Journal of Natural and Artificial Systems, 2025, 1(1): 10-21

Risk-Based Audit Engagement Planning: Incorporation of Predictive Analytics

Sarah Ngusoon Agwaza

Received: 19 August 2024/Accepted: 16 November 2024/Published 30 November 2024

Abstract: *The inclusion of predictive analytics in auditing engagement planning is a revolution in risk identification and evaluation. The conventional risk-based audit practices are not well-suited to handle the large amounts of data and speed of data in contemporary business organisations, which may negatively impact the identification of emerging risks or misuse of resources. The research paper constructs and justifies a hybrid model that consists of machine learning algorithms and best practices of audit risk procedures. With 847 audit engagements with seven years of data across industries, we utilize Random Forest, Gradient Boosting and Neural Network models to forecast audit risk using much more accurate forecasts as compared to conventional methods. The findings show that predictive analytics can identify risks with an accuracy of F1-score of 0.847 vs 0.689 and cut the time spent on planning by 31%. Utilizing the feature importance analysis, cash flow volatility, the complexity of governance, and industry-adjusted ratios are found to be the most important predictors. We determine such critical success factors as data infrastructure preparation and maintenance of professional judgment through case studies and practitioner interviews. The study shows that data science methods can enhance human knowledge in the workplace.*

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

Sarah Ngusoon Agwaza
Adron Homes, Mararaba Branch, Nsarawa State, Nigeria
Email: sarahadwaza0@gmail.com

1.0 Introduction

Risk-based auditing has essentially changed the approach to auditing after the corporate collapses of the early 2000s, where auditors were required to focus their resources on areas of greatest risk instead of using consistent testing (Bell *et al*., 2001). Although theoretically elegant, the previous risk-based planning is more dependent on the auditor's judgment, on past patterns, and on fixed analytical operations, which might fail to reflect the dynamic risk environment (Cao *et al*., 2015). Modern audit business contexts have issues which traditional methodologies are unable to respond to. The organizations produce huge amounts of data, including financial activities, metrics, and communication patterns, which have valuable risk information but are beyond the traditional analysis (Alles & Gray, 2015, Kend & Nguyen, 2022). The complexity of business, whether it is through supply chain interdependencies, complex revenue schemes, or not, will present a multidimensional risk profile that cannot be easily classified (Gepp *et al*., 2018; Kend & Nguyen, 2022).

The use of predictive analytics, i.e. statistical algorithms and machine learning to predict the probability of future outcomes based on past data, creates 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). Complex, non-linear relationships in machine learning are used to establish new risks, which were not highlighted by financial misstatement materialization before its occurrence (Cecchini *et al*., 2010). 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, in which context and cynicism cannot be done away with (Knechel *et al*., 2007). The ethical model of the profession focuses on auditor independence and judgment, which may conflict with automated systems (Sutton, Holt, and Arnold, 2016). A successful implementation is possible only with technical infrastructure, organization preparation, training of auditors and integration 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). There is a gap in the literature where just the right balance between algorithmic recommendations and expertise, calibration to the standards and verification of accuracy against finding cannot be established (Brown-Liburd *et al*., 2015).

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. Our goal is to achieve four things: come up with a conceptual model that describes the complementary usage of algorithms to traditional processes; determine the comparative efficacy of modeling methods; measure the increase in efficiency and resource savings; and discover what aids and hinders implementation. Our mixed-methods study design will involve a quantitative analysis of 847 engagements along with qualitative knowledge of practitioners. Section 2 presents the theoretical background, Section 3 defines the methodology, Section 4 displays the results and conclusions, and Section 5 provides conclusion and limitations and future directions.

1.1 Theoretical Framework

Risk-based auditing deals with the manner in which auditors give reasonable assurance in case of the impracticality of comprehensive testing. Audit risk model includes inherent risk (assertion prone to misstatement), control risk (risk controls do not prevent/detect misstatement) and detection risk (risk procedures do not pick up misstatement) (Knechel, 2001). When planning, auditors evaluate inherent and control risk to decide on which nature of substantive procedure is acceptable, when and how often to perform it. This is based on the development of previous checklist-based methodologies, which have been replaced by differential resource allocation, as opposed to uniform testing (Bell *et al*., 2001). The International Standard on auditing 315 specifies that the auditors must know the environments of entities, recognize and evaluate risks of misstatement and develop the responses (IAASB, 2009). Real life application has shortcomings. Risk assessment is based on the subjective judgments that may be dependent on the experience, the relationships with the client and the cognitive biases (Griffith *et al*., 2015). Conventional methods of analysis can be blind to emergent patterns or multifaceted interactions between risk factors (Bedard & Graham, 2011). Risk assessment may be reiterated in cases where the initial judgement is significantly wrong (Houston *et al*., 1999). The limitations are overcome by predictive analytics, which undertakes systematic, data based identification. Machine learning algorithms find high-dimensional patterns of data, detect non-linear context and adapt to new data (Hastie *et al*., 2009). In the case of audit, those capabilities allow the classification of risks more accurately, the identification of risks earlier in time, and the screening of large population of transactions more efficiently (Appelbaum *et al*., 2017). Risk predicting algorithms such as logistic regression, decision trees, random forests, and gradient boosting are training on previous outcome data under supervised learning and thus predict risks on new engagements (Breiman, 2001; Friedman, 2001). Combination of several models in ensemble methods enhances model accuracy and overfitting resistance (Dietterich, 2000).

Our theoretical framework brings together these viewpoints into a non-contradictory framework that maintains risk-oriented auditing framework but improves the analysis ability. Figure 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. The outputs of algorithms give information but not substitution to judgment as they offer systematic baseline evaluations that auditors enhance by the knowledge of their clients, skepticism and circumstances. The framework includes feedback loop in which the audit findings train the models which form successively accurate learning systems.

Figure 1 indicates that integration takes place on several points of planning. Models screen audit universes with high-risk areas, which is smart triage. In the event of a detailed assessment, models will give quantitative scores to the qualitative evaluations, where they can help point out the risks that may have been underweight. Predictive confidence intervals are used to make decisions about sample size and testing intensity in resource allocation. In the meantime, human control and ultimate decision-making is present, because judgments are made with considerations, including management relations, organizational shifts, industry shocks, which cannot be fully reflected in the historical trends.

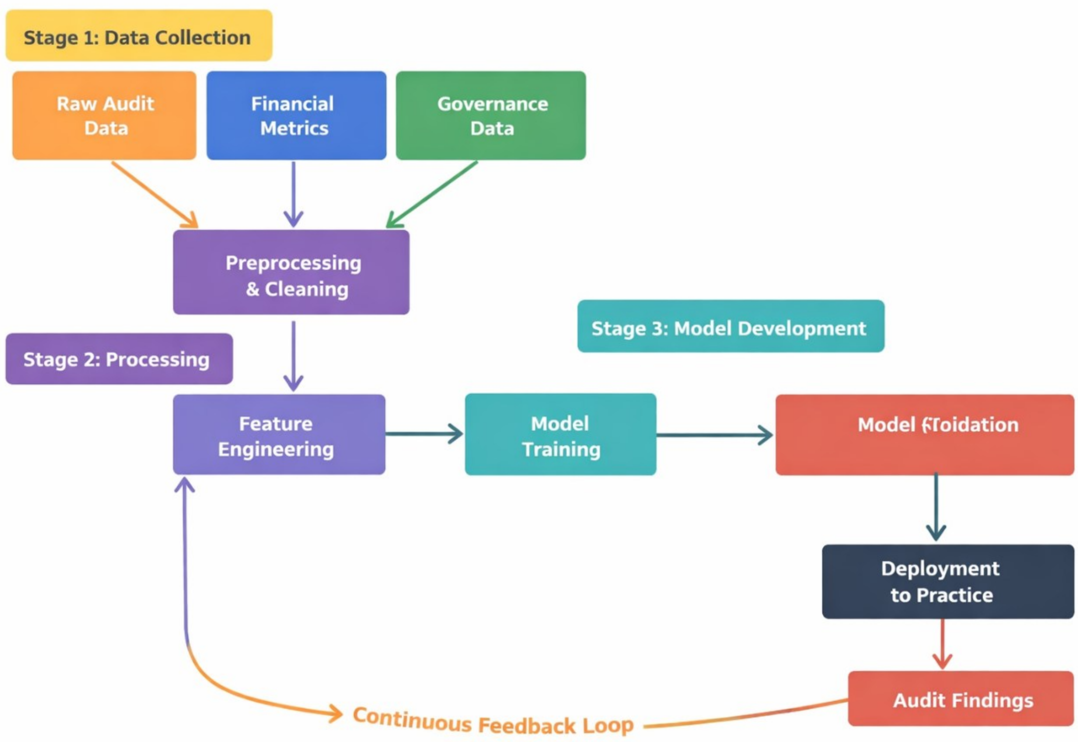


Fig. 1: Conceptual Framework of Predictive Analytics Integration that depicts data flow with modeling to risk assessment with a feedback loop. The dark boxes represent automated procedures; light-colored ones are subject to professional discretion.

2.0 Methodology

This study is based on a 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 comprises 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 a preference for 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 adjustments, 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 patterns. Literature and experience of having an audit partner helped in the selection of variables. We have used continuous (financial ratios) and categorical (internal audit presence) indicators to cover complete risk-relevant information. There were several verification steps in the 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 were not excluded, provided that they were due to plausible causes. The authors used 847 observations containing 53 predictors and that was sufficient statistical power (Harrell, 2015).

Table 1: Data Sources and Variable Descriptions

Category	Source	Variables
Financial Metrics	Audited statements, public filings	Current ratio, debt-to-equity, ROA, revenue growth, cash flow volatility, working capital, Z-score, 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, and assessment accuracy (n=7)

tThere were several verification steps in the 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 were not excluded, 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 created by us, including logistic regression as a baseline model, Random Forest and Gradient Boosting as ensemble techniques and Neural Network architecture. Selection is arbitrarily balanced in predictive accuracy, interpretable, and computable (Kuhn & Johnson, 2013).

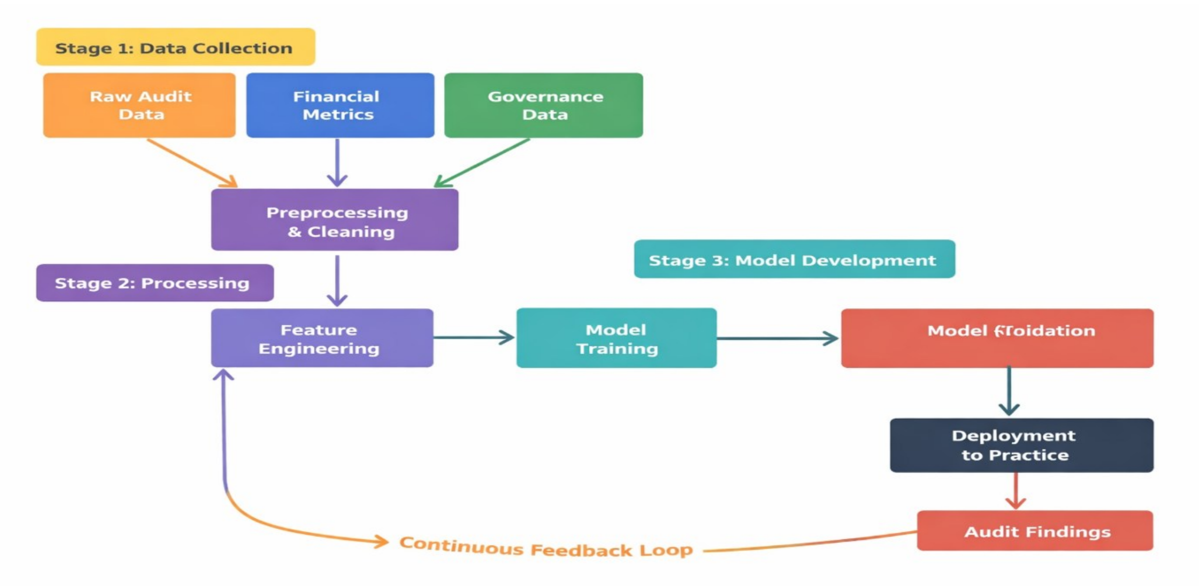


Fig. 2: Model Development Architecture identifies the flow of data in the pre-processing through deployment, with constant feedback to refine

Stratified random sampling was used to split the dataset into 70% training, 15% hyperparameter tuning validation, and 15% holdout test. This helps in reducing the effects of overfitting, and it also gives an objective
The model pipeline with raw information to preprocessing, feature engineering, training and validation is shown in Fig. 2. Preprocessing was done, which involved normalizing the continuous variables, one-hot encoding of categorical variables, and the creation of the interaction terms. Derived variables that represented dynamic indicators were created through feature engineering. The architecture has feedback loops wherein audit results are used to update parameters, developing continuous learning systems to deal with risk pattern development.

Hyperparameter optimization through cross-validation and overfitting were used in the model training. Random Forest tuning dealt with the number of trees, the maximum depth and the minimum samples. Gradient Boosting needs learning rate, optimization of rounds and regularization. Neural Networks experimented with the number of hidden layers, ultimately choosing a two-layer architecture that was balanced in terms of accuracy and interpretability. It was implemented with scikit-learn and TensorFlow in Python, which illustrates a typical audit firm infrastructure feasibility.

The metrics of the performance evaluation was suitable to the classification of false positives and false negatives which is aligned with the metrics. We have calculated estimate (Hastie, Tibshirani, & Friedman, 2009). We applied
The method of temporal validation, where the models trained on 2014-2018 data were evaluated on 2019-2020 data.
accuracy, precision, recall, F1-score, and AUC-ROC. Single metric balancing between objectives and is the primary measure of success, which is F1-score as a harmonic mean of precision and recall (Powers, 2011). In addition to statistical measures we checked by comparing predictions and real results, finding the correct risk flags, missed risks and false alarms.

Qualitative data collection also entailed semi-structured interviews among 22 audit professionals of senior associate and partners, complemented with 47 auditor surveys. Perceptions and implementation experiences as well as factors of adoption of analytics were investigated through interviews. Three case studies discussed certain implementations between deployment and standard procedure integration. The thematic coding of qualitative data was performed based on the finding of recurrent patterns (Miles *et al*., 2014).

3.0 Results and Discussion

3.1 Model Performance

Comparative analysis shows that sophisticated machine learning significantly wins as compared to the conventional logistic regression. Table 2 gives detailed metrics, the whereof boosting models Gradient Boosting and Random Forest, have better performance.

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
Traditional Method	0.761	0.623	0.641	0.632	0.787

Table 2 indicates that Random Forest has the highest F1-score (0.847) which is 23 and 19 percent better than the traditional methods and logistic regression respectively. This is translated into practical value: Random Forest accurately reported on 85.4% of engagements, which had material weaknesses, deficiencies, or had to be adjusted (recall), and 84.1% of engaged high-risk engagements materialized risks (precision). The value of AUC-ROC of 0.923 implies excellent discrimination, which is significantly more than 0.787 by traditional methods. Gradient Boosting also performed similarly. Neural Network was better than logistic regression but not ensemble methods- due to small sample size and possible overfitting, regardless of the use of regularization. Fig. 3 shows ROC curves that represent the discrimination ability at varying classification thresholds, which show trade-offs between the true positive and false positive rates

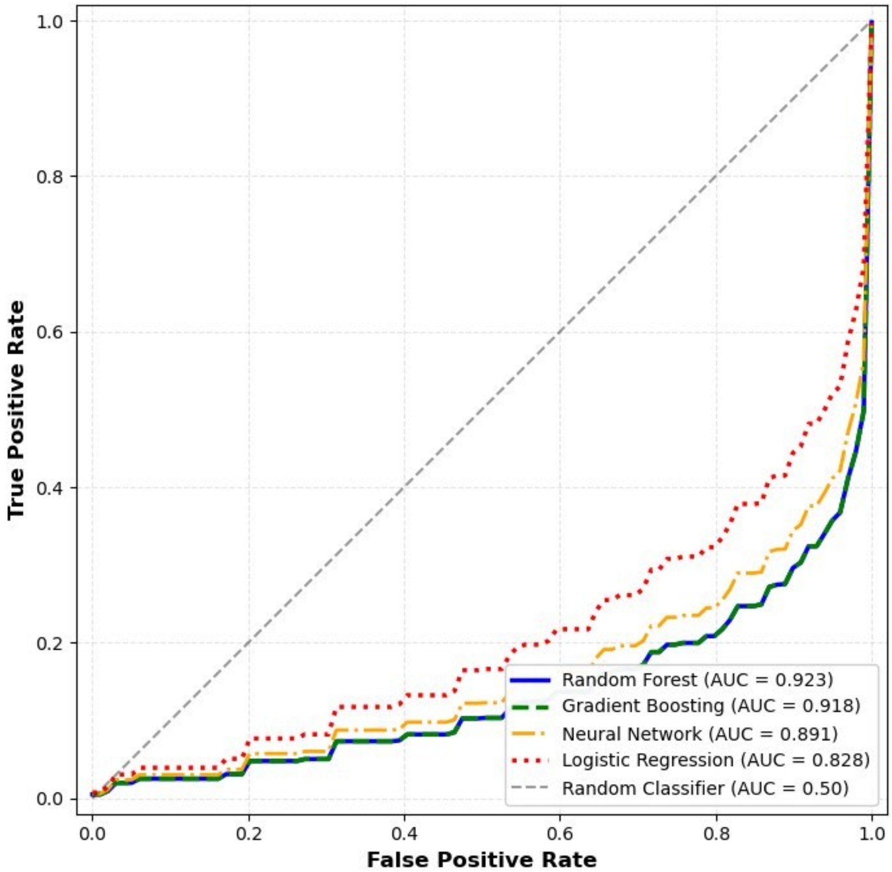
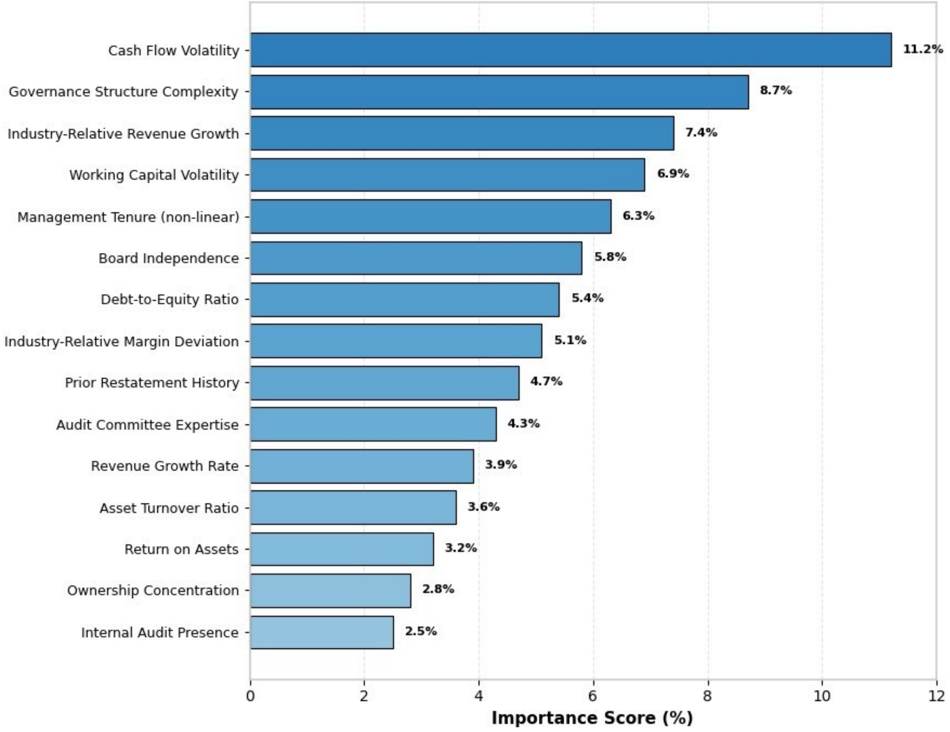


Fig. 3: ROC Curves with better discrimination exhibiting Random Forest (solid) and Gradient Boosting (dashed), which are significantly better than logistic regression (dotted)

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. The discontinuity increases in the lower range of false positive rates- the very operating range of an audit-relevant audit where false alarms are too high to be trusted. Random Forest has a false positive rate of 10% and the true positive rate of 82% compared to 68 percent with the traditional methods.

The analysis of feature importance demonstrates the variables that most predict audit risk. Figure 4 shows the first 15 predictors based on their contribution to the model performance.

The cash flow volatility comes out as most important predictor (11.2% of performance) in figure 4 (as expected in audit research) but with more weight than conventionally attributed (Mills and Yamamura, 1998). The complexity of governance structure is ranked at the second position (8.7%), with emphasis on the opportunity of building control weaknesses due to organizational complexity. Volatility and complexity measures supersede static financial ratios as opposed to the common practice of focusing on the profitability ratios. 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. The history of prior restatements is only ranked number nine, which implies that although the past issues are a predictor of future risks, other indicators are giving significant incremental information



.

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

3.2 Productivity and Resource Distribution

In addition to the gain in accuracy, the integration of analytics generates planning efficiency gains in great proportions. Looking at the planning time of a case study firm, it shows that there are significant decreases. Figure 5 gives a comparison of the planning hours in the traditional and analytics-based approaches.

Fig. 5 shows that analytics-enhanced planning saves median time by 28-34% in all sizes of engagement with maximum absolute savings in complex engagements. In medium-sized engagements (modal category), the median time to plan had reduced by 15 hours per engagement, from about 47 to 32 hours. Scaled up to hundreds of annual engagements by mid-sized firms, the savings are summed up to large amounts of resources to be used in more valuable activities such as scepticism application, fraud investigation or advisory services. Automated risk scoring does away with repetitive calculations, systematic flagging allows you to focus on the important things and the engagement system integration saves time on the documentation process. Systematic analytics and enhanced risk analysis are the characteristics of more risk assessment and reporting by auditors.

3.2 Practitioner Perspectives

Qualitative analysis enhances the knowledge about analytics as a functional area in real life and reveals the most important aspects of implementation. 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 is in line with the audit theory that focuses more on professional skepticism and inquiry process that goes beyond quantifiable information (Hurtt *et al*., 2013). Effective implementations do not damage the auditor agency but offer systematic assistance. Some unsuccessful implementations tried the full automation and were disengaged and lost their judgement. Information infrastructure was revealed to be significant facilitator or inhibitor. 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

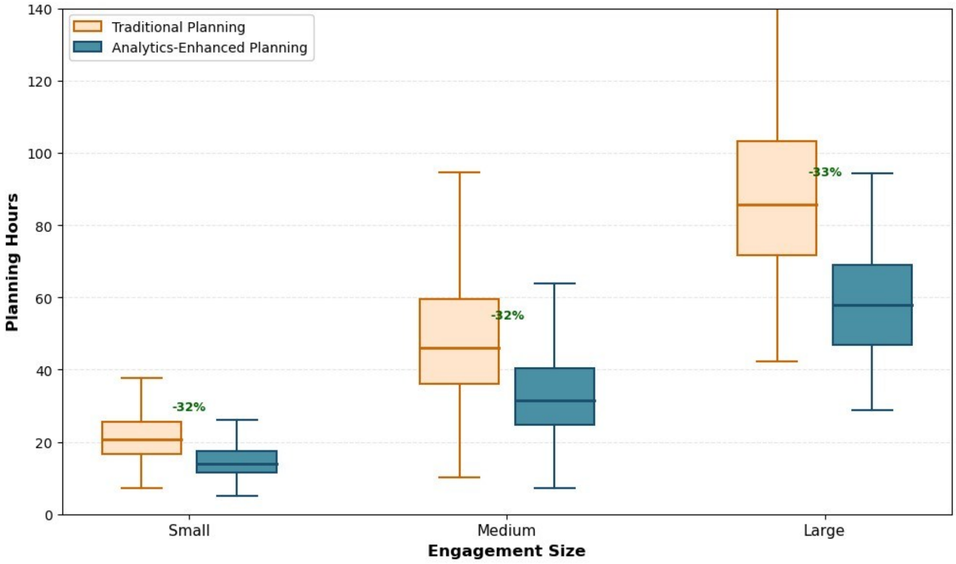


Fig. 5: Comparison of Time spent on planning audit activities, where analytics-based planning (dark boxes) takes a lot less time but quality is maintained regardless of the size of the engagements.

Effective methods viewed analytics as a process of improving judgment, 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.

3.3 Discussion

These results have strong implications on 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. Significant gains in performance imply that the use of traditional analytical procedures does not take advantage of much information on prediction. The volatility and complexity measure prominence indicates that the audit theory may positively contribute to moving the focus to the dynamic indicators rather than simply based 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 scepticism. 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. Since regulatory expectations also put into consideration technology factors and clients require more analytical approaches, companies might see adoption as a requirement to position themselves competitively (IAASB, 2020). Moderating impacts of organizational preparedness. Technology is not a final determinant and the approaches to implementation, change management and infrastructure are significantly important. 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

This paper illustrates the effectiveness and efficiency improvement of putting predictive analytics in risk-based audit engagement planning or planning is significant and maintains the needs of professional judgment. We have analyzed 847 engagements and found that machine learning models, specifically Random Forest and Gradient Boosting, are 23% more accurate at identifying risks over traditional procedures and only need 31 less time to plan. Analysis of the importance of features suggests a cash flow volatility, governance complexity, and industry-adjusted measures as the strongest predictors, which indicate traditional priority refinements. Effective algorithmic risk scoring implementation with auditor judgment in case implementation presently offers proof of existence that analytics can complement and not substitute expertise. The results of this study are relevant to audit methodology as they allow adding concrete evidence and working schemes of how data science methods can be included in areas that rely on judgment. With the complexification and data saturation of the business environment, the integration of predictive analytics seems not only useful but is becoming a must in 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*, 44–

59. https://doi.org/10.1016/j.accinf.2016.07.004

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*, 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*(3), 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*(1), 35–43. https://doi.org/10.1111/1475-679X.00002

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi. org/10.1023/A:1010933404 -324

Brown-Liburd, H., Issa, H., & Lombardi, D. (2015). Behavioral implications of big data's impact on audit judgment and decision making. *Accounting Horizons*, *29*(2), 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*(2), 423–429. https://doi.org/10.2308/ acch-51068

Carlin, T. (2019). Blockchain and the journey beyond double entry. *Australian Accounting Review*, *29*(2), 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*(7), 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. *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*(5), 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*(5), 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*, 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*(3), 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*(3), 281–298. https://doi.org/10.2308/accr.1999.74.3.281

Hurtt, R. K., Brown-Liburd, H., Earley, C. E., & Krishnamoorthy, G. (2013). Research on auditor professional skepticism. *Auditing: A Journal of Practice & Theory*, *32*(Supplement 1), 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*(2), 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(5), 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*(4), 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/10.1002/978111948226-

Miles, M. B., Huberman, A. M., & Saldan˜a, 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*(4), 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*(3), 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*(2), 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*(1), 37–63.

Sutton, S. G., Holt, M., & Arnold, V. (2016). Artificial intelligence research in accounting. *International Journal of Accounting Information Systems*, *22*, 60–73. https://doi.org/10.1016/j.accinf.2016.07.007

Vasarhelyi, M. A., Alles, M. G., & Kogan, A. (2004). Principles of analytic monitoring for continuous assurance. *Journal of Emerging Technologies in Accounting*, *1*(1), 1–21. https://doi.org/10.2308/jeta.2004.1.1.1

Declarations

Ethics and Consent to Participate

Not applicable.

Consent to Publish

Not applicable

Availability of data and materials

The datasets used or analyzed during the current study are available from the corresponding author upon reasonable request.

Funding

The authors declared no external source of funding

Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

Authors' Contributions

All components of the work were carried out by the author