..ISSN: 2040-0748 Vol-12 Issue-02 Nov 2023

Performance Evaluation of Machine Learning Algorithms for Credit Card Fraud Detection

Mr. V Naveen Kumar(Assistant Professor),
Katla Uday Kiran Reddy, Yamannagari Abhiram, Kotagiri Dinesh, Kammari Prashanth,
Department of CSE,
MALLA REDDY INSTITUTE OF TECHNOLOGY AND SCIENCE,Telangana, Hyderabad.

Abstract—

Credit card fraud and card fraud itself have grown in prevalence in the last few years. One common kind of internet fraud is the theft of credit card information and its subsequent use in fraudulent online purchases. Credit card companies and stores simply cannot process all of the real transactions in order to identify these fraudulent ones. The availability and collection of sufficient data will allow machine learning methods to be used to address this problem. In order to detect credit card fraud, this research employs popular supervised and unsupervised ML methods on a highly skewed dataset. When it comes to accurately classifying data and addressing skewness, unsupervised machine learning algorithms fared better than supervised ones.

I. INTRODUCTION

Online and offline shoppers both are increasingly turning to credit cards as a means of payment because of the convenience and flexibility they provide [1]. On the other hand, there have been certain problems associated with this convenience. Criminals, hackers, and others who perpetrate crimes have made credit card-based transactions a prime target. You don't even need to physically hold the card when using a credit card online; all you have to do is input the card details. The use of a One-Time Password (OTP) as an additional layer of security is being examined in some instances. In all other cases, especially those involving foreign transactions, it may be used for illicit acquisitions as it is not necessary. This kind of use is known as "Card-Not-Present" as it does not need the actual card itself but rather its information. Credit card information may be easily stolen with the rise of techniques like shoulder surfing, purchasing it, and sniffing site traffic. In a credit card fraud, one of three parties—the cardholder, the issuing bank, or the merchant—ends up paying the price. It is usually the responsibility of the cardholder to identify fraudulent activity and notify the bank that issued the card of any fraudulent transactions. After that, the bank looks into it, and if they find proof of fraud, they start the procedure to take the money out of the account. There is no assurance that this approach will fix the problem since it is not real-time [2].

The credit card firm is a major player because its reputation takes a major hit when the number of fraudulent charges on its cards rises. Therefore, a system for detecting and preventing fraud must be put in place by the issuer. Businesses often inform their consumers on the best practices for using their cards securely in an effort to reduce the occurrence of fraud. To prevent unauthorized use, additional authentication methods such as one-time passwords and security questions are sometimes used. Nevertheless, even with these safeguards in place, fraud incidents will always occur [3]. So, when a fraud has been revealed, the bank has to invest in post-mortem analysis, attempt to recover the funds, and punish the person or people responsible. The fact that this detection takes many days to process doesn't make it effective in discouraging fraud.

Credit card firms utilize automated systems based on machine learning called Fraud Detection Systems (FDS) to identify fraudulent transactions before the end customer ever notices [4]. Preventing fraudulent transactions from being committed to the database is the main goal of such a system. In an ideal FDS, the number of false positives, in which the end user is inconvenienced by the interruption of a legitimate transaction, would be minimized as well.

Algorithms based on machine learning identify future data observed in the domain by working with large amounts of example data from the underlying domain to develop a computation model. The example data classes must be pre-labeled for a class of algorithms known as Supervised Learning Algorithms. Data is grouped into identical groups and referred to as belonging to one class in another class of algorithms that employ Unsupervised Learning. The literature is replete with algorithms that combine the two methods [5–12]. FDS gathers a large amount of historical data in order to do calculations on it.

However, the amount of legitimate transactions usually much exceeds the amount of fraudulent ones

in transaction data sets. In this research, we present and assess a number of well-known ML algorithms on their accuracy in detecting fraudulent transactions in a real-world dataset that is unbalanced.

II. CREDIT CARD FRAUD DETECTION

Credit cards are rectangular pieces of plastic that a bank or other financial institution issues to a verified customer. The consumer may use it to buy things and services at physical point-of-sale terminals or on e-commerce websites. Credit card fraud encompasses both of these types of illegal card usage.

A. Fraud Prevention for Credit Cards

The nature of the scam determines the best method for detecting credit card fraud. A scam may be perpetrated via several channels [13]. Nevertheless, there are primarily two ways to classify them. The theft of physical cards via illicit means is the first kind of fraud. There are a number of ways to do this, including taking the card from its rightful owner either before or after delivery, or making a new cloned card that can pass itself off as the real one. The unauthorized acquisition of credit card details is the second kind of fraud. There are a number of ways to do this, including capturing card imprints at hotels, shoulder surfing, and phishing [14]. It is possible for the real allotted to attempt to fool the corporation into thinking he did not complete the deal. If an additional authentication method, such as an OTP, has been set up, then its protections must also be evaded. By distinguishing between the actions of a fraudster and those of a genuine user, a computational fraud detection system (FDS) may identify any kind of fraud.

I. Difficulties

If an FDS can identify every kind of credit card fraud, it would be ideal. The idea behind it is to understand how fraudsters use their cards and how they spend their money, rather than concentrating on the fraud vector [15]. Creating FDS turns into a binary classification challenge if there is access to long-term card use data of several individuals together with fraudulent transactions that occurred inside that time frame. "Normal" and "Fraudulent" transactions are the two categories that matter here. Thus, these datasets may be used with existing methods of Unsupervised and Supervised machine learning. Nevertheless, these algorithms have a few obstacles that prevent them from producing accurate categorization results. Class scenes, seasonal user behavior variations, domain metrics, a lack of truth labels, and the need for real-time categorization are a few of these problems [16].

III. MACHINE LEARNING ALGORITHMS FOR FRAUD DETECTION

We choose a few well-known supervised and unsupervised machine learning techniques to test for the core issue.

A. Supervised Learning: We have taken into account a wide range of supervised learning algorithms, from the most ancient to the most cutting-edge. Among them, you may find hybrid algorithms, neural networks (both deep and conventional), Bayesian methods, and algorithms based on trees.

• Random Forest (RF)—An essential approach classifier, Random Forest (RF) integrates many tree predictors. Using RF has the benefit of being resistant to outliers and noise. [17].

• NNs, or artificial neural networks, are the most popular tool for detecting fraud. Based on its associative memory of previously learned patterns, NN is able to distinguish related patterns and make value predictions [18]. Artificial neurons that are not being used make up an ANN. When training, ANNs employ the back propagation method, which is a technique used by Feed-Forward Neural Networks (NN). [16].

A multi-layer perception network trained using a stochastic gradient descent is the foundation of deep learning (DL). One unsupervised learning technique that uses back propagation by making the inputs equal to the outputs is the auto-encoder (AE), which the authors of [19] suggested as a foundation for deep learning.

When data is processed and categorized linearly, supervised learning makes use of Support Vector Machines (SVMs) [16] [18].

1. Naïve Bayes - Classifiers using the Naive Bayes (NB) method rely on Bayesian theory to make decisions using conditional probabilities [7]. 2. Logistic Regression - Using one or more features, Logistic Regression (LR) finds the optimal fitting parameter to estimate the probability of a binary response [7].

IV. PERFORMANCE EVALUATION

Here is Table A.

We used a publicly available, processed, and authentic dataset for this test. World line collaborated with Andrea Dal Pozzolo and colleagues from the Machine Learning Group at ULB (University Libre de Bruxelles) to study big data mining and fraud detection [20]. As a team, they collaborated to collect and evaluate the dataset. The data includes a total of 284,807 card transactions done throughout Europe in

..ISSN: 2040-0748 Vol-12 Issue-02 Nov 2023

September 2013. The dataset is severely imbalanced due to the 492 fake transactions included in it.

Due to concerns about individual privacy, only 28 attributes obtained from PCA of physical traits were made available. Time and quantity are the only pieces of data that have been preserved and shown without modification. To keep track of how much time has elapsed since the first transaction, the 'Time' feature is used to time-stamped each transaction in the dataset. As a property, the 'Amount' represents the total amount of the transaction. Also, when fraud is identified, the 'Class' element, which identifies the sort of transaction label, takes the value 1, while otherwise, it takes the value 0.

B. Purposes

When it comes to detecting fraud, the chosen algorithms attribute the problem to the classification difficulty. We have considered the confusion matrix in Figure 1 for the purpose of evaluating measures.

The genuine fraud detection rate cannot be captured by standard measurements such as confusion matrices and accuracy due to the skewness in each class's occurrences.

In Table 1 you can see a confusion matrix that is used to assess categorization.

Predicted	Actual	
	Normal	*Fraud*
Normal	True Negatives (TN)	False NEgatives (FN)
Fraud	False Positives (FP)	True Positives (TP)

find a happy medium between the two types of detection.

• Precision or Positive Predictive Value: PPV is characterized as

$$PPV = \frac{TP}{TP + FP}$$

One way to look at it is as a measure of the proportion of positive forecasts that really came true.

To define the negative predictive value, we use the following formula:

$$NPV = \frac{TN}{TN + FN}$$

NPV is the ratio of all negative projected values to the number of accurately detected negative values, which is '0' in this case.

The accuracy rate of identifying true negatives is known as specificity.

$$Specificity = \frac{TN}{TN + FP}$$

A sensitivity test determines how many false positives were really detected.

$$Sensitivity = \frac{TP}{TP + FP}$$

Accuracy that is balanced is the average rate of detection on both classes.

$$Balanced\ Accuracy = \frac{Sensitivity + Specificity}{2}$$

"Prevalence" refers to the frequency with which the condition "yes" really happens.

$$Prevalence = \frac{TP + FN}{TP + FP + FN + TN}$$

A word borrowed from the medical field is the Diagnostic Odd Ratio (DOR). Overall, it determines how well a categorization test performed.

$$DOR = \frac{PPV * NPV}{(1 - PPV) * (1 - NPV)}$$

C. Results

The outcomes of evaluating several supervised, hybrid, and unsupervised machine learning algorithms over specified metrics are shown in Tables II–IV. A subset was chosen at random for each run to serve as the test set.

TABLE II. Analyzing the Results of a Supervised Machine Learning System

Techniques Used	Classifier Metrics Values					
	Positive Predictive Value (Precision)	*Negative Predictive Value*	*Prevalence*	*True Negative Rate (Specificity)*	*True Positive Rate (Sensitivity / Recall)*	*Diagnostic Odd Ratio (DOR)*
NB	0.06	0.99	0.001	0.97	0.82	6.319
RF	0.99	0.16	0.998	0.91	0.99	18.857
K-NN	NaN	0.99	0.001	1.0	0.0	0
LR	0.99	0.63	0.99	0.87	0.99	168.56
XGBT	0.99	0.92	0.99	0.81	0.99	1138.5
SVM	NaN	NaN	NaN	0.92	0.93	0
ANN	0.99	0.84	0.99	0.77	0.99	462
DL	0.98	0.93	0.86	0.98	0.91	651
QDA	0.97	0.89	0.97	0.42	0.88	261.60
NN	0.99	NaN	0.99	0.0	1.0	0

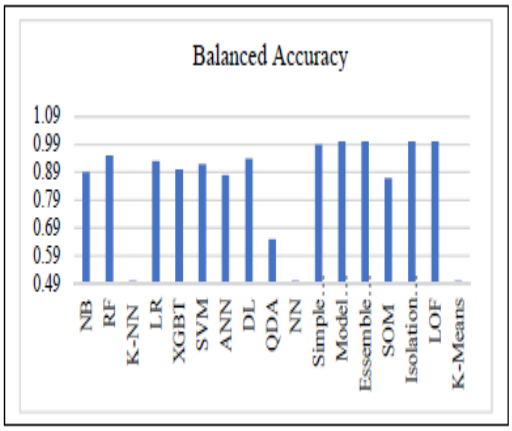
Section III: Hybrid Machine Learning Algorithm Performance Evaluation

Techniques Used	Classifier Metrics Values					
	Positive Predictive Value (Precision)	*Negative Predictive Value*	*Prevalence*	*True Negative Rate (Specificity)*	*True Positive Rate (Sensitivity/ Recall)*	*Diagnostic Odd Ratio (DOR)*
Simple Model	0.99	0.82	0.84	0.99	1.0	726
Model Outliers	0.99	0.82	0.84	1.0	1.0	726
Essemble Model	0.99	0.82	0.84	1.0	1.0	726

TABLE IV. UNIVISED MACHINE LEARNING ALGORITHM PERFORMANCE EVALUATION

Techniques Used	Classifier Metrics Values					
	Positive Predictive Value (Precision)	*Negative Predictive Value*	*Prevalence*	*True Negative Rate (Specificity)*	*True Positive Rate (Sensitivity/ Recall)*	*Diagnostic Odd Ratio (DOR)*
SOM Hybrid	0.92	0.84	0.99	0.83	0.92	60.375
Isolation Forest	0.99	0.99	0.99	1.0	1.0	9801
Local Outlier Factor	0.99	0.99	0.998	1.0	1.0	9801
K-Means	0.99	NaN	0.998	1.0	0.0	0

Quite a few rows in the table contain NaNs, meaning that the classifier failed to identify any positive or negative results.



The balanced accuracy of several machine learning algorithms is shown in Figure 2.

The outcomes of balanced accuracy are shown in Figure 2. The classifiers based on KNN, NB, and K-means have not been shown since their accuracy was very poor at 0.50, owing to almost no true positives.

D. Results Discussion

The results presented in Section C make it clear that no categorization method excels in every way. Thus, performance considerations pertinent to FDS users have been considered throughout the analysis and interpretation of the findings.

a) Is it a typical transaction if no alert goes off?

The NPV value may shed light on this matter; when it comes to supervised learning, NB & KNN produced an NPV of 0.99, NN & SVM produced an NPV of 0, and RF produced an NPV of 0.16. While K-means achieved a value of 0 and IF and LOF achieved a score of 0.99 for unsupervised learning,

all hybrid supervised learning approaches yielded an NPV of 0.82.

b) Is it really a scam if an alarm goes off?

The PPV was 0.99 for all supervised and unsupervised learning techniques with the exception of KNN, SVM, and SOM. They all came up with 0.82 using hybrid techniques. On the balanced accuracy front, supervised and unsupervised approaches are comparable, with ensemble techniques taking a distant second.

V. CONCLUSIONS

This research assesses the usefulness of machine learning algorithms for the purpose of identifying fraudulent charges on credit cards. The very significant imbalance between samples of legitimate and fraudulent transactions makes credit card fraud detection a unique classification issue. Various criteria were used to assess a variety of well-known algorithms across supervised, ensemble, and unsupervised categories. The results show that unsupervised algorithms do better than other methods on all measures, both in terms of absolute performance and relative performance, when dealing with dataset skewness.

REFERENCES

[1] The importance of credit cards: https://budgeting.thenest.com/importance-credit-cards-29514.html

[2] The chargeback process in a credit card: https://chargebacks911.com/chargeback-process/

[3] "Low and Slow" Is How the Credit Card Fraudsters Roll: https://www.threatmetrix.com/digital-identity-blog/fraudprevention/ low-and-slow-is-how-the-credit-card-fraudsters-roll/

[4] A. Dal Pozzolo, G. Boracchi, O. Caelen, C. Alippi and G. Bontempi, "Credit Card Fraud Detection: A Realistic Modeling and a Novel Learning Strategy," in IEEE Transactions on Neural Networks and Learning Systems, vol. 29, no. 8, pp. 3784-3797, Aug. 2018.

[5] L. Zheng, G. Liu, C. Yan and C. Jiang, "Transaction Fraud Detection Based on Total Order Relation and Behavior Diversity," in IEEE Transactions on Computational Social Systems, vol. 5, no. 3, pp. 796- 806, Sept. 2018.

[6] Vaishali. Article: Fraud Detection in Credit Card by Clustering Approach. International Journal of Computer Applications 98(3):29-32, July 2014.

[7] J. O. Awoyemi, A. O. Adetunmbi and S. A. Oluwadare, "Credit card fraud detection using machine learning techniques: A comparative analysis," 2017 International Conference on Computing Networking and Informatics (ICCNI), Lagos, 2017, pp. 1-9.

[8] L. Zheng et al., "A new credit card fraud detecting method based on behavior certificate," 2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC), Zhuhai, 2018, pp. 1-6.