JORMASS
Volume 11, Issue 2, Dec. 2025
JOURNAL OF RESEARCH IN MANAGEMENT AND SOCIAL SCIENCES
DOI: To be assigned

Machine Learning Algorithms for Detecting Tax Fraud: Application and Challenges in Nigeria

Obizuo, Chinwendu Judith, Ihendinihu, John Uzoma, Chigbo, Queendaline Ugochi, Alpheaus, Ogechi Eberechi, & Okoro, Chinonso Churchill
Department of Accounting, Michael Okpara University of Agriculture, Umudike
Corresponding author: obizuo.chinwendu@mouau.edu.ng

ABSTRACT:
The study focused on machine learning algorithms for detecting tax fraud especially its application and challenges in Nigeria. To achieve this survey research design was adopted. The population of the study is made up of 150 participants, including Data Analysts, Financial Accountants, Tax Consultants and Tax Auditors, who were chosen to offer a variety of perspectives on the practices of detecting tax fraud. The study adopted convenience sampling techniques. 150 questionnaire instruments were distributed evenly among the participants but only 115 were correctly completed and returned. The study used a descriptive field survey research design and adapted questionnaire as a research instrument. The data collected were analyzed using both descriptive statistics and simple regression analysis. The findings of objective one indicates that tax fraud is a significant issue in the Nigerian tax system, and that weak enforcement of tax laws and corruption among tax officials are the major causes of tax fraud in Nigeria. The findings from the analysis of objective two revealed that although machine learning techniques are seen to be effective tools for improving fraud detection, structural barriers hinder their broad implementation. These include limited access to quality data, ethical concerns, institutional resistance, insufficient government funding, and lack of regulatory clarity. Finally, the findings from the analysis of objective three demonstrate a strong and statistically significant effect of ML algorithms on improving tax fraud detection ($R^2 = 0.899$, $p < 0.001$), indicating that, when applied properly, technology adoption may significantly reduce fraud. Based on the findings, the study recommends that to guarantee that ML algorithms can operate with dependable input, the Nigerian government should make investments in digitizing tax records and enhancing data quality. The study also recommends that clear guidelines should be established to control the ethical use of AI in tax systems, with special emphasis to transparency, accountability, and privacy.

KEYWORDS: *Machine learning algorithms, Tax fraud detection, Tax administration and Tax policies*

MANUSCRIPT TYPE:
Research Paper

PUBLICATION DETAILS:
Received: XX Jul. 2025
Revised: XX Aug, XX Nov. 2025
Accepted: XX Nov. 2025

Publication College of Management Sciences, Michael Okpara University of Agriculture, Umudike Nigeria

INTRODUCTION

The current trend of international corporate globalization has led to an increase in tax fraud and significant challenges in detecting it. Tax fraud poses an imminent threat to a country's economic progress, particularly in developing countries like Nigeria where budgetary sustainability has long been threatened by tax evasion and misleading assertions. According to Nosiri et al. (2021) and Saxunova and Szarkova (2018), tax fraud is a criminal violation that carries penalties for intentionally avoiding taxes by making false claims or creating fake documents. The Federal Inland Revenue Service (FIRS) has regularly highlighted tax-related

irregularities, such as underreporting income, utilizing fraudulent documents, and neglecting to remit collected taxes (FIRS, 2023).The informal sector, inadequate infrastructure for tax compliance, and poor taxpayer data management have all contributed to Nigeria's systemic tax fraud, which costs the government billions of dollars in lost revenue annually (OECD, 2022; World Bank, 2023). Traditional methods of detecting tax fraud that tend to depend on manual audits and rule-based systems, are losing their effectiveness due to the complexity and volume of modern tax transactions (Chen & Lin, 2021). Machine learning (ML) algorithms are an innovative tool in this battle because of their ability to analyze big datasets, find hidden trends, and accurately prioritize high-risk scenarios.

Machine learning, a subfield of artificial intelligence, uses algorithms to analyze large amounts of data, identify patterns, and make judgments or predictions (Sarker, 2021). With the rise in digital transactions and data availability, machine learning (ML) offers revolutionary potential for tax fraud detection. Machine learning algorithms, as opposed to traditional rule-based systems, can identify complex patterns and anomalies in large datasets, enabling proactive and automated fraud detection. These algorithms, which can learn from historical data to predict and flag suspicious conduct, can improve the effectiveness of tax audits and investigations (Alves et al., 2022). Using machine learning (ML), which not only enhances fraud detection but also provides tools for better risk management, predictive analytics, and resource optimization, tax authorities should focus their efforts on where they can make the biggest difference (Olaniyi et al., 2023c; Olaniyi et al., 2023d).

One of the key benefits of using machine learning (ML) for tax fraud detection is its capacity to analyze enormous volumes of data and identify hidden patterns that may not be immediately apparent with conventional techniques. Models that can identify fraudulent activity in real time can be created by utilizing a variety of machine learning techniques, including anomaly detection, supervised learning, and unsupervised learning. These models can eventually enhance their detection skills by continuously learning and adapting to new fraud strategies. Support vector machines, decision trees, and random forests are a few examples of supervised learning algorithms that can be trained on historical data to identify trends connected to fraudulent transactions. These algorithms need labeled data to classify past transactions as either fraudulent or non-fraudulent. Once trained, these models can be used to predict the likelihood of fraud in new, unlabeled data. Unusual patterns or clusters of transactions that deviate from typical behavior can be found using unsupervised learning techniques like association rule mining and clustering. These methods are particularly helpful for identifying fraud categories that weren't previously identified because they don't require labeled data (Sharma & Kumar, 2021).

Tax authorities throughout the world, including those in the US and the UK, have used machine learning (ML)-based systems to increase the precision of fraud detection and reduce tax revenue losses (Melladoet al., 2021; HMRC, 2022). In Nigeria, however, machine learning is still a relatively new application in tax administration. Inadequate infrastructure, limited access to high-quality data, and legislative restrictions all hinder the deployment of advanced fraud detection systems. Nonetheless, the growing digitization of tax processes through initiatives like the Integrated Tax Administration System (ITAS) makes ML integration feasible (FIRS, 2023).

Despite several attempts aimed at improving tax collection and lowering fraud, tax evasion and manipulation remain persistent issues in Nigeria. Most of the fraud detection techniques used today rely on manual audits and rule-based systems, which are often reactive, time-consuming, and vulnerable to human error or corruption. This has led to significant revenue leakage, inefficient fraud detection, and delayed responses to suspicious activities.

Despite its success in identifying fraudulent activities in other sectors, such as finance and e-commerce, machine learning algorithms are still not commonly utilized in Nigeria's tax system. Low data quality, inadequate infrastructure, a lack of technological expertise, and ethical concerns about algorithmic bias and data privacy are some of the problems that make implementation more challenging (Afolabi et al., 2022; Uchenna & Olatunji, 2021).

Therefore, the primary problem is to understand how machine learning algorithms might be applied to detect tax fraud in Nigeria's unique institutional, technological, and economic context.

The main objective of this study is to explore the application and challenges of machine learning algorithms in detecting tax fraud in Nigeria, with the aim of strengthening the country’s tax administration system. This objective is justified by the persistent rise in tax-related offences—such as under-reporting, non-filing, and the manipulation of financial records—which continue to undermine revenue generation and fiscal sustainability. Understanding the current state of tax fraud is therefore essential for establishing a baseline for the problem and identifying the systemic loopholes exploited by taxpayers. Furthermore, examining the technical, institutional, and ethical constraints that hinder the adoption of machine learning is crucial, as these factors often determine the feasibility, accuracy, and acceptance of advanced analytical tools within public institutions. Finally, analyzing how machine learning algorithms can enhance fraud detection provides evidence on their potential contribution to improving compliance, automating risk assessment, and

supporting data-driven decision-making. Together, these interconnected aims offer a comprehensive foundation for evaluating both the prospects and challenges of integrating machine learning into Nigeria's tax administration framework.

This work significantly adds to the literature regarding enhancing tax administration in Nigeria by providing theoretical and empirical insights into the integration of machine learning (ML) algorithms for tax fraud detection. Unlike typical audit approaches, which are often manual and reactive, this study demonstrates how machine learning techniques can be leveraged to provide proactive fraud detection, anomaly analysis, and predictive risk profiling. In addition to improving operational efficiency, these advancements improve decision-making, compliance enforcement, and transparency in tax institutions (Odufisan et al., 2025; Ariyibi et al., 2024). The study bridges the gap between the potential of technology and practical implementation in developing nations by situating machine learning (ML) within Nigeria's unique physical, legal, and administrative issues. Furthermore, it underscores the need for regulatory reforms and capacity building to support the sustainable adoption of intelligent systems in public sector governance (Olabanji et al., 2024; Kraemer-Mbula et al., 2020).

LITERATURE REVIEW

Dynamics of Tax and Tax Fraud in Nigeria

Taxes are an essential tool for raising public funds, advancing social justice, and funding economic development. They play a vital role in budgetary policy as well. Nigeria has a complex tax structure with a mix of direct and indirect taxes managed by the federal and state governments. Notwithstanding its significance, the Nigerian tax system is marked by high non-compliance rates, poor revenue generation, and widespread tax fraud, all of which significantly impair the government's capacity to maintain macroeconomic stability and provide public services (Nosiri et al., 2021)

According to the Internal Revenue Service (2022), tax fraud is the deliberate, intentional act of concealing or fabricating information in order to reduce one's tax liability or evade paying taxes. It involves intentional deception and is considered a criminal act. In Nigeria, tax fraud is a serious offense that entails fines and possibly jail term as punishments (Adebayo, 2022). Underreporting income, overstating deductions, and failing to file tax returns are some of the most prevalent types of tax fraud in Nigeria (Adebayo, 2022). Poor management, tax revenue theft, and social, economic, and political issues are some of the causes of tax fraud (Bird & Zolt, 2008). Tax fraud can also result from a lack of civic duty, an unfair allocation of facilities and amenities, and taxpayers' inability to obtain government services (Torgler, 2007).

Empirical evidence indicates that Nigeria's tax system is severely constrained by institutional inefficiencies, low compliance, and widespread tax fraud. Intentional tax document falsification, underreporting of income, and noncompliance with remittance rules are all major causes of fraud in Nigeria's public tax system (Nosiri et al. 2021). These fraudulent activities are often made possible by inadequate enforcement procedures and corruption among tax authorities, according to Agbaje and Akinyemi (2021), who use the fraud triangle theory to show how economic pressure, opportunity, and rationality fuel tax theft.

The informal sector's dominance, which remains largely untaxed, exacerbates the situation. The OECD (2021) and World Bank (2023) report that developing countries like Nigeria lose billions annually due to tax-related irregularities rooted in poor taxpayer data management and manual processing systems. In line with this, Olayinka and Adedeji (2022) emphasize that Nigeria's economic and institutional conditions provide fertile ground for evasion and fraud to thrive.

Challenges of Tax Administration and the Adoption of Machine Learning in Nigeria

The Nigerian tax system suffers from structural inefficiencies that hinder effective administration and technological advancement. These include fragmented databases, inconsistent taxpayer records, low digital literacy among tax officers, and budgetary constraints (Olabanji et al., 2024). These challenges are amplified when attempting to implement advanced technologies such as machine learning (ML) for tax fraud detection. Machine learning algorithms are sets of instructions that enable computers to learn from data, make predictions, and progressively improve at what they do without the need for explicit programming.

The availability, consistency, and quality of data are significant obstacles. To provide reliable insights, machine learning algorithms require vast quantities of fast, precise, and structured data. However, Nigerian tax data is sometimes insufficient, erroneous, and inconsistent between authorities. The lack of established data formats and inadequate interagency collaboration complicate efforts to use data for efficient tax administration (Eze, Nwachukwu, & Daramola, 2024).

According to Ariyibi et al. (2024), other challenges include institutional resistance to change, a lack of technological proficiency, and ethical issues including algorithmic biases and data privacy. Many tax

authorities view technological advances as a threat to their informal revenue streams or job security. Reform is therefore not very common and attempts to implement it are often met with passive or hostile resistance (Okoye & Ani, 2022). Furthermore, as there is no mandate or incentive to encourage innovation in tax administration, this problem is exacerbated by the absence of clear legislative and regulatory frameworks for the use of AI in the public sector.

Cybersecurity and data privacy issues are also still unresolved. Machine learning requires processing large amounts of sensitive taxpayer data, which raises questions about data ownership, confidentiality, and potential misuse. Both taxpayers and authorities have reasonable worries over the security and moral use of data due to Nigeria's very low data protection legislation, even after the Nigeria Data Protection Act (NDPA) was passed in 2023. The deployment of machine learning (ML) in tax administration may find criticism from both inside and outside the government in the lack of adequate cyber security measures and open governance frameworks (NDPC, 2023). Onifade and Adeniran (2023) assert that the absence of robust legal and ethical frameworks, especially in light of Nigeria's Data Protection Regulation (NDPR), poses significant risks to the use of AI in the public sector.

Despite these formidable challenges, there are substantial potential benefits to integrating machine learning into Nigeria's tax administration. Machine learning algorithms provide revenue trend predictions, audit selection optimization, anomalies suggestive of fraud, and patterns in taxpayer behavior. Countries like Estonia, the UK, and India have already demonstrated how successful such technology is at improving operational efficiency, lowering tax evasion, and increasing tax compliance (OECD, 2023). If Nigeria wants to follow suit, it must prioritize strategic investments in digital infrastructure, capacity building, data governance, and legislative reform.

The Nexus between Machine Learning Algorithms and Tax Fraud Detection

Tax fraud is a significant issue for revenue authorities worldwide and poses a serious threat to fiscal sustainability, economic justice, and public trust in the government. The complexity of fraudulent schemes has made traditional audit and enforcement approaches less effective. In order to better detect, prevent, and deter tax fraud, governments and tax authorities are utilizing emerging technologies, such as machine learning (ML). Because machine learning algorithms have the innovative ability to analyze big, complex datasets, uncover hidden patterns, and generate predictive insights that would not be achievable with conventional methodologies, they are related to tax fraud detection (Ali et al., 2022).

Machine learning, a type of artificial intelligence (AI), makes use of computer algorithms that learn from historical data and progressively improve their performance without the need for explicit programming. Compared to rule-based systems that rely on predetermined conditions and static thresholds, machine learning (ML) offers significant advantages in the detection of tax fraud (Abubakar, Bello & Usman, 2023). Machine learning algorithms, as opposed to traditional methods, can identify previously undetected fraudulent activity, detect tiny abnormalities, and learn continually. The fact that fraudsters regularly alter their tactics to evade discovery makes this dynamic adaptability particularly crucial. Kotsiantis (2007) states that supervised, unsupervised, and semi-supervised learning models are the three general categories into which machine learning algorithms can be separated. There are specific applications for each of these models in fraud detection.

Numerous studies have demonstrated how well machine learning algorithms work to improve tax fraud detection through anomaly detection, predictive modeling, and real-time analysis. According to Odufisan et al. (2025), AI systems trained on tax data may effectively detect fraudulent activities using both supervised and unsupervised learning techniques. Muchuchuti (2024) confirms these results by demonstrating that ensemble machine learning models (such as logistic regression, decision trees, and support vector machines) provide superior fraud detection accuracy than single-algorithm models.

According to Murorunkwere et al. (2023), supervised machine learning models (such as random forests and artificial neural networks) are very good at spotting fraudulent transactions across a range of taxpayer profiles, especially when there is access to past audit data. These models not only improve detection accuracy but also reduce investigation time, helping tax authorities prioritize high-risk cases more efficiently.

Alsadhan (2023) introduces a multi-module ML framework that combines supervised and unsupervised components with behavioral scoring to assign fraud probabilities to tax returns. This system, tested on real data from the Saudi tax authority, reflects a scalable model that could be adapted for Nigeria, provided that the necessary data governance and infrastructure challenges are addressed.

Sociotechnical Systems Theory (STS)

This study is underpinned by the Sociotechnical Systems Theory (STS), which provides a comprehensive lens for examining the interaction between technological innovations and the social, institutional, and ethical systems into which they are introduced.

For technological systems to be properly deployed, their technical components and the social structures in which they operate must be harmoniously linked, according to STS, which was initially created by the Tavistock Institute in the 1950s (Baxter & Sommerville, 2011).

STS theory facilitates the understanding of tax fraud as a systemic issue influenced by socio-institutional dynamics and technological limitations. Nigeria's high prevalence of tax evasion is caused by a number of issues, including institutional weaknesses, a lack of trust in the government, dishonest behavior by tax authorities, and inadequate data analytics and enforcement techniques (Olaoye et al., 2023). Thus, from the perspective of STS, preventing tax fraud requires both technical and social measures.

One of the core realizations of STS is that technological innovations cannot succeed in isolation from the institutional and social contexts in which they are applied (Orlikowski, 1992). The application of machine learning algorithms in Nigeria is hindered by a number of sociotechnical issues, such as a lack of technical expertise, organizational reluctance to change, a lack of data infrastructure, a lack of regulatory oversight, and ethical worries about algorithmic bias and data privacy. These issues demonstrate a misalignment between the maturity of supporting institutions and the sophistication of technology (Dignum, 2019; Sarker et al., 2019).

STS enables the study of how machine learning algorithms, as technical tools, can enhance fraud detection only when included into well-established institutional procedures and ethical frameworks. Machine learning algorithms have several advantages that can significantly improve the detection of tax fraud and the enforcement of compliance, including pattern recognition, anomaly detection, and predictive analytics. However, these benefits can only be realized if the tools are part of an ecosystem that includes reliable data sources, knowledgeable personnel, transparent governance structures, and stakeholder support (Baxter & Sommerville, 2011).

According to this study, using machine learning (ML) algorithms to detect tax fraud in Nigeria is not only a technological endeavor but also a socio-technical one that involves institutional capabilities, ethical considerations, and cultural readiness. The concept supports a thorough understanding of how machine learning tools must be successfully incorporated into Nigeria's broader tax administration environment.

METHODOLOGY

This study adopted a survey research design to assess the role of machine learning algorithms in detecting tax fraud in Nigeria. A sample of 150 participants, including data analysts, financial accountants, tax consultants, and tax auditors, was selected using convenience sampling techniques. Out of the distributed questionnaires, 115 were correctly completed and returned. Data was collected through structured questionnaires using a four-point Likert scale and distributed via Google Forms. The reliability of the instrument was confirmed with a Cronbach's alpha score of 0.943, indicating strong internal consistency. Descriptive statistics (frequencies, percentages, and means) were used to analyze responses, while a simple regression model evaluated the effect of machine learning algorithms on tax fraud detection. The model for the study is stated thus:

Tax Fraud Detection = $\beta 0 + \beta 1$ML algorithms + μ 1

RESULT AND DISCUSSIONS

Data Presentation

The analysis focused on application and challenges of machine learning algorithm in detecting tax fraud. The study addressed three major questions (i) To examine the current state of tax fraud and existing detection mechanisms in Nigeria, (ii) To identify the key technical, institutional, and ethical challenges hindering the adoption of machine learning in Nigeria's tax administration and (iii) to analyze the effect of various machine learning algorithms on tax fraud detection. 150 questionnaires were distributed to the participants, including Data Analysts, Financial Accountants, Tax Consultants and Tax Auditors, and 115 questionnaires were correctly filled and returned.

Current state of tax fraud in Nigeria

The results in Table 1 provide descriptive statistics on respondents' perceptions of tax fraud in Nigeria. All items presented mean values above the benchmark threshold of 2.5, indicating strong consensus among participants about the prevalence of tax fraud and the inadequacies of existing detection frameworks.

The highest mean score (3.77) corresponds to the item stating that tax fraud is a major issue affecting Nigeria's tax system, emphasizing the perceived severity and widespread nature of the problem. Furthermore, the item indicating that digitalization of tax processes can help in reducing tax fraud in Nigeria received a mean score of 3.39.

Participants also largely agreed that tax fraud can be reduced in Nigeria (mean = 3.28), this implies that with machine learning, tax fraud can be reduced. Also, the responds of Weak enforcement of tax laws and

corruption among tax officials are the major causes of tax fraud in Nigeria received the mean score (3.27). Finally, the respondents are of the opinion that most tax fraud cases go undetected by the current system (mean = 2.98). This research supports the OECD's (2021) findings, which highlight the difficulties many nations, particularly developing ones like Nigeria, experience in detecting tax fraud due to outdated approaches. It promotes the use of cutting-edge technology, including AI and machine learning, to strengthen compliance and identify fraud more effectively. It also backs up the idea that strong frameworks are required because the current tools are inadequate.

Table 1: Result on current state of tax fraud in Nigeria

	SA	A	D	SD	MEAN	STD
Tax fraud is a major issue affecting Nigeria's tax system.	88	27	-	-	3.77	0.426
Most tax fraud cases go undetected by the current system.	40	44	20	11	2.98	0.955
Weak enforcement of tax laws and corruption among tax officials are the major causes of tax fraud in Nigeria.	45	40	12	3	3.2700	0.78951
Tax fraud can be reduced in Nigeria.	50	35	8	7	3.2800	0.88854
Digitalization of tax processes can help in reducing tax fraud in Nigeria.	46	48	5	1	3.3900	0.63397

Source: SPSS result extracted from Appendix 2A

Challenges hindering the adoption of machine learning in Nigeria's tax administration

Table 2: Result on challenges hindering the adoption of machine learning in Nigeria's tax administration

	SA	A	D	SD	MEAN	STD
Limited access to quality data is a major challenge in ML adoption.	46	67	2	-	3.38	0.523
Ethical concerns hinder the use of ML in tax fraud detection.	63	32	4	16	3.23	1.046
Institutional resistance to change hinders ML adoption.	53	44	11	7	3.24	0.864
There is insufficient government funding for ML initiatives.	77	26	9	3	3.54	0.753
Lack of clear policies and guidelines limits ML implementation.	37	51	18	9	3.01	0.893

Source: SPSS result extracted from Appendix 2B

The result in Table 2 presents a summary of respondents' perceptions regarding the challenges impeding the adoption of machine learning (ML) in tax fraud detection within the Nigerian context. All items recorded mean values significantly above the 2.5 criterion threshold, indicating general agreement among participants that multiple systemic and structural issues hinder effective ML implementation. The most prominent challenge identified is insufficient government funding, with a mean score of 3.54, suggesting that inadequate financial commitment remains a major impediment to technological innovation in public sector tax administration. This is closely followed by limited access to quality data (mean = 3.38), which underscores the foundational role of reliable and structured data in the successful deployment of ML

algorithms. Furthermore, institutional resistance to change (mean = 3.24) and ethical concerns (mean = 3.23) were recognized as significant non-technical barriers. These findings reflect bureaucratic inertia, a lack of technical awareness, and growing apprehension surrounding data privacy, transparency, and accountability in AI systems. Lastly, the lack of clear policies and guidelines (mean = 3.01) was also acknowledged, pointing to a regulatory gap that complicates the integration of AI-driven solutions in public governance structures. While this aligns with Kshetri (2018), Olabanji et al. (2024), and Kraemer-Mbula et al. (2020) who noted that many developing countries lack legal and ethical frameworks for AI, it contrasts with more optimistic views like UNCTAD (2021), which argue that some African countries are making significant progress in drafting AI policies.

Table 3: Regression Analysis

Parameters	Coefficient	Std Error	T-statistics	P-value
Constant	-1.018	0.143	-7.146	0.000
ML_ALGORITHMS	0.253	0.008	31.691	0.000
R-Square	0.899			
Adjusted R-Square	0.898			
F-statistics	1004.302			
P-value	0.0000			

Source: SPSS result extracted from Appendix 2C

Table 3 presents the regression result on the effect of machine learning (ML) algorithms on tax fraud detection in Nigeria. From the model summary table above, the following information can be distilled.
The R2 which measure the level of variation of the dependent variable caused by the independent variables stood at 0. 899.The R-squared value suggests that approximately 89.9% of the variance in tax fraud detection can be explained by the adoption and application of ML algorithms. This high explanatory power underscores the substantial impact ML can have in improving fraud detection outcomes.
The adjusted R2 of approximately 0.898 indicates that if other factors are considered in the model, this result will deviate from it by only 0.001 (i.e. 0.899 – 0.898). This result shows that there will be a further deviation of the variation caused by the independent factors to be included by 0.001%. This confirms the robustness of the model, even when adjusted to degrees of freedom.
The unstandardized coefficient (B) for ML algorithms is 0.253 with a standard error of 0.008. This positive coefficient indicates that for every one-unit increase in the application of ML algorithms, there is a 0.253-unit improvement in tax fraud detection. Furthermore, the t-value of 31.691 and the significance value of 0.000 provide strong evidence that the relationship between ML algorithms and tax fraud detection is statistically significant.
Finally, the result shows that the significant variation of Fisher's statistics (1004.302) has the probability value of 0.0000 which implies that model is statistically significant at 5%.

Test of Hypothesis
H0: Machine learning algorithms have no significant and positive effect on tax fraud detection in Nigeria.
HA: Machine learning algorithms have a significant and positive effect on tax fraud detection in Nigeria.
The t-statistics of 31.691 has the P-value of 0.000. Since the p-value is less than 5% (0.05), the study rejects and null hypothesis and accepts the alternative hypothesis which implies that machine learning algorithms have a significant and positive effect on tax fraud detection in Nigeria. These findings align with studies such as those by Baesens et al. (2016) and Zhou & Kapoor (2011), which emphasized the utility of AI/ML and data-driven techniques in uncovering financial irregularities.

CONCLUSION AND RECOMMENDATION

The analysis focused on machine learning algorithms for detecting tax fraud: application and challenges in Nigeria. Data was collected using questionnaires administered to participants through social media platforms. The data collected were analyzed using both descriptive and simple regression analysis. The findings of objective one indicates that tax fraud is a significant issue in the Nigerian tax system, and the existing detection techniques are usually regarded as ineffective. The findings from the analysis of objective two revealed that although machine learning techniques are seen to be effective tools for improving fraud detection, structural barriers hinder their broad implementation. These include limited access to quality data, ethical concerns, institutional resistance, insufficient government funding, and lack of regulatory clarity. Finally, the findings from the analysis of objective three demonstrate a strong and statistically significant

effect of ML algorithms on improving tax fraud detection ($R^2 = 0.899$, $p < 0.001$), indicating that, when applied properly, technology adoption may significantly reduce fraud.

To enhance the effectiveness of machine learning (ML) algorithms in tax fraud detection, the study recommends a multifaceted approach. First, the Nigerian government should invest in digitizing tax records and improving data quality to ensure that ML systems are fed with reliable inputs. It is also necessary to establish clear ethical guidelines to govern the use of artificial intelligence in tax administration, with a focus on ensuring transparency, accountability, and privacy. Furthermore, increased funding is needed to support digital tax initiatives, including research, pilot projects, and the acquisition of relevant AI technologies. To build institutional readiness, training programs should be developed for tax administrators, IT specialists, and policymakers to improve their understanding and acceptance of ML tools. Finally, resistance to technological change must be tackled through strong leadership, strategic change management, and inclusive decision-making practices.

Policy Implications

The findings of this study have direct implications for tax policy in Nigeria. Policymakers must recognize the transformative potential of artificial intelligence and machine learning in improving tax governance, especially in fraud detection and revenue assurance. There is a pressing need to enact policies that support the digitization of tax administration and create enabling legal frameworks for ethical and transparent use of data-driven technologies. Additionally, investment in institutional capacity building, data protection regulation, and public-private partnerships will be crucial in aligning Nigeria's tax infrastructure with global best practices in technological adoption.

Limitations

While the study provides valuable insights into the applicability of machine learning for tax fraud detection in Nigeria, it is not without limitations. The reliance on self-reported data through questionnaires introduces the possibility of respondent bias. The use of convenience sampling may limit the generalizability of findings across broader tax institutions or regions. Furthermore, the study is exploratory in nature and does not involve real-time implementation or technical evaluation of ML models, which would require more advanced tools, datasets, and computational resources. Future research should consider experimental designs and partnerships with tax authorities to validate model effectiveness in real-world settings.

REFERENCES

Abubakar, M. A., Bello, M. U., & Usman, A. (2023). Application of machine learning in detecting financial crimes in Nigeria. *Journal of Digital Finance*, 4(1), 45–58.

Adebayo, A. A. (2022). Tax evasion and avoidance in Nigeria: A review of the literature. *Journal of Taxation and Economic Development*, 21(1), 1-15.

Afolabi, B., Adebayo, K., &Onifade, T. (2022). Technological disruptions and the future of tax administration in Nigeria. *Journal of African Taxation*, 8(2), 45-59.

Agbaje, A. A., & Akinyemi, B. T. (2021). Rationalizing tax evasion in Nigeria: Insights from the Fraud Triangle. *Journal of Forensic Accounting Research*, 6(2), 120–138.

Ali, A., AbdRazak, S., Othman, S.H., Eisa, T.A.E., Al-Dhaqm, A., Nasser, M., Elhassan, T., Elshafie, H., & Saif, A. (2022). Financial fraud detection based on machine learning: A Systematic Literature *Review. Appl. Sci*., 12, 9637.

Alm, J., & Martinez-Vazquez, J. (2021). Tax evasion and tax avoidance: A review of the literature. *Journal of Economic Surveys*, 35(2), 412-434.

Ariyibi, K.O., Bello, O.F., Ekundayo, T.F., Oladepo, O., Wada, I.U., & Makinde, E.O. (2024). Leveraging artificial intelligence for enhanced tax fraud detection in modern fiscal systems. *GSC Advanced Research and Reviews*, 21(02), 129-137

Baghdasaryan, V., Davtyan, H., Sarikyan, A., & Navasardyan, Z. (2022). Improving tax audit efficiency using machine learning: The role of taxpayer ' s network data in fraud detection. *Applied Artificial Intelligence*, 36(1).

Baxter, G., & Sommerville, I. (2011). Socio-technical systems: From design methods to systems engineering. *Interacting with Computers*, 23(1), 4–17.

Bird, R. M., & Zolt, E. M. (2008). Tax policy and the informal economy. *International Tax and Public Finance*, 15(4), 449-464.

Bird, R. M., & Zolt, E. M. (2008). Technology and taxation in developing countries: From hand to mouse. *National Tax Journal*, 61(4), 791–821.

Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.

Breiman, L. (2011). Random forests. *Machine Learning*, 45(1), 5-32.

Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (2014). *Classification and regression trees*. CRC Press.

Chaudhary, R., & Sharma, M. (2022). Tax fraud detection using data mining techniques: A Review. *International Journal of Computer Applications*, 184(9), 22–30.

Cobham, A., & Janssens, W. (2020). Illicit financial flows and the challenge of measuring tax evasion. *Journal of Economic Policy Reform*, 23(1), 1-16.

Cortes, C., & Vapnik, V. (2015). Support-vector networks. *Machine Learning*, 20(3), 273-297.

Cressey, D. R. (1953). *Other People's Money: A Study in the social psychology of embezzlement*. Free Press.

De Roux, D., Pérez, B., Moreno, A., Del Pilar Villamil, M., & Figueroa, C. (2018). *Tax fraud detection for under-reporting declarations using an unsupervised machine learning approach*. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 215–222.

Dignum, V. (2019). *Responsible artificial intelligence: How to develop and use AI in a responsible way*. Springer.

Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. arXiv preprint arXiv:1702.08608.

Eze, N. J., Nwachukwu, C. N., & Daramola, O. S. (2024). Corruption, ethics, and tax morale in sub-saharan Africa: Nigeria in focus. *Journal of African Governance and Ethics*, 3(1), 40–55.

Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI Magazine*, 17(3), 37–54.

Federal Inland Revenue Service (FIRS). (2023). *Annual tax administration report*. Abuja: FIRS.

Friedman, J. H. (2011). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29(5), 1189-1232.

Han, J., Pei, J., & Kamber, M. (2011). *Data Mining: Concepts and Techniques* (3rd ed.). Elsevier.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction*. Springer.

HMRC. (2020). *Annual report and accounts 2019-2020. her majesty's revenue and customs*, United Kingdom.

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression*. John Wiley & Sons.

Jain, A. K., Murty, M. N., & Flynn, P. J. (2019). Data clustering: A review. *ACM Computing Surveys*, 31(3), 264-323.

Journal of Taxation. (2022). *Using machine learning to detect tax evasion*.

Kotsiantis, S. B. (2007). Supervised machine learning: A review of classification techniques. *Informatica*, 31(3), 249-268.

Kshetri, L. (2018) Artificial intelligence in developing countries: Challenges and opportunities. *IT Professional*, 20(2), 18–23.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444.

MacQueen, J. (2007). *Some methods for classification and analysis of multivariate observations.* Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, 1, 281-297.

Mellado, A., Romero, A., & Chen, J. (2021). Artificial intelligence for tax compliance: A Survey of machine learning applications. *International Journal of Digital Governance*, 6(1), 22-34.

Muchuchuti, K.C. (2024). An ensemble machine learning model to detect tax fraud: Conceptual framework. *International Journal of Research and Innovation in Social Science (IJRISS)* 8(6), 2276-2283

Murorunkwere, B. F., Haughton, D., Nzabanita, J., Kipkogei, F., & Kabano, I. (2023). Predicting tax fraud using supervised machine learning approach. *African Journal of Science, Technology, Innovation and Development*, 15(6), 731–742.

Ngai, E. W. T., Hu, Y., Wong, Y. H., Chen, Y., & Sun, X. (2011). The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature. *Decision Support Systems*, 50(3), 559–569.

Nosiri, H., Abiahu, M. F. C., & Uwakwe, T. N. (2021). *Forensic accounting and tax fraud detection in the Nigerian public sector*. In forensic accounting and tax fraud detection in the Nigerian public sector. In taxation, social contract and economic development. Proceedings of 4th Annual International Academic Conference of the Chartered Institute of Taxation of Nigeria.

Odufisan, O.I., Abhulimen, O.V., & Ogunti, E.O. (2025). Harnessing artificial intelligence and machine learning for fraud detection and prevention in Nigeria. *Journal of Economic Criminology* 7, 1-9

OECD. (2021). *Tax administration 3.0: The digital transformation of tax administration*. Organisation for Economic Co-operation and Development.

OECD. (2021). *Technology tools to tackle tax evasion and tax fraud*. OECD Publishing.

OECD. (2022). *Tax administration 2022: Comparative information on OECD and other advanced and emerging economies.* OECD Publishing.

OECD. (2023). *Tax administration 2023: Comparative information on OECD and other advanced and emerging economies.* OECD Publishing

Okoye, E. I., & Ani, W. U. (2022). Tax administration and the informal sector: The Case of Nigeria. *African Research Review,* 16(1), 102–117.

Olabanji, O.S., Olaniyi, O.O., & Olaoye, O.O. (2024). Transforming tax compliance with machine learning: reducing fraud and enhancing revenue collection. *Asian Journal of Economics, Business and Accounting* 24 (11):503-13.

Olaniyi, O. O., Shah, N. H., & Bahuguna, N. (2023d). Quantitative analysis and comparative review of dividend policy dynamics within the banking sector: Insights from global and U.S. financial data and existing literature. *Asian Journal of Economics, Business and Accounting*, 23(23), 179–199

Olaniyi, O.O., Asonze, C.U., Ajayi, S.A., Olabanji, S. O., & Adigwe, C. S. (2023c). A regressional study on the impact of organizational security culture and transformational leadership on social engineering awareness among bank employees: The interplay of security education and behavioral change. *Asian Journal of Economics, Business and Accounting*, 23(23), 128–143.

Olaoye, S. A., Ayodele, O. I., & Olatunji, I. (2023). Enhancing tax revenue generation through Artificial Intelligence in Nigeria. *Journal of Tax Administration and Innovation*, 4(1), 71–85.

Olayinka, I. K., & Adedeji, O. A. (2022). Tax compliance and evasion in Nigeria: The Role of economic and institutional factors. *African Journal of Accounting and Finance*, 14(1), 34–49.

Onifade, T. T., & Adeniran, S. (2023). Data privacy and machine learning in Nigeria's public sector: risks and remedies. *Nigerian Law and Technology Review*, 5(1), 100–115.

Orlikowski, W. J. (1992). The duality of technology: Rethinking the concept of technology in organizations. *Organization Science*, 3(3), 398–427.

Pearson, K. (2021). On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, 2(11)

Sarker, I.H., (2021). Machine learning: algorithms, real-world applications and research directions. SN *Comput. Sci*. 2, 160.

Sarker, S., Chatterjee, S., Xiao, X., & Elbanna, A. (2019). The Socio technical axis of cohesion for the is discipline: Its historical legacy and its continued relevance. *MIS Quarterly*, 43(3), 695–719.

Saxunova, D., & Szarkova, R. (2018). Global efforts of tax authorities and tax evasion challenge. *Journal of Eastern Europe Research in Business and Economics*, 2018, 1-14.

Sharma, A., & Kumar, K. (2021). Detecting Gst fraud through machine learning techniques. *Journal of Emerging Technologies and Innovative Research (JETIR)*, 8(3), 320-337

Torgler, B. (2007). *Tax compliance and tax morale: A theoretical and empirical analysis*. Edward Elgar Publishing.

Uchenna, O., & Olatunji, A. (2021). Digital transformation and tax evasion in developing economies. *Nigerian Journal of Economic Policy*, 28(1), 90-107.

World Bank. (2023). *Nigeria public finance review: Fiscal reforms for inclusive growth*. World bank group