Engineering and Technology Journal e-ISSN: 2456-3358
Volume 09 Issue 10 October-2024, Page No.- 5370-5378
DOI: 10.47191/etj/v9i10.17, I.F. – 8.227

Evaluating The Performance of Machine Learning Models in Audit Opinion Prediction – A Study in Vietnam

Dang Dinh Tan
Ho Chi Minh University of Banking

ABSTRACT: This study investigates the effectiveness of machine learning models in predicting audit opinions using a dataset from the FiinPro-X platform, comprising 9,783 audited consolidated financial statements from public companies listed on Vietnamese stock exchanges from 2016 to 2023. The dataset spans various industries, excluding banks and financial institutions, and focuses on identifying key financial, non-financial, and qualitative variables that influence audit opinions. Six supervised learning algorithms were applied—Logistic Regression, K-Nearest Neighbors (KNN), Decision Trees, Random Forests, Support Vector Machines (SVM), and Naive Bayes—evaluated based on their ability to predict both fully acceptable (unqualified) and non-fully acceptable audit opinions. All data processing and model training were implemented in a Python environment. The Random Forest model demonstrated the best overall performance, achieving an accuracy of 0.868 and an AUC-ROC of 0.87, though its F1 score for predicting non-fully acceptable audit opinions was lower (0.585). This suggests that while machine learning models can improve prediction accuracy, challenges remain in handling imbalanced data and non-linear relationships among input variables. The study also reduced the number of features by 30%, improving the models' performance. Future research should further refine data and feature construction processes to ensure comparability and practical applicability.

KEYWORDS: Audit opinion prediction, machine learning, Random Forest, financial statements, Vietnam

1. INTRODUCTION

In the era of digital transformation, the integration of data science methods, particularly data mining based on machine learning algorithms, into auditing activities is becoming increasingly urgent. As businesses digitize, the volume and complexity of financial data are growing rapidly, rendering traditional auditing methods less effective. Previous studies have demonstrated that data mining techniques using machine learning algorithms can predict audit opinions (e.g., unqualified audit opinions or non-unqualified opinions such as qualified opinions, adverse opinions, or disclaimers of opinion) by training models on historical data. This capability is crucial as the demand for timely and accurate financial information increases, especially with the rise of continuous and real-time auditing. Machine learning models not only reduce the time required for auditors to reach conclusions but also enhance transparency and audit quality. However, previous studies have identified several limitations that need to be addressed (Stanišić et al., 2019).

First, most research focuses primarily on data from large-scale listed enterprises. For example, in Vietnam, studies typically use data from companies listed on the HOSE, while there is a lack of studies involving small and medium-sized enterprises or those listed on other exchanges. Second, the research time frame is often limited to 3-5 years, which is insufficient to capture economic and financial fluctuations. Third, the complexity of machine learning models poses challenges for auditors in applying and explaining the results, particularly for those unfamiliar with the terminology of data science and artificial intelligence. This creates a challenge for auditors who must adhere to auditing standards that emphasize transparency and the clear explanation of audit results.

To address these limitations, this study expands the dataset both temporally and geographically. It includes consolidated financial statement data from listed companies on all three major exchanges in Vietnam (HOSE, HNX, UPCoM) during the 2016-2023 period. Additionally, the study employs optimal feature selection methods to minimize model training costs and make it easier for auditors to verify and explain the results. The findings not only assist auditors in using traditional accounting and auditing terminology to evaluate the models but also open up possibilities for the practical application of machine learning tools in the auditing process.

2. LITERATURE REVIEW AND RESEARCH METHOD

2. 1 Literature review

Since the 2000s, numerous academic studies have significantly advanced the understanding of how data mining techniques, particularly those based on machine learning algorithms, can improve the accuracy of audit opinion predictions and enhance the overall efficiency of audits (Stanišić et al., 2019; Barr-Pulliam et al., 2022). Early work by Kirkos et al. (2007) demonstrated the use of data mining classification algorithms, including C4.5 Decision Tree, Multilayer Perceptron Neural Network, and Bayesian Belief Network, to assist auditors in distinguishing between unqualified and modified audit opinions. Using a dataset of 450 financial statements from non-financial companies in the UK and Ireland, this study showed how both qualitative and quantitative variables, such as profitability ratios, firm size, and audit firm type, could be used to predict non-unqualified audit reports.

Saif et al. (2012) furthered this line of research by exploring the applicability of advanced data mining techniques, specifically combining Support Vector Machines (SVM) and Decision Tree models, to improve audit opinion prediction accuracy and interpretability. Analyzing 1,018 observations from publicly listed Iranian companies, the study introduced 30 specific decision rules linking financial criteria to audit opinions, providing a rule-based framework for auditors to evaluate financial statements. The interpretability of these decision rules enhanced practical application for auditors. Similarly Yaşar et al. (2015) focused on predicting non-unqualified audit opinions using discriminant analysis, logistic regression, and Decision Tree C5.0 models based on twelve financial ratios. Their findings, based on 110 observations from Turkish companies, revealed that the "retained earnings to total assets" ratio was the most effective predictor across all models, with the C5.0 decision tree achieving a 98.2% accuracy rate.

More recent studies have introduced innovative approaches to address gaps in previous research. Sánchez-Serrano et al. (2020) developed a multilayer perceptron artificial neural network model to predict audit opinions for consolidated financial statements, filling a gap where previous studies focused only on parent company reports. The study, which included 298 Spanish corporations from 2017, achieved over 86% accuracy and highlighted differences between predictors for consolidated versus parent company statements, using a combination of financial, non-financial, and qualitative variables. Zarei et al. (2020) also examined the effectiveness of financial and non-financial variables in predicting non-unqualified audit opinions for companies listed on the Tehran Stock Exchange, using a probit model to analyze 480 observations. The study demonstrated that financial ratios and audit firm type significantly influence the likelihood of receiving a non-unqualified opinion, with a 72.9% prediction accuracy.

Zeng et al. (2022) introduced a novel combination of Batch Sparse Principal Component Analysis (BSPCA) and Kernel Fuzzy Clustering (KFCM) with SVM to address issues like class imbalance and feature selection in audit opinion prediction models. By categorizing 448 variables into financial indicators, corporate governance, and market transactions, the study provided a comprehensive assessment of corporate performance and significantly improved classification accuracy. The SKFCM-SVM model outperformed other methods, showing its potential for broader financial predictions beyond audit opinions.

Expanding the scope to the Vietnamese market, Pham Thi Thu Oanh et al. (2024) studied factors affecting audit opinions of non-financial companies listed on the Vietnamese stock exchange. Their model, using machine learning algorithms like Decision Trees and Random Forest, achieved a 97% accuracy rate, with the previous year's audit opinion emerging as the most critical predictor for the current year's opinion. This research provided valuable tools for auditors, helping them focus on high-risk companies and reduce costs.

Despite these advances, several research gaps remain. Many studies focus on relatively small sample sizes and short time periods, limiting the generalizability of their findings. For instance, datasets often cover only 3-5 years, which is insufficient to capture long-term economic fluctuations or changing business environments. Furthermore, most research is centered on large, publicly listed companies, excluding small and medium-sized enterprises (SMEs) that might face different financial reporting challenges. Another key issue is the complexity and interpretability of machine learning models. While techniques such as Decision Trees provide transparency, more advanced methods like neural networks and SVM, despite their accuracy, are often opaque and difficult for auditors without data science expertise to interpret. This lack of interpretability raises concerns about the practical application of these models in audit settings, where transparency and compliance with auditing standards are crucial. Future research should address these gaps by expanding data scope, including SMEs, extending the time frames studied, and enhancing the transparency of machine learning models to improve their practical utility in auditing.

2.2 Research method

2.2.1 Research samples

The data for this study was collected from FiinPro-X [1], the latest version of the FiinPro Platform, a core provider of financial data in Vietnam since 2015. Consistent with previous studies, this research focuses on audited consolidated annual

[1] https://fiinpro.com/fiinpro-x

financial statements, excluding banks, financial companies, and insurance firms. However, to meet the study’s objectives, the dataset includes audited consolidated annual financial statements from 2016 to 2023 for public companies in various sectors listed on the HOSE, HNX, and UPCoM exchanges [2]. The industry classification is based on the Level I classification used by these exchanges, covering the following sectors: “Industry,” “Information Technology,” “Pharmaceuticals and Healthcare,” “Oil and Gas,” “Consumer Services,” “Consumer Goods,” “Raw Materials,” “Community Utilities,” “Finance” (limited to real estate companies), and “Telecommunications.” After data pre-processing, the study collected 9,783 observations (audited consolidated financial statements) by industry for the 2016–2023 period, as shown in Table 1.

Table 1: Description of annual observation sample by industry

Year	**Industry**	**IT**	**Pharmaceuticals and Health care**	**Oil and Gas**	**Consumer Services**	**Consumer Goods**	**Raw Materials**	**Community Utilities**	**Finance (Real estate)**	**Telecommunications**	**Total**
2016	346	21	43	8	78	142	119	83	71	5	916
2017	499	26	50	10	103	177	152	117	93	8	1235
2018	484	25	46	10	100	184	155	121	92	7	1224
2019	495	23	52	10	107	196	161	128	96	8	1276
2020	489	23	54	10	109	195	154	129	94	8	1265
2021	495	23	54	11	110	203	156	130	100	8	1290
2022	495	21	55	12	111	202	160	136	101	8	1301
2023	476	19	55	11	109	208	158	137	95	8	1276
Total	3779	181	409	82	827	1507	1215	981	742	60	9783

2.2.2 Machine learning models used in research

This study employs a range of supervised learning algorithms to construct models for predicting audit opinions, including Logistic Regression, K-Nearest Neighbors (KNN), Decision Trees, Random Forests, Support Vector Machines (SVMs), and Naive Bayes. These algorithms were chosen for their varied methodologies and their ability to handle different data types, offering a comprehensive analysis of the factors influencing audit opinions.

- **Logistic Regression** is a linear model used for binary classification, providing the probability that a given input belongs to a specific class.
- **K-Nearest Neighbors (KNN)** is a non-parametric algorithm that classifies data based on the majority class among its “k” nearest neighbors in the feature space.
- **Decision Trees** generate a hierarchical structure of decision rules, which are easily interpretable and effective in capturing non-linear relationships between features and the target variable.
- **Random Forests** are an ensemble method that combines multiple decision trees to improve the accuracy and stability of predictions.
- **Support Vector Machines (SVMs)** are particularly effective for classification tasks in high-dimensional spaces, finding the hyperplane that maximizes the margin between support vectors from different classes.
- **Naive Bayes** is a probabilistic classifier based on Bayes’ theorem, operating under the assumption of independence between input features.

2.2.3 Evaluation methods of machine learning models

To assess the performance of these models, several evaluation metrics are employed, each offering unique insights into model effectiveness:

- **Confusion Matrix**: This matrix summarizes prediction results by showing true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN).
  - **Accuracy**: Accuracy measures the ratio of correctly predicted observations to the total observations, calculated as (TP + TN) / (TP + TN + FP + FN).

[2] Vietnam currently has three stock exchanges: the Ho Chi Minh City Stock Exchange (HOSE), established in 2000; the Hanoi Stock Exchange (HNX), established in 2005; and the UPCoM (Unlisted Public Company Market), which began trading in 2009 and is managed by HNX. All securities traded on HOSE, HNX, and UPCoM are registered with the Vietnam Securities Depository (VSD).

- **Precision**: Precision is the ratio of true positives to the total number of predicted positives, calculated as TP / (TP + FP).
- **Recall**: Recall (also known as sensitivity) measures the ratio of true positives to the total actual positives, calculated as TP / (TP + FN).
- **F1 Score**: The F1 score is the harmonic mean of precision and recall, calculated as 2 × (Precision × Recall) / (Precision + Recall). It balances precision and recall, making it especially useful for imbalanced datasets.
- **AUC-ROC (Area Under the Receiver Operating Characteristic Curve)**: AUC-ROC evaluates the performance of classification models, particularly in binary classification. A model with an AUC-ROC closer to 1 demonstrates higher predictive ability. Like the F1 score, AUC-ROC is beneficial in handling imbalanced datasets.

2.2.5 Models variables

a. Output (label) variable

The output variable in this study is a categorical variable representing the type of audit opinion. Auditors can issue two types of opinions based on their evaluation of sufficient and appropriate audit evidence: (1) an unmodified audit opinion, or (2) a modified audit opinion, which encompasses three specific types: (a) qualified opinion, (b) adverse opinion, and (c) disclaimer of opinion (IAASB, 2022). Table 2 presents the distribution of observations by audit opinion type for the annual consolidated financial statements.

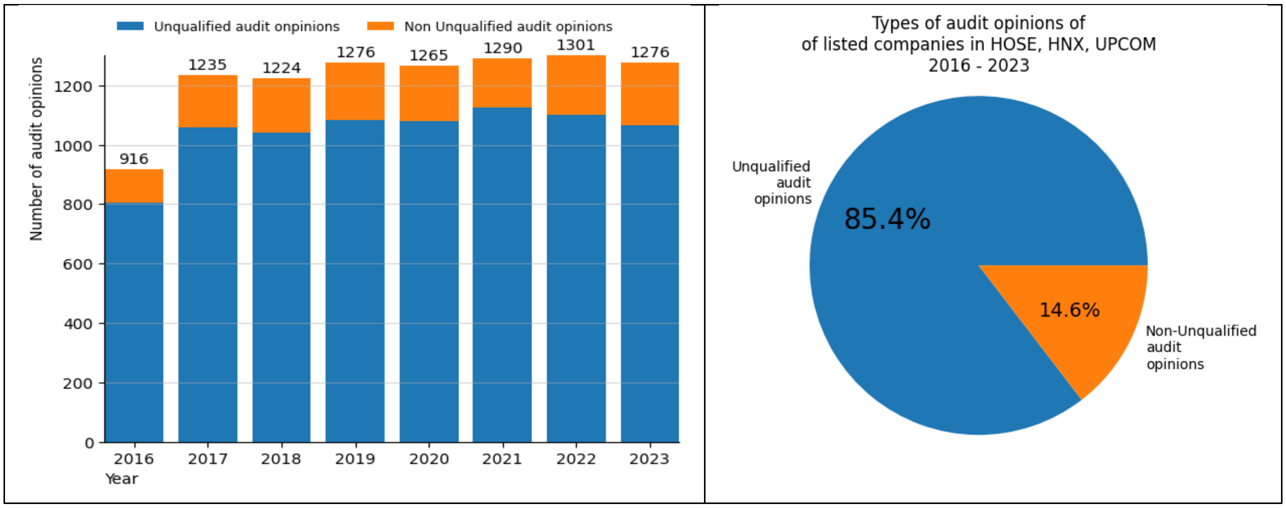
Table 2: Description of the sample observations by type of annual audit opinion

Year	Unqualified opinions	Unqualified opinions with emphasis	Qualified	Adverse opinions	Disclaimer of opinions	Total
2016	719	85	104	1	7	916
2017	952	108	157	1	17	1235
2018	936	104	163	0	21	1224
2019	972	110	173	0	21	1276
2020	995	85	155	1	29	1265
2021	1042	82	142	0	24	1290
2022	1002	100	166	1	32	1301
2023	1040	25	177	1	33	1276
Total	7658	699	1237	5	184	9783

Both practice and sample analysis show that audit opinion types such as “qualified,” “adverse,” and “disclaimer of opinion” consistently represent a small proportion of the total, resulting in data imbalance. This imbalance complicates the development of accurate machine learning models. To address this issue, many previous studies have grouped these opinion types into a single category called “unqualified opinions.” This approach increases the sample size of the smaller category, reducing the effects of data imbalance and improving model accuracy (Zeng et al., 2022).

Furthermore, this grouping is highly practical for stakeholders, such as investors and auditors, who are generally more concerned with whether a company has received an unqualified opinion rather than differentiating between types of unqualified opinions. By simplifying the classification into two groups— “unqualified” and “non-unqualified” —the decision-making process becomes more straightforward, and the information more accessible for practical use (Saif et al., 2012). Figure 1 presents a comparison of the proportion of unqualified audit opinions versus other opinions from 2016 to 2023 after data processing.

Figure 1. Comparison of types of annual audit opinions



b. Input variables (features)

As discussed, previous studies in this field demonstrate significant diversity and, consequently, inconsistency in the selection of input features for audit opinion prediction models. While this diversity has enhanced understanding of the role of independent variables in prediction models, it also limits practical applicability. To address this issue, the present study builds upon the feature selection approaches of previous research, incorporating financial, non-financial, and qualitative features. However, instead of pre-selecting features, this study employs a "gradual elimination" method, beginning with a comprehensive "prototype" set of all financial coefficients related to the annual consolidated financial statements of companies, collected from the Fiinpro-X platform.

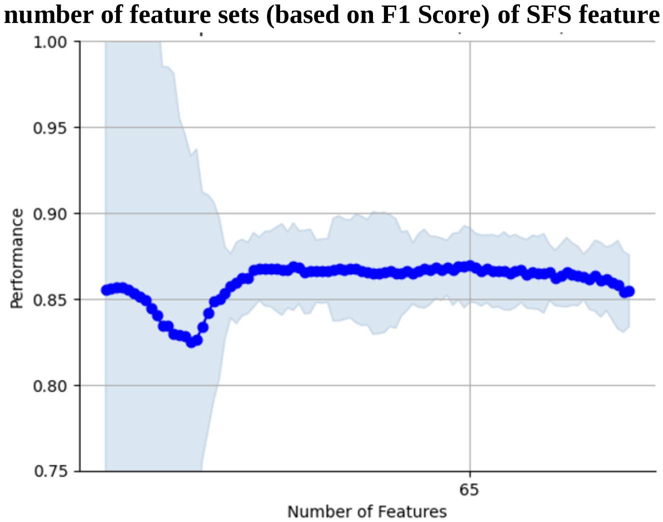
The practical advantage of this method is that when auditors and audit firms train predictive models using reliable and consistently processed data from commercial sources (with copyright protection), the resulting model predictions are more likely to meet auditing standards for the sufficiency and appropriateness of audit evidence.

3. RESEARCH RESULTS AND DISCUSSIONS

3.1 Selection of input variables (features)

The Sequential Forward Selection (SFS) algorithm, in conjunction with the Decision Tree algorithm, identified 65 optimal features (highlighted in bold in Table 3) from an initial set of 93 features (Figure 2), based on the F1 Score metric.

Figure 2. Optimal number of feature sets (based on F1 Score) of SFS feature selection algorithm



Nguồn: Kết quả nghiên cứu.

The number of features in the models has been significantly reduced to approximately 70% of the original set, which decreases the cost of training the models. Additionally, the importance of these 65 selected variables (highlighted in bold) provides valuable insights into the financial, non-financial, and qualitative indicators that influence audit opinions, reflecting the true and fairness of the consolidated financial statements of public companies in Vietnam (Table 3).

Table 3: List of input variables (features) and features selected by the SFS algorithm

X1: Revenue/share X2: Cash flow generated/share X3: Basic EPS X4: Diluted EPS **X5: Free cash flow to the enterprise (FCFF)** **X6: Free cash flow (FCF)** X7: Basic P/E **X8: Diluted P/E** X9: P/B **X10: P/S** **X11: P/Tangible Book** X12: Enterprise value/Sales X13: Enterprise value/EBITDA X14: Enterprise value/EBIT **X15: Enterprise value (EV)** X16: ROE % X17: ROA % X18: ROIC % X19: ROCE % **X20: Receivables turnover ratio** **X21: Inventory turnover ratio** **X22: Payables turnover ratio** **X23: Cash payment ratio face** **X24: Quick ratio** **X25: Current ratio** **X26: Long-term debt/Equity** **X27: Long-term debt/Total assets** X28: Total debt/VASCH **X29: Total debt/Total assets** **X30: Short-term debt/Equity** **X31: Short-term debt/Total assets** **X32: Total debt/Equity** **X33: Total debt/Total assets** **X34: Intangible asset ratio/Total assets** **X35: Total assets/Equity** **X36: Revenue/Total assets** **X37: EBIT** **X38: EBITDA**	X39: Gross profit margin X40: EBITDA margin X41: EBIT margin **X42: Pre-tax profit margin** **X43: Net profit margin** **X44: Asset turnover ratio** **X45: Equity turnover ratio** **X46: Operating profit margin** **X47: (NET - CFO)/ Revenue** **X48: Charter capital growth (YoY) (Quarter, Year)** **X49: Net revenue growth (YoY) (Quarter, Year)** **X50: Parent company shareholder profit growth (YoY) (Quarter, Year)** **X51: Gross profit growth (YoY) (Quarter, Year)** **X52: EBITDA growth (YoY) (Quarter, Year)** **X53: EBIT growth (YoY) (Quarter, Year)** **X54: Pre-tax profit growth (Quarter, Year)** **X55: Capital Expenditure Growth (CAPEX) (YoY) (Quarter, Year)** X56: Total Asset Growth (YoY) (Quarter, Year) X57: Equity Growth (YoY) (Quarter, Year) X58: 3-year net revenue growth - CAGR (Y) X59: 5-year net revenue growth - CAGR (Y) **X60: 3-year gross profit growth - CAGR (Y)** **X61: 5-year gross profit growth _CAGR (Y)** **X62: 3-year pre-tax profit growth - CAGR (Y)** X63: 5-year pre-tax profit growth - CAGR (Y) **X64: 3-year net profit growth - CAGR (Y)** X65: 5-year net profit growth - CAGR (Y)	X66: 3-year EBIT growth - CAGR (Y) **X67: 5-year EBIT growth - CAGR (Y)** **X68: 3-year EBITDA growth - CAGR (Y)** **X69: 5-year EBITDA growth - CAGR (Y)** **X70: 3-year total assets growth - CAGR (Y)** **X71: 5-year total assets growth - CAGR (Y)** X72: 3-year equity growth - CAGR (Y) X73: 5-year equity growth - CAGR (Y) **X74: 3-year charter capital growth - CAGR (Y)** **X75: 5-year charter capital growth - CAGR (Y)** **X76: 3-year fixed asset investment (CAPEX) growth - CAGR (Y)** X77: 5-year fixed asset investment (CAPEX) growth - CAGR (Y) **X78: EPS growth (YoY)** X79: % of planned revenue X80: % of planned profit before tax X81: % of planned profit after tax **X82: HOSE** **X83: UPCoM** **X84: auditor_Non Big4** **X85: Sector - ICB L1_Information Technology** **X86: Sector - ICB L1_Pharmaceuticals and Healthcare** **X87: Sector - ICB L1_Oil and Gas** **X88: Sector - ICB L1_Consumer Services** **X89: Sector - ICB L1_Consumer Goods** **X90: Sector - ICB L1_Raw Materials** **X91: Sector - ICB L1_Community Utilities** **X92: Sector - ICB L1_Finance** **X93: Sector - ICB L1_Telecommunications**

The feature selection results align with previous studies, which also identified both quantitative and qualitative variables as significant predictors of audit opinion (Stanišić et al., 2019). Notably, the algorithm highlighted company efficiency and performance growth as important factors for determining audit opinions and the accuracy of consolidated financial statements in Vietnam. Furthermore, this method preserves the original features (reflecting various performance indicators), rather than abbreviating or synthesizing them. This allows auditors to more easily assess the relationship between these input features and the resulting audit opinion, facilitating professional judgment in accounting and auditing.

3.2 Models evaluation

The training dataset, comprising 90% of the observations and corrected for class imbalance between fully acceptable audit opinions and other opinions using the SMOTE method, was utilized to train models with the 65 selected features. Classification algorithms applied included logistic regression, KNN, Decision Tree, Random Forest, Support Vector Machine, and Naïve Bayes. Grid search, with five-fold cross-validation, was employed to optimize model parameters, with the F1 score used as the performance metric (Table 4).

Table 4: Results of finding optimal parameters of the models

Classifier	Optimum parameters with 5-fold cross-validation	F1 scores
Logistic Regression	C:10. max_iter:500. solver: liblinear	0.747
KNN	n_neighbors:11	0.855
Decision Tree	criterion: gini, max_depth: 17	0.869
Random Forest	criterion: gini, max_depth: 15, n_estimators: 200	0.929
SVM	kernel: poly	0.741
Naïve Bayes		0.678

The model with the optimal parameters for each algorithm was then used to predict and evaluate performance on the test dataset (10% of observations), comparing the AUC-ROC coefficient. Among the six models, the Random Forest model demonstrated the best overall performance, both for overall prediction accuracy (covering both majority and minority classes) and for the minority class (i.e., audit opinions other than unqualified opinions), making it the most optimal model for this dataset (Figures 3 and 4).

Figure 3. Comparison of models performances by F1 Scores

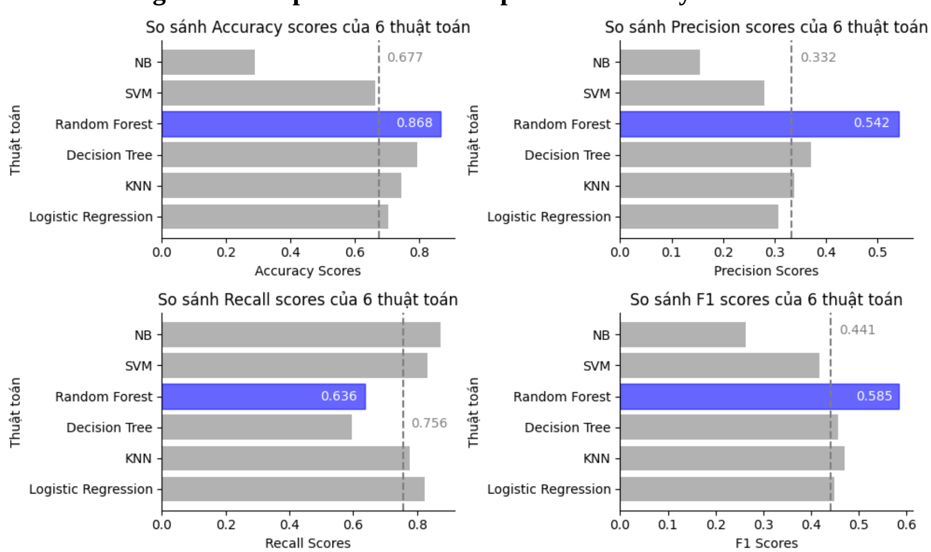
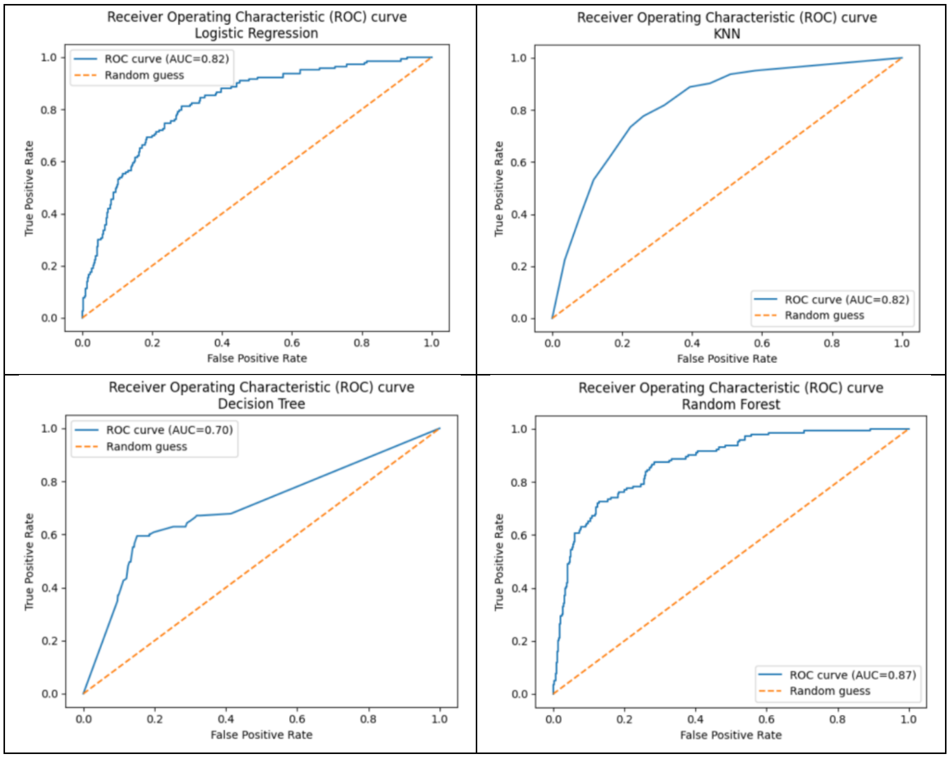


Figure 4. Comparison of models performances by AUC-ROC



The findings of this study align with previous research, where Decision Tree and Random Forest models were identified as the best performers for predicting audit opinions. This is likely due to the lack of linear correlations among the financial dataset's input variables (Saif et al., 2012; Pham Thi Thu Oanh et al., 2024). However, a notable difference is that, despite the Random Forest model achieving an overall prediction accuracy of 0.868 and an AUC-ROC of 0.87, its ability to accurately predict non-unqualified audit opinions only reached an F1 score of 0.585 (Figure 4). These differences from prior studies may stem from variations in the temporal and spatial scope of the data, as well as differences in feature construction techniques. Given the nature of machine learning approaches, future studies in this field should provide more detailed descriptions of data and feature construction processes to ensure comparability of results.

4. CONCLUSION AND RECOMMENDATIONS

The primary objective of an audit is to provide the auditor with reasonable assurance that the financial statements are free from material misstatement, whether due to fraud or error (IAASB, 2022). To achieve this, the auditor must plan and conduct the audit with professional skepticism, apply professional judgment, and adhere to ethical standards when forming an opinion on the financial statements. The auditor and the audit firm bear the responsibility of ensuring that the audit opinion aligns with the "substance" of the financial statements, and of minimizing audit risk to an acceptably low level. This requires comprehensive audit planning, thorough risk assessment, effective audit procedures, and strict compliance with auditing standards and ethical guidelines.

In this context, machine learning models offer valuable tools for auditors, particularly in the increasingly digital business environments of audited entities. This study developed machine learning models to predict audit opinions using data from the Fiinpro-X platform, which included 9,783 observations (based on audited consolidated financial statements) from 1,454 publicly traded companies in Vietnam, spanning various industries from 2016 to 2023 (Table 1). Of these observations, 85.4% represented fully acceptable audit opinions, while 14.6% were non-fully acceptable (Figure 1).

The research supports the view that Decision Tree and Random Forest models remain the most effective for predicting audit opinions, especially in cases where financial characteristics lack strong linear correlations. However, this study revealed that while the Random Forest model achieved high overall accuracy (0.868) and an AUC-ROC of 0.87, its F1 score for predicting non-fully acceptable audit opinions was only 0.585. This discrepancy may be due to differences in the temporal and spatial scope of the data, as well as feature construction techniques. Future research should clearly describe data and feature construction methods to ensure comparability and feasibility of results.

Additionally, this study successfully reduced the number of features used in the machine learning models by 30%, leading to cost savings in model training. It identified 65 key financial, non-financial, and qualitative variables that are important for predicting audit opinion types. The feature selection process aligned with previous studies, highlighting the significance of quantitative and qualitative variables related to business performance and growth. A notable advantage of this method is that it preserves the original input features, facilitating auditors’ ability to evaluate and present the relationship between business performance indicators and audit opinion types in a clear and applicable manner.

In conclusion, this study not only verifies the effectiveness of machine learning models on a dataset of Vietnamese publicly traded companies but also presents a feature construction method that retains the meaning of the original data. This approach enables auditors and audit firms to better explain and evaluate the predictions generated by machine learning models, while also presenting these results transparently in audit reports.

5. REFERENCES

1. Barr-Pulliam, D., Brown-Liburd, H. L., & Munoko, I. (2022). The Effects of Person-specific, Task, and Environmental Factors on Digital Transformation and Innovation in Auditing: A Review of the Literature. Journal of International Financial Management & Accounting, 33 (2):, 33(2), 337-374.e3
2. IAASB. (2022). Handbook of International Quality Management, Auditing, Review, Other Assurance,and Related Services Pronouncements. Retrieved from https://www.iaasb.org/publications/2022-handbook-international-quality-management-auditing-review-other-assurance-and-related-services
3. Kirkos, E., Spathis, C., Nanopoulos, A., & Manolopoulos, Y. (2007). Identifying Qualified Auditors’ Opinions: A Data Mining Approach. Journal of Emerging Technologies in Accounting, 4, 183-197. doi:10.2308/jeta.2007.4.1.183
4. Pham Thi Thu Oanh, Dang Ngoc Hung & Vu Thi Thuy Van. (2024). Forecasting audit opinions on financial statements: statistical algorithm or machine learning. Electronic Journal of Applied Statistical Analysis, 17(01), 133-152. doi:10.1285/i20705948v17n1p133
5. Saif, S. M., Sarikhani, M., & Ebrahimi, F. (2012). Finding rules for audit opinions prediction through data mining methods. European Online Journal of Natural and Social Sciences, 1(2), 28-36.
6. Sánchez-Serrano, J. R., Alaminos, D., & García-Lagos, F. (2020). Predicting Audit Opinion in Consolidated Financial Statements with Artificial Neural Networks. Mathematics, 8(1288). doi:10.3390/math8081288
7. Stanišić, N., Radojević, T., & Stanić, N. (2019). Predicting the Type of Auditor Opinion: Statistics, Machinelearning, or a Combination of the Two? 16(2), 1-58.
8. Yaşar, A., Yakut, E., & Gutnu, M. M. (2015). Predicting Qualified Audit Opinions Using Financial Ratios: Evidence from the Istanbul Stock Exchange. International Journal of Business and Social Science, 6(8(1)), 57-67.
9. Zarei, H., Yazdifar, H., Ghaleno, M. D., & Azhmaneh, R. (2020). Predicting auditors’ opinions using financial ratios and non-financial metrics: evidence from Iran. Journal of Accounting in Emerging Economies, 10(3), 425-446. doi:10.1108/JAEE-03-2018-0027
10. Zeng, S., Li, Y., & Li, Y. (2022). Research on Audit Opinion Prediction of Listed Companies Based on Sparse Principal Component Analysis and Kernel Fuzzy Clustering Algorithm. Mathematical Problems in Engineering. doi:10.1155/2022/4053916