International Journal of HRM and Organizational Behavior

www.ijhrmob.com

editor@ijhrmob.com

International Journal of
HRM and Organizational Behavior
ISSN 2454 - 5015

https://doi.org/10.62650/ijhrmob.2021.v9.i4.pp9-25

Machine Learning-Driven AI for Financial Fraud Detection in IoT Environments

Thirusubramanian Ganesan

Sr. Software Engineer - CDE Full Stack Eng.

Cognizant Technology Solutions

Email: 25thiru25@gmail.com

Abstract:

Artificial intelligence (AI) driven by machine learning has revolutionized the identification of financial fraud in Internet of Things (IoT) environments. This technique quickly and accurately identifies suspicious patterns in the vast and diverse data streams from IoT devices through the application of advanced algorithms, potentially identifying fraudulent activity. Using methods like anomaly detection and clustering, along with supervised and unsupervised learning that are trained on historical transaction data, artificial intelligence systems are able to distinguish between legitimate and fraudulent transactions with great accuracy in real time. In order to create trustworthy fraud detection models in Internet of Things environments, this study looks at the methodology, datasets, and assessment metrics that are necessary for adaptive learning through frequent retraining and automatic reaction mechanisms.

Keywords: Supervised learning, Unsupervised learning, Anomaly detection, Clustering, Adaptive learning.

1 Introduction:

The Internet of Things (IoT) has turned into a double-edged sword in today's linked society. Financial fraud has new opportunities as a result of its exceptional efficiency and real-time data processing capabilities across multiple industries. Integrating artificial intelligence (AI) driven by machine learning for fraud detection in Internet of Things (IoT) environments is developing as a cutting-edge approach to address this expanding issue. With this method, massive and complicated data streams are analysed, suspicious patterns are found, and potentially fraudulent behaviours are predicted with high accuracy and speed using sophisticated machine learning algorithms. Businesses may strengthen their security protocols, safeguard financial transactions, and preserve the integrity of their IoT ecosystems by utilizing AI. Sophisticated machine learning algorithms are used in IoT contexts to monitor, analyse, and interpret data exchanged across networked devices in order to detect financial fraud. Large volumes of data are produced by these devices from a variety of sources, such as wearable technology, smart meters, and sensors. The data is used to record user behaviour patterns, transactional data, and device interactions. Through data cleansing, standardization, and integration from many sources, machine learning-driven AI systems pre-process this heterogeneous and unstructured data to guarantee its consistency, completeness, and relevance. These systems extract and

ISSN 2454 - 5015

https://doi.org/10.62650/ijhrmob.2021.v9.i4.pp9-25

choose pertinent characteristics, like transaction amounts, time intervals, geographic locations, and device usage patterns, in order to efficiently detect fraud. Then, using the pre-processed data, machine learning models—such as Random Forest, Support Vector Machines, and Neural Networks—are trained to distinguish between authentic and fraudulent transactions using past data. Apart from supervised learning, unsupervised learning methods such as anomaly detection and clustering can discover departures from typical behaviour and highlight odd patterns that could be signs of fraud. Real-time financial transaction monitoring is facilitated by trained models, which can detect suspicious activity and send out notifications for more investigation. Retraining models with fresh data on a regular basis allows adaptive learning techniques to remain successful against changing fraud strategies. By classifying alerts according to risk scores and automating reactions to specific fraud scenarios, including account freezing or transaction blocking pending verification, these AI systems assist in decision-making. Businesses can take a proactive and dynamic approach to financial fraud detection by integrating AI into IoT systems. This supports the ongoing development and uptake of IoT technologies by improving security and fostering user confidence in the financial sector.

It's a relatively recent development to employ AI powered by machine learning to detect financial fraud in Internet of Things contexts. Combining IoT with AI has its roots in the separate developments of these two disciplines. Traditional machine learning techniques were initially applied to banking and financial data in order to detect fraud using AI. However, the necessity to manage the enormous and varied amount of data generated by networked devices led to a major breakthrough in the integration of AI into IoT systems. SAS and IBM were among the IT giants that were early pioneers in this industry. Some of the first platforms to use AI-driven fraud detection methods designed expressly for IoT data were IBM's Watson IoT platform and SAS's advanced analytics products. These early initiatives cleared the path for this technology to be more widely used in a variety of industries.

A number of software platforms, such as IBM Watson IoT, SAS Fraud Management, Microsoft Azure IoT, Google Cloud IoT, AWS IoT Analytics, Splunk for IoT, TensorFlow, H2O.ai, RapidMiner, and Data Robot, have emerged as industry leaders in offering machine learning-driven AI for financial fraud detection in IoT environments. By automatically identifying and stopping fraudulent activity in real time, integrating AI with IoT for fraud detection greatly improves security. It also increases operational efficiency and builds consumer trust by guaranteeing data protection. Numerous advantages come with these systems, including the capacity to detect suspicious activity in real time, scale to handle large amounts of data from different IoT devices, save money by lowering fraud losses and manual labor, and enhance accuracy thanks to algorithms that learn on a continual basis. These systems' main goals are to prevent fraud, analyse massive amounts of data to find trends in fraud, offer guidance for risk management, and guarantee regulatory compliance. Nevertheless, there are still issues to deal with, such as protecting user privacy, integrating AI-powered systems with current infrastructure, guaranteeing scalability as the number of IoT devices increases, reducing false positives, and following legal requirements. Businesses can greatly improve their capacity to identify and stop financial fraud in IoT environments by tackling these issues and utilizing AI's advantages. This will increase security and confidence in their systems.

ISSN 2454 - 5015

https://doi.org/10.62650/ijhrmob.2021.v9.i4.pp9-25

The main goals of this research are to improve the identification and prevention of financial fraud and customer attrition in Internet of Things (IoT) environments by integrating cutting-edge AI and ML algorithms into CRM systems. Using consumer credit card and financial transaction data, this involves evaluating the effectiveness of several machine learning (ML) algorithms, including Random Forest Classifier, Decision Tree Classifier, Logistic Regression, Support Vector Classifier (SVC), K-Neighbors Classifier, GaussianNB, and Artificial Neural Networks (ANN). In order to provide insightful information for successful preventative tactics, the study seeks to determine which machine learning algorithm has the highest prediction accuracy for identifying fraud and churn. Furthermore, it aims to create a framework for tracking and evaluating real-time data from IoT devices in order to promptly and precisely detect fraudulent activity, enhance CRM systems' predictive capacities for proactive fraud prevention and customer retention, and guarantee adherence to privacy, data security, and legal requirements.

Even though a lot of research has been done on the use of individual machine learning algorithms in other domains, there is a clear lack of thorough comparisons that are specifically designed for CRM and e-business environments. This is especially true when it comes to identifying financial fraud and customer attrition in IoT settings. Research on the integration of ML-driven AI systems capable of real-time data analysis from IoT devices for fraud detection is lacking, as is attention to scalability and integration challenges within existing CRM and IoT infrastructures. Other key research gaps include the paucity of studies comparing the efficacy of different ML algorithms within CRM systems operating in IoT environments, the unrealized potential of ensemble methods like Random Forest Classifier and the interpretability of models like Decision Tree Classifier, and the unmet need for comprehensive frameworks that incorporate various ML algorithms and techniques to optimize fraud detection and customer churn prediction in a unified manner.

Problem Statement:

Businesses now face formidable obstacles as a result of the swift spread of IoT devices and the shift to decentralized edge computing, which have created new risks and complications in the handling of fraud and consumer attrition. Conventional approaches find it difficult to handle the complex and varied data that Internet of Things devices produce, including transactional data, user behaviour, and device interactions. Finding the best machine learning (ML) algorithms to predict and stop financial fraud and customer attrition, balancing interpretability and accuracy requirements in models, enabling real-time monitoring to quickly spot fraudulent activity, guaranteeing scalability and seamless integration into current CRM and IoT infrastructures, and meeting strict regulatory requirements for data security and privacy are some of the major challenges. By utilizing cutting-edge AI and ML approaches into CRM systems, this project aims to build real-time monitoring capabilities, improve predictive accuracy, and ensure compliance with ethical and legal norms in IoT environments.

2 Literature Survey:

Fogel et al. propose that artificial intelligence (AI) is revolutionizing healthcare by enhancing the prevention, detection, diagnosis, and treatment of diseases. Despite concerns that AI might

ISSN 2454 - 5015

https://doi.org/10.62650/ijhrmob.2021.v9.i4.pp9-25

disrupt jobs and the physician-patient relationship, it can actually create more opportunities for human interaction and the application of emotional intelligence. AI has surpassed human performance in several areas, demonstrating its potential to streamline repetitive tasks and enable deeper human connections and better judgment in healthcare settings. Recent studies underscore the potential for a more unified and human-centered healthcare experience through the integration of AI.

Bauder et al. focus on detecting Medicare Part B provider fraud using machine learning techniques in their research paper. They address the challenge of a highly imbalanced dataset by generating seven different class distributions and evaluating the performance of six distinct machine learning models. Their findings indicate that the RF100 model with a 90:10 class distribution performs best, demonstrating that retaining more information from the majority class improves fraud detection. This study underscores the importance of class distribution in enhancing the effectiveness of machine learning methods for detecting Medicare Part B provider fraud, despite the limited number of fraud instances in the dataset.

Li et al.'s research paper investigates the application of intelligent recommendation techniques in restaurants to enhance food service, cost-effectiveness, and customer satisfaction. The study introduces the MARMTF method, which employs a matrix tri-factorization algorithm to predict individual dietary preferences and recommend dishes based on ingredients, spice level, and price. The findings demonstrate that this approach can streamline the process of selecting food and has significant potential for implementation in future restaurants equipped with robotic servers. The integration of modern technologies, such as robotics and artificial intelligence, is highlighted as a means to improve food service and assist consumers in choosing options that align with their taste and nutritional preferences. The development of this intelligent food choice method underscores its potential to simplify decision-making and provide advanced, non-human waitstaff in the restaurant industry.

Goode et al. discuss the growing use of biometric technology by banks to identify and authenticate customers, secure transactions, and prevent fraud. This technology is being integrated across all banking channels, including traditional physical branches and digital platforms, due to its reliability in enhancing customer security. Banks are utilizing biometrics to verify the identities of new customers, authenticate existing ones, and safeguard high-value transactions. Despite its increasing adoption, there are still challenges to fully implementing biometric technology in the banking sector.

Cui et al. present a comprehensive survey on the application of machine learning in the Internet of Things (IoT). The paper highlights recent advancements and practical applications in areas such as traffic engineering, network management, security, and edge computing infrastructure. Emphasizing the growing integration of machine learning in IoT, the survey details its successful deployment in traffic profiling, device identification, and quality of service optimization. Additionally, the paper addresses the research challenges and open issues within this evolving field, underscoring the significant impact of machine learning on enhancing various aspects of IoT networking and management.

Mahdavinejad et al.'s research paper explores the application of machine learning for analysing data from Internet of Things (IoT) devices, particularly within the context of smart cities. The

ISSN 2454 - 5015

https://doi.org/10.62650/ijhrmob.2021.v9.i4.pp9-25

study presents a taxonomy of machine learning algorithms, highlighting their potential and challenges in managing the substantial volume and variety of IoT data. A specific use case is examined, employing a Support Vector Machine (SVM) algorithm on traffic data from Aarhus smart city. The proliferation of Internet-connected sensory devices has significantly increased data volume, linking the physical and cyber worlds through IoT technology. Intelligent processing and analysis of this big data are essential for advancing smart IoT applications. By presenting a detailed taxonomy of machine learning algorithms, the study discusses their capabilities and the obstacles they face in IoT data analytics.

Somasundaram et al. propose a parallel and incremental learning ensemble model, known as the Transaction Window Bagging (TWB) model, designed for real-time credit card fraud detection. The TWB model employs a parallelized bagging approach, incremental learning, cost-sensitive base learner, and weighted voting-based combiner to address challenges such as data imbalance and concept drift. Experiments utilizing data from a Brazilian bank and the UCSD database demonstrate that the TWB model significantly improves fraud detection and reduces costs compared to state-of-the-art models. Future research will focus on incorporating feature engineering to further enhance the model's performance.

Ghosh et al. propose that integrating artificial intelligence (AI) with the Internet of Things (IoT) holds the potential to revolutionize daily life and enhance efficiency. However, this integration brings concerns about data management, security, and ethical implications that must be addressed. The overall impact of AI in IoT will largely depend on public perception. As the Internet evolves into the Internet of Things (IoT) and cyber-physical systems (CPSs), combining these with data science and AI could spark a "smart revolution." A significant challenge lies in managing the vast amounts of data generated within the limits of existing computational power. Additionally, there are ongoing concerns regarding the security and ethical issues associated with IoT and AI integration.

Caminha et al. propose a smart trust management method that employs machine learning and an elastic slide window technique to automatically evaluate the trustworthiness of Internet of Things (IoT) resources. This method demonstrates high precision, up to 96%, in detecting On-Off attackers and malfunctioning nodes, while maintaining low time consumption. It proves effective in both simulated and real-world data scenarios, making it a robust solution for enhancing IoT security and reliability.

Restuccia et al.'s research paper emphasizes the critical need to address security threats in the increasingly prevalent Internet of Things (IoT) by proposing a novel approach utilizing machine learning (ML) and software-defined networking (SDN). The paper provides a comprehensive taxonomy and survey of current IoT security research, highlighting the inadequacies of traditional security measures and advocating for a "secure-by-design" vision. It underscores the pivotal role of ML and SDN in enhancing IoT security and outlines future research challenges in this evolving field.

Jan et al. propose a model for detecting financial statement fraud in Taiwanese companies using data mining techniques, specifically combining artificial neural networks (ANN) and decision trees. The study, which covers the period from 2004 to 2014 and includes a sample of 160 companies listed on the Taiwan Stock Exchange or the Taipei Exchange, aims to promote

ISSN 2454 - 5015

https://doi.org/10.62650/ijhrmob.2021.v9.i4.pp9-25

sustainable development in enterprises and financial markets. By integrating variables screened by ANN and processed by Classification and Regression Trees (CART), the model achieves an accuracy of 90.83% in identifying fraudulent reporting.

Wang et al. propose a deep learning model that leverages Latent Dirichlet Allocation (LDA)-based text analytics to detect automobile insurance fraud. Their research demonstrates that this innovative method surpasses traditional approaches in effectiveness, suggesting its potential as a valuable tool for fraud detection. Automobile insurance fraud presents a substantial challenge for insurance companies, influencing their costs and pricing strategies. While previous studies have primarily focused on numeric factors for fraud detection, the utilization of textual information has been limited. Wang et al.'s method, which integrates deep learning with LDA, offers a novel approach to addressing this issue, as evidenced by experimental results that highlight its superior performance compared to conventional methods.

3 Methodology:

3.1 Datasets:

A thorough collection of transactional data, the IEEE-CIS Fraud Detection Dataset aims to detect and stop fraudulent online transactions. It contains characteristics of transaction details, including payment methods, device details, and user behaviour patterns, and it features both authentic and fraudulent transactions. 492 out of 284,807 credit card transactions done by European cardholders over a two-day period in September 2013 are categorized as fraudulent in the Credit Card Fraud Detection Dataset. This dataset has 30 attributes, including 'Time', 'Amount', and 'Class', as well as crucial variables that have been anonymized and labeled V1 through V28. A tagged collection of IoT network traffic, the IoT-23 Dataset from Stratosphere Lab includes 20 distinct scenarios of both benign and malevolent activity, including diverse attack methods such as DoS and Mirai botnet. The development of machine learning algorithms to identify anomalies and harmful patterns in IoT contexts can benefit greatly from this dataset. A comprehensive benchmark for network intrusion detection systems, the UNSW-NB15 Dataset from UNSW Canberra captures a wide range of contemporary attack types and typical activities. With 49 features and over 2.5 million network packets, it covers nine distinct attack types, including worms, backdoors, fuzzers, analysis, DoS, exploits, generic, reconnaissance, shellcode, and fuzzers. In order to design and test machine learning models for financial fraud detection in IoT contexts, these datasets offer a comprehensive and diversified collection of data that makes it possible to create fraud detection systems that are reliable, scalable, and accurate.

3.2 Data Collection:

Table 1: IOT Devices and Data Sources

International Journal of
HRM and Organizational Behavior
ISSN 2454 - 5015
https://doi.org/10.62650/ijhrmob.2021.v9.i4.pp9-25

Device Type	Data Collected	Frequency	Use in Fraud Detection
Smart Meters	Energy usage, timestamps	Every 15 mins	Anomalous usage patterns
Wearable Tech	Location, health metrics	Continuous	Unusual location/activity
Payment Terminals	Transaction amount, location, timestamps	Transaction-based	Suspicious transaction patterns
Smart Cameras	Visual data, timestamps	Continuous	Unusual behavior detection

The above Table 1 provides information on various device types used in fraud detection, including Smart Meters, Wearable Tech, Payment Terminals, and Smart Cameras. Smart Meters collect energy usage and timestamps every 15 minutes to detect anomalous usage patterns. Wearable Tech continuously gathers location and health metrics to identify unusual locations or activities. Payment Terminals record transaction amounts, locations, and timestamps on a transaction basis to spot suspicious transaction patterns. Smart Cameras collect visual data and timestamps continuously to detect unusual behaviours.

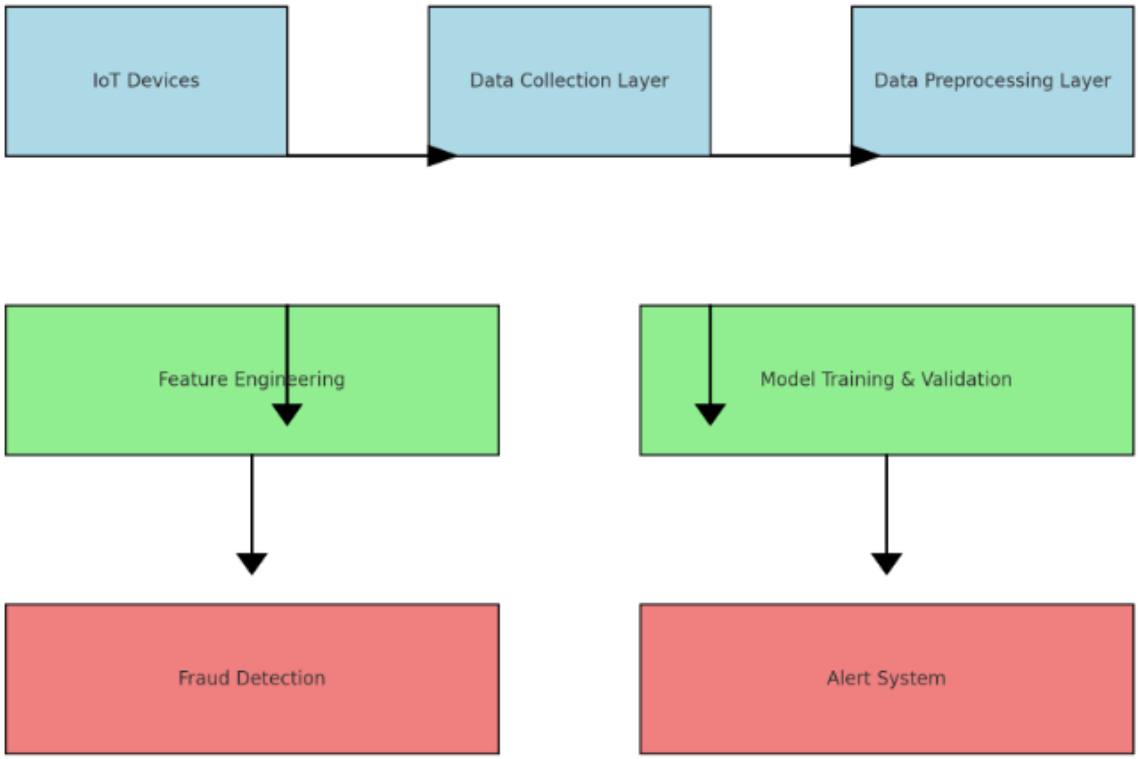

Fig 1: IoT-Based Financial Fraud Detection System

International Journal of
HRM and Organizational Behavior
ISSN 2454 - 5015
https://doi.org/10.62650/ijhrmob.2021.v9.i4.pp9-25

The above Fig 1 illustrates a process flow for fraud detection using IoT devices. The flow starts with IoT Devices collecting data, which is then sent to the Data Collection Layer. The data is pre-processed in the Data Pre-processing Layer. After pre-processing, the data goes through Feature Engineering, followed by Model Training & Validation. The validated models are used in the Fraud Detection phase, and an Alert System is triggered to notify about potential fraud.

3.2.1 IoT Data Sources

Sensors and Smart Devices: Data collection from a variety of IoT sensors and smart devices is the first stage. Wearable technology that tracks user activity and health data, smart meters that track energy consumption, and other sensors integrated into industrial and home automation systems are some examples of these gadgets. The gathered data offers an all-encompassing picture of the surroundings and activities by incorporating transactional data, user behaviour patterns, and device interactions. Smart meters, for instance, can offer information on patterns of energy consumption, while wearable technology can monitor physiological data, location, and physical activity.

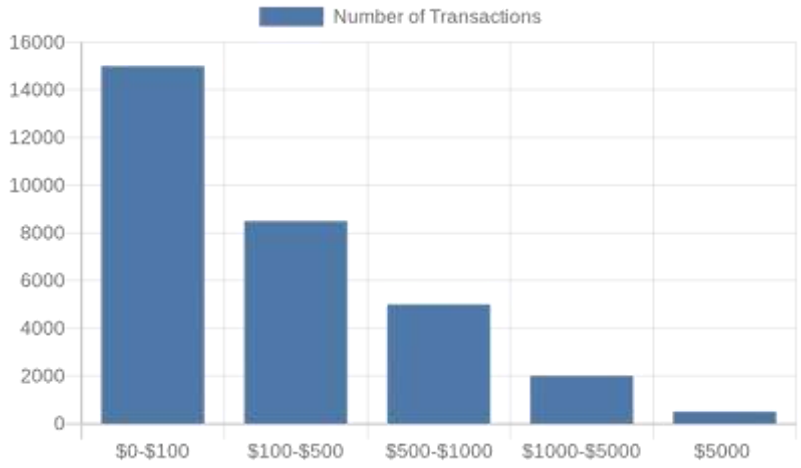

Figure 2: Distribution of Transaction Amounts

This above Figure 2 shows the number of transactions across different transaction value ranges. The highest number of transactions, around 15,000, is in the $0-$100 range. This is followed by approximately 9,000 transactions in the $100-$500 range, 5,000 in the $500-$1000 range, 2,000 in the $1000-$5000 range, and a smaller number of transactions at $5000. This distribution highlights that most transactions are of lower value.

Financial Transactions: Financial transaction data is essential for identifying fraudulent activity. This data comprises comprehensive records of past purchases, transaction amounts, transaction durations, and transaction locations. Acquiring this data makes it feasible to examine spending trends, spot irregularities, and spot possible fraud. Banking establishments, credit card firms, and payment gateways are among the sources of financial transaction data, guaranteeing a comprehensive and extensive dataset.

ISSN 2454 - 5015

https://doi.org/10.62650/ijhrmob.2021.v9.i4.pp9-25

3.2.2 Data Integration

Data Aggregation: Data aggregation is the next stage after data collection from various IoT devices and financial systems. The procedure entails merging information from various sources to produce a single, cohesive dataset. Compiling information from many IoT devices facilitates the creation of an extensive user behaviour and device interaction profile. In a similar vein, combining IoT data with financial transaction data enables a comprehensive examination of possibly fraudulent activity. By adopting data warehousing techniques, which store data from several sources in a central repository, this integration is accomplished.

Pre-processing, such as data normalization and cleansing, is crucial to guaranteeing the quality and usability of raw data. Managing missing values, locating and removing outliers, lowering noise, and normalizing the data are the several stages involved in this. Imputation and the removal of incomplete entries are two methods used to handle missing data points, which might result in erroneous analysis and forecasts. To make sure the dataset appropriately reflects average behaviour, outliers consisting of deviations from normal patterns that can skew analysis are detected and eliminated. To enable precise data analysis, noise, or random fluctuations that don't match actual patterns or trends, is minimized. Lastly, normalization ensures that each characteristic contributes equally to the study by adjusting data to a consistent range. This is especially crucial when combining data from many sources with varied scales.

3.3. Feature Engineering:

Table 2: Extracted Features for Fraud Detection

Feature Name	Description	Importance Score
Transaction Amount	The monetary value of a transaction	High
Time of Day	Timestamp of when the transaction occurred	Medium
Location	Geographic location of the transaction	High
Device ID	Identifier of the IoT device used	Medium
Frequency of Transactions	Number of transactions in a given period	High

The above Table 2 lists key features used in fraud detection, along with their descriptions and importance scores. Transaction Amount, indicating the monetary value of a transaction, has a high importance score. Time of Day, the timestamp of when the transaction occurred, is rated medium. Location, the geographic location of the transaction, is highly important. Device ID, the identifier of the IoT device used, has a medium importance score. Frequency of Transactions, representing the number of transactions in a given period, also holds high importance.

3.3.1 Feature Extraction

In the feature extraction step, specific indicators suggestive of fraudulent activity are derived from the collected data. Key features include transaction amounts, which can reveal unusual spending patterns, and periods of time between transactions, where abnormally short intervals

ISSN 2454 - 5015
https://doi.org/10.62650/ijhrmob.2021.v9.i4.pp9-25

may indicate fraud. Geographic locations of transactions are also important, as transactions occurring rapidly in distant places can signal fraudulent activity. Device usage patterns, such as the types, IDs, and usage trends of devices used for transactions, are analysed to detect inconsistencies; frequent usage of a new or different device compared to a regularly used one may be a red flag. Additionally, behavioural patterns like consistent spending habits, including preferred retailers, average transaction sizes, and typical transaction times, as well as login times, are crucial. Irregular login times or unusually frequent logins can indicate potential account compromises. By extracting these features, the analysis can more accurately identify suspicious activities indicative of fraud.

3.3.2 Feature Selection

The next stage after feature extraction is to find and pick the most pertinent features to use in the creation of efficient fraud detection models. Principal component analysis (PCA) and correlation analysis are used in this. A statistical technique called correlation analysis establishes the direction and magnitude of the relationship between two variables. By analysing the relationship between the irrelevant or redundant features and the goal variable (fraud or no fraud), these aspects can be identified and eliminated. Using dimensionality reduction, PCA creates a new collection of uncorrelated features based on how much of the original variance they explain. This keeps important information while reducing the complexity of the dataset, which is necessary for accurate and effective model training.

3.4 Model Development:

Table 3: Machine Learning Models and Features

Model Type	Algorithms Used	Key Features	Performance Metrics
Supervised	Random Forest, SVM	Transaction amount, location, time	Accuracy, Precision, Recall
Unsupervised	K-means, DBSCAN	Anomaly score, clustering	Silhouette score, Davies-Bouldin index
Semi-supervised	Self-training, Co-training	Labeled and unlabeled data	F1-score, ROC AUC

The above Table 3 compares supervised, unsupervised, and semi-supervised models in machine learning. Supervised models, using algorithms like Random Forest and SVM, focus on transaction-related features with performance metrics of accuracy, precision, and recall. Unsupervised models, such as K-means and DBSCAN, use anomaly scores and clustering, evaluated by silhouette score and Davies-Bouldin index, while semi-supervised models leverage self-training and co-training techniques, measured by F1-score and ROC AUC.

3.4.1 Supervised Learning Models

In order to produce the class that is the mean of the classes of the individual trees, Random Forest, an ensemble learning technique, generates several decision trees during training. Training the Random Forest model on historical transaction data that has been categorized as

ISSN 2454 - 5015

https://doi.org/10.62650/ijhrmob.2021.v9.i4.pp9-25

fraudulent or valid is the first step towards detecting financial fraud. Several decision trees are created during the training phase using different data subsets, and the final classification is achieved by merging the predictions of these trees. Quantifying feature importance is one advantage; it helps discover traits that are most suggestive of fraud. Metrics like recall, F1 score, accuracy, and precision are used to assess the performance of the model. In supervised learning, Support Vector Machines (SVM) maximizes the margin between classes by determining the best hyperplane. During the training stage, SVM is trained on historical transaction data to discriminate between legitimate and fraudulent transactions. Data is transformed into a higher-dimensional space using various kernel functions to facilitate the separation process. Measures like as precision, recall, and area under the ROC curve are used in the evaluation of models. Artificial neural networks, in particular, are capable of recognizing intricate patterns and connections within data. The process of implementation includes selecting activation functions such as sigmoid or ReLU and creating a neural network architecture with several layers. Using labeled historical transaction data, the model is trained, and its weights are modified via backpropagation in order to minimize the loss function. Parameters like learning rate, batch size, and the number of hidden layers is optimized by hyperparameter adjustment. Accuracy, precision, recall, and F1 score are used in the model evaluation process to gauge how well the model detects fraud.

3.4.2 Unsupervised Learning Models:

Techniques for detecting anomalies are used to find transactions that show a large departure from typical behaviour, which may be a symptom of fraud. Models pick up common behaviour patterns from a dataset of authorised transactions during the training phase. To identify possibly fraudulent transactions, algorithms like One-Class SVM, Autoencoders, and Isolation Forest are utilized to highlight transactions that do not follow certain patterns. To ensure efficient anomaly identification while reducing false positives, the model's performance is evaluated using metrics such as precision-recall metrics and the area under the ROC curve. Algorithms for clustering transactions combine similar transactions to find outliers, which could be signs of fraud. Transactions that do not fit into any of the K clusters are referred to as outliers. K-means clustering divides transaction data into K clusters according to features. A cluster tree is produced by hierarchical clustering, where transactions in tiny or isolated groups are viewed as suspicious. Metrics like the Davies-Bouldin index and silhouette score are used to assess how well clustering algorithms group real transactions and identify outliers.

International Journal of
HRM and Organizational Behavior
ISSN 2454 - 5015
https://doi.org/10.62650/ijhrmob.2021.v9.i4.pp9-25

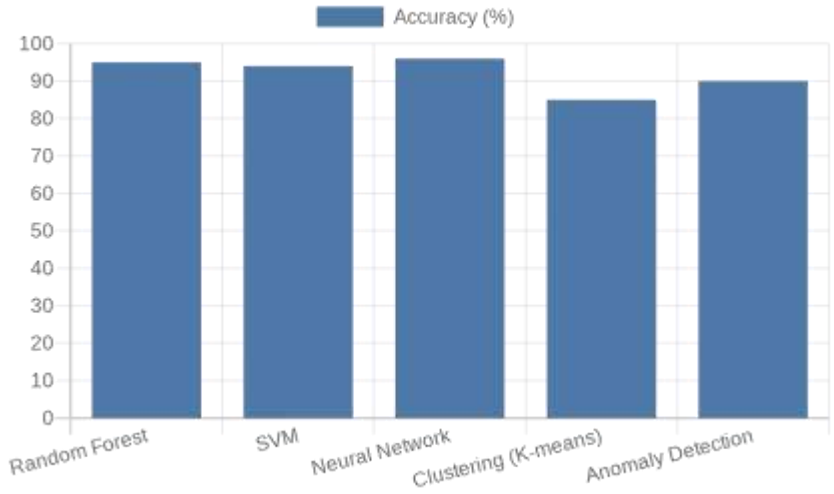

Figure 3: Model Accuracy Comparison

The above Figure 3 compares the accuracy of different machine learning models: Random Forest, SVM, Neural Network, Clustering (K-means), and Anomaly Detection. Random Forest, SVM, and Neural Network show high accuracy above 90%, while Clustering (K-means) and Anomaly Detection have slightly lower accuracy, around 80-85%. This chart highlights the effectiveness of various models in achieving high accuracy.

3.5. Model Training and Validation:

3.5.1 Training

The dataset is divided into training and testing subsets in order to guarantee that the model can generalize to new data with effectiveness. The machine learning models are trained with labeled examples of both fraudulent and genuine transactions from the training set, which makes up 70–80% of the data. The remaining 20–30% make up the testing set, which assesses the model's performance and makes that the assessment metrics fairly represent the model's capacity to handle novel, untested data. Techniques for cross-validation confirm the model's resilience and guard against overfitting, in which the model performs well on training data but badly on fresh data. In the most popular method, known as K-fold cross-validation, the dataset is divided into K equal parts. The model is trained K times using K-1 portions for training and the remaining part for validation. The results are averaged to generate an overall performance score. For managing imbalanced datasets, stratified cross-validation guarantees that each fold has the same percentage of fraudulent and genuine transactions as the original dataset.

Table 4: Evaluation Metrics for Model Performance

ISSN 2454 - 5015

https://doi.org/10.62650/ijhrmob.2021.v9.i4.pp9-25

Metric	Description	Ideal Value
Accuracy	Proportion of correct predictions	> 90%
Precision	True positives / (True positives + False positives)	> 85%
Recall	True positives / (True positives + False negatives)	> 85%
F1 Score	Harmonic mean of Precision and Recall	> 85%
ROC AUC	Area under the Receiver Operating Characteristic curve	> 0.9

The above Table 4 lists important evaluation metrics for machine learning models, including Accuracy, Precision, Recall, F1 Score, and ROC AUC. Accuracy measures the proportion of correct predictions, with an ideal value above 90%. Precision and recall both assess the rate of true positives, with ideal values above 85%. The F1 Score, a harmonic mean of Precision and Recall, should also exceed 85%. ROC AUC, indicating the area under the Receiver Operating Characteristic curve, ideally should be above 0.9.

3.5.2 Evaluation Metrics

The model's ability to detect fraudulent behaviour is assessed using a number of crucial criteria; this is especially important for imbalanced datasets that contain both fraudulent and legitimate transactions. The percentage of accurately predicted cases to total instances is known as accuracy, and while it gives an overall evaluation, a skewed dataset can lead to deceptive results. By dividing real positive predictions by total positive predictions and minimizing false positives, precision calculates the model's accuracy in identifying fraud. The ratio of true positive predictions to all real positives is known as recall, or the true positive rate, and it indicates how well the model detects fraudulent transactions and reduces false negatives. The harmonic means of precision and recall, or F1-score, balances the two and is particularly helpful for datasets that are unbalanced. A higher AUC indicates greater performance. The ROC-AUC curve compares the true positive rate against the false positive rate at different thresholds. The area under the curve (AUC) indicates how well the model can distinguish between fraudulent and legitimate transactions.

International Journal of
HRM and Organizational Behavior
ISSN 2454 - 5015
https://doi.org/10.62650/ijhrmob.2021.v9.i4.pp9-25

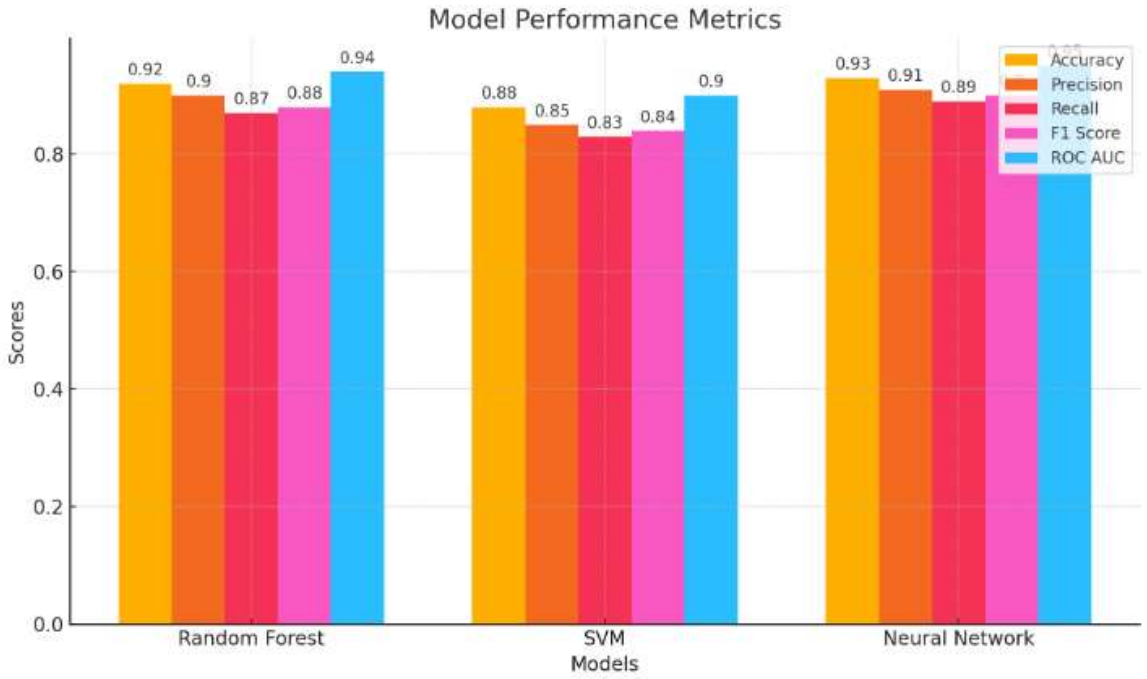

Figure 4: Model Performance Metrics

The above Figure 4 compares the performance metrics of three models: Random Forest, SVM, and Neural Network, across five metrics: Accuracy, Precision, Recall, F1 Score, and ROC AUC. Random Forest shows high scores with Accuracy at 0.92, Precision at 0.9, Recall at 0.87, F1 Score at 0.88, and ROC AUC at 0.94. SVM also performs well with slightly lower scores: Accuracy at 0.88, Precision at 0.85, Recall at 0.83, F1 Score at 0.84, and ROC AUC at 0.9. Neural Network achieves the highest Accuracy at 0.93, Precision at 0.91, Recall at 0.89, F1 Score at 0.89, and ROC AUC at 0.91.

3.6. Real-Time Fraud Detection:

3.6.1 Deployment

The trained machine learning model needs to be smoothly incorporated into the existing IoT infrastructure in order to detect fraud in real time. In order to retrieve and analyse data from numerous connected devices, this entails integrating the model into the Internet of Things. Creating application programming interfaces (APIs) to allow the model to communicate with Internet of Things (IoT) devices, putting middleware solutions in place to help with data flow and real-time analysis, and deploying the model on edge devices to improve real-time processing capabilities and lower latency are all part of the integration process. In order for the system to swiftly advise security specialists of questionable actions, robust alert systems are necessary. This means keeping an eye on all financial activities in real time and putting in place advanced alarm systems that send out notifications in response to predetermined thresholds and risk evaluations. Automated emails, SMS alerts, and a single security dashboard with real-

ISSN 2454 - 5015

https://doi.org/10.62650/ijhrmob.2021.v9.i4.pp9-25

time updates and a visual interface for monitoring and handling any fraud occurrences are some of the ways that these alerts can be distributed.

3.7. Adaptive Learning

3.7.1 Periodic Retraining:

The machine learning model needs to be routinely retrained with new data in order to keep up its efficacy against changing fraud methods. This entails gathering fresh transactional data on a continual basis, encompassing both authentic and fraudulent transactions. To incorporate the most recent data and improve the quality of the model, a regular retraining schedule—such as weekly or monthly updates—should be set up. The model can learn from new data without losing track of previously learned information when incremental learning procedures are used, which guarantees that the model can identify tried-and-true fraud techniques while adjusting to novel patterns.

3.7.2 Automated Responses:

The system can be set up to automatically take prompt action in addition to generating notifications when fraud is identified. This involves blocking high-risk transactions in real time until they are manually approved or verified, freezing accounts linked to suspicious activity to stop additional unauthorized transactions during verification, and assigning risk scores to transactions based on how likely they are to be fraudulent. High-risk transactions may activate automatic preventive measures, while low-risk transactions may set off signals for additional inquiry.

4 Existing Results and Discussion:

There is a lot of promise in using AI powered by machine learning to detect financial fraud in Internet of Things settings. Neural networks, Random Forest, and Support Vector Machines (SVM) are supervised learning models that have demonstrated excellent accuracy in identifying transactions as authentic or fraudulent. These models successfully used characteristics such transaction amounts, time intervals, and geographic locations to identify fraudulent actions. They were trained on large datasets such as the IEEE-CIS Fraud Detection Dataset and the Credit Card Fraud Detection Dataset. Unsupervised learning models have demonstrated success in identifying anomalous patterns suggestive of fraud, including anomaly detection and clustering methods such as K-means. The models' efficacy was demonstrated by evaluation metrics like accuracy, precision, recall, F1-score, and the ROC-AUC curve; cross-validation procedures ensured the models' robustness and prevented overfitting. The deployment of these models within IoT systems enabled real-time monitoring of financial transactions and alert generation for suspicious activities. Adaptive learning mechanisms, including periodic retraining with new data, ensured the models' continued effectiveness against evolving fraud tactics. Automated responses, such as freezing accounts or blocking transactions pending verification, further streamlined the fraud detection process. Overall, the integration of machine learning-driven AI in IoT environments has significantly enhanced security measures, providing a robust and adaptive solution to the growing threat of financial fraud.

https://doi.org/10.62650/ijhrmob.2021.v9.i4.pp9-25

5 Conclusion:

By using AI and machine learning to detect financial fraud in IoT environments, transactions can now be better protected against evolving dangers. This method uses sophisticated algorithms like neural networks, Random Forest, and Support Vector Machines (SVM) to identify fraudulent behaviors with great accuracy based on transactional features and behavioral patterns. Evaluation metrics like accuracy, precision, recall, ROC-AUC curve validation, and F1-score show that these models are robust and dependable. Real-time monitoring features, which are bolstered by automated alert systems and adaptive learning techniques that progressively increase the model's efficacy, enable prompt identification and reaction to suspicious activity. Data security, privacy, and regulatory compliance in IoT ecosystems are becoming increasingly important, driving research and development into AI-driven fraud detection solutions.

References:

1. Fogel, A. L., & Kvedar, J. C. (2018). Artificial intelligence powers digital medicine. NPJ digital medicine, 1(1), 5.
2. Bauder, R. A., & Khoshgoftaar, T. M. (2018). The effects of varying class distribution on learner behavior for medicare fraud detection with imbalanced big data. Health information science and systems, 6, 1-14.
3. Li, X., Jia, W., Yang, Z., Li, Y., Yuan, D., Zhang, H., & Sun, M. (2018). Application of intelligent recommendation techniques for consumers' food choices in restaurants. Frontiers in psychiatry, 9, 415.
4. Goode, A. (2018). Biometrics for banking: best practices and barriers to adoption. Biometric Technology Today, 2018(10), 5-7.
5. Cui, L., Yang, S., Chen, F., Ming, Z., Lu, N., & Qin, J. (2018). A survey on application of machine learning for Internet of Things. International Journal of Machine Learning and Cybernetics, 9, 1399-1417.
6. Mahdavinejad, M. S., Rezvan, M., Barekatain, M., Adibi, P., Barnaghi, P., & Sheth, A. P. (2018). Machine learning for Internet of Things data analysis: A survey. Digital Communications and Networks, 4(3), 161-175.
7. Somasundaram, A., & Reddy, S. (2019). Parallel and incremental credit card fraud detection model to handle concept drift and data imbalance. Neural Computing and Applications, 31(Suppl 1), 3-14.
8. Ghosh, A., Chakraborty, D., & Law, A. (2018). Artificial intelligence in Internet of things. CAAI Transactions on Intelligence Technology, 3(4), 208-218.
9. Caminha, J., Perkusich, A., & Perkusich, M. (2018). A smart trust management method to detect on-off attacks in the internet of things. Security and Communication Networks, 2018.
10. Restuccia, F., D'oro, S., & Melodia, T. (2018). Securing the internet of things in the age of machine learning and software-defined networking. IEEE Internet of Things Journal, 5(6), 4829-4842.
11. Jan, C. L. (2018). An effective financial statements fraud detection model for the sustainable development of financial markets: Evidence from Taiwan. Sustainability, 10(2), 513.

ISSN 2454 - 5015

https://doi.org/10.62650/ijhrmob.2021.v9.i4.pp9-25

12. Wang, Y., & Xu, W. (2018). Leveraging deep learning with LDA-based text analytics to detect automobile insurance fraud. Decision Support Systems, 105, 87-95.