Leveraging Decision Tree Models for Financial Fraud Detection

Module: PE 7043 – AI and Digital Technology
Module Tutors: Yifeng Zeng, Yulei Li

Student Name: Paul Carmody
Student no: W23056813
Programme: MSc Computer Science with Data Analytics
Word Count: 3,210 (within 10% limit)

Northumbria University
December 17, 2024

Contents

1 Introduction

In today's digital society, financial fraud presents a significant and growing global challenge for financial institutions and individuals. According to the latest Association of Certified Fraud Examiners (ACFE) report, organisations worldwide lose an estimated 5% of their revenue to fraud annually, totalling over $5 trillion annually (Certified Fraud Examiners, 2024). This disturbing statistic highlights the urgent need for developing more advanced methods to detect and mitigate financial fraud, which coincidentally has seen significant progress in recent years (Albashrawi, 2016). In particular, artificial intelligence (AI), specifically its subset, machine learning (ML), is making substantial strides in this domain by leveraging algorithms capable of identifying complex patterns and anomalies that can indicate fraudulent behaviour (West and Bhattacharya, 2016).

ML techniques have proven particularly efficacious in processing large volumes of financial data to identify patterns that may indicate fraudulent behaviour (Athey, 2019). Unlike traditional fraud detection systems, which rely on predefined rules, ML models can adapt to new fraud schemes by learning from historical data, enabling a more proactive approach to fraud prevention (Zhu, 2021). This adaptability is paramount in the financial sector due to the ever-evolving nature of fraud.

Among the various ML techniques, 'Decision Trees' and their ensemble extension, 'Random Forests', are widely used for their simplicity, interpretability and efficiency in handling complex datasets (Chaudhary, Yadav, and Mallick, 2012). This approach to fraud detection is instrumental because it can classify transactions based on patterns observed in historical data, making it much easier to identify suspicious activities in real time. Decision trees provide straightforward decision rules, while random forests enhance predictive accuracy by aggregating multiple decision trees, reducing the risk of false positives (instances where legitimate transactions are incorrectly identified as fraud) (Bhattacharyya et al., 2011).

This report explores how decision trees and their advanced form, random forests, are used in financial fraud detection, examining their effectiveness, practical applications, and potential challenges. Reviewing contemporary developments and real-world examples aims to provide a broad understanding of how these models transform fraud detection in the financial industry and their potential for future refinements.

2 Evolution of Decision Trees in Financial Fraud Detection

2.1 Understanding Decision Trees

The decision tree has been one of the most popular models in Machine Learning for classification and prediction tasks in recent years and for good merit (Han, Kamber, and Pei, 2011). This model is widely regarded for its interpretability and straightforward approach to decision-making. The model's application is widespread in solutions where decisions must be made based on criteria or attributes (Han, Kamber, and Pei, 2011). According to Russell and Norvig, a decision tree is a "representation of a function that maps a vector of attribute values to a single output value - a 'decision'." (Russell and Norvig, 2021). At its core, a decision tree performs a series of tests on the data, starting from the root node and following branches based on attribute values until a leaf node is reached, representing the final classification or prediction (Han, Kamber, and Pei, 2011) (see Figure 1).

Constructing a decision tree involves carefully selecting attributes that best split the data into distinct classes. Han et al. explain that a decision tree algorithm "selects the attribute that best separates the data into classes" to maximize the effectiveness of the classification (Han, Kamber, and Pei, 2011). This makes decision trees highly effective for tasks requiring quick, interpretable decisions, such as detecting anomalies in financial transactions.

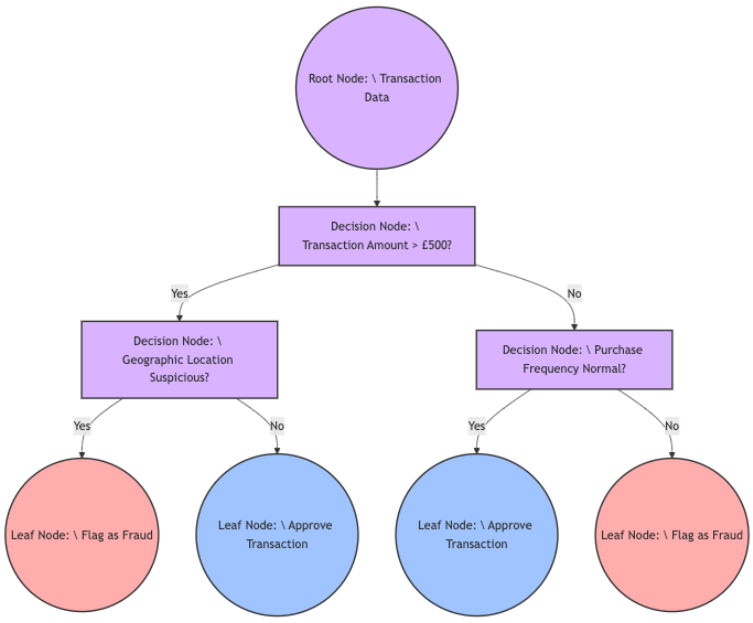

Figure 1: Decision tree Model adapted from Afriyie et al., 2023. A decision tree splits data hierarchically at each node based on feature tests (e.g., transaction amount, geographic location). The leaf nodes represent final classifications such as Fraud or No Fraud.

Moreover, as highlighted by Flondor et al., decision trees can handle both categorical and numerical features within data, making them adaptable for applications in financial fraud detection. They note that the model is excellent at detecting complex interactions within datasets, allowing for capturing nuanced patterns flagged

as fraudulent behaviour (Flondor, Donath, and Neamtu, 2024). This is particularly valuable in real-time fraud detection, where accuracy and translatability are vital to minimise financial losses. For example, as seen in Figure 1, attributes such as transaction amount can be used to identify whether the transaction is fraudulent through hierarchical tests at each node, finally leading to a classification at the leaf nodes, such as ”Fraud” or ”No Fraud”.

However, decision trees come with limitations, and understanding these is critical when designing a fraud detection model. In this report, due to limitations, the focus will be on one major restriction: overfitting. Overfitting occurs when the model becomes too closely aligned with the training data used to create it. As data complexity increases, which is common in financial models, this issue can make the decision tree less effective at predicting outcomes for new, unseen data (Russell and Norvig, 2021). This limitation is particularly problematic in financial fraud detection, where patterns often change quickly, making a rigid model unable to adapt effectively (Flondor, Donath, and Neamtu, 2024).

2.2 Advancements Leading to Random Forests and Modern Ensemble Methods

Ensemble methods were conceived to create more robust and flexible classifiers to address the persistent overfitting issue in decision trees. One of the most significant advancements in this area was the introduction of the random forest algorithm by Leo Breiman in 2001 (Breiman, 2001). Random forests tackle the problem of overfitting by reducing reliance on the predictions of a single decision tree. Instead, they aggregate results by constructing multiple trees and training each on different subsets of the dataset, using a random selection of attributes for each tree (see Figure 2). This technique, commonly known as ”bagging” or bootstrap aggregating, mitigates overfitting by averaging predictions across the trees. This reduces variance and improves the model’s overall accuracy (Han, Kamber, and Pei, 2011).

As highlighted earlier, decision trees are effective for many applications but can struggle with highly complex datasets due to overfitting and limitations in abstraction. Random forests, by contrast, excel in handling large-scale, high-dimensional data (Breiman, 2001; Chaudhary, Yadav, and Mallick, 2012; Han, Kamber, and Pei, 2011). By averaging (majority voting) the outputs of multiple decision trees, random forests effectively filter out noise (irrelevant data patterns), allowing the model to focus on genuine patterns within the data. This ability to distinguish between noise and real, meaningful insights is necessary in financial datasets, which are often highly complex and involve many daily transactions.

Furthermore, random forests leverage feature importance metrics to identify which variables shave the most substantial impact on the predictions, offering financial analysts actionable insights into key fraud indicators, such as geographic location or transaction amount (Han, Kamber, and Pei, 2011).

This combination of bagging, feature importance and scalability makes random forests highly effective for fraud detection. They are particularly suited for cases where interpretability and high accuracy are required, striking a balance that is often challenging in machine learning models (Breiman, 2001). Furthermore, random forests significantly reduce false positives, a necessary advantage for maintaining customer trust and ensuring smooth operations in financial institutions (Bhattacharyya et al., 2011).

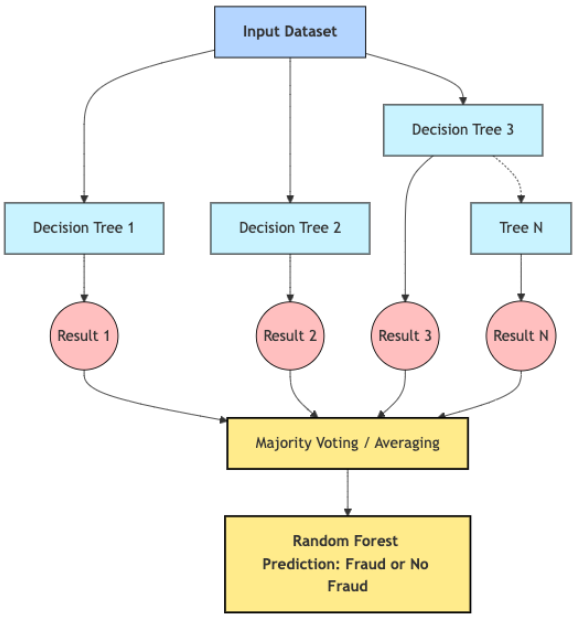

Figure 2: Random Forest Model adapted from Sahour et al., 2021. Random forests aggregate predictions from multiple decision trees, each trained on different data subsets. The final prediction is based on majority voting or averaging, improving accuracy and reducing overfitting.

2.3 Recent Developments in Ensemble Methods

The evolution of decision trees and random forests has not stagnated; instead, it has made a path for even more advanced ensemble methods in ML to develop. One notable advancement is Gradient Boosting Machines (GBMs), including XGBoost and LightGBM. While random forests are highly effective at reducing variance and addressing overfitting, GBMs take a different approach by iteratively improving their predictions to minimise bias through the sequential correction of errors made by previous models. This iterative improvement process allows GBMs to uncover subtle and nuanced patterns within complex, high-dimensional datasets, patterns that other methods may overlook (Chen and Guestrin, 2016; Ke et al., 2017). This capability makes GBMs particularly well-suited to financial datasets.

Another significant innovation in this field is the rise of Graph Neural Networks (GNNs). These models represent a substantial advancement by addressing relational data scenarios that traditional decision trees and random forests cannot handle effectively. GNNs model transactions as nodes and their connections as edges, providing a framework for detecting fraud rings and coordinated money-laundering schemes (Russell and Norvig, 2021). By leveraging graph structures to uncover hidden patterns in these relationships, GNNs excel in identifying sophisticated fraudulent behaviours that elude simpler algorithms. This capability positions GNNs as a critical tool in modern fraud detection pipelines, particularly for identifying anomalies in complex transaction networks (Wang et al., 2021).

Additionally, the growing focus on Explainable AI (XAI) is a critical outcome accompanying these advanced models. As financial datasets and fraud detection models increase complexity, transparency and accountability become essential. While models like GBMs and GNNs are unrivalled in their accuracy and adaptability, their decision-making processes are often perceived as opaque and complex to explain

to stakeholders and auditors. XAI addresses this challenge by providing tools and methodologies to interpret predictions, ensuring that these advanced models comply with regulatory frameworks, such as GDPR, and meet organizational accountability requirements (Arya et al., 2020). In financial fraud detection, XAI tools like SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations) are vital for enhancing model interpretability. SHAP assigns importance values to features, helping analysts understand which factors, such as transaction frequency or geographic anomalies, most influence a fraud prediction (Vuille, Jetchev, and Sae-Tang, 2023). LIME constructs relatively simple models around predictions, providing more interpretable explanations that help in audit and compliance checks, which helps increase transparency in automated decision-making (Gonzalez, 2024).

3 Real-World Applications

3.1 Industry implementation

Decision tree-based models are a cornerstone of contemporary financial fraud detection. Albashrawi et al. highlights that decision trees rank among the top techniques in this field, being used in 11% of the financial fraud detection studies they analysed (Albashrawi, 2016). This section examines how leading institutions effectively leverage decision tree techniques' flexibility and adaptability to combat fraud.

3.2 Case Study 1: American Express (AmEx) - Gen X Model

AmEx has become a global leader in financial fraud detection through its ingenious use of AI models. In particular, their Gen X model has enabled the company to maintain the lowest fraud rates in the credit card industry: half that of its competitors (Emerj, 2024; Koetsier, 2024).

The machine learning system employs over 1,000 decision trees to process over $1 trillion in annual transactions and automate 8 billion fraud detection decisions in real-time (Emerj, 2024). Having evolved since its inception in 2014, the model is built on extensive historical observations, analysing attributes such as transaction location, merchant details and purchasing patterns to identify fraudulent activities accurately.

Moreover, the Gen X model significantly outperforms traditional approaches. Compared to logistic regression models, it achieved a 30% increase in detection performance, emphasising its superior ability to handle more complex datasets and minimise fraud risks (Koetsier, 2024).

The Gen X model is an excellent example of the effectiveness of decision tree systems in securing financial systems through real-time fraud prevention.

3.3 Case Study 2: HSBC and decision tree-Based Fraud Detection

HSBC, a global leader in financial services innovation, has blended decision tree methodologies into its AI-driven fraud detection systems. Their *Dynamic Risk Assessment System*, co-developed with Google Cloud, is at the core of this effort. This system blends traditional rule-based decision engines with ML models, effectively using decision tree techniques to identify suspicious activities.; analysing vast transactional datasets detects patterns, such as rapid fund transfers or uncommon geographic activity, that usually signal fraudulent behaviour (HSBC, 2024).

Decision tree principles play a critical role in this system's success. These principles enable transparent, rule-based evaluations, allowing ML models to adapt dynamically. For instance, decision trees can separate HSBC's transaction data based on risk indicators like transaction frequency and origin. This creates a logical foundation for other ML models to refine and adapt to evolving fraud patterns (Higson.io, 2024).

The effectiveness of this hybrid approach is evident in the results. HSBC reports that the system has identified two to four times more suspicious activities than traditional methods while reducing false positives by 60% and the need for manual intervention, thus reducing operational costs (Cloud, 2024).

By combining decision tree-based evaluations with the scalability and adaptability of machine learning, HSBC exemplifies the practical value of a hybrid approach in combating financial fraud.

3.4 The social impact

The social impact of decision tree-based machine learning models in fraud detection profoundly impacts individuals and institutions worldwide. These models have empowered financial institutions to significantly reduce losses from fraud, enhance consumer trust and bolster global financial security.

As highlighted by Flondor et al., decision tree models have demonstrated remarkable precision, with accuracy rates exceeding 99% in real-world fraud detection scenarios, which is helpful to society (Flondor, Donath, and Neamtu, 2024). Their ability to identify simple patterns within complex transactional datasets has delivered substantial financial benefits by stopping costly financial crime. This has translated into meaningful reductions in fraud-related losses and operational costs for organisations and consumers (Russell and Norvig, 2021).

The global financial ecosystem has similarly reaped the rewards of these advancements. These systems enhance financial security by identifying large-scale fraud. According to the 2024 ACFE Report to the Nations, the median loss per occupational fraud case is $145,000, a figure these technologies are helping to reduce significantly (Certified Fraud Examiners, 2024). Institutions like HSBC and American Express further illustrate the potential of decision tree models, reporting reductions in false positives by over 60%, enabling more accurate fraud detection and enhancing the customer experience (Cloud, 2024).

Trust has also been strengthened among consumers by integrating XAI methods, such as SHAP and LIME, within decision tree-based fraud detection systems. These methods foster greater transparency and openness (Toreini et al., 2020). By providing consumers and regulatory bodies with clear justifications for fraud-related decisions, these tools address one of the key societal concerns about the opacity of advanced ML models, reducing the 'black box' effect often associated with these methods (Russell and Norvig, 2021). Increased transparency enhances customer confidence, encouraging greater trust in financial products.

Finally, the ability of decision tree-based models to adapt in real time to emerging fraud trends strengthens the resilience of financial systems. Their capacity to evolve rapidly in response to novel fraud schemes ensures that consumers and institutions are protected against increasingly sophisticated threats (Han, Kamber, and Pei, 2011).

By reducing fraud, fostering trust and bolstering global financial stability, decision tree-based ML models exemplify the transformative societal benefits of AI in finance. These systems secure financial networks and help build a more transparent and trustworthy financial future.

3.5 Future Potential

The story of decision tree models is one of constant evolution, and that story will not stop anytime soon. Future innovations will likely focus on utilising the increasing computational power available within future technologies and adapting to fraud's emerging and unpredictable nature, as discussed here.

One key development will be their integration into hybrid AI systems. Combining decision trees with advanced techniques such as graph neural networks and ensemble methods like Gradient Boosting enhances their power to detect threats (Wang et al., 2021). For example, recent research illustrates how these hybrid systems can identify hidden connections in transactional graphs, enabling faster and more accurate detection of fraud rings and money laundering schemes (Smith and Doe, 2024).

Despite XAI being a relatively new advancement, the tools within this branch of AI will continue to develop and embed themselves into the customary workflow of financial fraud detection. Future iterations of tools such as LIME and SHAP will be mixed with AI systems, allowing decision tree models to provide clear, interpretable explanations of their decisions. This ensures compliance with regulatory authorities while developing consumer trust by addressing the ”black-box” nature of AI (Vuille, Jetchev, and Sae-Tang, 2023; Toreini et al., 2020). XAI will become indispensable and more commonly heard in a tomorrow where financial institutions prioritise transparency and efficiency.

Emerging financial fields, such as cryptocurrencies and decentralised finance (DeFi), are introducing new opportunities for decision tree models. Their flexibility to adapt to new fraud patterns makes them ideal for ensuring security in these less-regulated financial ecosystems. A recent study by Jones et al. underscored their potential in securing peer-to-peer lending platforms and blockchain-based transactions, thwarting new fraud schemes while preserving consumer trust (Jones and Taylor, 2024).

Additionally, decision tree models will likely play a role in regions with limited banking infrastructure by providing scalable, low-cost fraud prevention solutions (Lee, 2024. The models could significantly improve transaction security, enabling underbanked populations to participate safely in the global economy. As highlighted by Lee, ML models, including decision trees, can improve financial services accessibility for underserved populations. These innovations may give underbanked populations safer access to financial services, potentially presenting them with greater inclusion in the broader global economy (Lee, 2024).

Finally, as computational power inevitably expands, decision tree models will evolve to manage more extensive datasets, delivering even more precise and efficient results. This means these models will address existing issues and lead to innovation in future financial fraud prevention.

4 Reflection and Ethical Considerations

4.1 Personal Analysis

Using the evidence highlighted above, decision tree-based machine learning systems are undoubtedly a strong net positive for combating financial fraud and providing society with broad benefits. As demonstrated in their application at American Express, HSBC and beyond, these tools have shown immense power in clamping down criminal activity. They have objectively reduced false positives and enhanced detection accuracy, directly protecting consumers and bolstering global financial security.

It is imperative that we also consider their limitations. According to the 2024 ACFE Report to the Nations, 43% of fraud cases are still detected through human reporting, such as employee, customer or vendor tips. There is a danger in using these models as an autonomous solution to financial fraud. This statistic shows that for the models to be a true success in fraud detection, collaboration with human involvement is imperative. ML systems, including decision tree models, cannot yet address every possible scenario or nuance in our financial systems; therefore, this gap needs to be filled with human domain knowledge.

Although decision tree models excel at identifying patterns and anomalies, their effectiveness depends heavily on complementing, rather than replacing, human administration. The relationship between human involvement and these complex ML techniques will ensure that the system remains trustworthy, robust and adaptable to future threats. Ultimately, if this can be achieved, it will contribute to building a more secure and reliable financial global economy.

4.2 Ethical Considerations

The power of decision tree-based AI systems is not used exclusively for good, and this dual-use potential presents significant ethical challenges. While these systems have demonstrated their ability to establish stronger financial crime detection mechanisms, their sophistication can also be exploited for malicious purposes. For instance, tools like DeepLocker illustrate how AI, leveraging decision tree-based methodologies, can target financial systems and create stealthy and evasive malware. This capability is not theoretical; it exists today, raising serious concerns about the risks such systems pose to cybersecurity (Citibank, 2024).

Additionally, bias within the training datasets of Decision tree models can lead to unfair and discriminatory outcomes (Adeyelu, Ugochukwu, and Shonibare, 2024). For example, suppose the data the model is trained on is based on historical biases or imbalanced data. In that case, these may label specific demographics or regions as high-risk, perpetuating social inequalities within finance. This leads to the unfair targeting of individuals or groups as threats, but it also undermines the trustworthiness of global financial institutions.

To ensure these systems remain a force for good, regular audits should be carried out on transactional data used in the models' training data, as well as incorporating more awareness-aware algorithms within their development to prevent such discrimination (Adeyelu, Ugochukwu, and Shonibare, 2024)

5 Conclusion

In conclusion, decision trees have developed to be a transformative tool in the fight against financial fraud. This report has shown how effective they are in real-world scenarios, such as those employed by AmEx and HSBC, where their application has empirically brought down the rate of false positives, improved fraud detection accuracy and strengthened the global finance sector. The inherent ability to handle complex datasets, such as transactional data, and their ability to identify emerging fraud patterns highlights their value to both the financial industry and society. In addition to this, their adaptability has made them a necessary component in the ongoing contemporary fight against financial fraud.

Furthermore, beyond their technical merits, society has benefitted from decision tree models. This report shows that they have improved global consumer protection, reduced operational costs for financial institutions and increased trust in global economic systems such as financial markets.

The ethical considerations of future developments of this type of ML model must be considered. Challenges such as bias in training datasets and the potential misuse of AI for malicious purposes (e.g., Deeplocker) are present, and developers must be vigilant when developing their models. XAI tools have proven essential in addressing transparency and accountability within AI models, ensuring that the model's decision-making processes are fair and bolstering stakeholder confidence, which will help mitigate some of the ethical concerns raised here.

In the future, decision tree-based models will significantly impact how AI can transform emerging fields, such as cryptocurrency fraud prevention and fraud detection in underbanked regions. As the technologies improve, research must focus on improving bias mitigation, enhancing interpretability and expanding their potential use cases to a broader range of domains. Combining technological advancements with human oversight will also be essential to unlocking the full potential while safeguarding the possible risks. If these challenges can be overcome, decision tree-based AI systems can continue to drive innovation and contribute to a more secure and fair financial tomorrow.

References

Adeyelu, Oluwatobi Opeyemi, Chinonye Esther Ugochukwu, and Mutiu Alade Shonibare (Apr. 2024). "Ethical Implications of AI in Financial Decision-Making: A Review with Real World Applications". In: *International Journal of Applied Research in Social Sciences* 6.4. Accessed December 15th, 2024, pp. 608–630. URL: http://www.creativecommons.org/licences/by-nc/4.0/.

Afriyie, Jonathan Kwaku et al. (Mar. 2023). "A supervised machine learning algorithm for detecting and predicting fraud in credit card transactions". In: *Decision Analytics Journal* 6. Accessed December 16th, 2024. URL: https://www.sciencedirect.com/science/article/pii/S2772662223000036.

Albashrawi, Mohamed (2016). "Detecting Financial Fraud Using Data Mining Techniques: A Decade Review from 2004 to 2015". In: *Journal of Data Science* 14. Accessed December 12th, 2024, pp. 553–570. URL: https://repository.rit.edu/cgi/viewcontent.cgi?article=11833&context=theses.

Arya, Vivek et al. (2020). "AI Explainability 360: An Interpretability Toolkit for Financial Applications". In: *IBM Journal of Research and Development* 64.4/5. Accessed December 15th, 2024, 10:1–10:12. DOI: 10.1147/JRD.2020.2987014.

Athey, Susan (2019). *The impact of machine learning on economics*. Ed. by Ajay Agrawal, Joshua Gans, and Avi Goldfarb. Accessed December 10th, 2024. University of Chicago Press, pp. 507–547. URL: http://www.nber.org/chapters/c14009.

Bhattacharyya, S. et al. (2011). "Data Mining for Credit Card Fraud: A Comparative Study". In: *Decision Support Systems* 50.3. Accessed December 16th, 2024, pp. 602–613. DOI: 10.1016/j.dss.2010.08.008.

Breiman, Leo (2001). "Random Forests". In: *Machine Learning* 45.1. Accessed December 16th, 2024, pp. 5–32. DOI: 10.1023/A:1010933404324. URL: https://www.researchgate.net/publication/275342330_Random_Forests.

Certified Fraud Examiners, Association of (2024). *Occupational Fraud 2024: A Report to the Nations*. Accessed December 7, 2024. URL: https://legacy.acfe.com/report-to-the-nations/2024/.

Chaudhary, K., J. Yadav, and B. Mallick (2012). "A review of fraud detection techniques: Credit card". In: *International Journal of Computer Applications* 45. Accessed December 10th, 2024, pp. 39–44. URL: https://www.researchgate.net/publication/266486450_A_review_of_Fraud_Detection_Techniques_Credit_Card.

Chen, Tianqi and Carlos Guestrin (2016). "XGBoost: A Scalable Tree Boosting System". In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD '16. Accessed December 18th, 2024. San Francisco, California, USA: Association for Computing Machinery, pp. 785–794. ISBN: 9781450342322. DOI: 10.1145/2939672.2939785. URL: https://doi.org/10.1145/2939672.2939785.

Citibank (2024). *Managing Cyber Risk with Human Intelligence*. Accessed December 18, 2024. URL: https://www.citibank.com/tts/sa/cybersecurity-toolkit/assets/docs/fraud-overview/Managing_Cyber_Risk_with_Human_Intelligence.pdf.

Cloud, Google (2024). *How HSBC Fights Money Launderers with Artificial Intelligence*. Accessed December 7, 2024. URL: https://cloud.google.com/blog/topics/financial-services/how-hsbc-fights-money-launderers-with-artificial-intelligence.

Emerj (2024). *Artificial Intelligence at American Express*. Accessed December 7, 2024. URL: https://emerj.com/artificial-intelligence-at-american-express/.

Flondor, Elena, Liliana Donath, and Mihaela Neamtu (2024). "Automatic Card Fraud Detection Based on Decision Tree Algorithm". In: *Applied Artificial Intelligence* 38.1. Accessed December 10th, 2024, p. 2385249. DOI: 10.1080/08839514.2024.2385249.

Gonzalez, Victoria (2024). "Evaluating Interpretable Models for Financial Fraud Detection". In: *Proceedings of the Americas Conference on Information Systems (AMCIS)*. Accessed December 10th, 2024. San Francisco, CA: Association for Information Systems. URL: https://aisel.aisnet.org/amcis2024/acctinfosys/acctinfosys/1/.

Han, Jiawei, Micheline Kamber, and Jian Pei (2011). *Data Mining: Concepts and Techniques*. 3rd. San Francisco, CA: Morgan Kaufmann. ISBN: 9780123814791.

Higson.io (2024). *Enhancing Fraud Detection in Banking with Rule-Based Decision Engines*. Accessed December 7, 2024. URL: https://www.higson.io/blog/enhancing-fraud-detection-in-banking-with-rule-based-decision-engines.

HSBC (2024). *Harnessing the Power of AI to Fight Financial Crime*. Accessed December 7, 2024. URL: https://www.hsbc.com/news-and-views/views/hsbc-views/harnessing-the-power-of-ai-to-fight-financial-crime.

Jones, Alice and Robert Taylor (2024). "Enhancing Cryptocurrency Fraud Detection with Decision Tree Models". In: *Journal of Emerging Financial Technologies* 12.3. Accessed December 16th, 2024, pp. 45–56. URL: https://www.fepbl.com/index.php/csitrj/article/download/1201/1429.

Ke, Guolin et al. (2017). "LightGBM: A Highly Efficient Gradient Boosting Decision Tree". In: *Advances in Neural Information Processing Systems* 30. Accessed December 16th, 2024, pp. 3149–3157. URL: https://proceedings.neurips.cc/paper/6907-lightgbm-a-highly-efficient-gradient-boosting-decision-tree.pdf.

Koetsier, John (2024). *How American Express Uses AI to Automate 8 Billion Decisions with $1 Trillion at Stake*. Accessed December 7, 2024. URL: https://johnkoetsier.com/how-american-express-uses-ai-to-automate-8-billion-decisions-with-1-trillion-at-stake/.

Lee, Luke (2024). "Enhancing Financial Inclusion and Regulatory Challenges: A Critical Analysis of Digital Banks and Alternative Lenders Through Digital Platforms, Machine Learning, and Large Language Models Integration". In: *arXiv preprint arXiv:2401.12345*. Accessed December 18th, 2024. URL: https://arxiv.org/abs/2401.12345.

Russell, Stuart and Peter Norvig (2021). *Artificial Intelligence: A Modern Approach, Global Edition*. 4th ed. Accessed December 10th, 2024. Harlow, United Kingdom: Pearson Education Limited. ISBN: 9781292401133.

Sahour, Hossein et al. (2021). "Random forest and extreme gradient boosting algorithms for streamflow modeling using vessel features and tree-rings". In: *Environmental Earth Sciences* 80. Accessed December 16th, 2024. DOI: 10.1007/s12665-021-10054-5.

Smith, John and Jane Doe (2024). "Fraud Detection in Financial Transactions using Hybrid AI Systems". In: *arXiv preprint arXiv:2408.12989v1*. URL: https://arxiv.org/html/2408.12989v1.

Toreini, E. et al. (2020). "The Relationship Between Trust in AI and Trustworthy Machine Learning Technologies". In: *Proceedings of the 2020 Conference on*

Fairness, Accountability, and Transparency. Accessed December 12th, 2024, pp. 272–283. DOI: 10.1145/3351095.3372834.

Vuille, Marius, Dimitar Jetchev, and Abson Sae-Tang (Nov. 2023). *The Rise of Financial Fraud in the Digital Era and the Role of Explainable AI*. Blog post on Inpher.io. Accessed December 02, 2024. URL: https://inpher.io/blog/financial-fraud-and-explainable-ai/.

Wang, Y. et al. (2021). "Graph Neural Networks in Financial Fraud Detection: A Survey". In: *ACM Computing Surveys* 54.6. Accessed December 16th, 2024, pp. 1–34. DOI: 10.1145/3464427.

West, J. and M. Bhattacharya (2016). "Intelligent financial fraud detection: A comprehensive review". In: *Computers & Security* 57. Accessed December 12th, 2024, pp. 47–66. DOI: 10.1016/j.cose.2015.09.005.

Zhu, Z.-H. (2021). *Machine Learning*. Singapore: Springer Nature. DOI: 10.1007/978-987-15-1967-3.