Detecting Financial Fraud in the Digital Age: The AI and ML Revolution

Dr. Sathisha H K[1], Dr. Sowmya G S[2]

[1]Assistant Professor, Govt. R C College of Commerce and Management, Bengaluru
[2]Associate Professor, Centre for PG Studies, Sindhi College, Bengaluru

Abstract:

Along with innovations in technology, swindlers are becoming smarter. Fraud is a major concern for the financial services industry. Earlier frauds in this industry were limited to robbery and loan defaults as the customers had limited ways to interact with their financial institutions. The technology innovation opened up various ways of interacting and transacting with financial services providers, which has led to a significant increase in several types of financial fraud. Technology is making a positive impact across many industries, including the financial sector. Technology is redefining what is possible in combating fraud and financial crimes. Artificial Intelligence (AI) and Machine Learning (ML) have grown drastically from being mere technology buzzwords to the commercial reality it is today. The aim of the present study is to understand the types of financial fraud and the role of AI and ML in financial fraud detection. Secondary data collected from websites and reports are used for the study.

Keywords: Digital Age, Money laundering, AI, ML, Financial Fraud, Predictive models

Introduction

Fraud exists in all walks of life and technological advancement in the field of commerce at the global level have evidenced new opportunities for online usage and on the other hand more opportunities for new financial fraud and abuse. With the growth of big data, artificial intelligence, and machine learning, new opportunities have arisen in using artificial intelligence and machine learning models to detect fraud.

Artificial Intelligence and Machine Learning are redefining the way companies do fraud detection these days. Before the advent of AI and ML, fraud prevention systems relied on set patterns and analyzed only restricted fraud patterns. But things have changed now, all credit to Artificial Intelligence and Machine Learning. By combining supervised learning algorithms trained on previous data with unsupervised learning, Artificial Intelligence helps to gain a better understanding of the customers' behaviors thus making it a good option for the effective detection of new and emerging fraud attacks.

Literature Review

Fraud has been a major issue in the financial service industry for years. Due to the increase in online financial transactions, fraudulent activities have also increased. With the advent of new technology, Artificial Intelligence and Machine Learning play a greater role in detecting financial crimes and frauds.

The following reviews showcase the usage of artificial intelligence, and machine learning in financial fraud detection, which are used as a base for conducting the present study.

Choi & Lee (2018) in their research work an artificial intelligence approach to financial fraud detection under IoT environment: a survey of implementation surveyed the financial fraud methods using machine learning and deep learning methodology and also proposed a process for accurate fraud detection. Based on the advantages and limitations of the existing methods, they have developed a model for fraud detection and validated the model with actual financial transactions data of the year 2015 from Korea. West & Bhattacharya (2016) in their research work intelligent financial fraud detection: a comprehensive review provided a review of financial fraud detection methods using data mining methods and computational intelligence-based techniques. Scientific literature for the period 2004-2014, was analyzed in the study. They have explored some types of fraud and data mining methods. They have also mentioned that there is an opportunity to examine the performance of existing methods and cost-benefit analysis of computational fraud detection.

Bolton & Hand (2002) in their research work statistical fraud detection: a review, described the tools available for statistical fraud detection and the areas in which fraud detection technologies are most used. The researchers mentioned that statistics and machine learning provide effective technologies for fraud detection and have been applied successfully to detect activities such as money laundering, e-commerce credit card fraud etc.

Statement of the Problem

Financial fraud is a growing concern in the modern financial system, causing significant losses to individuals, organizations, and economies. Traditional methods of detecting financial fraud, such as rule-based systems, statistical analysis, expert systems, and others, have limitations in terms of accuracy, speed, and adaptability to changing fraud patterns. As a result, there is a growing demand for more advanced and sophisticated methods of financial fraud detection. Artificial intelligence (AI) and machine learning (ML) offer potential solutions to address the limitations of traditional fraud detection methods. However, the effective implementation of AI and ML in financial fraud detection requires a thorough understanding of the algorithms, techniques, and tools used, as well as the challenges and limitations associated with their use. The present study aims at exploring the use of AI and ML algorithms for financial fraud detection.

Objective and Methodology

The objectives of the present study are to understand the different types of financial fraud, to investigate the current methods of detecting financial fraud and their limitations and to explore the potential of using AI and ML algorithms for financial fraud detection. To achieve these objectives the study follows descriptive methodology and secondary data collected from published reports and articles are used for the study.

Research Discussion

Frauds are the biggest challenge for the finance industry and its customers which results in huge losses. With the increase in the number of transactions happening across the globe, the threat of financial

fraud has increased too. The finance industry is now utilizing state of art artificial intelligence and machine learning technologies to capture these frauds as early as possible and prevent them from taking place.

Types of Financial Fraud

Financial frauds are of many different types. Some of the major types of frauds selected from the list provided in the Federal Bureau of Investigation, Financial Crimes Report (2010-2011), United States are listed below (Figure 1)
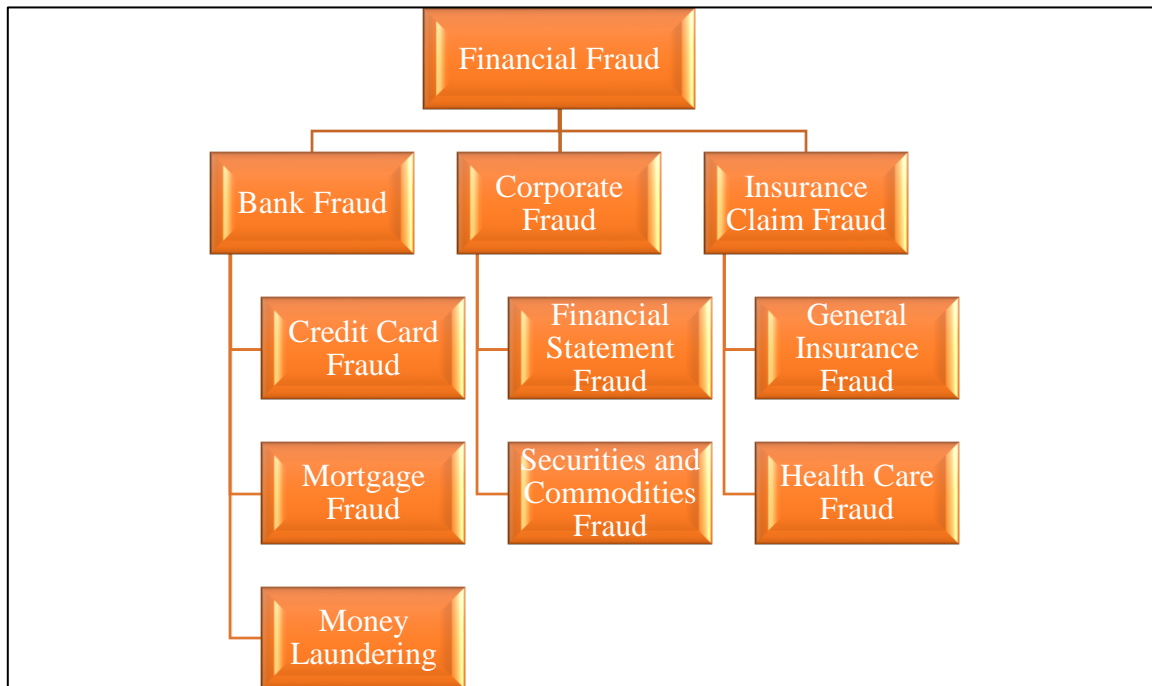

Figure 1: Types of Financial Fraud

❖ **Credit Card Fraud:** Credit card fraud refers to the unauthorized use of a person's credit card to perform fraudulent transactions without the user's knowledge. The transactions can be performed using the physical card, where the card was either lost or stolen, but is often performed remotely. The cardholder's information may be acquired by phishing, which involves a fraudster impersonating a finance official to convince the user to divulge their details, swipers or skimmers provide an interface to an ATM or POS device which can read the card directly, or entire databases of user's information can be obtained if the fraudster is able to breach the financial institution's network security or enlist the help of an assistant within the company. The ambiguity and availability of these remote methods have given rise to the prevalence of organized crime in credit card fraud.

❖ **Mortgage Fraud:** Mortgage fraud is a specific form of financial fraud that refers to the manipulation of property or mortgage documents. It is often committed to misrepresent the value of a property for the purpose of influencing a lender to fund a loan for it.

❖ **Money Laundering:** Money laundering is a method used by criminals to insert proceeds obtained from criminal ventures into valid businesses. This hides the origin of the money, giving them the appearance of legal income and making it difficult to track their crimes.

❖ **Financial Statement Fraud:** Financial statements are the documents published by a company that elucidates details such as their expenses, loans, income, and profits. The various financial statements

that the company releases give an overall picture of the company's status, and can be used to indicate how successful the company is, how it influences stock prices, etc.

Financial statement fraud, also known as corporate fraud, involves falsifying these statements to make the company appear more profitable. it is difficult to diagnose financial statement fraud because of a general lack of understanding of the field, the infrequency in which it occurs, and the fact that it is usually committed by knowledgeable people within the industry who are capable of covering their dishonesty.

❖ **Securities and Commodities Fraud**: Securities and commodities fraud refers to a variety of methods by which a person is misled for investing in a company stock/commodity based on false information. It includes Pyramid Schemes, Ponzi Schemes, Hedge Fund Fraud, Foreign Exchange Fraud, Embezzlement, etc.

❖ **Insurance Fraud:** Insurance fraud is a fraud that can be committed at any point during the insurance process, and by any people in the chain. Insurance claims fraud occurs when a customer submits a fraudulent insurance claim as a result of an exaggerated injury or loss of assets, or a completely fraudulent event. A common form of claims fraud is automobile insurance fraud, which is often committed by faking or intentionally committing accidents that result in excessive repair and injury costs.

Financial Fraud Trend in India

Financial fraud trends in India according to a report titled Experian India Fraud Report 2018-19 by Experian Services India Pvt. Ltd., a credit information services company is shown in Figure 2. Identity theft accounted for 28 percent of all frauds, while Market Alert Fraud (MAF) accounted for another 28 percent, followed by fraud contact information at 25 percent and document fabrication at 10 percent.
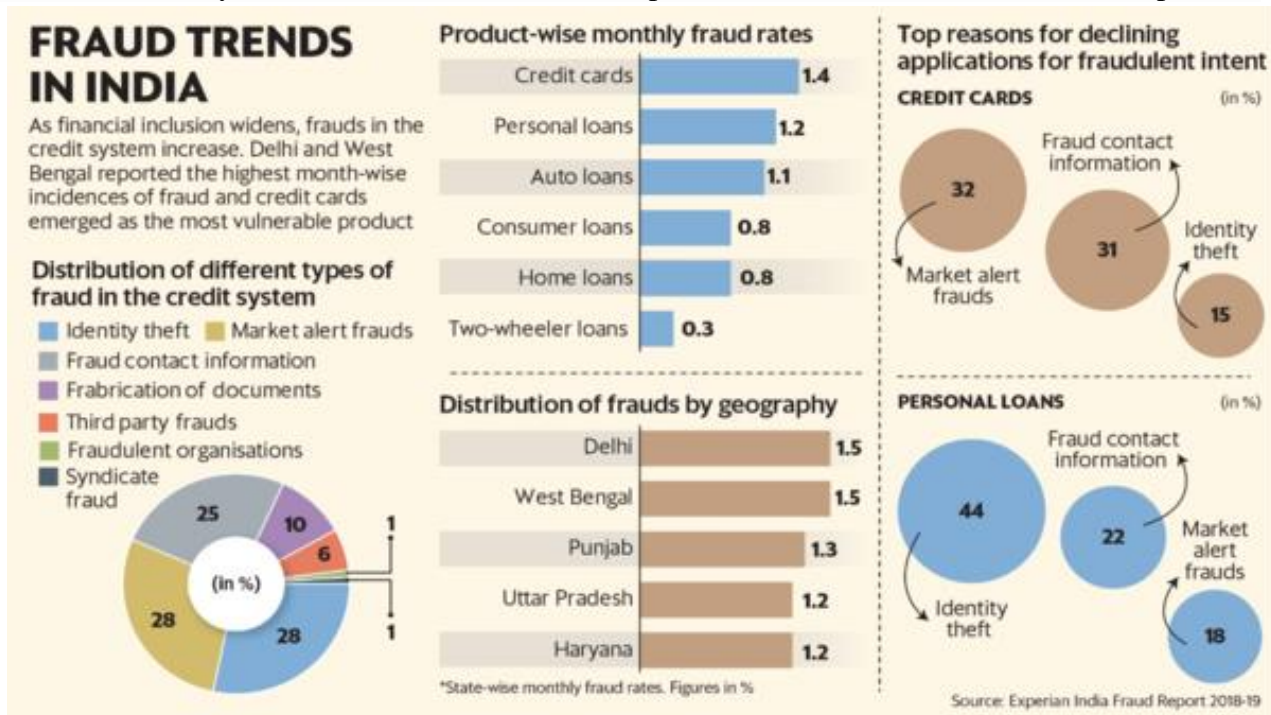
Source: livemint.com
Figure 2: Financial Fraud Trends in India

Current methods of detecting financial fraud and their limitations

Current methods of detecting financial fraud and their limitations are:

- *Rule-based systems*: These systems use predefined rules and criteria to identify suspicious transactions. The limitations of these systems include being prone to false alarms, being unable to detect new and unknown forms of fraud, and being unable to adapt to changes in fraudulent activities.
- *Statistical analysis*: This method involves analyzing large amounts of financial data to identify patterns and anomalies that may indicate fraud. Limitations of this method include being unable to detect fraud in real-time and the need for large amounts of historical data to be effective.
- *Expert systems:* These systems rely on human expertise and knowledge to identify fraud. Limitations include being prone to human error, being limited by the expertise of the individuals involved, and being time-consuming.

Overall, current methods of detecting financial fraud have limitations in terms of accuracy, speed, and adaptability to changing fraud patterns. This highlights the need for a more advanced and sophisticated approach using AI and machine learning.

Artificial Intelligence (AI)

Artificial Intelligence refers to the simulation of human intelligence in machines that are programmed to think like humans and imitate their actions. The ideal characteristic of artificial intelligence is its ability to rationalize and take actions that have the best chance of achieving a specific goal.

Machine Learning (ML)

A subset of artificial intelligence is machine learning. Machine learning (ML) is the study and practice of designing and implementing algorithms that can learn from the past. Without arousing the suspicions of people completing transactions. Deep learning techniques enable this automatic learning through the absorption of huge amounts of unstructured data such as text, images, or video. Machine learning can combat financial fraud by utilising huge data more effectively and rapidly than humans ever could.
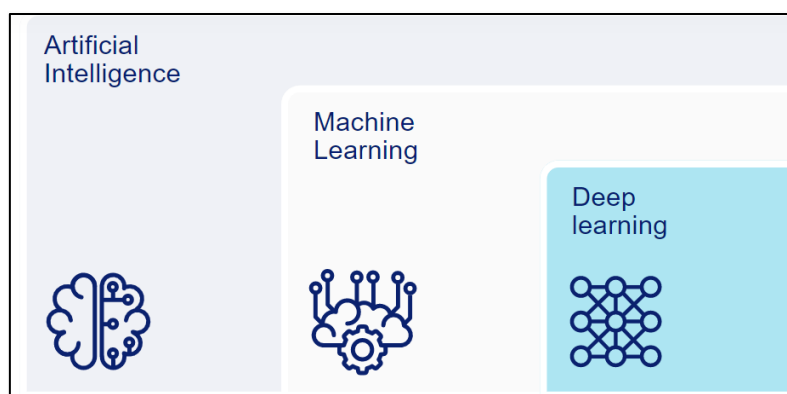
Figure 3: Artificial Intelligence vs Machine Learning

Role of AI and ML in Fraud Detection

Artificial Intelligence and Machine Learning have popped up as powerful technology that has the potential to prevent financial fraud. Fraud Detection with Machine Learning becomes possible due to the ability of ML algorithms to learn from historical fraud patterns and recognize them in future transactions.

Machine Learning algorithms appear more effective than humans when it comes to the speed of information processing. Also, ML algorithms are able to find sophisticated fraud traits that a human simply cannot detect.
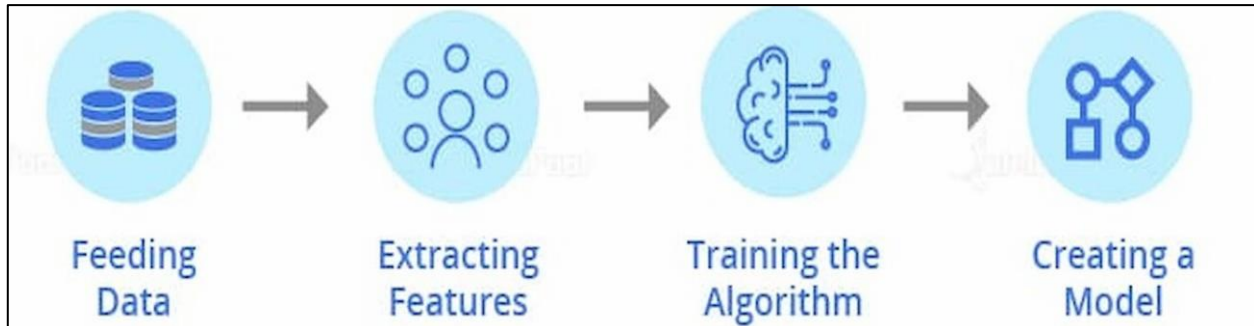

Source: intellipaat blog

Figure 4: Basic structure of the working of fraud detection algorithms using Machine Learning

Figure 4 depicts the basic structure of the working of fraud detection algorithms using machine learning. The first step is to feed the data into the model. The next step is to extract the features of each and every thread associated with the transaction process. Once the fraud detection algorithm is created, it has to be trained by providing customers data, to help the algorithm to learn how to distinguish between fraud and genuine transactions. Once the training is completed with the specific data set, the model is ready for detecting fraudulent and non-fraudulent transactions.

Artificial Intelligence (AI) and Machine Learning (ML) play a significant role in financial fraud detection. They help organizations identify and prevent fraudulent activities more effectively and efficiently. AI algorithms can analyze large amounts of data and identify patterns and anomalies that may indicate fraudulent behavior. ML algorithms, in particular, can be trained to detect fraud based on historical data and make predictions about the likelihood of future fraud. Some specific ways in which AI and ML are used for financial fraud detection are:

- *Behavioral analysis*: AI algorithms can analyze customer transactions and behaviors to identify any unusual or suspicious activities.
- *Risk scoring:* AI and ML algorithms can analyze customer data to determine the risk level associated with a particular transaction, helping organizations prioritize their investigations.
- *Fraud detection models:* ML algorithms can be trained to detect fraud based on historical data, and then applied to new transactions to identify potential fraud in real time.
- *Natural language processing (NLP):* AI algorithms can analyze large amounts of text data, such as emails, chat logs, and customer feedback, to identify any instances of fraudulent behavior.
- *Fraud network detection:* AI algorithms can analyze large amounts of data to identify relationships between different fraudulent activities, helping organizations understand the extent of fraud and identify the key players involved in these activities.

Fraud detection is the most important use case of artificial intelligence. Artificial intelligence improves fraud detection by combining supervised learning algorithms with unsupervised learning to the effect of gaining a better understanding of customers' behaviors. A better understanding of customers' behaviors allows organizations to better identify and prevent unauthorized activity.
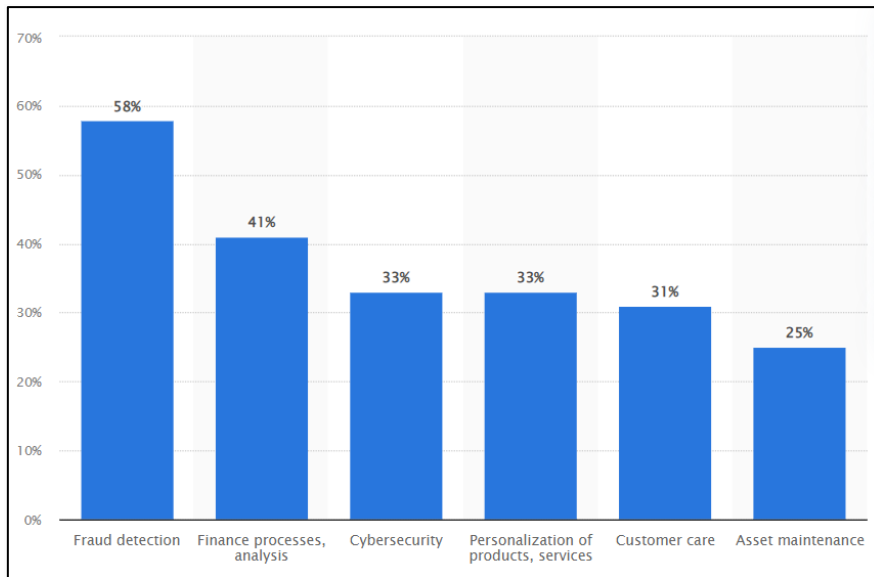
Source: Published by Bergur Thormundsson, Mar 17, 2022

Figure 5: AI use cases in financial services industry worldwide as of 2020

AI Techniques in Fraud Detection:

The main AI techniques used for fraud detection include:

- **Data mining** to classify, cluster, and segment the data and automatically find associations and rules in the data that may signify interesting patterns, including those related to fraud.
- **Expert systems** to encode expertise for detecting fraud in the form of rules.
- **Pattern recognition** to detect approximate classes, clusters, or patterns of suspicious behavior either automatically (unsupervised) or to match given inputs.
- **Machine learning** techniques to automatically identify characteristics of fraud.
- **Neural nets** to independently generate classification, clustering, generalization, and forecasting that can then be compared against conclusions raised in internal audits or formal financial documents such as 10-Q.

Other techniques such as link analysis, Bayesian networks, decision theory, and sequence matching are also used for fraud detection. A new and novel technique called the System properties approach has also been employed wherever rank data is available.

Statistical analysis of research data is the most comprehensive method for determining if data fraud exists. Data fraud as defined by the Office of Research Integrity (ORI) includes fabrication, falsification, and plagiarism.

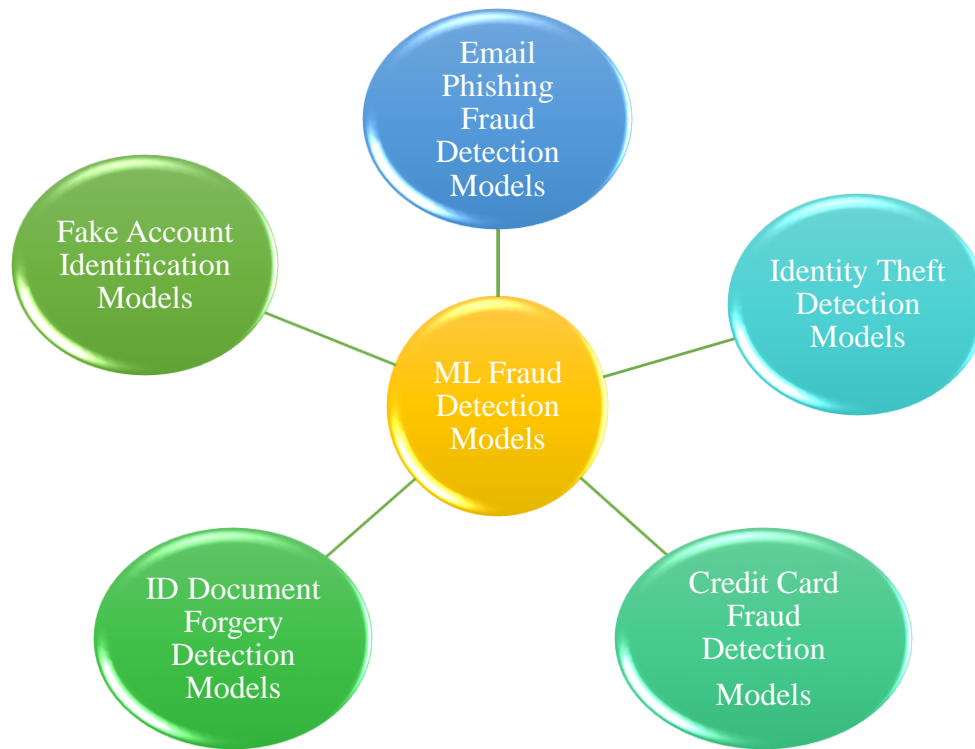**Machine Learning Fraud Detection Models**

Figure 6: Machine Learning Fraud Detection Models

Some of the machine learning fraud detection models are listed below:

❖ **Email Phishing Fraud Detection Models**: Phishing emails represent spam letters that have falsified intentions. Phishers make fake websites and their URLs are very similar both visually and semantically to the originals. They are mostly threats to the Banking sector, multinational companies, and medical establishments.

Logistic regression is one of the machine learning algorithms used for phishing detection. Other ways are of using traditional machine learning classification models such as SVM, Naive Bayes, and Extreme Learning Machines

❖ **Identity Theft Detection Models:** Identity theft detection is considered an anomaly detection challenge, so various state-of-art unsupervised Machine Learning algorithms such as LOF, PCA, one-class SVM, and Isolation Forest help find abnormal patterns of a user's behavior in order to detect unauthorized actions. They work as a litmus test to find anomalies in the field of normal behavior. These algorithms group abnormal behavior data points together in a dense cluster than differs from clusters of normal behaviors.

❖ **Credit Card Fraud Detection Models:** Fraud models can be tackled with both supervised and unsupervised Machine Learning algorithms.

❖ **ID Document Forgery Detection Models:** ID document forgery detection deals with image processing. Certain techniques are used to make sense of the visual information that an image carries. CNN models are usually trained to perform this task, whereas neural networks are built in a way to minimize losses. Forgery detection technique relies on hyperspectral image analysis. This method implies building an electromagnetic spectrum map to obtain the spectrum for each pixel in the image.

❖ **Fake Account Identification Models:** Fake account identification is a classification problem, therefore the first step is selecting the profile that needs to be classified as fake. The most important part of classification is feature selection, i.e., parameters such as rate of engagement, activity, number of followers compared to the number of people the account is following, and the relevancy of comments. After the feature matrix is built, it is fed into the classification model — which may be one of the most efficient binary classifiers, such as Naive Bayes, SVM, Decision Trees, Logistic Regression, etc. The classifier can be continually trained with new data on fake and real accounts, which helps increase the accuracy of its predictions.

Real-World Examples of Fraud Detection Using Machine Learning

Businesses across industries such as e-commerce, banking, online gaming, and healthcare are using artificial intelligence and machine learning to detect financial fraud. No matter what industry you are in, you can always benefit from the power of AI and ML to process large amounts of data and detect patterns to protect against fraud. For example, ML and AI in finance detect account takeovers, unauthorized access, and other fraud by detecting patterns in customer behavior. Some real-world examples of companies that are already leveraging the power of financial fraud detection through machine learning are:

- Compliance.ai
- PayPal
- MasterCard
- Feedzai

Compliace.ai: Compliance.ai compliance management software leverages adaptive machine learning models in the financial industry to automate research and track financial regulatory updates on a single platform.

PayPal: PayPal, a leading global fintech company, is also using machine learning to improve its fraud detection and risk management capabilities. Using a combination of linear neural networks and deep learning techniques, PayPal's risk management engines can determine the risk levels associated with a customer within milliseconds.

MasterCard: MasterCard, the world's second-largest payments company, integrates AI and ML to track and process variables such as transaction time, size, location, device, and purchase data. MasterCard's fraud detection, based on ML, assesses customer account behavior for each transaction and provides real-time insight into whether a transaction is genuine or fraudulent.

Feedzai: Feedzai, a fintech company, develops real-time machine learning solutions to detect fraudulent payment transactions in finance, retail, e-commerce and other industries. The company believes that a fine-tuned machine learning tool can detect up to 95% of all fraud while reducing human labor in the investigation phase, which accounts for 25% of fraud expenditure.

Brief statistics where fraud detection worked:

- Highmark Inc. has saved hundreds of millions of dollars in 2019 due to fraud, waste, and abuse, amounting to around $250 million. Moreover, they have saved over $850 million in the previous five years by deploying AI for fraud detection (Artificial Intelligence).

- Microsoft spends 33 percent of its yearly sales to secure cloud storage from any type of fraud, generating up to 44.5 percent additional profit annually.
- A scam involving sim swapping across Europe was spotted by an AI, which saved over 100 victims and 3.5 million euros, and caught 26 fraudsters.
- A teenager was discovered hacking Twitter and committing fraud totaling more than $110,000, and the money was returned by the fraud detector.
- By detecting the Wirecard Meltdown scam, an AI has saved €70.74 million.

Benefits of using AI and ML in Fraud Detection:

When processing massive datasets, machines are far superior to humans. They can discover and distinguish thousands of purchase journey patterns, as opposed to the handful caught by rules. Some of the benefits of fraud detection using machine learning are shown in Figure 7.
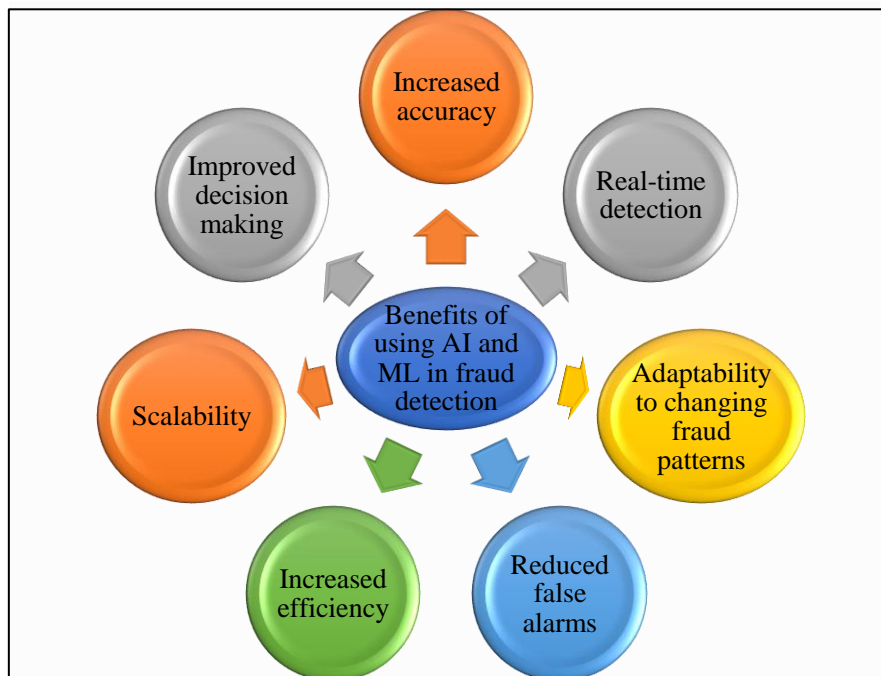

Figure 7: Benefits of AI and ML in Fraud Detection

The benefits of using Artificial Intelligence (AI) and Machine Learning (ML) in fraud detection include:

- *Increased accuracy*: AI and ML algorithms can analyze large amounts of financial data and identify patterns and anomalies that may indicate fraud, leading to improved accuracy in detecting fraud.
- *Real-time detection*: AI and ML algorithms can process large amounts of financial data in real-time, allowing for the early detection of fraud.
- *Adaptability to changing fraud patterns*: AI and ML algorithms can learn and adapt to changes in fraudulent activities, allowing them to evolve and improve over time.
- *Reduced false alarms:* AI and ML algorithms can reduce false alarms by analyzing large amounts of data and making more accurate predictions.
- *Increased efficiency*: AI and ML algorithms can automate and streamline the fraud detection process, reducing the need for manual intervention and freeing up resources for other tasks.

- *Scalability:* AI and ML algorithms can be easily scaled to handle large amounts of data, making them suitable for use in large financial systems.
- *Improved decision-making:* AI and ML algorithms can provide insights and recommendations for decision-making, helping organizations make more informed decisions about their financial activities.

Conclusion

The emergence of artificial intelligence and machine learning technologies has ushered in a transformative era in the battle against financial fraud in the digital age. This research study has provided a comprehensive overview of the significant strides made in fraud detection through AI and ML, highlighting their potential to revolutionize the financial industry's approach to security. From sophisticated anomaly detection algorithms to predictive modeling and real-time monitoring, these advancements offer powerful tools for safeguarding financial systems and protecting the interests of individuals and organizations alike. As we move forward, it is imperative that policymakers, financial institutions, and researchers continue to collaborate and adapt to the evolving landscape of financial fraud, harnessing the full potential of AI and ML to stay one step ahead of increasingly sophisticated fraudulent activities. This synergy between technology and human expertise holds the promise of a more secure, efficient, and resilient financial ecosystem for the digital age.

References:

1. Benson Edwin Raj, S. and Annie Portia, A. (2011). Analysis on credit card fraud detection methods. In: *2011 International Conference on Computer, Communication and Electrical Technology (ICCCET)*. March 2011. ieeexplore.ieee.org. pp.152–156.
2. Choi, D., & Lee, K. (2018). An Artificial Intelligence Approach to Financial Fraud Detection under IoT Environment: A Survey and Implementation. *Security and Communication Networks*, 1–15.
3. Das, S. et al. (2015). Applications of artificial intelligence in machine learning: review and prospect. *International Journal of Computer*. [Online]. Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.695.5829&rep=rep1&type=pdf
4. Paruchuri, H. (2017). Credit Card Fraud Detection using Machine Learning: A Systematic Literature Review. *ABC Journal of Advanced Research*, 6 (2), pp.113–120.
5. Richard J. Bolton. David J. Hand. (2002). Statistical Fraud Detection: A Review. *Statist. Sci*. 17 (3), 235 – 255.
6. Ryman-Tubb, N. F., Krause, P. and Garn, W. (2018). How Artificial Intelligence and machine learning research impacts payment card fraud detection: A survey and industry benchmark. *Engineering applications of artificial intelligence*, 76, pp.130–157.
7. Smeureanu, I., Ruxanda, G. and Badea, L. M. (2013). Customer segmentation in private banking sector using machine learning techniques. *Journal of business*. [Online]. Available at:https://www.tandfonline.com/doi/abs/10.3846/16111699.2012.749807
8. Tripathy, A. and Rath, S. K. (2014). Application of Natural Language Processing in Object-Oriented Software Development. In: *2014 International Conference on Recent Trends in Information Technology*. April 2014. ieeexplore.ieee.org. pp.1–7.
9. Tsai, C.-F. and Chen, M.-L. (2010). Credit rating by hybrid machine learning techniques. *Applied soft computing*, 10 (2), pp.374–380

10. Viaene S, Ayuso M, Guillen M, Van Gheel D, and Dedene G. Strategies for detecting fraudulent claims in the automobile insurance industry. *European Journal of Operational Research,* 176, 565-83, 2007

11. West, J., & Bhattacharya, M. (2016). Intelligent financial fraud detection: A comprehensive review. *Computers & Security*, 57, 47–66.