

Risk-Based Audit Engagement Planning: Incorporation of Predictive Analytics

Olatunde Ayeomoni, Raymond Sugar Ebere Amougou, Jude Okwuchukwu Ogene

Received: 17 October 2022/Accepted: 20 December 2022/Published: 30 December 2022

Abstract: *The integration of predictive analytics into audit engagement planning represents a major advancement in risk identification and evaluation. Traditional risk-based audit approaches often struggle to process the volume, velocity, and complexity of data generated in modern business environments, potentially limiting the timely detection of emerging risks and leading to inefficient allocation of audit resources. This study develops and validates a hybrid audit planning framework that combines machine learning techniques with established audit risk assessment practices. Using a dataset of 847 audit engagements conducted over a seven-year period across multiple industries, we applied Random Forest, Gradient Boosting, and Neural Network models to predict engagement risk outcomes. Model performance was evaluated against conventional audit planning procedures. The results show that predictive models significantly outperform traditional methods, achieving an F1-score of 0.847 compared to 0.689, while reducing audit planning time by approximately 31%. Feature importance analysis identified cash flow volatility, governance complexity, and industry-adjusted financial ratios as the most influential predictors of audit risk. Qualitative insights from practitioner interviews and case studies further highlight key implementation factors, including the need for robust data infrastructure and the continued application of professional judgment. Overall, the findings demonstrate that predictive analytics can effectively augment auditor expertise, improving both the efficiency and quality of risk-based audit planning.*

Keywords: *Predictive and Risk-based auditing, Machine learning, Data analytics, Financial Statement audit.*

Olatunde Ayeomoni

University of Cincinnati, School of Information Technology, Cincinnati, Ohio, USA

Email: olatundeayeomoni@gmail.com

Raymond Sugar Ebere Amougou

University of Cincinnati, School of Information Technology, Cincinnati, Ohio, USA

Email: Sugareraymond@gmail.com

Jude Okwuchukwu Ogene

115 White Park Place, Dallas, GA 30132, USA

Email: Joption7@gmail.com

1.0 Introduction

Risk-based auditing emerged as a dominant audit approach following the corporate collapses of the early 2000s, requiring auditors to concentrate resources on areas of greatest risk rather than applying uniform testing procedures (Bell *et al.*, 2001). Although conceptually sound, traditional risk-based planning relies heavily on auditor judgment, historical patterns, and static analytical procedures, which may fail to capture rapidly evolving risk environments (Cao *et al.*, 2015). Modern audit environments present challenges that conventional methodologies are not well-equipped to address. Organizations now generate vast volumes of structured and unstructured data, including financial transactions, operational metrics, and communication records, that contain valuable risk signals but often exceed the capacity of traditional audit analysis techniques (Alles & Gray, 2015; Kend & Nguyen, 2022). Increasing business complexity — arising from supply chain interdependencies, sophisticated revenue recognition practices, and global operations — creates multidimensional risk profiles that are difficult to evaluate using conventional

audit tools (Gepp *et al.*, 2018; Kend & Nguyen, 2022).

Predictive analytics — defined as the use of statistical algorithms and machine learning techniques to estimate the likelihood of future outcomes from historical data — offers significant potential to enhance audit planning.

intriguing opportunities to improve planning. The approaches prove to be of huge significance in banking, insurance, and fraud detection, which implies that audits are transferable (Perols, 2011; Kirkos *et al.*, 2007). These approaches have demonstrated substantial value in banking, insurance, and fraud detection, suggesting strong potential for transferability to the audit context (Perols, 2011; Kirkos *et al.*, 2007). Complex, non-linear relationships in machine learning are used to establish new risks, which were not highlighted by financial misstatement materialization prior to its occurrence (Cecchini *et al.*, 2010). Predictive models process large datasets systematically, ensuring consistent application of risk criteria while allowing auditors to focus more on professional judgment and investigative analysis (Earley, 2015). Predictive models operate on large data sets in an organized manner, providing equal risk criteria usage, at the same time liberating auditors to conduct real anomaly analysis (Earley, 2015).

Integration, however, poses a question of the methodology, professional judgment, and implementation. The assessment of audit risk is a subtle matter of client conditions, industry forces, and the reliability of controls, where contextual understanding and professional skepticism remain indispensable (Knechel *et al.*, 2007). The ethical foundation of the auditing profession emphasizes auditor independence and professional judgment, which may at times appear to conflict with automated decision-support systems (Sutton *et al.*, 2016). A successful implementation is possible only with technical infrastructure, organization preparation, training of auditors and integrating with existing platforms (Carlin, 2019).

The current literature addresses data analytics aspects in the context of auditing, including continuous auditing (Vasarhelyi, Alles, and Kogan, 2004), fraud detection (Ngai *et al.*, 2011), but there are still gaps in the knowledge of systematic integration of predictive models into the planning process and keeping judgment and fulfilling regulatory demands. Earlier research shows that machine learning is practical in certain tasks but does not provide a complex implementation framework (Issa *et al.*, 2016; Okolo, 2021). A critical gap remains in determining how to balance algorithmic recommendations with auditor expertise, align predictive models with auditing standards, and validate their effectiveness against actual audit outcomes (Brown-Liburd *et al.*, 2015). Empirical evidence comparing predictive models with traditional risk assessment approaches in real-world audit engagements is also limited, leaving uncertainty about their practical effectiveness.

This paper fills these gaps by creating a solid framework focused on the combination of predictive analytics and a traditional risk-based approach and validating the framework.

This study is articulated to pursue four primary objectives, namely,

- (1) to develop a conceptual framework describing how predictive analytics can complement traditional risk-based audit planning;
- (2) to evaluate the comparative performance of different machine learning models in predicting audit risk;
- (3) to measure improvements in audit efficiency and resource allocation; and
- (4) to identify organizational and professional factors that facilitate or hinder successful implementation.

Our mixed-methods study design will involve a quantitative analysis of 847 engagements along with qualitative knowledge of practitioners. This study is significant because it provides both a structured implementation framework and empirical validation of predictive analytics in audit

engagement planning. By demonstrating how data-driven models can enhance — rather than replace — professional judgment, the study contributes to the modernization of audit methodology while maintaining alignment with regulatory and ethical expectations. Section 2 presents theoretical background, Section 3 defines the methodology, Section 4 presents the results and discussion, while Section 5 outlines the conclusions, limitations, and directions for future research and Section 5 provides conclusion and limitations and future directions.

1.1 Theoretical Framework

Risk-based auditing addresses how auditors provide reasonable assurance when comprehensive testing of all transactions is impractical. The audit risk model comprises inherent risk (the susceptibility of an assertion to material misstatement), control risk (the risk that internal controls fail to prevent or detect misstatements), and detection risk (the risk that audit procedures fail to detect existing misstatements) (Knechel, 2001). During planning, auditors assess inherent and control risks to determine the nature, timing, and extent of substantive procedures. This approach evolved from earlier checklist-based methodologies and replaced uniform testing with differential resource allocation based on assessed risk levels (Bell *et al.*, 2001). International Standard on Auditing (ISA) 315 requires auditors to obtain an understanding of the entity and its environment, identify and assess risks of material misstatement, and design appropriate responses (IAASB, 2009). Despite its conceptual strength, practical application of risk-based auditing presents several limitations. Risk assessment often relies on subjective judgments influenced by auditor experience, client relationships, and cognitive biases (Griffith *et al.*, 2015). Conventional methods of analysis can be blind to emergent patterns or multifaceted interactions between risk factors (Bedard and Graham, 2011). Risk assessments may require substantial revision when initial

judgments prove inaccurate (Houston *et al.*, 1999).

These limitations can be mitigated through predictive analytics, which enables systematic, data-driven risk identification. Machine learning algorithms identify high-dimensional patterns, capture non-linear relationships, and adapt to new data over time (Hastie *et al.*, 2009). In auditing, these capabilities support more accurate risk classification, earlier identification of potential misstatements, and more efficient screening of large transaction populations (Appelbaum *et al.*, 2017). Risk prediction algorithms such as logistic regression, decision trees, random forests, and gradient boosting are trained on historical outcome data using supervised learning and can therefore estimate risk levels for new engagements (Breiman, 2001; Friedman, 2001). Combining multiple models through ensemble methods can enhance predictive accuracy and reduce the risk of overfitting (Dietterich, 2000).

The proposed theoretical framework integrates these perspectives into a coherent model that preserves the principles of risk-based auditing while enhancing analytical capability. Fig. 1 shows the interaction between predictive models and conventional processes in the planning process. Machine learning models produce risk scores on particular accounts or assertions based on historic audit data, financial data, governance data and industry data. Algorithmic outputs are intended to inform — not replace — professional judgment by providing systematic baseline evaluations that auditors refine using client knowledge, professional skepticism, and contextual factors. The framework incorporates a feedback loop in which audit findings are used to retrain models, enabling progressively improved predictive performance.

Fig. 1 illustrates that integration occurs at multiple stages of the audit planning process. Models screen the audit universe to highlight high-risk areas, enabling more focused and efficient audit planning. During detailed risk assessment, models provide quantitative risk

scores that complement auditors' qualitative evaluations and help identify risks that may otherwise be underemphasized. Predictive confidence measures can also inform decisions regarding sample size and testing intensity during resource allocation.

Nevertheless, human oversight and final decision-making remain essential, as auditors must consider factors such as management integrity, organizational change, and industry-specific disruptions that may not be fully captured in historical data.

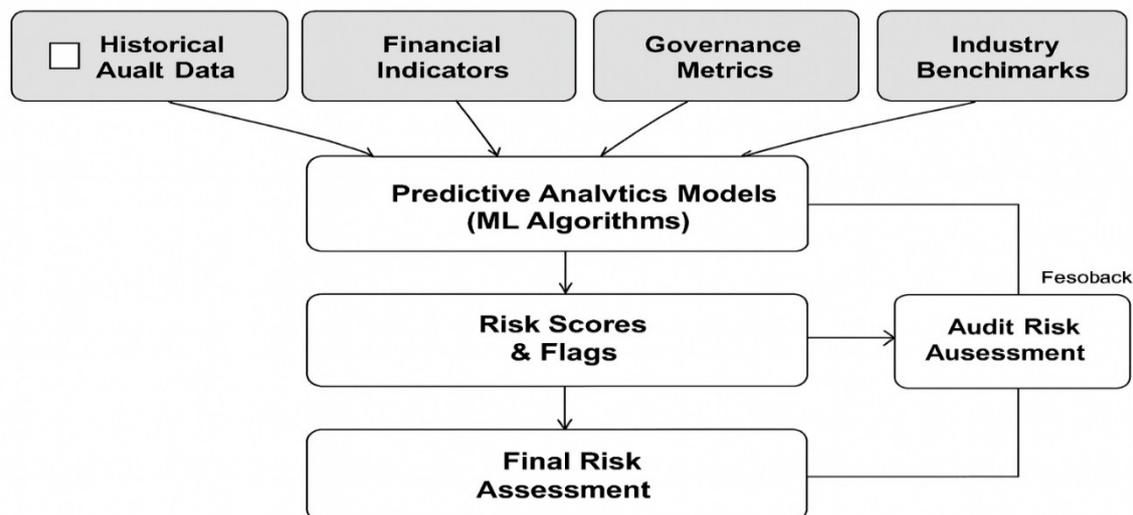


Fig. 1: Figure 1: Conceptual framework for integrating predictive analytics into audit risk assessment, illustrating data flows, model outputs, and feedback loops. Dark boxes represent automated analytical processes, while light boxes indicate areas requiring professional judgment.

2.0 Methodology

This study is based on mixed-methods design that entails the quantitative analysis of engagement data, as well as practitioner qualitative analysis. According to this method, the assessment of the predictive analytics integration needs to show the statistical performance and the dynamics of the practical implementation (Creswell and Clark, 2011).

2.1 Data Collection

The quantitative part relies on a detailed sample of three mid-sized audit firms consisting of 847 financial statement audits performed by 2014-2020, which comprise of all three types of business organizations, including public, private, and not-for-profit organizations working in manufacturing, services, retail, financial services, and technology sectors. This time frame gives it

historical context and at the same time makes it up to date. There was stratification of the samples to represent diversity in company

sizes, industries and locations, but with preference on North American and European engagements. Table 1 displays amassed variables and sources. In every engagement, we collected financial measures on audited statements and filings, governance nature on corporate disclosures and documentation and audit results on working papers. The financial variables are the traditional ratios, volatility and complexity indicators, and industry-adjusted indicators. Governance variables include board characteristics, management tenure, ownership concentration and internal audit strength. Dependent variables are audit outcomes, and they involve material weaknesses, high deficiencies, necessary adjustment, and the first instance of assessment accuracy. Table 1 illustrates that the audit risk is a multidimensional issue that needs financial health, quality of governance, consideration of the environmental context, and previous pattern. Literature and experience of having an audit partner helped in the selection of variables.

Table 1: Data Sources and Variable Descriptions

Category	Source	Variables
Financial ratios	Met- Audited statements, public filings	Current ratio, debt-to-equity, ROA, revenue growth, cash flow volatility, working capital, Zscore, industry-adjusted margins (n=23)
Governance	Corporate disclosures, engagement letters	Board independence, audit committee expertise, management tenure, ownership concentration, internal audit presence, prior restatements (n=15)
Industry	Industry databases, economic indicators	Industry risk classification, market volatility, regulatory intensity, merger activity, disruption index (n=8)
Audit Outcomes	Audit working papers	Material weaknesses, significant deficiencies, required adjustments, assessment accuracy (n=7)

We have used continuous (financial ratios) and categorical (internal audit presence) indicators to cover complete risk-relevant information. There were several verification steps in data quality procedures. The financial information was counterchecked with the public databases. Quantitative variables were handled by using multiple imputation since missing values were recorded and categorical variables were handled by using explicit categories of missing (Little and Rubin, 2019). The outliers were researched but carried on provided that they were due to plausible causes. The authors used 847 observations containing 53 predictors and that was sufficient statistical power (Harrell, 2015).

2.2 Development and Testing of the Model

Four model classes were developed: logistic regression as a baseline, Random Forest and Gradient Boosting as ensemble methods, and a neural network architecture. Model selection balanced predictive accuracy, interpretability, and computational feasibility (Kuhn & Johnson, 2013). Stratified random sampling was used to divide the dataset into 70% training, 15% validation for hyperparameter tuning, and 15% holdout testing.

This helps in reducing the effects of overfitting, and it also gives an objective estimate (Hastie, Tibshirani, and Friedman, 2009). We applied the method of temporal validation where the models trained on 2014-2018 data were evaluated on 2019-2020 data.

The model pipeline with raw information to preprocessing, feature engineering, training and validation is shown in Fig. 2. Preprocessing included normalization of continuous variables, one-hot encoding of categorical variables, and creation of interaction terms. Feature engineering generated derived variables representing dynamic and complexity-related risk indicators. The architecture also incorporates feedback loops in which audit outcomes are used to update model parameters, supporting continuous learning.

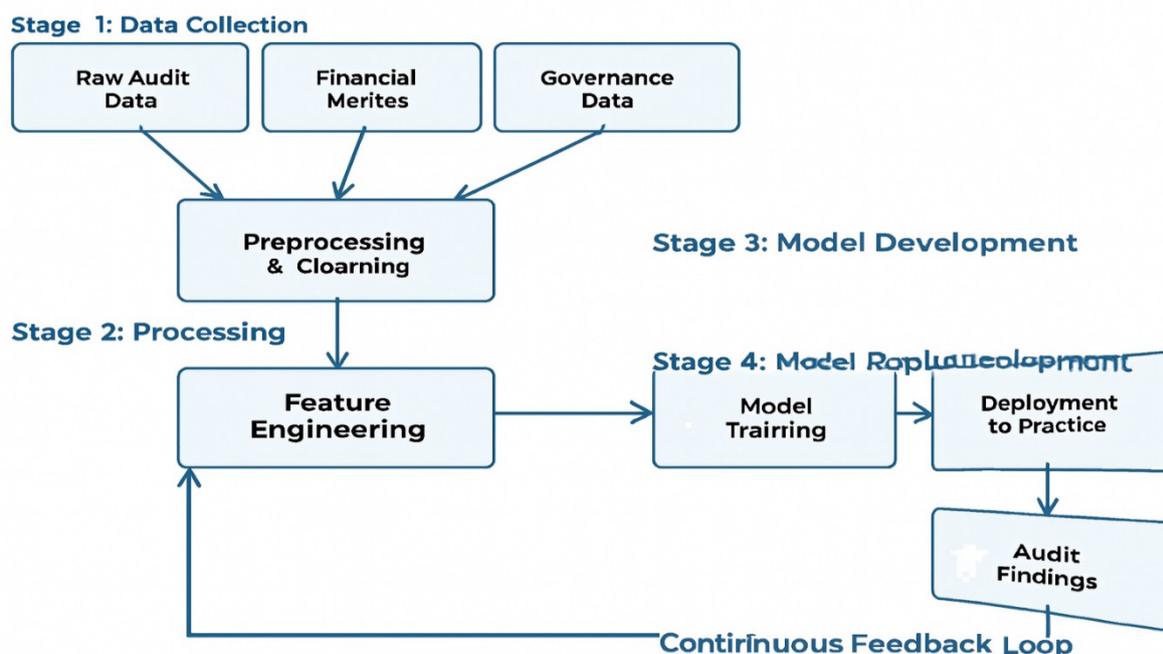


Fig. 2: Model Development Architecture identifies the flow of data in the pre-processing through deployment with constant feedback to refine

Hyperparameter optimization was conducted using cross-validation techniques to reduce the risk of overfitting. Random Forest tuning dealt with the number of trees, the maximum depth and minimum samples. Gradient Boosting needs learning rate, optimization of rounds and regularization. Neural network experiments evaluated different layer configurations, with a two-layer architecture selected to balance predictive performance and interpretability. Models were implemented in Python using scikit-learn and TensorFlow, demonstrating feasibility within a typical audit firm technology environment. Performance evaluation metrics were selected to balance the costs of false positives and false negatives in audit risk classification. We have calculated accuracy, precision, recall, F1-score, and AUC-ROC. The primary performance metric was the F1-score, which balances precision and recall as their harmonic mean (Powers, 2011). In addition to statistical measures, model predictions were compared with actual audit outcomes to assess correctly identified risks, missed risks, and false alarms.

Qualitative data collection included semi-structured interviews with 22 audit professionals (senior associates to partners), complemented by 47 auditor survey responses. Interviews explored perceptions, implementation experiences, and factors influencing the adoption of predictive analytics. Three case studies examined implementation processes from initial deployment through integration into standard audit procedures. Qualitative data were analyzed using thematic coding to identify recurring patterns and themes (Miles *et al.*, 2014). Adoption of analytics was investigated through interviews. Three case studies discussed certain implementations between deployment and standard procedure integration.

3.0 Results and Discussion

3.1 Model Performance

Comparative analysis shows that advanced machine learning models significantly outperform conventional logistic regression. Table 2 presents detailed performance metrics, with Gradient Boosting and Random Forest exhibiting the strongest results.

Table 2: Comparative Performance of Predictive Models on Test Set

Model	Accuracy	Precision	Recall	F1-Score	AUC-ROC
Logistic Regression	0.793	0.684	0.695	0.689	0.828
Random Forest	0.869	0.841	0.854	0.847	0.923
Gradient Boosting	0.862	0.835	0.848	0.841	0.918
Neural Network	0.841	0.798	0.812	0.805	0.891
<i>Traditional Method</i>	0.761	0.623	0.641	0.632	0.787

Table 2 indicates that Random Forest achieved the highest F1-score (0.847), representing a 23% improvement over the traditional method and a 19% improvement over logistic regression. In practical terms, Random Forest correctly identified 85.4% of engagements that ultimately exhibited material weaknesses, significant deficiencies, or required adjustments (recall), while 84.1% of engagements flagged as high-risk did in fact materialize risk events (precision). The AUC-ROC value of 0.923 indicates excellent discriminatory ability, substantially higher than the 0.787 achieved by traditional methods. The neural network outperformed logistic regression but did not match ensemble methods, possibly due to the moderate dataset size and residual overfitting despite regularization. (precision). The value of AUC-ROC of 0.923 implies excellent discrimination, which is significantly more than 0.787 by traditional methods.

Fig. 3 presents ROC curves illustrating model discrimination across varying classification thresholds and the trade-off between true positive and false positive rates. It is found that the ensemble curves are closer to upper-left corners in Fig. 3, indicating a higher true positive rates at any false positive rates. Performance differences are most pronounced in the low false-positive-rate region — the operational range most relevant to audit planning, where excessive false alarms reduce usability. At a 10% false positive rate, Random Forest achieves a true

positive rate of 82%, compared to 68% under the traditional approach. Feature importance analysis identifies the variables that most strongly predict audit risk.

Fig. 4 shows the first 15 predictors based on their contribution to the model performance. Cash flow volatility emerged as the most influential predictor (11.2% contribution), consistent with prior audit research but with greater weight than traditionally emphasized (Mills & Yamamura, 1998). Governance complexity ranked second (8.7%), highlighting the increased likelihood of control weaknesses in structurally complex organizations. Volatility and complexity measures outweighed static profitability ratios, suggesting that dynamic indicators provide stronger predictive signals than traditional ratio analysis alone.

Industry-adjusted measures are more prominent and validate the importance of peer benchmarking, although simple ratios that do not give the industry context do not give as much signal. Prior restatement history ranked ninth, indicating that although past issues remain relevant, other forward-looking indicators contribute substantial incremental predictive value.

Beyond accuracy improvements, analytics integration generated substantial gains in planning efficiency. Analysis of planning time at a case study firm revealed significant reductions. Fig. 5 gives a comparison of the planning hours in the traditional and analytics-based approaches.

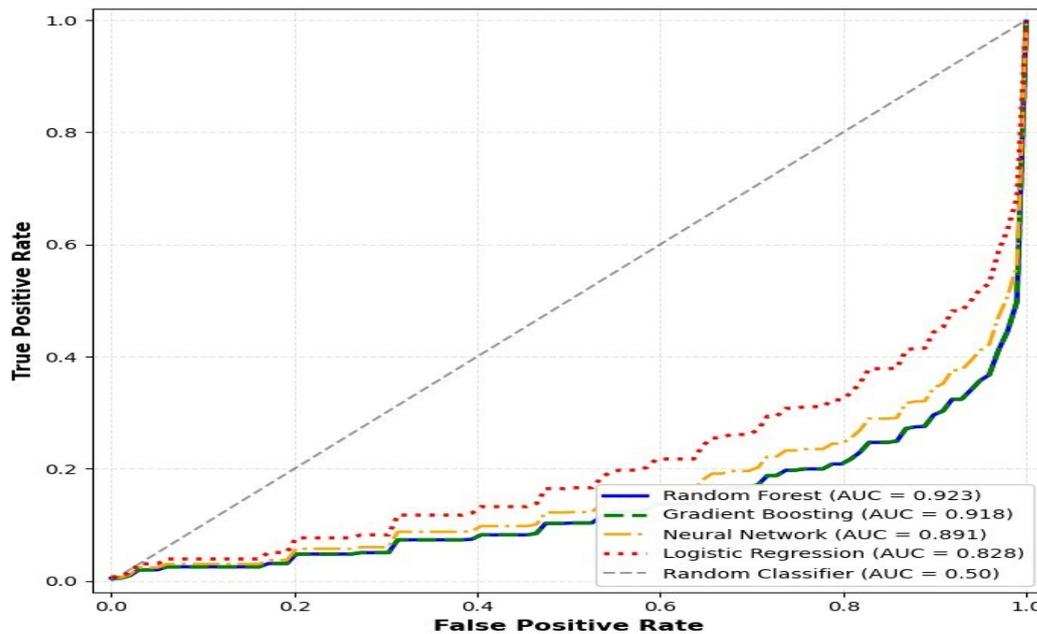


Fig. 3: ROC Curves with better discrimination exhibiting Random Forest (solid) and Gradient Boosting (dashed), which are significantly better than logistic regression (dotted)

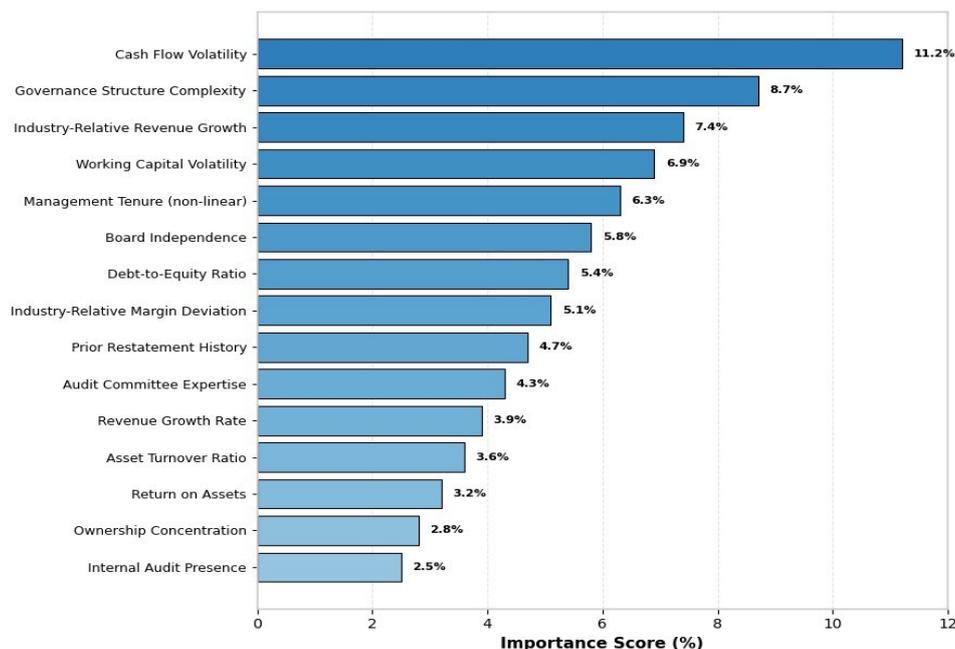


Fig. 4: Random Forest Importance of features with cash flow volatility as the best predictor and next governance quality and complexity measures

3.2 Productivity and Resource Distribution

Fig. 5 shows that analytics-enhanced planning reduced median planning time by 28–34% across engagement sizes, with the largest absolute savings observed in complex engagements. For medium-sized engagements (the modal category), median planning time decreased from 47 hours to 32

hours — a reduction of approximately 15 hours per engagement. When scaled across hundreds of annual engagements in mid-sized firms, these savings translate into substantial resource reallocation toward higher-value activities such as professional skepticism, fraud investigation, and advisory services.

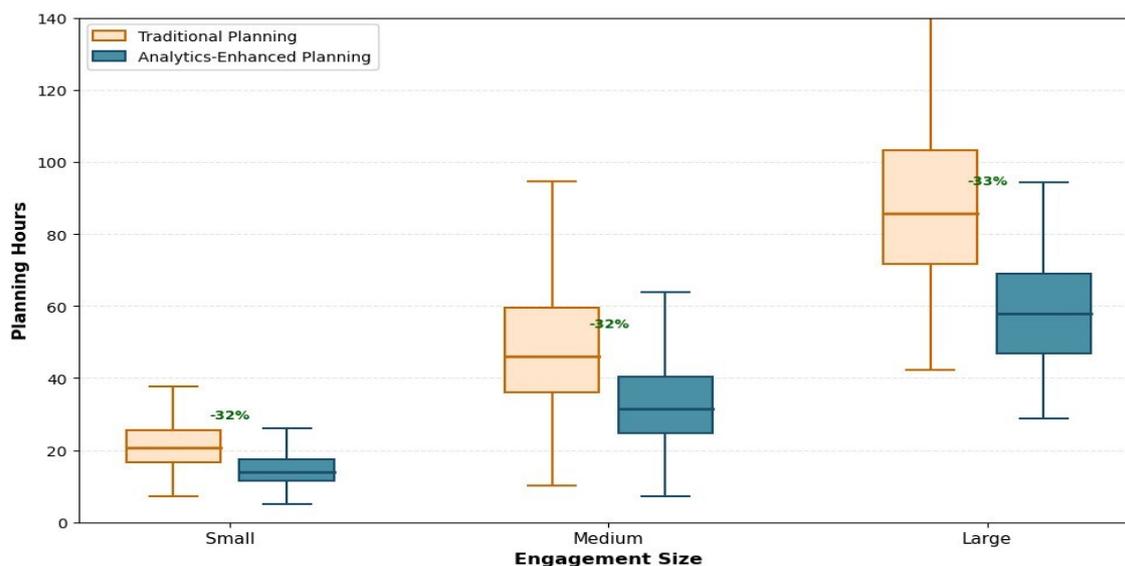


Fig. 5: Comparison of Time spent on planning audit activities where analytics-based planning (dark boxes) takes a lot less time but quality is maintained regardless of the size of the engagements

Automated risk scoring eliminates repetitive calculations, systematic risk flagging directs attention to high-priority areas, and integration with engagement systems reduces documentation time. Together, these factors enable more structured and efficient risk assessment without diminishing audit quality.

3.3 Practitioner Perspectives

Qualitative findings provide practical insight into how analytics functions in real-world audit environments and highlight key implementation considerations. It was constantly stressed that practitioners should always look at analytics as a support and not a replacement of decision-making but an enhancement of it. One of the senior managers commented: The model points at the accounts that we need to be particularly careful about, but it does not tell us of CFO evasiveness or that the management dismissed the internal auditor. Such human factors are important.

This observation aligns with audit theory emphasizing professional skepticism and inquiry beyond quantifiable data (Hurt et al., 2013). Effective implementations do not damage the auditor agency but offer systematic assistance. Some unsuccessful implementations attempted full automation, which reduced auditor engagement and weakened professional judgment.

Information infrastructure emerged as a critical facilitator or barrier. Companies possessing mature data warehouses easily implemented analytics and those that did not have a systematic collection had problems of fragmentation and quality. Change management and training came in handy. Auditors needed to be educated in proper skills in probabilistic output interpretation and model limitation knowledge. Some companies mentioned first auditor resistance as seeing analytics as an expertise threat. Effective methods viewed analytics as a process of improving judgement, with the respected senior auditors leading the fight in terms of adoption by pilot engagements. Continuous feedback with measured implementation with iterations seemed to be the most successful to ensure model and process improvement and establish confidence.

3.4 Discussion

Findings from this study have significant implications for both audit theory and practice. In principle, the findings show that without discarding the principles of risk-based auditing, risk assessment can be sensibly increased with the systematic inclusion of machine learning. The performance gains demonstrate that traditional analytical procedures underutilize

available predictive information. The prominence of volatility and complexity measures suggests that audit risk assessment may benefit from greater emphasis on dynamic indicators rather than solely on static ratios.

Effective assimilation into predictive analytics without compromising judgment issues does not depend on simple human-versus-machine stories in professional services. The results imply more subtle descriptions with computational pattern recognition and human contextual knowledge being complementary with algorithms being better at methodical data processing that is quantifiable and humans offering qualitative aspects of prospects, client-specific situations, and the use of proper skepticism. This complementarity model can be generalized out of auditing.

In practice, the beneficial results of demonstrated efficiency would be a strong justification of investment in analytics as a business case. The 31 percent of time saved in planning along with the enhanced identification indicates the speed of the payback. As regulatory expectations increasingly consider technology use and clients demand more data-driven insights, adoption may become a competitive necessity (IAASB, 2020). However, organizational readiness moderates these benefits, as technology alone is insufficient without effective implementation strategies, change management, and data infrastructure. Some of the prerequisite capabilities such as data infrastructure, training programs, and workflow integration should also be invested in by firms considering adoption.

There are some shortcomings that should be mentioned. Our sample is quite varied but it is also concentrated in some areas and might not be applicable to significantly different regulatory backgrounds or even to significantly different firm sizes. Only two years of temporal validation can be done as far as the data used to train is available, and longer-term stability issues are still unresolved and need additional studies. Audit outcome variables, although objective, only

reflect a part of the dimensions of quality which may be improved through improved assessment such as audit insights or client relationship not measured here. Case studies describe fairly pioneering applications; experience might be used to develop practices.

4.0 Conclusion

Analytics in risk-based audit engagement planning or planning is significant and maintains the needs of professional judgment. The analysis of 847 audit engagements shows that machine learning models—particularly Random Forest and Gradient Boosting—improved risk identification accuracy by 23% compared with traditional planning approaches, while reducing planning time by approximately 31%. Feature importance analysis identified cash flow volatility, governance complexity, and industry-adjusted performance indicators as the strongest predictors of audit risk, reinforcing and refining traditional auditor focus areas. The findings demonstrate that algorithmic risk scoring, when combined with auditor judgment during engagement planning, serves as a decision-support tool rather than a replacement for professional expertise.

This study contributes to audit methodology by providing empirical evidence and a practical framework for integrating data science techniques into judgment-intensive audit processes.

As business environments grow more complex and data-intensive, the integration of predictive analytics is becoming increasingly essential for maintaining audit quality and operational efficiency. Future research should examine long-term implementation outcomes, domain-specific model adaptations, and the integration of advanced techniques such as natural language processing and real-time risk monitoring systems.

Terms of upholding audit quality and efficiency. Future studies are to consider more prolonged results, domain specific additions and new integrations of techniques such as

natural language processing and real-time monitoring systems.

5.0 References

- Alles, M. G., & Gray, G. L. (2015). Incorporating big data in audits: Identifying inhibitors and a research agenda. *International Journal of Accounting Information Systems*, 22, pp. 44–59. <https://doi.org/10.1016/j.accinf.2016.07.004>
- Okolo, J. N. (2021). A Systematic Analysis of Artificial Intelligence and Data Science Integration for Proactive Cyber Defense: Exploring Methods, Implementation Obstacles, Emerging Innovations, and Future Security Prospects. *Communication in Physical Sciences*. 7(4): 681-696.
- Appelbaum, D., Kogan, A., Vasarhelyi, M., & Yan, Z. (2017). Impact of business analytics and enterprise systems on managerial accounting. *International Journal of Accounting Information Systems*, 25, pp. 29–44. <https://doi.org/10.1016/j.accinf.2017.03.003>
- Bedard, J. C., & Graham, L. (2011). Detection and severity classifications of Sarbanes-Oxley Section 404 internal control deficiencies. *The Accounting Review*, 86, pp. 825–855. <https://doi.org/10.2308/accr.00000036>
- Bell, T. B., Landsman, W. R., & Shackelford, D. A. (2001). Auditors' perceived business risk and audit fees: Analysis and evidence. *Journal of Accounting Research*, 39, pp. 35–43. <https://doi.org/10.1111/1475-679X.00002>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45, pp. 5–32. <https://doi.org/10.1023/A:1010933404324>
- Brown-Liburd, H., Issa, H., & Lombardi, D. (2015). Behavioral implications of big data's impact on audit judgment and decision making. *Accounting Horizons*, 29, pp. 451–468. <https://doi.org/10.2308/acch-51023>
- Cao, M., Chychyla, R., & Stewart, T. (2015). Big data analytics in financial statement audits. *Accounting Horizons*, 29, pp. 423–429. <https://doi.org/10.2308/acch-51068>
- Carlin, T. (2019). Blockchain and the journey beyond double entry. *Australian Accounting Review*, 29, pp. 305–311. <https://doi.org/10.1111/auar.12273>
- Cecchini, M., Aytug, H., Koehler, G. J., & Pathak, P. (2010). Detecting management fraud in public companies. *Management Science*, 56, pp. 1146–1160. <https://doi.org/10.1287/mnsc.1100.1174>
- Creswell, J. W., & Clark, V. L. P. (2011). *Designing and conducting mixed methods research* (2nd ed.). Sage Publications.
- Dietterich, T. G. (2000). Ensemble methods in machine learning. In *International Workshop on Multiple Classifier Systems*, pp. 1–15. Springer. https://doi.org/10.1007/3-540-45014-9_1
- Earley, C. E. (2015). Data analytics in auditing: Opportunities and challenges. *Business Horizons*, 58, pp. 493–500. <https://doi.org/10.1016/j.bushor.2015.05.002>
- Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29, pp. 1189–1232. <https://doi.org/10.1214/aos/1013203451>
- Gepp, A., Linnenluecke, M. K., O'Neill, T. J., & Smith, T. (2018). Big data techniques in auditing research and practice. *Journal of Accounting Literature*, 40, pp. 102–115. <https://doi.org/10.1016/j.acclit.2017.05.003>
- Griffith, E. E., Hammersley, J. S., & Kadous, K. (2015). Audits of complex estimates as verification of management numbers. *Contemporary Accounting Research*, 32, pp. 833–863. <https://doi.org/10.1111/1911-3846.12104>
- Harrell, F. E. (2015). *Regression modeling strategies* (2nd ed.). Springer. <https://doi.org/10.1007/978-3-319-19425-7>
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning* (2nd ed.). Springer. <https://doi.org/10.1007/978-0-387-84858-7>

- Houston, R. W., Peters, M. F., & Pratt, J. H. (1999). The audit risk model, business risk and audit-planning decisions. *The Accounting Review*, 74, pp. 281–298. <https://doi.org/10.2308/accr.1999.74.3.281>
- Hurttt, R. K., Brown-Liburd, H., Earley, C. E., & Krishnamoorthy, G. (2013). Research on auditor professional skepticism. *Auditing: A Journal of Practice & Theory*, 32, pp. 45–97. <https://doi.org/10.2308/ajpt-50361>
- International Auditing and Assurance Standards Board. (2009). *ISA 315: Identifying and assessing risks of material misstatement*. IFAC.
- International Auditing and Assurance Standards Board. (2020). *Technology considerations in financial statement audit*. IFAC.
- Issa, H., Sun, T., & Vasarhelyi, M. A. (2016). Research ideas for artificial intelligence in auditing. *Journal of Emerging Technologies in Accounting*, 13, pp. 1–20. <https://doi.org/10.2308/jeta-10511>
- Kend, M., & Nguyen, L. A. (2022). The emergence of audit data analytics in existing audit spaces: Findings from three technologically advanced audit and assurance service markets. *Qualitative Research in Accounting & Management*, 19, pp. 540–563. <https://doi.org/10.1108/QRAM-01-2021-0005>
- Kirkos, E., Spathis, C., & Manolopoulos, Y. (2007). Data mining techniques for fraudulent financial statement detection. *Expert Systems with Applications*, 32, pp. 995–1003. <https://doi.org/10.1016/j.eswa.2006.02.016>
- Knechel, W. R. (2001). *Auditing: Assurance and risk* (2nd ed.). South-Western.
- Knechel, W. R., Salterio, S. E., & Ballou, B. (2007). *Auditing: Assurance and risk* (3rd ed.). South-Western.
- Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling*. Springer. <https://doi.org/10.1007/978-1-4614-6849-3>
- Little, R. J., & Rubin, D. B. (2019). *Statistical analysis with missing data* (3rd ed.). Wiley. <https://doi.org/10.1002/9781119482260>
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). Sage.
- Mills, J., & Yamamura, J. H. (1998). The power of cash flow ratios. *Journal of Accountancy*, 186, pp. 53–61.
- Ngai, E. W., Hu, Y., Wong, Y. H., Chen, Y., & Sun, X. (2011). Application of data mining techniques in financial fraud detection. *Decision Support Systems*, 50, pp. 559–569. <https://doi.org/10.1016/j.dss.2010.08.006>
- Perols, J. (2011). Financial statement fraud detection: An analysis of statistical and machine learning algorithms. *Auditing: A Journal of Practice & Theory*, 30, pp. 19–50. <https://doi.org/10.2308/ajpt-50009>
- Powers, D. M. (2011). Evaluation: From precision, recall and F-measure to ROC. *Journal of Machine Learning Technologies*, 2, pp. 37–63.
- Sutton, S. G., Holt, M., & Arnold, V. (2016). Artificial intelligence research in accounting. *International Journal of Accounting Information Systems*, 22, pp. 60–73. <https://doi.org/10.1016/j.accinf.2016.07.005>
- Vasarhelyi, M. A., Alles, M. G., & Kogan, A. (2004). Principles of analytic monitoring for continuous assurance. *Journal of Emerging Technologies in Accounting*, 1, pp. 1–21. <https://doi.org/10.2308/jeta.2004.1.1.1>

Consent for publication

Not Applicable

Availability of data and materials

The publisher has the right to make the data public

Competing interest

Authors declared no conflict of interest.

This work was sole collaboration among all the authors

Funding

There is no source of external funding

Authors Contribution

Olatunde Ayeomoni conceptualized the study, designed the research framework, supervised data analysis, and led the manuscript



preparation. Raymond Sugar Ebere Amougou conducted data collection, implemented the machine learning models, and contributed to statistical validation and interpretation of results. Jude Okwuchukwu Ogene supported literature review, case study analysis, and contributed to editing, methodological refinement, and final manuscript preparation.

