Credit Card Transactions Fraud Detection for Multiple Consumer Behaviors

Baker Al Smadi
Computer Science and Digital Technologies
Grambling State Univeresity
Grambling, Louisiana
bakir_smadi@hotmail.com

William B. Glisson
Computer Science Department
Louisiana Tech University
Ruston, Louisiana
glisson@latech.edu

Majd Tahat
Cyberspace Engineering
Louisiana tech University
Ruston, Louisiana
Latech.tht@gmail.com

Hosam Alamleh
Computer Science
University of North Carolina Wilmington
Wilmington, North Carolina
hosam.amleh@gmail.com

Ali Abdullah S. AlQahtani
Computer Systems Technology
North Carolina A&T State University
Greensboro, North Carolina
alqahtani.aasa@gmail.com

***Abstract*—In today's digital era, credit card fraud is a prevalent issue that costs financial institutions and individuals billions of dollars. To prevent such fraud, fraud detection systems are implemented that use machine learning algorithms to analyze patterns and detect transaction anomalies. However, these systems heavily rely on historical data, and if the data is limited or biased, the system's accuracy decreases significantly. This study addresses this issue by investigating a real credit card transaction dataset and determining different consumer behaviors. Synthetic datasets are generated based on consumer behaviors to enhance the accuracy of detecting fraudulent activities. According to the findings, the Logistic Regression (LR) model exhibited superior performance in both experiments. It achieved an impressive accuracy of 96.4% with remarkable time efficiency in the first experiment, and 94.5% accuracy in the second experiment, while still maintaining excellent time efficiency. This research goes through the procedures involved in analyzing a real dataset, understanding consumer behaviors, and generating synthetic datasets.**

***Index Terms*—Fraud, Credit Card, Security System, Fraud Detection, Fraud Prevention.**

I. Introduction

Credit card fraud is a serious issue in the United States, affecting both consumers and businesses [1] A study has revealed that credit card fraud is on the rise, with an increasing number of people feeling violated and unsafe.The business sector lost a significant amount of money amounting to 28.65 billion in 2018 due to fraud through data breaches and method of contact [2]. Most cases of identity theft, and credit card fraud occurred either through phone calls or websites [3]. This has led to a growing need for a suitable solution to eliminate these cybercrimes.

The severity of credit card fraud is apparent from the most devastating data breaches ever recorded in history resulting from credit card fraud that took place between 2005 and 2014 [4]. These breaches impacted a large number of accounts, emphasizing the importance of improving credit card security systems. This is essential to establish a safe environment and provide enhanced security for credit card users. Notably, there has been a surge in identity theft cases, with credit card fraud emerging as a prevalent form of identity theft since January 2020. [5]

Identity theft crimes is a severe issue since credit card fraud accounts for 29% of all identity theft reported in 2018 [6]. Based on a 2018 study, the United States has witnessed a surge in reports of identity theft fraud. Instances of credit card fraud have doubled within the last year and are projected to continue growing [7]. These developments have inflicted significant harm on financial systems. Unauthorized financial transactions can result in bankruptcy, making it extremely challenging for businesses to receive timely payments.

Types of fraudulent activities that can be identified include credit card fraud, computer intrusion, telecommunication fraud, theft/counterfeit fraud, bankruptcy fraud, and application fraud. [8]. Offline and online credit card fraud are different. When the card is present and physically stolen, it is considered offline fraud. Otherwise, when the credit card information is used for unauthorized purchases made through the internet, it is considered online fraud. A security system must be developed to prevent such fraud, especially online fraud [9].

Credit card fraud detection can be challenging because users spend differently, which makes it difficult to implement a one-size-fits-all fraud detection model [10]. As consumer behavior varies widely [11], it can be challenging to establish fraud patterns, making it difficult to distinguish legitimate transactions from fraudulent ones. Therefore, it is necessary to adopt a more personalized approach to fraud detection, utilizing advanced technology and machine learning algorithms to detect unusual patterns and behavior specific to each user.

This paper aims to address the issue of fraud in multiple consumer behaviors. The proposed solution involves devel-

oping a system that employs machine learning techniques to detect fraudulent activities. To achieve this, the system is trained on various spending behaviors, and each user is assigned a behavior based on their spending habits. Different models for fraud detection are then employed for each type of behavior. This allows for a personalized approach to fraud detection, which is more effective in identifying suspicious activity specific to each user. Employing different models for fraud detection for each type of behavior allows for more accurate detection and reduces the number of false positives, which can be costly and time-consuming to investigate.

II. Related Work

Numerous studies have been conducted worldwide with the objective of mitigating credit card fraud. These studies have explored various approaches, such as employing fraud detection algorithms and monitoring consumer behavior patterns to identify potentially fraudulent credit card transactions. [12]–[14]. In a study conducted by Saxena and Ponnapalli [15], a novel system was developed to generate offline, serverless, one-time credit card numbers for users. This innovative approach eliminates the need for online communication with the server. The system utilizes a shared key to generate unique credit card numbers and employs a private key for each customer to sign and record transaction information, ensuring non-repudiation of online transactions. By leveraging the existing credit card numbering structure, the authors enabled online transactions using traditional infrastructure.

Meredith et al. [16] undertook a project focused on credit card fraud detection, leveraging mobile device location tracking. Their system incorporated a processor capable of determining a fraud percentage by monitoring the location of the user's device, which is linked to the credit card account. Specifically, it considers the initial registered area location where the user establishes the credit card account.. Rajasekaran and Varadarajan [17] built a model to decrease the potential credit card fraud by generating a one-time credit card number at the user's machine and sending it to the credit card issuer bank/company and merchant. The card issuer then applies authentication operations, such as a one-way password, or a string of letters, to be able to verify the user's identity. If the numbers match, the user is authenticated.

Other studies proposed different models in an effort to minimize credit card fraud [18], [19]. For example, Essebag et al [20] have developed a comprehensive system that provides a dynamic security code that can change the security code CVV (Card Verification Value) of a prepaid, debit, or credit card, providing dynamic security code values that have limited use to one online transactions only.

In their proposal, McDonald [21] introduced a system that enables secure online purchases without the need for physical cards. This system utilizes a Personal Digital Identity Token (PDIT) as a biometric identifier for the cardholder, establishing a connection to a reliable set of civil identity credentials. On a similar note, Barbour [22] developed a system that facilitates financial transactions over a communication medium. Users who possess an account linked to a consistent account number can utilize this system. A unique, one-time-use number, derived from the account holder's information, is generated and authorized for transfer as part of the transaction process.

The majority of the research on credit card fraud detection primarily focuses on using anomaly detection to analyze users' behavior [23] However, this is often done in two ways - either by training one model for all users [24], [25] or by training a model for each user [26], [27]. Both these approaches have their limitations. The first approach is less accurate as it fails to fit each category of behavior, while the second requires more computing resources and may perform poorly when working with small data sets.

III. Datasets

This section explains the multiple datasets that have been utilized in this paper.

A. Real Dataset

The initial set of information is highly popular among scholars globally. The dataset includes 284,807 credit card transactions carried out online by European credit card holders over a period of two days in September 2013 [28]. The data offered in this set has been subject to Principal Component Analysis (PCA) transformation due to reasons of privacy and confidentiality.

B. Artificially Generated Datasets

This study generated multiple synthetic datasets that reflect distinct patterns of user behavior based on their spending habits and online payment transactions across a range of stores or websites. The datasets incorporate novelty characteristics to better capture these patterns.

Researchers have created these datasets with the intention of simulating a variety of hypothetical scenarios. Consequently, they have produced six distinct datasets, each categorized under one of six different categories. Each category of the six categories' specification are listed as follows:

1) The first category of datasets portrays a typical spending pattern, i.e., that of an average individual. This category includes three different users, each with a total of 1000 transactions. These users have only a few IMEI/IP locations and addresses (longitudes and latitudes) and make purchases from a limited number of online stores. On average, they spend $150 per transaction and make purchases in the period from 8 am to 11:59 pm.
2) The second category of datasets pertains to multiple locations, that is, individuals who make online purchases from multiple websites. This category includes three different users, each with a total of 1000 transactions. These users have several IMEI/IP addresses associated with varying locations, and they purchase from a limited number of online stores. On average, they spend $500 per transaction and shop at various times throughout the day, without any discernible pattern.

3) The third category of datasets represents high spending behavior, that is, individuals who spend a significant amount of money on online purchases. This category includes three different users, each with a total of 1000 transactions. These users have a limited number of IMEI/IP locations and addresses, and they shop from multiple e-markets. On average, they spend $3000 per transaction and shop at various times throughout the day, following a specific pattern.
4) The fourth category of datasets relates to extensive usage of locations, that is, individuals who make online purchases from over 20 different locations. This category includes three different users, each with a total of 1000 transactions. These users have a large number of various locations and IMEI/IP addresses, and they shop from a moderate number of different online stores. On average, they spend $500 per transaction and make purchases between 8 am and 11:59 pm.
5) The fifth category of datasets pertains to individuals who purchase items from multiple online stores, that is, different stores buyers. This category includes three different users, each with a total of 1000 transactions. These users have a moderate number of IMEI/IP addresses and locations, and they shop from a wide variety of online stores. On average, they spend $1000 per transaction and make purchases between 8 am and 11:59 pm.
6) The final category of datasets combines all five aforementioned categories. It includes three different users, total of 1000 transactions each. These users shop from various locations using multiple IMEI/IP addresses and purchase from a diverse range of online stores. On average, they spend $3000 per transaction and make purchases at various times throughout the day, without any particular pattern.

The above six categories were designed based on different variables that reflect various online payment behaviors. The six synthetic datasets were created to mimic real human spending behavior while maintaining the most likely situations for each user that are included in the synthetic datasets can be predicted. The six categories represent different levels of difficulty in guessing the user's pattern. The first two categories represent the simplest user behavior, while the third and fourth categories are more complex, with an extended range of variables. These cases are expected to be more challenging for ML algorithms to specify a clear pattern or behavior, leading to a decrease in accuracy and precision. The last two categories, 5 and 6, are very sophisticated, and ML algorithms may encounter significant difficulties when trying to determine the spending behavior of individual users. Hence, the weakest performance is anticipated in these categories.

In 2020, Tugba Sabanoglu conducted a study on credit card transactions [29]. The study revealed that most customers engage in online transactions, with 31% of respondents making at least one online transaction per month. Additionally, 24% of cardholders make online purchases twice every 14 days, while 20% of cardholders use their credit cards for online transactions once per week. These percentages are relatively similar to each other. Therefore, to represent the actual data in the real world, the study considered two transactions per week.

The original synthetic dataset consists of nine variables or columns: UserAccountNumber, UserName (email address), IP address, TransactionTime, TransactionAmount, TransactionStore, Latitude, Longitude, and Status (a variable indicating fraudulence, with a value of zero for non-fraudulent transactions and one for fraudulent transactions). While the data may be unmbalanced, it is free from any inappropriate or missing values. To ensure compatibility with machine learning algorithms, the generated datasets need to go through adjustments and normalization to achieve balance.

Before integrating the machine learning algorithms, the synthetic datasets undergo the following processing steps:

1) Transform the transaction's daytime into a numerical representation that corresponds to the hour of the day when the transaction occurred
2) Transform the date of each transaction into the number of days that have passed between consecutive transactions.
3) To make the Transaction IP address numerical and readable by machine learning algorithms, you can split it into four groups based on its components. Each group can represent one part of the IP address, such as the octets in the IPv4 format. This will enable the algorithms to process the IP address effectively.
4) The Username column, which contains nominal data, should be removed from the dataset as machine learning algorithms typically cannot handle nominal variables
5) Normalizing all columns to fit the ML models

After completing the necessary operations, the generated datasets now contain eleven columns and are prepared for analysis. While there were no fraudulent transactions in the initial datasets, researchers added fraudulent transactions to represent 1% of the all transactions. Additionally, they analyzed the spending patterns of legitimate and fraudulent transactions and considered the online purchasing pattern of users, taking into account the possibility of IP addresses and locations being spoofed. To generate the fraudulent transactions, also they used 10% of the user's genuine location and IP address. The fraudulent transactions were designed similarly to legitimate transactions, as fraudsters often use genuine accounts to commit fraud but with different consumption patterns based on various factors such as the amount, time, and location. The study uses the most efficient ML algorithms to find fraudulent transactions in their datasets. which will lead to generating unbiased synthetic datasets that mimic the distribution of real-world data and represent various consumer spending behaviors. The synthetic datasets comprise nine variables:

1) The User Account Number: A distinct identifier assigned to each user in the dataset, allowing for numerical

identification. Herein, the distribution of values for this variable does not indicate any bias since it represents a unique identification number. Moreover, the dataset comprises three distinct users, with each user having 1000 recorded transactions.

2) Username: Researchers utilized email addresses as usernames for each user in the dataset. However, this variable is a string value. Thus, it's not incorporated into the integration of the machine learning algorithm.
3) Time gap: To calculate the time difference between consecutive transactions, the system examines the number of days. The researchers took care to generate values that follow a normal distribution for this feature, ensuring that there is no bias present.
4) Transaction amount: The distribution of this variable is heavily skewed to the left due to the prevalence of small transaction amounts and a few instances of larger amounts. Based on the specifications of Case six dataset, which signifies high consumption behavior, the average transaction amount in Case 6 is $3000, with a maximum value of $11,900 and a minimum value of $13.5. The standard deviation for this variable is $2,390, indicating the extent of variability in transaction amounts. It is designed to closely resemble the data distribution of transaction amounts in the real dataset. In Case 6, all users (user #0, user #1, user #2) have the same mean and are evenly distributed.
5) Transaction Store: The variable representing the merchant's website for each transaction is a normalized number. This normalization operation was implemented to avoid any bias in the data generation. Note that this feature exhibits the same data distribution for every user in the dataset, with a mean value of 10 and a normal distribution. Hence, there is no bias in the generation of this feature.
6) Transaction IP: This variable shows the IP address for each transaction is treated as a string variable. Therefore, the data distribution is not considered significant. As an address, it doesn't carry any inherent bias.
7) Location (two variables "latitude, longitude"): This variable is used to determine the precise location of each transaction. Real values were utilized by the researchers to populate this variable. The generation process involved assigning distinct locations to different users within the same dataset.
8) Time/h: This variable denotes the hour of the day in which each transaction occurred. The value of this variable was generated according to the specifications of each artificial dataset, as suggested by the researchers. Different time ranges were considered in different scenarios to appropriately reflect the desired conditions.
9) Status: The fraudulent determination variable distinguishes between fraudulent and non-fraudulent transactions. While, a value of "Zero" indicates a non-fraudulent transaction, while a value of "One" indicates a fraudulent transaction. Note that the statistical analysis incorporates data from both fraudulent and non-fraudulent transactions.

IV. Research methodology

Due to the growing interest in consumer behavior and the widespread use of credit card fraud detection, a personalized approach to fraud detection is crucial. By utilizing advanced technology and machine learning algorithms to identify unusual patterns and behavior unique to each user. Consequently, this research focuses on creating a system that uses machine learning techniques to detect fraudulent activities.

This paper proposes a novel approach to address this issue. It trains machine learning models for different categories of user behavior. By categorizing users based on their credit card usage habits, the model assigns each user to a specific category. This results in more personalized and precise fraud detection, with reduced classification errors and without requiring massive computing resources. Overall, this new approach aims to improve the accuracy and efficiency of credit card fraud detection while minimizing resource requirements.

The study is divided into three stages, outlined below: the selected machine learning algorithms are applied to the first dataset (the real dataset) to obtain the outcomes of each algorithm. Those results will be used to calibrate the algorithm in the next stage. Secondly,the most efficient machine learning algorithms are executed on the artificially generated datasets to evaluate their performance in each category. Finally, the outcomes of each algorithm in all experiments will ultimately determine the most suitable and best-fitting machine learning algorithm to be used with the generated datasets.

V. Experiments and Results

In this section, we will explore the selection and rationale behind the machine learning algorithms employed in this study. Specifically, five ML algorithms have been chosen to effectively detect fraudulent transactions within the real dataset, which are Decision Tree (DT), Support Vector Machine (SVM), Random Forest (RF), item Logistic Regression (LR), and Naïve Bayes (NB).

In this experiment, investigators opted for machine learning algorithms because these algorithms incorporate a binary classifier, which is essential for the study. Also, the investigators are interested in 2 possible outcomes: a fraudulent transaction or a non-fraudulent transaction. Therefore, investigators have two distinct classes that they need to differentiate from each other.

A. ML Experiment on the Real Dataset

This section evaluates the performance of 5 machine learning algorithms using a real dataset to distinguish between fraudulent and non-fraudulent credit card transactions. The research presents a comparison model that helps us analyze how each machine learning algorithm performed on the dataset. One effective way to compare the models is to use the Receiver Operating Characteristic (ROC) curve, which measures

the performance of classification problems. The ROC curve provides a visual representation of each model's ability to predict True-Positive Cases (TPC) and False-Positive Cases (FPC), enabling us to compare their results more effectively. See Figure 1.

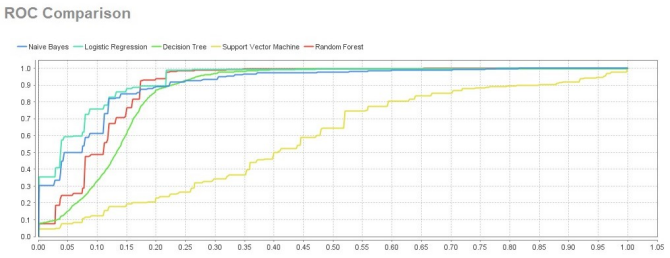

Fig. 1: ROC comparison [30]

The ROC curve plot is a summary of the confusion matrices of all the models used. The TPR is plotted on the Y-axis, and the FPR is plotted on the X-axis. The closer a point is to the top-left corner of the graph, the better the prediction ratio of that model. In Figure 1, the ROC curve shows that the LR model outperformed all other tested models. The light blue line that represents the LR model in the graph is closest to the top-left corner, indicating that it had fewer false-positive cases and more true positives. In simpler terms, it can be said that the Y-axis indicates the sensitivity of the model, while the X-axis shows (1-specificity). Observation revealed that the models had similar sensitivity ratios, with the exception of the NB model, which had a sensitivity ratio of 91%. Among all the models tested, the LR and NB models exhibited the best specificity ratios, with specificity rates of 79% and 80%, respectively. Table I shows a review of all measures for all machine learning models.

TABLE I: Performance Comparison- Real Dataset [30]

Model/Measure	Accuracy	AUC	Precision	Recall	F-Measure
NB	89.6%	0.908	96.7%	91.9%	93.8%
LR	96.4%	0.935	96.8%	99.1%	97.9%
DT	93.9%	0.867	93.9%	99.4%	96.5%
RF	94.6%	0.895	99.5%	99.5%	96.9%
SVM	86.2%	0.582	99.9%	100%	92.6%

According to Table I, The LR model exhibited the best overall performance, although some models, such as RF and DT, performed similarly in certain metrics. The LR model achieved the highest accuracy 96.4% and F-Measure 97.9% scores, and its AUC ratio 0.935 was the best among all the machine learning models. On the other hand, the NB and SVM models had the poorest performance among all the models tested and will not be considered in future experiments. However, there is another crucial criterion that the table did not cover, which is time efficiency. Table II compares the time efficiency among all ML models.

The LR and NB models demonstrated superior time efficiency in this research. As a result, the LR model was deemed the best model in terms of both performance and

TABLE II: Time Efficiency- Real Dataset

Model/Time Measure	Training Time (1000 Rows)	Scoring Time (1000 Rows)	Total Time
NB	10 ms	198 ms	17 s
LR	23 ms	219 ms	17 s
DT	50 ms	172 ms	18 s
RF	282 ms	1 s	2 min 25 s
SVM	4 s	12 s	27 min 1 s

time efficiency. Table II indicated that the SVM and RF models performed well, but their execution times were considerably longer. Thus, the outcomes must be balanced between both time and performance to determine the most appropriate model. Researchers executed these machine learning algorithms on a standard machine with a Core i5 1.6 GHz processor, 8 GB RAM, and Windows 10 operating system using RapidMiner Studio. It's important to note that using machines with better system specifications may result in different processing times.

B. ML Experiment on the synthetic datasets

This section shows the results of the top-performing machine learning algorithms (DT, LR, and RF) on the actual dataset used in the initial experiment. Table III illustrates the performance of the selected ML algorithms on the synthetic dataset used in the second experiment.

TABLE III: Average Performance- Synthetic Dataset

Model/Measure	Accuracy	AUC	Precision	Recall	F-Measure
LR	94.5%	0.875	94.10%	97%	95.50%
DT	87.2%	0.5605	93.6%	88.70%	87.00%
RF	90.3%	0.82	95.35%	51.80%	85.20%

Table IV summarizes all ML algorithms' time efficiency in each case of the artificial dataset. In the table, researchers observe that the machine learning algorithms behave differently with different inputs and variables.

TABLE IV: Time efficiency of the six cases

Model/ Time Measure	Training Time (1000 Rows)	Scoring Time (1000 Rows)	Total Time
LR	177.5 ms	112.5 ms	2 s
DT	170.6 ms	160 ms	2 s
RF	129.5 ms	601 ms	14 s

The synthetic experiment involved using three ML algorithms with a binary classifier to distinguish fraudulent transactions from non-fraudulent ones. As expected, different performances were observed between the real and synthetic datasets due to their varying sizes. However, the LR algorithm still demonstrated an average accuracy of 94.5% and precision ratio of 96%, compared to 96.4% in the real dataset experiment. Consequently, it can be concluded that the LR algorithm remains the best choice for the system as it has the best performance, accuracy, precision, F-measure, and time efficiency among all models.

However, due to the variation of inputs in each case in the synthetic dataset, the machine learning models took more time

to process data in complex cases. The RF model's complexity resulted in constructing numerous DTs based on the input variables, making it the least time-efficient model among others.

C. Real dataset Experiment Vs. the Synthetic dataset Experiment

In this section, the results of each experiment will be examined, and a comprehensive comparison will be made to evaluate the performance and efficiency across all experiments.

1) Performance Comparison : This section compares the results of the real and synthetic datasets in terms of different metrics to evaluate the performance of each machine learning algorithm with varying dataset sizes and variables. The ML algorithm with the highest accuracy, precision, F-measure, and AUC percentages in the shortest time is considered the best performer. Table V shows that in the first experiment, the five ML algorithms achieved an average F-measure of 89.23% and precision of 94.35%. The SVM and NB models showed the worst performance and were not used in the synthetic dataset experiment. The SVM algorithm was excluded from the experiment because it takes too much time to detect fraudulent transactions. The LR model performed the best in the first experiment, with 96.4% accuracy and 96.8% precision, and an F-measure of 97.9% based on all performance measures.

TABLE V: Performance Comparison- Synthetic vs. Real Datasets

(a) Real Dataset

Model/ Measure	Accuracy	AUC	Precision	Recall	F-Measure
NB	89.6%	0.908	96.7%	91.9%	93.8%
LR	96.4%	0.935	96.8%	99.1%	97.9%
DT	93.9%	0.867	93.9%	99.4%	96.5%
RF	94.6%	0.895	99.5%	99.5%	96.9%
SVM	86.2%	0.582	99.9%	100%	92.6%

(b) Synthetic Dataset

Model/ Measure	Accuracy	AUC	Precision	Recall	F-Measure
LR	94.5%	0.875	94.10%	97%	95.50%
DT	87.2%	0.5605	93.6%	88.70%	87.00%
RF	90.3%	0.82	95.35%	51.80%	85.20%

The second experiment investigates six distinct synthetic datasets containing different variables, each containing 3000 rows that correspond to three users. the proposed unsupervised machine learning model learns from past transactions and trains itself accordingly. In contrast, the dataset used in the first experiment (the real dataset) has 284,804 rows. Thus, achieving 94.5% accuracy and 96% precision using the LR algorithm in the second experiment is a significant accomplishment, as the model has limited data to rely on.

2) Time Efficiency Comparison: This study highlights a disparity in the number of variables employed between the two experiments. The real dataset encompasses 21 variables post data cleaning, whereas the synthetic datasets comprise only 11 variables. Consequently, it is anticipated that the ML algorithms will exhibit superior performance in detecting fraudulent transactions within the real dataset. Table VI provides a summary of the time efficiency exhibited by the ML algorithms in both experiments. The table reveals that the SVM model demonstrated the poorest time efficiency in the first experiment. Conversely, the LR model displayed good time efficiency in both experiments, while the DT model exhibited a comparable level of time efficiency to the LR model.

TABLE VI: Time Efficiency (Real Vs. Synthetic Datasets)

(a) Real Dataset

Model/ Time Measure	Training Time (1000 Rows)	Scoring Time (1000 Rows)	Total Time
NB	10 ms	198 ms	17 s
LR	23 ms	219 ms	17 s
DT	50 ms	172 ms	18 s
RF	282 ms	1 s	2 min 25 s
SVM	4 s	12 s	27 min 1 s

(b) Synthetic Dataset

Model/ Time Measure	Training Time (1000 Rows)	Scoring Time (1000 Rows)	Total Time
LR	177.5 ms	112.5 ms	2 s
DT	170.6 ms	160 ms	2 s
RF	129.5 ms	601 ms	14 s

The LR model continues to demonstrate superior performance and time efficiency in both experiments. When comparing the time efficiency in both experiments, it becomes apparent that the scoring time in the synthetic dataset is 50% faster than in the real dataset, resulting in quicker identification of fraudulent transactions. Therefore, it can be inferred that utilizing fewer critical features leads to better time efficiency while maintaining a similar level of performance.

VI. Discussion

This section covers the technical details of the fraud detection process, which is executed on the server side without requiring user intervention. This experiment suggests some new features to enhance the current online credit card system, including:

1) User's location (longitude, latitude)
2) IP address or the IMEI number
3) Transaction's store
4) Transaction time period (time difference between every two consecutive transactions)
5) Time (the time in hours of the day)
6) Transaction's amount

This section will delve into the underlying operations of the proposed system. The system performs the primary fraud detection operations on the server-side without the user's involvement. The machine learning algorithm is applied by the server to carry out fraud detection using the proposed features in this study. To test the model, researchers used

several datasets, including a real dataset and artificially generated datasets that simulate the real dataset. Additionally, they evaluated their approach by integrating multiple ML algorithms on all datasets to observe their performance and their interaction with different situations or user behavior. As a result, they selected the LR ML algorithm as it demonstrated great performance in both experiments.

Trivedi et al. [31] conducted a comparative study using the same real dataset used in this research. They found that the RF algorithm performed best with 95% precision, but they did not consider the time required for this model. In contrast, the proposed system achieved better performance with the LR algorithm, achieving 96% precision while being very time-efficient.

In another study by Lakshmi et al. [32], they implemented three ML algorithms (DT, LR, and RF) on different variables in two datasets with five and ten variables. Their results showed the following accuracies for each model: for the first dataset, LR 87.2%, DT 89%, and RF 90.1%, and for the second dataset, LR 88.6%, DT 92.5%, and RF 93.6%. The developed system uses only six features, and researchers obtained better results than their approach, achieving 94.5% accuracy for the LR model.

VII. Conclusion

This research examined several ML algorithms to identify the most suitable one for the proposed system based on six hypothetical cases. In the first experiment, five ML algorithms are used, including DT, LR, RF, NB, and SVM, on a real dataset. researchers then selected the top three models to use in the second experiment. After conducting both experiments, the study found that the LR model performed the best. In the first experiment, the LR model achieved 96.4% accuracy with excellent time efficiency, and in the second experiment, it reached 94.5% accuracy while maintaining great time efficiency. The second experiment also demonstrated a remarkable reduction in scoring time by more than 50% compared to the first experiment by utilizing only six crucial features. In the third experiment, researchers found that combining all six features led to better performance than using some of them.

References

[1] F. Hayashi, "Payment card fraud rates in the united states relative to other countries after migrating to chip cards," *Economic Review*, vol. 104, no. 4, pp. 23–40, 2019.

[2] N. Report, "Card fraud losses reach $28.65 billion," Dec 2020.

[3] R. Jain, B. Gour, and S. Dubey, "A hybrid approach for credit card fraud detection using rough set and decision tree technique," *International Journal of Computer Applications*, vol. 139, no. 10, pp. 1–6, 2016.

[4] C. Silver, "Capital one data breach impacts 106 million customers," May 2021.

[5] X. Hu, X. Zhang, and N. P. Lovrich, "Forecasting identity theft victims: Analyzing characteristics and preventive actions through machine learning approaches," *Victims & Offenders*, vol. 16, no. 4, pp. 465–494, 2021.

[6] "Shift credit card processing, credit card fraud statistics," Jan 2021.

[7] S. Al Balawi and N. Aljohani, "Credit-card fraud detection system using neural networks,"

[8] P. Roszkowska, "Fintech in financial reporting and audit for fraud prevention and safeguarding equity investments," *Journal of Accounting & Organizational Change*, vol. 17, no. 2, pp. 164–196, 2021.

[9] K. Modi and R. Dayma, "Review on fraud detection methods in credit card transactions," in *2017 International Conference on Intelligent Computing and Control (I2C2)*, pp. 1–5, IEEE, 2017.

[10] N. Kasa, A. Dahbura, C. Ravoori, and S. Adams, "Improving credit card fraud detection by profiling and clustering accounts," in *2019 Systems and Information Engineering Design Symposium (SIEDS)*, pp. 1–6, IEEE, 2019.

[11] S. Goel, J. M. Hofman, S. Lahaie, D. M. Pennock, and D. J. Watts, "Predicting consumer behavior with web search," *Proceedings of the National academy of sciences*, vol. 107, no. 41, pp. 17486–17490, 2010.

[12] P. Tiwari, S. Mehta, N. Sakhuja, J. Kumar, and A. K. Singh, "Credit card fraud detection using machine learning: a study," *arXiv preprint arXiv:2108.10005*, 2021.

[13] B. Al Smadi and M. Min, "A critical review of credit card fraud detection techniques," in *2020 11th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, pp. 0732–0736, IEEE, 2020.

[14] R. Sailusha, V. Gnaneswar, R. Ramesh, and G. R. Rao, "Credit card fraud detection using machine learning," in *2020 4th international conference on intelligent computing and control systems (ICICCS)*, pp. 1264–1270, IEEE, 2020.

[15] A. Saxena and H. Ponnapalli, *U.S. Patent Application No. 13/109,946*. 2012.

[16] D. P. S. Meredith, D. Kent, *Fraud detection via mobile device location tracking." U.S. Patent No. 9,858,575. 2*. 2018.

[17] S. Rajasekaran and R. Varadarajan, "U.s. patent no. 6,908,030." U.S. Patent 6,908,030, 2005.

[18] D. Varmedja, M. Karanovic, S. Sladojevic, M. Arsenovic, and A. Anderla, "Credit card fraud detection-machine learning methods," in *2019 18th International Symposium INFOTEH-JAHORINA (INFOTEH)*, pp. 1–5, IEEE, 2019.

[19] B. Al Smadi, A. A. S. AlQahtani, and H. Alamleh, "Secure and fraud proof online payment system for credit cards," in *2021 IEEE 12th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, pp. 0264–0268, IEEE, 2021.

[20] J. Essebag, S. Pochic, and C. Lalo, "U.s. patent no. 10,032,169." U.S. Patent 10,032,169, 2018.

[21] G. McDonald, "System and method for cardless secure on-line credit card/debit card purchasing," Sept. 23 2010. US Patent App. 12/408,325.

[22] P. Barbour, *Method and system for conducting financial transactions using single-use credit card numbers.* 2004.

[23] Y. Jain, N. Tiwari, S. Dubey, and S. Jain, "A comparative analysis of various credit card fraud detection techniques," *Int J Recent Technol Eng*, vol. 7, no. 5S2, pp. 402–407, 2019.

[24] I. Sadgali, N. Sael, and F. Benabbou, "Human behavior scoring in credit card fraud detection," *IAES International Journal of Artificial Intelligence*, vol. 10, no. 3, p. 698, 2021.

[25] Q. Li and Y. Xie, "A behavior-cluster based imbalanced classification method for credit card fraud detection," in *Proceedings of the 2019 2nd International Conference on Data Science and Information Technology*, pp. 134–139, 2019.

[26] I. Sadgali, N. Sael, and F. Benabbou, "Human behavior scoring in credit card fraud detection," *IAES International Journal of Artificial Intelligence*, vol. 10, no. 3, p. 698, 2021.

[27] F. Carcillo, Y.-A. Le Borgne, O. Caelen, Y. Kessaci, F. Oblé, and G. Bontempi, "Combining unsupervised and supervised learning in credit card fraud detection," *Information sciences*, vol. 557, pp. 317–331, 2021.

[28] N. Report, "Credit card fraud detection dataset," Mar 2018.

[29] T. Sabanaglu, "Online shopping frequency according to online shoppers worldwide as of october 2018," 2018.

[30] B. Al-Smadi, *Credit Card Security System and Fraud Detection Algorithm*. PhD thesis, Louisiana Tech University, 2021.

[31] N. K. Trivedi, S. Simaiya, U. K. Lilhore, and S. K. Sharma, "An efficient credit card fraud detection model based on machine learning methods," *International Journal of Advanced Science and Technology*, vol. 29, no. 5, pp. 3414–3424, 2020.

[32] S. Lakshmi and S. Kavilla, "Machine learning for credit card fraud detection system," *International Journal of Applied Engineering Research*, vol. 13, no. 24, pp. 16819–16824, 2018.