016). Recent researches show a shift of the methodology used in financial statement fraud detection towards machine learning and analysis-based techniques from traditional accounting analysis

(Beemamol, 2024). This shift highlights the need for more advanced methods to deal with financial data.

The parallel use of new digital technologies brings benefits and risks to audit firms. For instance, technologies like artificial intelligence, automation, and analytics can help in improving audit planning and risk assessment but demand additional skills, professional judgment, and interpretation capabilities (Vitali & Giuliani, 2024). From the perspective of comprehensive literature reviews on artificial intelligence in auditing, the audit industry is heading towards a stage when auditors have to combine their subject-matter knowledge with intelligent analysis tools (Binh, 2025). When focusing on the area of fraud detection and prevention, the role of data analytics has increased since complex fraud can be concealed in vast databases using conventional techniques (Gkegkas et al., 2025).

With this foundation, this study aims to determine the ability of audit risk indicators in distinguishing between risky and non-risky firms. This study uses audit risk indicators quantitatively to examine risk trends, contrast risky firms from non-risky firms, and evaluate the effectiveness of analytics-based classification in making audit decisions. The objectives of this study are:

- To examine the descriptive profile of major audit-risk indicators associated with firm-level risk classification.
- To compare risky and non-risky firms based on monetary exposure, inherent risk, control risk, and audit parameter indicators.
- To assess the usefulness of business analytics techniques for audit risk prediction and fraudulent firm classification.

2. Methodology

2.1 Research Design

The methodology employed in this study is quantitative in design. This research will attempt to predict audit risk and classify firms with fraudulent-risk potential through business analytics. Empirical research design has been chosen since the study aims at identifying audit risk factors that help in classifying risks and determining the usefulness of business analytics in aiding audit decision-making processes. The research will utilize a predictive analytic approach supplemented with descriptive and associative analyses. Descriptive analysis explains the general features of the dataset, while correlation and comparative analysis will be used to establish associations among audit factors and risk classification. The predictive classification process will then be used to establish whether audit risk factors can discriminate between fraudulent-risk and non-risk firms.

2.2 Data Source and Sample Description

This research uses an audit-risk dataset that is publicly available and contains both firm-level audit observations and risk-related variables (Siddhartha, 2019). In this dataset, the following variables have been observed: Sector Score, Audit Parameters, Money, District Loss, Historical Indicators, Risk Scores, Inherent Risk, Control Risk, Detection Risk, Audit Risk, and Classification of Risk. The dataset was chosen because it provides systematic quantitative variables that can be used in the prediction of audit risk and fraudulent-risk classification. In this dataset, the sample size has been divided into two categories: risky and non-risky. In this case, Risk will be the dependent variable because it represents the classification of audit risk. Audit parameters, money, historical indicators, risk scores, and audit risk will be the independent variables.

2.3 Data Preprocessing

Before any analysis, the dataset was preprocessed for dealing with missing values, duplicates, inconsistencies, and problems arising from formatting errors. Missing values were imputed using an appropriate technique to make sure that observations were not excluded unnecessarily from the analysis. Duplicate data observations were deleted to avoid having the same observation affect the statistical results of models adversely. The "location" variable was considered as categorical since it consisted of both numeric and non-numeric data. Names of variables were evaluated, and changes made if necessary. Such preprocessing ensured the quality and consistency of data analysis.

2.4 Analytical Framework

The analytical model was based on three main steps. In the first step, descriptive statistics were calculated to describe the distribution of audit risk indicators. The statistics that included the mean value, median, minimum value, maximum value, standard deviation, and frequency distribution were applied to describe the structure of data and the share of observations that belong to the category of risky and non-risky firms. In the second step, comparisons and correlation studies were conducted. The observations were divided into two groups depending on the result of the risk assessment. Then, the average values of audit indicators were compared for the risky group and the non-risky group. Correlation analysis was further applied to estimate the relationship between audit risk indicators and the outcome of the risk assessment, finding out which factors have a stronger influence on risky firms. Predictive classification was conducted in the third phase. Explanatory variables consisted of audit risk indicators to classify firms as risky or non-risky. These predicted classifications were then compared with actual classifications to determine the accuracy of the model in identifying firms classified as risky.

2.5 Model Development

This study has formulated the classification model with the use of audit-risk indicators as predictors and risk classification as the response variable. Since the response variable is dichotomous, it has been modeled in such a way that it enables classification into two categories: risky and non-risky. Input variables used in the modeling process include audit-related factors, monetary amount, risk indicators of components, historical indicators, risk of control, risk of detection, among others. Attention has been paid to developing a model whose output would be meaningful in relation to auditing activities. Variable selection has taken into account their theoretical relationship to audit risk and association with the output variable. The classification model has been created using the existing audit-risk indicators with an objective of verifying their ability to differentiate risky from non-risky cases.

2.6 Model Evaluation

The model was tested using the confusion matrix with the help of performance measures for classification models. The use of the confusion matrix helped compare the actual vs. predicted classifications and hence identify the observations which were correctly and incorrectly classified. This helped test the ability of the model to classify between risky and non-risky firms. Accuracy and precision, among other performance measures, were considered. Accuracy denoted the percentage of observations classified correctly, whereas precision signified the percentage of risky firm predictions that turned out to be correct. Recall measures how well the model can detect true risky firms, whereas specificity evaluates the capability of detecting true non-risky firms. F1-score is an appropriate balance between the precision and recall scores. These metrics have been chosen due to the need to carefully analyze both categories in audit risk detection.

2.7 Feature Relevance and Interpretation

The relevance of each feature was established based on its connection to the final risk classification variable. If there were a strong connection between a risk classification predictor and a final risk classification variable and there was a significant difference between risky and non-risky firms, then this predictor was considered to be more important and a good indicator of firm fraudulent classification. The study focused not only on predicting risk but also on identifying important audit indicators. Understanding the importance of these indicators is crucial from practice-related perspectives.

3. Results

3.1 Descriptive Statistics of Audit-Risk Indicators

The results from the descriptive analysis show high levels of differences in the levels of audit risk. There are indicators with mean values significantly greater than the median value, thus meaning that most of the firms had very low audit risks, while few had very high audit risk. Such a situation is common in audit risk, since not many firms have very high risks. In Table 1 below, monetary risk, inherent risk, and total audit risk are found to have very significant differences in terms of minimum and maximum values. The high standard deviation values further indicate that audit-risk intensity varied considerably across firms.

Table 1. Descriptive Statistics of Major Audit-Risk Indicators

Indicator	Mean	Median	Standard deviation	Minimum	Maximum
Audit parameter exposure A	2.491	0.900	5.718	0.000	85.000
Audit parameter exposure B	10.983	0.440	50.489	0.000	1264.630
Total audit parameter exposure	13.443	1.420	51.720	0.000	1268.910
Monetary exposure	14.360	0.110	67.105	0.000	935.030
Risk component A	1.374	0.180	3.465	0.000	51.000
Risk component B	6.442	0.088	30.317	0.000	758.778
Risk component D	8.406	0.022	40.296	0.000	561.018
District-level loss indicator	2.509	2.000	1.231	2.000	6.000
Historical risk indicator	0.106	0.000	0.535	0.000	9.000
Inherent risk	17.957	2.238	55.164	1.400	801.262
Control risk	0.575	0.400	0.447	0.400	5.800
Detection risk	0.500	0.500	0.000	0.500	0.500
Overall audit risk	7.284	0.590	38.986	0.280	961.514

From the data illustrated in Table 1, the average level of total audit parameter exposure is 13.443, whereas the median is 1.420. Equally, the mean and median levels of monetary exposure are 14.360 and 0.110 respectively. This shows a huge difference between the mean and the median and indicates the presence of risky firms whose audit and monetary exposure are relatively high. Detection risk did not vary throughout the sample and had zero standard deviation, suggesting that there was no variance due to detection risk.

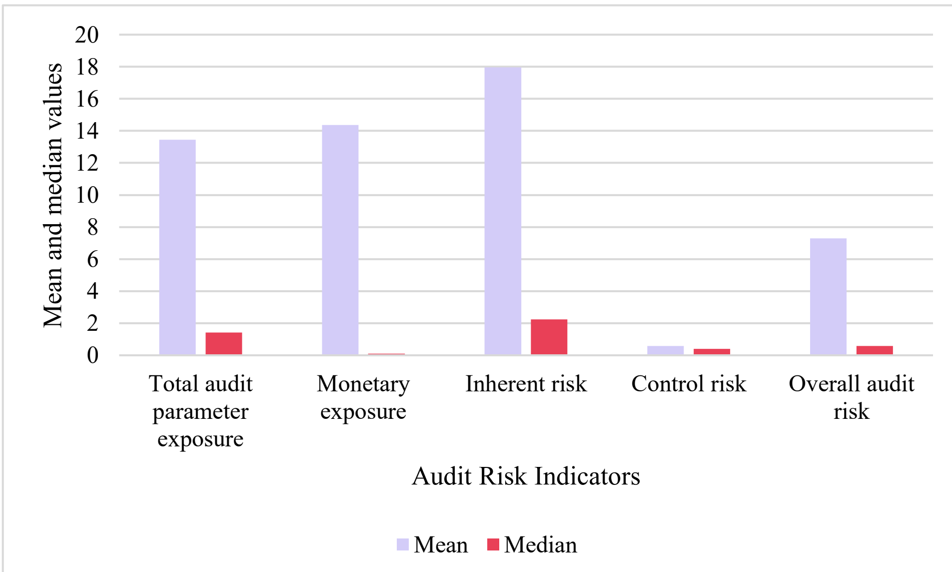

Figure 1. Mean and Median Values of Major Audit-Risk Indicators

As can be seen from Figure 1, there were huge differences between the mean and median values for both monetary exposure, risk, and audit risk. This means that the dataset contained few firms with very high audit-risk levels.

3.2 Comparative Analysis of Risky and Non-Risky Firms

From the above analysis, it can be observed that there are clear differences between risky and non-risky firms. Risky firms have higher average values for almost all aspects of audit risk. It can be observed that the classification of an organization in terms of its risk is truly related to its quantified audit-risk characteristics. As seen from Table 2, risky firms have significantly higher levels of audit-parameter risk, monetary risk, inherent risk, and audit risk overall. These findings indicate that firms classified as risky were associated with greater audit complexity, higher monetary involvement, and stronger underlying risk conditions.

Table 2. Comparison of Audit-Risk Indicators between Non-Risky and Risky Firms

Indicator	Non-risky firms mean	Risky firms mean	Mean difference
Audit parameter exposure A	0.738	5.138	4.400
Audit parameter exposure B	0.458	26.875	26.416
Total audit parameter exposure	1.197	31.932	30.735
Monetary exposure	0.366	35.489	35.123
Risk component A	0.293	3.006	2.713
Risk component B	0.164	15.920	15.756
Risk component D	0.092	20.960	20.868
District-level loss indicator	2.109	3.112	1.003
Historical risk indicator	0.002	0.263	0.261
Inherent risk	1.978	42.082	40.104
Control risk	0.424	0.801	0.377
Detection risk	0.500	0.500	0.000
Overall audit risk	0.420	17.648	17.227

The data provided by Table 2 shows that risky firms have an average inherent risk of 42.082, while non-risky firms have an average inherent risk of 1.978. There is a huge difference in terms of monetary exposure among these two types of firms. The average monetary exposure of risky firms is 35.489, while that of non-risky firms is only 0.366. This means that one of the defining factors of risky firms is their extensive monetary exposure. Another difference that can be seen among these two groups is the average total audit parameter exposure. Risky firms have a total average audit parameter exposure of 31.932, while non-risky firms have an average total audit parameter exposure of 1.197. This indicates that accumulated audit exposure plays a major role in distinguishing firms with higher fraudulent-risk potential.

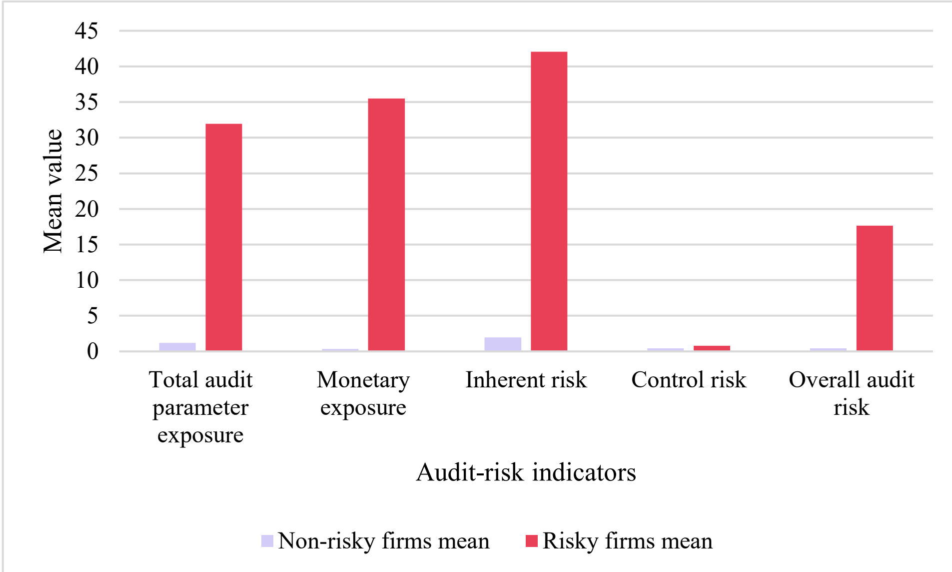

Figure 2. Comparison of Major Audit-Risk Indicators by Risk Category

Figure 2 illustrates that the risky firms always had larger values for all the main changing variables. The visual evidence supports the finding that the classification of risky firms was strongly related to the audit-risk profile characteristics.

3.3 Association between Audit Indicators and Risk Classification

The correlation analysis was done to identify the audit risk indicators which are highly correlated to risk classification. It was found that the audit composite scores, money exposure scores, audit parameters scores, control risk, loss at district level, and component risk indicators showed a positive correlation with risk classification. From Table 3 below, it is clear that the strongest correlation was seen in case of the audit composite scores, followed by money exposure scores, audit parameters B scores, and audit parameters A scores. It can thus be concluded that both the scoring indicators and audit-risk indicators worked effectively.

Table 3. Major Indicators Associated with Risk Classification

Indicator	Correlation with risk classification
Composite audit score	0.785
Monetary exposure score	0.688
Audit parameter B score	0.635
Audit parameter A score	0.618

Control risk	0.413
Risk component E	0.407
District-level loss indicator	0.399
Risk component A	0.384
Audit parameter exposure A	0.377
Inherent risk	0.356
Audit parameter C score	0.353
Risk component C	0.341

As can be observed from Table 3, it is evident that risk classification is not based on one variable alone. Instead, risk classification is influenced by several auditing factors, including audit scores, monetary risk, risk of controls, losses at district level, and risk of component. In other words, the use of business analytics for risk classification is appropriate since several variables are considered.

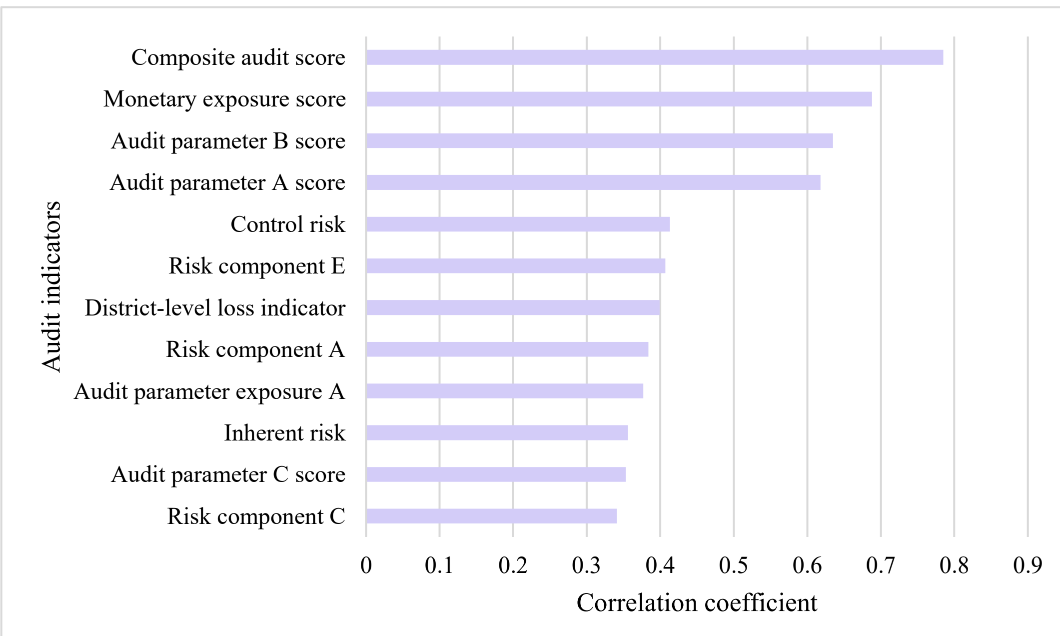

Figure 3. Correlation Strength of Audit Indicators with Risk Classification

As shown in Figure 3, scoring-based measures were found to have the highest relationship with risk classification. However, control risk, loss by district, and component risk also showed significant relationships, which suggest that these factors must be considered in audit risk assessment.

3.4 Audit-Risk Classification Performance

Classification test was conducted to find out if there exists an ability for the indicators used for measuring audit risks to discriminate between risky and non-risky firms. In this case, accuracy, precision, recall, specificity, and F1-score were utilized to evaluate the results obtained. These measures have been used due to their requirement for successful audit risk prediction, including the ability to predict risky firms.

As presented in Table 4, the classification assessment produced very strong performance across all evaluation measures. The overall accuracy was 1.000, indicating complete agreement between the predicted and observed risk categories.

Table 4. Classification Performance of Audit-Risk Assessment

Performance metric	Value
Accuracy	1.000
Precision	1.000
Recall / sensitivity	1.000
Specificity	1.000
F1-score	1.000

From the results shown in Table 4, it can be observed that the classification assessment showed complete agreement between the predicted and observed risk categories within the available audit-risk structure since all risky and non-risky firms could be distinguished accurately from one another. A precision value of 1.000 means that all firms predicted as risky matched the observed risk category. A value of 1.000 for recall means that all risky firms were classified accurately, whereas a value of 1.000 for specificity means that all non-risky firms were also classified accurately. The complete classification agreement suggests a high level of consistency among the audit-risk factors and the ultimate risk classification. However, this should be considered relative to the set of audit-risk factors that is currently available. In particular, the result may suggest that the risk classification is consistent with the overall audit-risk profile.

3.5 Confusion Matrix of Risk Classification

A confusion matrix has been used to find out the number of wrongly and correctly classified firms. This process is important because simple accuracy may not show if any of the high-risk or low-risk firms were wrongly included into another category. As can be seen from Table 5, all 459 low-risk firms have been classified properly as non-risky and all 304 risky firms have been classified properly as risky. There were no false positive or false negative classifications.

Table 5. Confusion Matrix of Audit-Risk Classification

Actual classification	Predicted non-risky	Predicted risky
Non-risky firms	459	0
Risky firms	0	304

The results in Table 5 confirm that the classification process produced no misclassification. The absence of false negatives is especially important in audit-risk prediction because false negatives represent risky firms that are incorrectly treated as non-risky. Similarly, the absence of false positives indicates that non-risky firms were not unnecessarily flagged as risky.

3.6 Key Indicators Contributing to Fraudulent Firm Classification

The identification of the most significant risk indicators was based on the assessment of group differences and the magnitude of associations of those indicators to risk categorization. It can be seen that total audit parameter risk, monetary risk, inherent risk, total audit risk, control risk, and component risks play an important role in the identification of risky firms. Based on the data shown in Table 6, it is clear that inherent risk had the highest mean difference among risky firms compared to non-risky firms, followed by monetary risk and total audit parameter risk exposure, risk component D, and overall audit risk.

Table 6. Key Indicators Distinguishing Risky and Non-Risky Firms

Rank	Indicator	Mean difference between risky and non-risky firms
1	Inherent risk	40.104
2	Monetary exposure	35.123
3	Total audit parameter exposure	30.735
4	Audit parameter exposure B	26.416
5	Risk component D	20.868
6	Overall audit risk	17.227
7	Risk component B	15.756
8	Audit parameter exposure A	4.400
9	Risk component A	2.713
10	District-level loss indicator	1.003

According to the results provided in Table 6, the first prominent distinguishing factor that can be singled out is inherent risk, since risky firms have an average difference of 40.104 compared to non-risky firms. Second in line is monetary risk, which shows that firms with increased monetary exposure are more likely to be regarded as risky firms. Thirdly, total audit parameter exposure demonstrates a very high difference in the average value, confirming the importance of audit exposure as an important determinant of the risky firms' classification. Finally, taking into account the importance of component D and audit risk in general, it can be argued that both factors play an essential role.

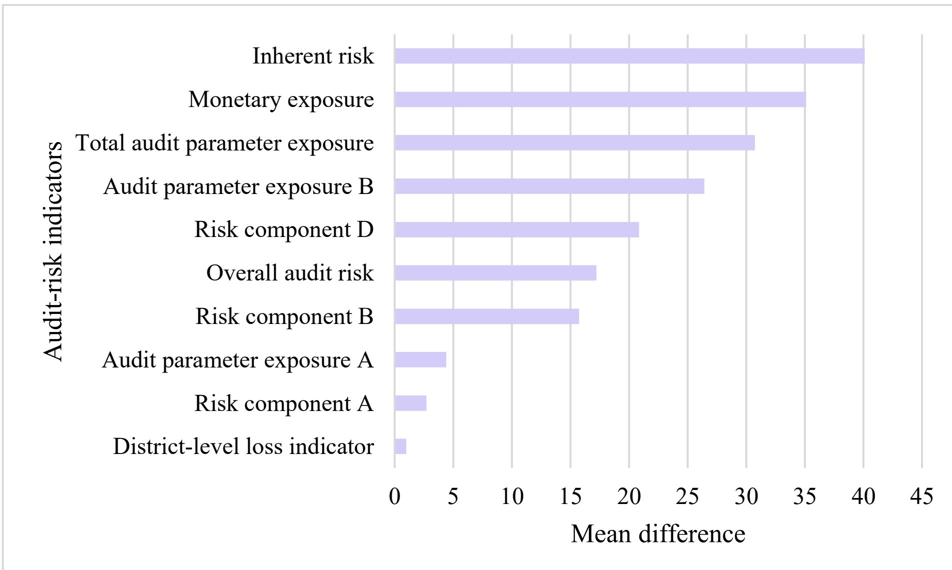

Figure 4. Key Audit Indicators Distinguishing Risky and Non-Risky Firms

As shown in Figure 4 below, the key distinguishing elements identified were inherent risk, monetary risk, and aggregate audit parameter risk. This suggests that auditors and other governance actors need to focus their attention on firms with high levels of inherent risk, significant monetary risk, and high aggregated audit parameters during the risk assessment process.

4. Discussion

Based on the results, audit-risk indicators provide a strong analytical basis for identifying risky firms and supporting fraudulent-risk classification. From the obtained results, it is clear that firms identified as risky have significantly high values with respect to inherent risk, money at stake, total audit risk exposure, risk elements, and total audit risk. Therefore, the classification of firms as risky appears to be systematic and based on some measurable characteristics that can be established using business analytics. Since the detection of audit fraud involves detecting patterns among indicators of firm-level risk, use of audit analytics, which has a machine learning orientation, makes sense as per previous literature on audit-fraud data prediction (Tiwari & Hooda, 2018). The clear differences between risky and non-risky firms therefore confirm that audit risk assessment can be improved by using systematic and data-based methods, especially when audit teams have to screen firms before carrying out extensive investigation work.

One of the main insights generated from the analysis is the significant difference in inherent risk between risky firms and non-risky firms. Risky firms demonstrated considerably higher levels of inherent risk, which implies that the possibility of audit risk or material irregularity may increase if there is an inherently risky situation within the business. In fact, in practice, inherent risk is often measured through professional judgement. Nevertheless, this analysis proves that quantitative measures help professionals to improve their decision-making process since such an approach makes the screening process objective and uniform. Indeed, this finding fits the wider view that audit analytics enable accountants to overcome purely judgmental decision-making and even the sampling problem if large volumes of audit data become available (Huang et al., 2022). Thus, inherent risk must be regarded as one of the key screening measures in audits.

It is apparent that there is a significant difference in terms of monetary exposure between the two categories of firms, which are labeled as risky firms and non-risky firms. The risky firms demonstrate more monetary exposure in comparison to non-risky firms. However, monetary exposure itself does not define the reason for being categorized into one category or another because this metric becomes very useful together with other variables used in an audit process. For instance, while identifying financial fraud, it is useful to use machine learning techniques due to the possibility of analyzing multiple indicators at once, which may not be possible by hand (Ali et al., 2022). Thus, high monetary exposure is especially dangerous if accompanied by parameter exposure and control risk indicators.

The outcome of the correlation analysis supports the interpretation that classification of risky firms relates to several dimensions of audit risk. The most significant correlations have been found for composite audit risk, monetary risk, and audit parameter risk. However, there are still some significant correlations in control risk, district risk, component risk indicator, and inherent risk. This evidence suggests that an accurate audit risk prediction model cannot be built using one indicator. Instead, the auditor should use various indicators that together demonstrate financial risk, internal control problems, past trends, and cumulative audit parameter risk. This interpretation is consistent with the financial statement fraud research, which states that fraud detection usually requires multiple financial indicators rather than just one measure (Ashtiani & Raahemi, 2021).

Predictive analysis shows the utility of business analytics in the classification of audit risks. The performance of the classification model was impressive, which means that indicators of audit risk have the ability to distinguish between firms that belong to the category of high-risk versus low-risk. From the practical perspective, such a model can help auditors and regulatory specialists identify situations of need for audit emphasis and improve planning of auditing activities. Hence, predictive analysis can help audit planning through a rational identification of firms that should receive more attention from the auditor. Recent studies on fraud detection reveal that the application of machine learning is becoming common practice to enhance screening and classification, but the usefulness of this approach lies in interpreting the outcomes from an accounting and governance perspective (Hernandez Aros et al., 2024). Thus, analytics-based classification should be viewed as decision support rather than automatic proof of fraudulent behaviour.

Another valuable outcome of the study is the establishment of the essential criteria that distinguish firms at risk from those that are not. In particular, intrinsic risk, financial exposure, total audit parameter exposure, audit parameter exposure B, and risk factor D turned out to be the strongest discriminators. The practical value of this finding is evident, as it helps auditors have a more solid rationale for prioritizing indicators while assessing initial risks. Instead of giving equal importance to all audit parameters, they could use this information to focus on those that demonstrate higher discrimination power. The rising trend towards using intelligent fraud detection tools in finance also highlights the need for developing flexible systems that can detect potential risks in sophisticated financial settings (Zhu et al., 2021). Therefore, the findings support a more targeted and risk-sensitive audit approach.

The results have implications regarding governance and internal controls. The consideration of control risk as an aspect of the analysis suggests that weak conditions of control can increase the likelihood of errors of classification. Hence, audit-risk prediction cannot be reduced to a mere technical task but also entails an organizational component. An organization categorized as high-risk based on audit analytics may require further assessment of its governance, stronger controls, and more thorough surveillance on behalf of management and audit committees. Thus, audit analytics serve not only to facilitate assurance procedures but also help managers make decisions.

Since modern firms have to deal with dynamic markets, they must be flexible, and adaptability is related to better firm performance in such settings (Suder et al., 2025). This reinforces the relevance of analytics-based audit systems for modern organizational governance.

All things considered, this study shows that the theory and application of predicting audit risks using business analytics are relevant and useful in practice. Risky firms have differences with non-risky firms with regard to various criteria related to audit risk, including such factors as inherent risk, monetary risk, overall risk of audit parameters, and control risk. These results suggest the use of business analytics in the process of developing an audit plan, identifying potentially fraudulent firms, and assessing risks associated with corporate governance. These results strengthen the role of analytics-based evidence in audit planning, fraudulent firm screening, and governance-related risk assessment.

5. Conclusion

In this study, the issue of auditing risk prediction and fraudulent-risk classification has been analyzed using business analytics techniques. The findings confirm that the audit-risk factors have proven effective in distinguishing firms classified as risky from those that are not. Indeed, firms that are considered risky have significantly higher values of intrinsic risk, monetary risk, total audit parameter risk, component risk factors, and audit risk. These differences indicate that fraudulent-risk classification is more associated with the presence of certain audit factors than just random chance. Further, this study finds that there are significant relationships between the composite audit risk score, monetary risk score, audit parameter score, control risk, district loss, and intrinsic risk. From the results of the classification analysis, one can conclude that audit-risk indicators are very effective in distinguishing between firms characterized by high risk and those with low risk. This means that business analytics can serve as the basis for providing decision support in making choices that can prove useful for auditors, managers, regulators, and other involved parties. The implications of the findings from the analysis performed during the course of the project include an improved ability to plan audit procedures by focusing on firms characterized by high inherent risk, large financial exposure, high accumulated audit parameter values, and high control-risk indicator values. An improved approach to the planning of audit can help achieve greater efficiency in audit work and detect risks more efficiently while allocating audit resources to firms with higher risks. In further research, one can focus on expanding the scope of analysis using different approaches, including using bigger datasets, industry samples, longitudinal audit data, and explainable analytics techniques. In conclusion, the analysis shows that business analytics can be used effectively to predict audit risk and support fraudulent-risk classification.

References

1. Abdelwahed, A. S., Abu-Musa, A. A., Badawy, H. A., & Moubarak, H. (2025). Unleashing the beast: the impact of big data and data analytics on the auditing profession—Evidence from a developing country. Future Business Journal, 11(1), 12.
2. Ali, A., Abd Razak, S., Othman, S. H., Eisa, T. A. E., Al-Dhaqm, A., Nasser, M., ... & Saif, A. (2022). Financial fraud detection based on machine learning: a systematic literature review. Applied Sciences, 12(19), 9637.
3. Alvarez-Foronda, R., De-Pablos-Heredero, C., & Rodríguez-Sánchez, J. L. (2023). Implementation model of data analytics as a tool for improving internal audit processes. Frontiers in Psychology, 14, 1140972.
4. Ashtiani, M. N., & Raahemi, B. (2021). Intelligent fraud detection in financial statements using machine learning and data mining: a systematic literature review. IEEE Access, 10, 72504-72525.
5. Beemamol, M. (2024). Mapping the trends of Financial Statement Fraud detection research from the historical roots and seminal work. Journal of Economic Criminology, 6, 100096.
6. Binh, N. T. T. (2025). Transforming auditing in the AI era: a comprehensive review. Information, 16(5), 400.
7. Chen, S. (2016). Detection of fraudulent financial statements using the hybrid data mining approach. SpringerPlus, 5(1), 89.
8. Fedyk, A., Hodson, J., Khimich, N., & Fedyk, T. (2022). Is artificial intelligence improving the audit process?. Review of Accounting Studies, 27(3), 938-985.
9. Gkegkas, M., Kydros, D., & Pazarskis, M. (2025). Using data analytics in financial statement fraud detection and prevention: A systematic review of methods, challenges, and future directions. Journal of Risk and Financial Management, 18(11), 598.
10. Hernandez Aros, L., Bustamante Molano, L. X., Gutierrez-Portela, F., Moreno Hernandez, J. J., & Rodríguez Barrero, M. S. (2024). Financial fraud detection through the application of machine learning techniques: a literature review. Humanities and Social Sciences Communications, 11(1), 1-22.

11. Huang, F., No, W. G., Vasarhelyi, M. A., & Yan, Z. (2022). Audit data analytics, machine learning, and full population testing. The Journal of Finance and Data Science, 8, 138-144.
12. Leocadio, D., Malheiro, L., & Reis, J. (2024). Artificial intelligence in auditing: A conceptual framework for auditing practices. Administrative Sciences, 14(10), 238.
13. Noordin, N. A., Hussainey, K., & Hayek, A. F. (2022). The use of artificial intelligence and audit quality: An analysis from the perspectives of external auditors in the UAE. Journal of Risk and Financial Management, 15(8), 339.
14. Siddhartha, M. (2019). Audit data: Audit risk dataset for classifying fraudulent firms [Dataset]. Kaggle. https://www.kaggle.com/datasets/sid321axn/audit-data15.
15. Suder, M., Kusa, R., Duda, J., & Okręglicka, M. (2025). Mediating or moderating? Innovative approach to the role of flexibility in the relationship between entrepreneurial orientation and firm growth under different market conditions. Journal of Innovation & Knowledge, 10(2), 100658.
16. Tiwari, A., & Hooda, N. (2018). Machine learning framework for audit fraud data prediction. Res Gate Publ, 7(6).
17. Vitali, S., & Giuliani, M. (2024). Emerging digital technologies and auditing firms: Opportunities and challenges. International Journal of Accounting Information Systems, 53, 100676.
18. Zhu, X., Ao, X., Qin, Z., Chang, Y., Liu, Y., He, Q., & Li, J. (2021). Intelligent financial fraud detection practices in post-pandemic era. The Innovation, 2(4).