Reproducible Machine Learning Models with R Implementation for Credit Card Fraud Analysis

Antrang Agrawal – 20BCI0262
School of Computer Science and Engineering
Vellore Institute of Technology, Vellore, India
antrang.agrawal2020@vitstudent.ac.in

Abstract:

One of the widely used methods of payment for online transactions in both developed and developing nations is the credit card. Credit cards' invention has made internet transactions simple, pleasant, and convenient. However, it has also given fraudsters new chances, which has raised the rate of fraud. The global impact of credit card fraud is significant; many businesses and people have lost millions of US dollars as a result. Furthermore, hackers frequently develop advanced strategies; as a result, it is vital to create new, dynamic procedures that can quickly react to changing fraudulent trends. An analysis of more effective methods for detecting credit card fraud is provided in this research. This research is focused on recently proposed credit card fraud detection strategies based on machine learning and nature inspiration. This paper gives an overview of current developments in the detection of credit card fraud. Additionally, this study covers some of the drawbacks and contributions of the current credit card fraud detection methods. It also gives researchers in this field the background knowledge they need. Additionally, this evaluation acts as a roadmap and stepping stone for both individuals and financial organizations. As a result, during the past ten years, the development of payment card fraud detection systems has increasingly centered on machine learning (ML)-based strategies that automate the process of recognizing fraudulent patterns from massive amounts of data.

Keywords - Credit card Fraud - Electronic transactions - Machine learning - Algorithms - Nature-inspired techniques – Cybercriminals

I. Introduction:

For business owners, payment card issuers, and transactional services businesses, payment card fraud poses a significant challenge and costs them a lot of money each year. Card fraud losses worldwide climbed from 9.84 billion in 2011 to 27.85 billion in 2018, and they are anticipated to surpass 40 billion in 2027, according to the 2019 Nilsson Report.

It is well recognized that identifying fraud trends in payment card transactions is a highly challenging task. A human analyst can no longer detect fraudulent patterns
in transaction datasets, which are frequently characterized by a huge number of samples, multiple dimensions, and online updates, due to the exponential growth in the amount of data created by credit card transactions.

Credit card fraud can be defined as illegal use of credit card information for online purchase. Credit card transactions are done physically or virtually.

Physical interactions with the seller are considered transactions that involve physical exchanges. At the moment of sale, customers must physically display their cards. Virtual transactions are transactions carried out over the phone or the internet. For online purchases, it is necessary for consumers to give specific card information (such as the CVV number, password, security question, etc). The invention of credit cards has not only made online transactions seamless, easier, comfortable and convenient, it has also provided new fraud opportunities for criminals, and increased the rate of fraud.

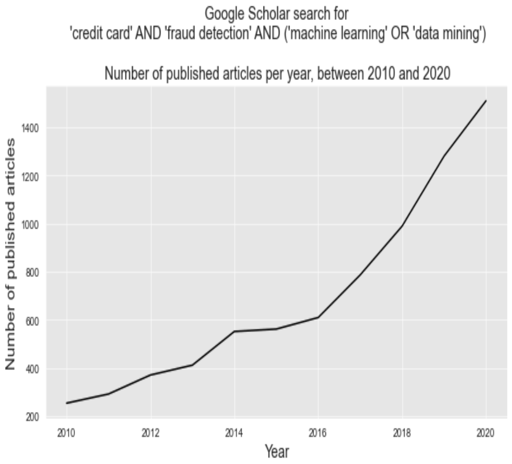
Credit card fraud detection is a classification problem It is now clear that machine learning techniques can provide effective solutions to the problem of credit card fraud detection, and the research literature on the topic has grown exponentially in the last decade.

Credit card fraud detection (CCFD) is like looking for needles in a haystack. It requires identifying the fraudulent transactions among the millions that occur every day. It is currently practically difficult for a human specialist to identify significant patterns in transaction data due to the exponential growth in data. This is why information extraction from huge datasets is necessary in the field of fraud detection, where machine learning techniques are now widely used.

II. Credit Card Fraud Detection Using ML:

The use of ML approaches in payment card fraud detection systems has significantly increased their capacity to identify frauds more quickly and help payment processing intermediaries spot unauthorized transactions. Although there have been more fraudulent transactions in recent years, in 2016 the percentage of losses attributable to fraud began to decline. Implementing ML-based fraud detection systems is now a requirement for institutions and businesses in order to gain the trust of their customers, in addition to helping them save money.

The study of algorithms that get better automatically as they gain experience is known as machine learning (ML). Data mining, pattern recognition, and statistics are all strongly related to machine learning. The algorithmic portion of the knowledge extraction process receives particular emphasis in this emerging area of computer science and artificial intelligence. ML is important to many scientific fields, and we use its applications every day. It is utilized, for instance, to filter spam emails, predict the weather, diagnose illnesses, recommend products, find faces, and detect fraud.



(Fig 1)
(Source: [1])
Number of published articles on the topic of machine learning and credit card fraud detection between 2010 and 2020

Credit card fraud detection (CCFD) is a challenging problem, which requires analyzing large volumes of transaction data to identify fraud patterns. The large volumes of data, together with the evolving techniques of fraudsters, make it impossible for human investigators to efficiently address this problem. In the last decade, CCFD has been increasingly complemented with computer algorithms known as *Machine Learning* (ML), which allows searching and detecting patterns from large amounts of data. ML algorithms have been shown to significantly improve the efficiency of fraud detection systems, and assist fraud investigators in detecting fraudulent transactions.

III. **Literature Survey**:

Author	Contribution	Limitation and Results
Wong et al. (2016)	An improved AIS-based credit card detection	Because of low detection rate, this model can’t be used for production.

	technique was proposed	After profiling the data, I can see that the model had low Classification Accuracy.
Khan et al. (2018b)	Real time fraud detection technique was introduced.	Misclassification rate is fairly low and time taken to classify is very less. Also, Dataset size taken was low and Training took too much time.
Khan et al. (2016)	A new and Robust Model for Fraud Detection was introduced.	Low impact transactions were not properly considered. Multiple low impact transactions can compromise the system
Seeja and Zareapoor (2017)	An improved and robust model capable of handling data imbalance was introduced. System was tested on large number of data instance.	The proposed method is not dynamic, it has a slow classification rate, and it cannot identify transactions with comparable fraud and legal patterns.
Potamitis (2018)	An ontology-based expert system for fraud detection was proposed.	Proposed expert system is static, it requires regular updates
Carminati et al. (2016)	A semi-supervised and unsupervised decision support system for handling fraud and anomaly detection was proposed.	The suggested technique's clustering step uses a lot of storage space, and synthetic data was utilized to create the model.
Mahmoudi and Duman (2015)	A novel technique based on modified version of Fisher Discriminant Function was proposed.	Proposed method underperforms ANN, decision tree, NB, and normal Fisher and cannot handle false negatives well.
Soltani et al. (2017)	A novel model capable of handling misuse and anomaly detection was proposed.	FP is too high and Classification speed can be affected by generating detectors for all transactions.
Zareapoor and Shamsolmoali (2015)	A novel technique based on bagging ensemble classifier was introduced.	Bagging ensemble classifier involves classification of different datasets; hence it is slow.
Carminati et al. (2015)	A semi-supervised and unsupervised decision support system for handling fraud and anomaly detection was proposed.	The suggested technique's clustering step uses a lot of storage space, and synthetic data was utilized to create the model.
Stolfo et al. (2019)	A technique based on meta-learning was proposed.	FP rate of 13, 16, 16 and 23% was achieved but Classification speed is slow, it

		consists of a combination of several classifiers.
Soltani Halvaiee and Akbari (2018)	A modified method for negative selection was introduced for Fraud.	Memory generation phase and calculation of affinity are time-consuming.
Lu and Ju (2021)	Proposed technique is capable of handling data imbalance in Fraud Detection.	CA is low and Dataset is highly imbalanced.
Maes et al. (2012)	ANN-based and BN-based technique was proposed.	Proposed ANN technique can only handle discrete variables.

IV The General Process of Credit Card Transaction:

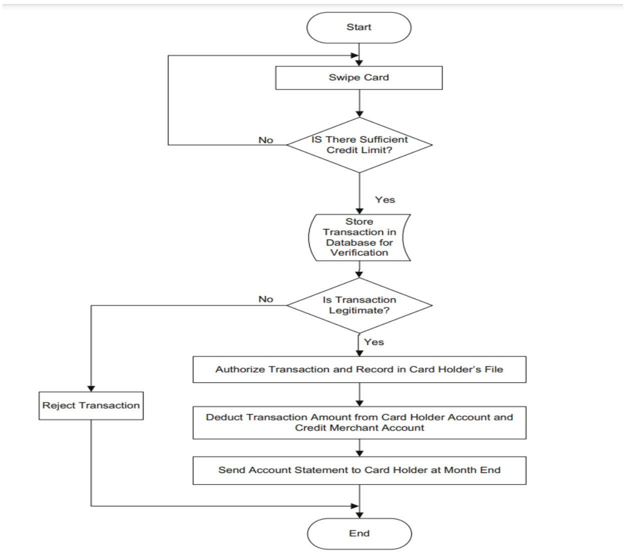

Start
Swipe Card
No
IS There Sufficient Credit Limit?
Yes
Store Transaction in Database for Verification
No
Is Transaction Legitimate?
Yes
Authorize Transaction and Record in Card Holder's File
Reject Transaction
Deduct Transaction Amount from Card Holder Account and Credit Merchant Account
Send Account Statement to Card Holder at Month End
End

Fig 2.
(Source: [17])

(a) Firstly, user is expected to swipe a card (for virtual transaction) or enter card details (for physical transactions).

(b) Furthermore, the transaction is approved if the card information is verified ok and if there is sufficient credit limit.

(c) Afterwards, the transaction request is sent to a file. The request stays in the file for 5 days. During this period, the transaction is verified by the merchant.

(d) Afterwards, if the transaction is legitimate, it is authorized and recorded in the cardholder’s file. Transaction authorization is an important part of a transaction. It is also the first level of security. The credit card limit of cardholders is regulated at this level.

(e) After authorization, transaction amount is deducted from the account balance of cardholder. The deducted amount will be credited to merchant’s account.

(f) At the end of the month, the financial institution will send a statement of account containing the list of transactions that have been performed by the user. The statement contains the outstanding balance. The user is then expected to pay the total balance.

V. Process Diagram and Architecture in The Real-World Credit Card Fraud Detection System:

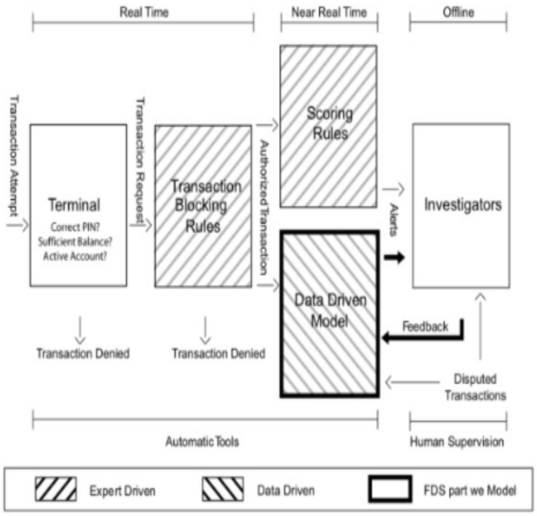


Fig 3.
Diagram illustrating the layers of control in an FDS
(Source: [1])

VI. Challenges in Credit Card Fraud Detection using ML:

Class imbalance: Transaction data contain much more legitimate than fraudulent transactions: The percentage of fraudulent transactions in a real-world dataset is typically well under 1%. Learning from imbalanced data is a difficult task since most learning algorithms do not handle well large differences between classes.

Concept drift: Transaction and fraud patterns change over time. On the one hand, the spending habits of credit card users are different during weekdays, weekends, vacation periods, and more generally evolve over time. On the other hand, fraudsters adopt new techniques as the old ones become obsolete.

Near real-time requirements: Fraud detection systems must be able to quickly detect fraudulent transactions. Given the potentially high volume of transaction data (millions of transactions per day), classification times as low as tens of milliseconds may be required.

Categorical features: Transactional data typically contain numerous *categorical* features, such as the ID of a customer, a terminal, the card type, and so on. Categorical features are not well handled by machine learning algorithms and must be transformed into numerical features.

Sequential modeling: Each terminal and/or customer generates a stream of sequential data with unique characteristics. An important challenge of fraud detection consists in modeling these streams to better characterize their expected behaviors and detect when abnormal behaviors occur.

Class overlaps: The last two challenges can be associated with the more general challenge of overlapping between the two classes. With only raw information about a transaction, distinguishing between a fraudulent or a genuine transaction is close to impossible. This issue is commonly addressed using feature engineering techniques, that add contextual information to raw payment information.
Performance measures: Standard measures for classification systems, such as the mean misclassification error or the AUC ROC, are not well suited for detection problems due to the class imbalance issue, and the complex cost structure of fraud detection.

Lack of public datasets: For obvious confidentiality reasons, real-world credit card transactions cannot be publicly shared. There exists only one publicly shared dataset, which was made available on Kaggle by our team in 2016.

VII. Proposed Methodology:

We will Be Using these Steps and Implementing 4 ML Algorithms on The Credit Card Dataset.

- Dataset Information
- Data Visualization and Exploration
- Model Selection
- Decision Tree Model
- Logistic Regression Model
- Random Forest
- XGBoost
- Gradient Boosting (GBM)

1.) Dataset Information:

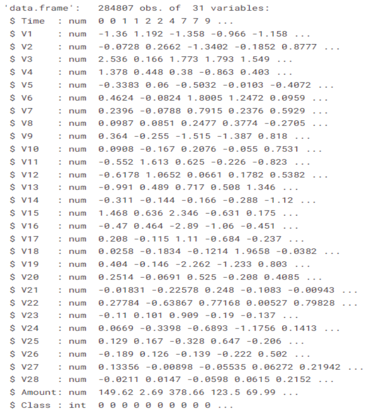
The dataset includes credit card transactions made by European cardholders in September 2013.
We have 492 frauds out of 284,807 transactions in this dataset of transactions that took place over the course of two days. The dataset is very skewed, with frauds making up 0.172% of all transactions in the positive class.

It only has numeric input variables that have undergone PCA transformation. Unfortunately, we are unable to offer the original characteristics and additional context for the data due to confidentiality concerns. The major components obtained with PCA are features V1, V2, ….V28. The only features that have not been changed with PCA are "Time" and "Amount." The seconds that passed between each transaction and the dataset's first transaction are listed in the feature "Time."

The feature 'Amount' is the transaction Amount, this feature can be used for example-dependent cost-sensitive learning. Feature 'Class' is the response variable and it takes value 1 in case of fraud and 0 otherwise.

Given the class imbalance ratio, we recommend measuring the accuracy using the Area Under the Precision-Recall Curve (AUPRC). Confusion matrix accuracy is not meaningful for unbalanced classification.

Summary of Dataset used



```
'data.frame':   284807 obs. of  31 variables:
 $ Time  : num  0 0 1 1 2 2 4 7 7 9 ...
 $ V1    : num  -1.36 1.192 -1.358 -0.966 -1.158 ...
 $ V2    : num  -0.0728 0.2662 -1.3402 -0.1852 0.8777 ...
 $ V3    : num  2.536 0.166 1.773 1.793 1.549 ...
 $ V4    : num  1.378 0.448 0.38 -0.863 0.403 ...
 $ V5    : num  -0.3383 0.06 -0.5032 -0.0103 -0.4072 ...
 $ V6    : num  0.4624 -0.0824 1.8005 1.2472 0.0959 ...
 $ V7    : num  0.2396 -0.0788 0.7915 0.2376 0.5929 ...
 $ V8    : num  0.0987 0.0851 0.2477 0.3774 -0.2705 ...
 $ V9    : num  0.364 -0.255 -1.515 -1.387 0.818 ...
 $ V10   : num  0.0908 -0.167 0.2076 -0.055 0.7531 ...
 $ V11   : num  -0.552 1.613 0.625 -0.226 -0.823 ...
 $ V12   : num  -0.6178 1.0652 0.0661 0.1782 0.5382 ...
 $ V13   : num  -0.991 0.489 0.717 0.508 1.346 ...
 $ V14   : num  -0.311 -0.144 -0.166 -0.288 -1.12 ...
 $ V15   : num  1.468 0.636 2.346 -0.631 0.175 ...
 $ V16   : num  -0.47 0.464 -2.89 -1.06 -0.451 ...
 $ V17   : num  0.208 -0.115 1.11 -0.684 -0.237 ...
 $ V18   : num  0.0258 -0.1834 -0.1214 1.9658 -0.0382 ...
 $ V19   : num  0.404 -0.146 -2.262 -1.233 0.803 ...
 $ V20   : num  0.2514 -0.0691 0.525 -0.208 0.4085 ...
 $ V21   : num  -0.01831 -0.22578 0.248 -0.1083 -0.00943 ...
 $ V22   : num  0.27784 -0.63867 0.77168 0.00527 0.79828 ...
 $ V23   : num  -0.11 0.101 0.909 -0.19 -0.137 ...
 $ V24   : num  0.0669 -0.3398 -0.6893 -1.1756 0.1413 ...
 $ V25   : num  0.129 0.167 -0.328 0.647 -0.206 ...
 $ V26   : num  -0.189 0.126 -0.139 -0.222 0.502 ...
 $ V27   : num  0.13356 -0.00898 -0.05535 0.06272 0.21942 ...
 $ V28   : num  -0.0211 0.0147 -0.0598 0.0615 0.2152 ...
 $ Amount: num  149.62 2.69 378.66 123.5 69.99 ...
 $ Class : int  0 0 0 0 0 0 0 0 0 0 ...
```

Fig 4.

(Source: Own Work)

2.) Data Visualization:

Distribution of Class Labels

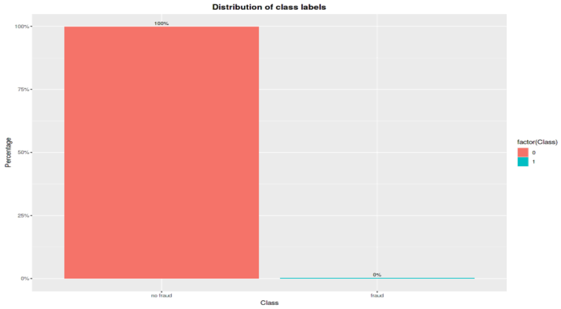


Fig 5
(Source: Own Work)

With non-fraudulent transactions accounting for 99.8% of cases, the dataset is obviously excessively unbalanced. A straightforward metric like accuracy is inappropriate in this situation because even a classifier that labels all transactions as legitimate will be over 99% accurate. AUC would be a suitable indicator of model performance in this case (Area Under the Precision-Recall Curve)

Distribution of
Transaction Amount by Class

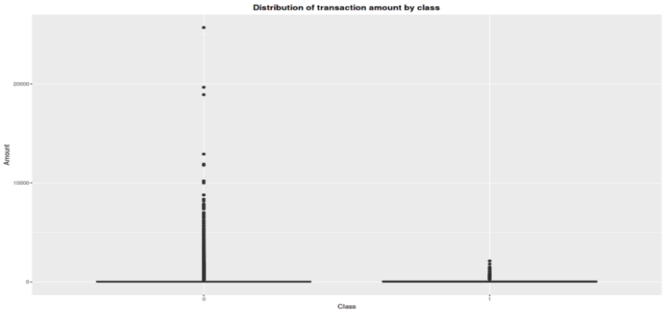


Fig 6.
(Source: Own Work)

There is clearly a lot more variability in the transaction values for non-fraudulent transactions.

Correlation Graph:

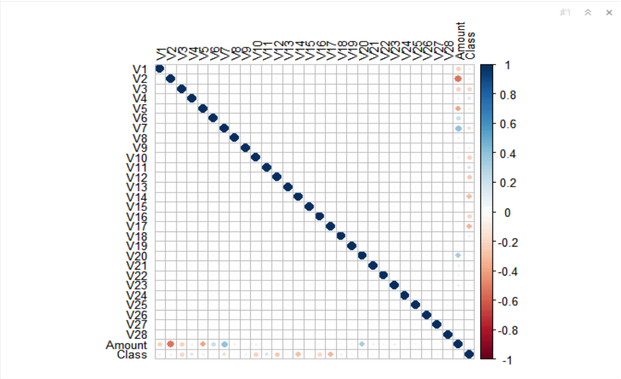


Fig 7.
(Source: Own Work)

The majority of the data features are not connected, as we can see. This is due to the fact that a Principal Component Analysis (PCA) method was given with the majority of the data prior to publication. The Principal Components that were most likely produced after propagating the actual features using PCA are the features V1 through V28. We are unsure if the relevance of the Principal Components is reflected in the numbering of the features.

t-SNE Visualization of Transactions

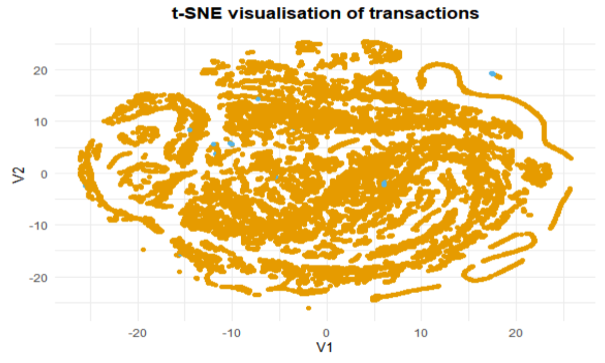
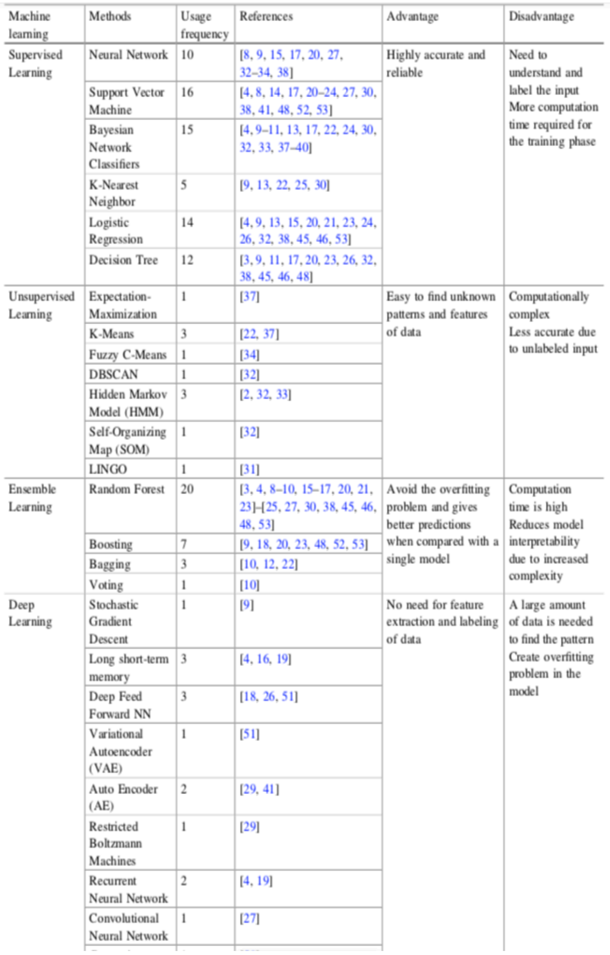


Fig 8.
(Source: Own Work)

t-distributed stochastic neighbor embedding (t-SNE) is a statistical method for visualizing high-dimensional data by giving each datapoint a location in a two or three-dimensional map. It is based on Stochastic Neighbor Embedding where Laurens van der Matten proposed the *t*-distributed variant. It is a nonlinear dimensionality reduction technique well-suited for embedding high-dimensional data for visualization in a low-dimensional space of two or three dimensions. Specifically, it models each high-dimensional object by a two- or three-dimensional point in such a way that similar objects are modeled by nearby points and dissimilar objects are modeled by distant points with high probability.

3.) Model Selection:

Usage Frequency of Various Machine Learning Models on Credit Card Fraud Detection:



Machine learning	Methods	Usage frequency	References	Advantage	Disadvantage
Supervised Learning	Neural Network	10	[8, 9, 15, 17, 20, 27, 32–34, 38]	Highly accurate and reliable	Need to understand and label the input More computation time required for the training phase
	Support Vector Machine	16	[4, 8, 14, 17, 20–24, 27, 30, 38, 41, 48, 52, 53]		
	Bayesian Network Classifiers	15	[4, 9–11, 13, 17, 22, 24, 30, 32, 33, 37–40]		
	K-Nearest Neighbor	5	[9, 13, 22, 25, 30]		
	Logistic Regression	14	[4, 9, 13, 15, 20, 21, 23, 24, 26, 32, 38, 45, 46, 53]		
	Decision Tree	12	[3, 9, 11, 17, 20, 23, 26, 32, 38, 45, 46, 48]		
Unsupervised Learning	Expectation-Maximization	1	[37]	Easy to find unknown patterns and features of data	Computationally complex Less accurate due to unlabeled input
	K-Means	3	[22, 37]		
	Fuzzy C-Means	1	[34]		
	DBSCAN	1	[32]		
	Hidden Markov Model (HMM)	3	[2, 32, 33]		
	Self-Organizing Map (SOM)	1	[32]		
	LINGO	1	[31]		
Ensemble Learning	Random Forest	20	[3, 4, 8–10, 15–17, 20, 21, 23]–[25, 27, 30, 38, 45, 46, 48, 53]	Avoid the overfitting problem and gives better predictions when compared with a single model	Computation time is high Reduces model interpretability due to increased complexity
	Boosting	7	[9, 18, 20, 23, 48, 52, 53]		
	Bagging	3	[10, 12, 22]		
	Voting	1	[10]		
Deep Learning	Stochastic Gradient Descent	1	[9]	No need for feature extraction and labeling of data	A large amount of data is needed to find the pattern Create overfitting problem in the model
	Long short-term memory	3	[4, 16, 19]		
	Deep Feed Forward NN	3	[18, 26, 51]		
	Variational Autoencoder (VAE)	1	[51]		
	Auto Encoder (AE)	2	[29, 41]		
	Restricted Boltzmann Machines	1	[29]		
	Recurrent Neural Network	2	[4, 19]		
	Convolutional Neural Network	1	[27]		

Fig 9.
(Source: [2])

Model selection consists in selecting the model that is expected to provide the best prediction performances on future data. For a fraud detection system, the best model can be defined as the model that has the highest expected fraud detection performances on the next block of transactions.

The estimation of model performances on future data is obtained by a validation procedure.

Validation procedures are however computationally intensive tasks. They require to repeat the training procedures many times in order to assess the performances of prediction models with different hyperparameters and using different sets of data.

A key challenge for model selection consists in efficiently exploring the space of model hyperparameters in order to best address the trade-off between fraud detection performances and computation times

In our Project we are going to compare performances of 5 Models:

- Decision Tree (CART)
- Logistic Regression
- Random Forest
- XGBoost
- Gradient Boosting

4.) Logistic Regression:

Despite its name, logistic regression is more of a classification model than a regression model. For situations involving binary and linear classification, logistic regression is a straightforward and more effective approach. It's a classification model that's incredibly simple to implement and performs admirably with linearly separable classes. It is a widely used categorization method in business. Similar to the Adaline and perceptron, the logistic regression model is a statistical technique for binary classification that can also be applied to multiclass classification.

5.) Decision Trees:

A decision tree model's ability to give findings that are simple to grasp in terms of the predictor factors and target variables is one of its advantages. An induced rule set may even be preferable because it describes the splits in the decision tree in terms of simple IF-THEN-ELSE rules that managers may easily comprehend. The findings of a neural net model, on the other hand, might be more predicative, but it's harder to comprehend the outcomes in terms of the predictor factors. The results' accuracy and interpretability are frequently traded off. This trade-off could be affected by the modeling algorithm selected. For some data sets, some algorithms perform better than others.

6.) Random Forest:

An ensemble technique known as RF classifier trains several decision trees concurrently with bootstrapping, aggregation, and bagging. Bootstrapping describes the parallel training of many individual decision trees on various subsets of the training dataset using various subsets of the available characteristics. Bootstrapping makes ensuring that every decision tree in the random forest is distinct, which lowers the RF classifier's total variance. RF classifier exhibits strong generalization since it aggregates individual trees' decisions into the final determination. In terms of accuracy and without overfitting problems, RF classifier typically outperforms the majority of other classification techniques. RF classifier doesn't require feature scaling, just as DT classifier. RF classifier is more resistant to training dataset noise and training sample selection than DT classifier is. In comparison to DT classifier, RF classifier is more difficult to read but simpler to tune the hyperparameter.

7.) Gradient Boosting:

By combining several weak prediction models, gradient boosting develops prediction-based models. Parameters with weak hypotheses perform only marginally better than selections picked at random. When employed with the right cost functions, boosting can be an optimization algorithm, according to American statistician Leo Bierman. Iteratively selecting weak hypotheses or a function with a somewhat negative gradient is how one optimizes cost functions. Numerous improvements to the gradient boosting method have been made in order to optimize the cost functions.

8.) XGBoost:

Extreme Gradient Boosting (XGBoost) is a distributed, scalable gradient-boosted decision tree (GBDT) machine learning framework. The top machine learning library for regression, classification, and ranking issues, it offers parallel tree boosting. XGBoost is a scalable and extremely accurate gradient boosting solution that pushes the limits of computing power for boosted tree algorithms. It was created primarily to enhance the performance and computational speed of machine learning models. Trees are constructed using XGBoost in parallel as opposed to GBDT's sequential method. It employs a level-wise approach, scanning over gradient values and assessing the quality of splits at each potential split in the training set using these partial sums.

VIII. Results:

Logistic Regression

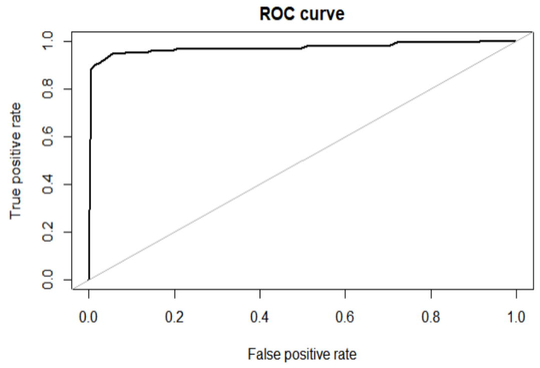


Fig 10.
(Source: Own Work)

Area under the curve (AUC): 0.971

Decision Trees

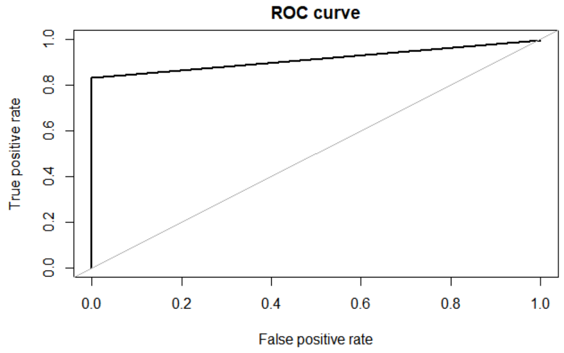


Fig 11.
(Source: Own Work)

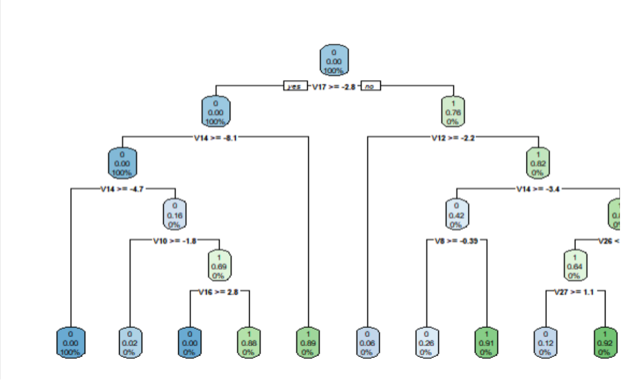


Fig 12.
(Source: Own Work)

AUC Under the Curve: 0.912

Random Forest:

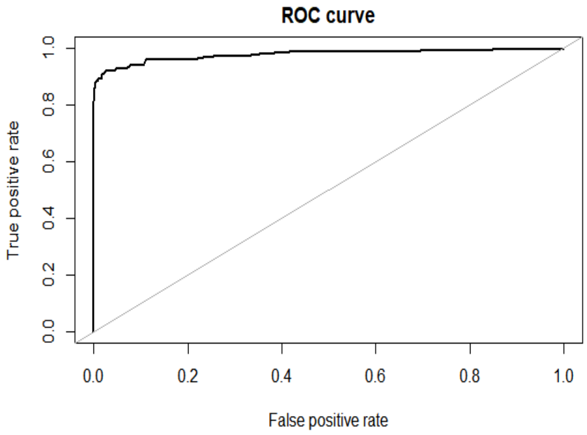


Fig 13.
(Source: Own Work)
Area Under the Curve: 0.977

Gradient Boosting:

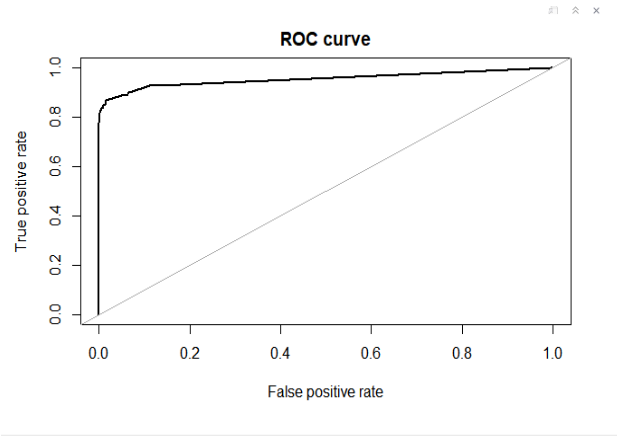


Fig 14.
(Source: Own Work)
Area under the Curve: 0.955

XGBoost:

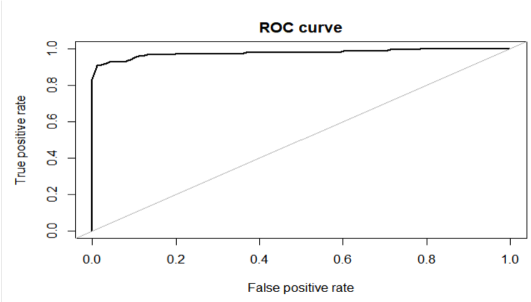


Fig 15.
(Source: Own Work)
AUC Under the Curve: 0.977

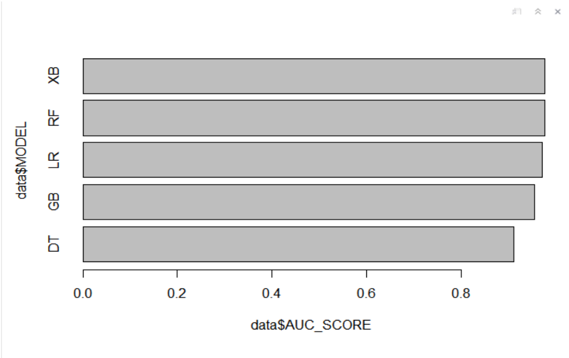


Fig 16.
(Source: Own Work)

IX. Discussion:

Overall, ensemble approaches produced the best prediction performances. Logistic Regression and balanced random forest training times could have been slightly faster with the use of unbalanced learning approaches.

In the majority of Project, XGBoost showed the best performance, demonstrating its resilience to data imbalance conditions across all performance criteria. The most likely explanation is that the minority class naturally receives higher weight in the residuals, operating as a cost-sensitive method.

X. Conclusion:

To create a Reproducible credit card fraud detection (CCFD) model using machine learning. In order to develop this model. We also presented the performance curves for each model and learned how to distinguish fraudulent transactions from other forms of data by analyzing and visualizing data.

The field of credit card detection is exciting. There is still room for more investigation in this field. It is highly challenging to design new strategies because few authors who have worked in this field have released little to no information on the datasets utilized, the characteristics used, and the outcomes of their investigations. In addition, a lot of authors used an unbalanced dataset.

Credit card detection methods employed ML methods when it was surveyed for this paper. However, several of them produced results with low FP rates, False Negative.

rates, and classification accuracy. The lack of a good and efficient feature selection and parameter optimization technique is probably to blame for this. Future research should concentrate on developing classification models that can manage variables with various misclassification costs. The effectiveness of solutions for credit card detection will probably improve as a result.

XI. Future of Credit Card Fraud Detection:

ML algorithms have been shown to significantly improve the efficiency of fraud detection systems, and assist fraud investigators in detecting fraudulent transactions.ML for CCFD has become an active research field.

Although we were unable to achieve our original aim of 100% accuracy in fraud detection, we did manage to develop a system that can, given enough time and data, come very near to it. Just like any

There is some space for improvement with this project. Due to the nature of the project, it is possible to integrate many algorithms as modules and combine their outputs to improve the final result's accuracy.

More algorithms can be incorporated into this model to further enhance it. The output of these algorithms must, however, follow the same format as that of the others. The modules are simple to add once that criterion is met, as seen in the code. This offers a significant amount of modularity and Adding flexibility to the project

The dataset contains more opportunities for development. As was previously shown, as dataset size grows, algorithmic precision also grows. Consequently, more data will undoubtedly improve the model's ability to identify frauds and decrease the number of false positives. However, the banks themselves must formally support this.

XII. References:

1. Andrea Dal Pozzolo, Olivier Caelen, Reid A. Johnson and Gianluca Bontempi. Calibrating Probability with Undersampling for Unbalanced Classification. In Symposium on Computational Intelligence and Data Mining (CIDM), IEEE, 2015
2. Dal Pozzolo, Andrea; Caelen, Olivier; Le Borgne, Yann-Ael; Waterschoot, Serge; Bontempi, Gianluca. Learned lessons in

credit card fraud detection from a practitioner perspective, Expert systems with applications, IEEE,2020

3. Dal Pozzolo, Andrea; Boracchi, Giacomo; Caelen, Olivier; Alippi, Cesare; Bontempi, Gianluca. Credit card fraud detection: a realistic modeling and a novel learning strategy, IEEE transactions on neural networks and learning systems,2018, IEEE
4. Dal Pozzolo, Andrea Adaptive Machine learning for credit card fraud detection ULB MLG PhD thesis (supervised by G. Bontempi)
5. Carcillo, Fabrizio; Dal Pozzolo, Andrea; Le Borgne, Yann-Aël; Caelen, Olivier; Mazzer, Yannis; Bontempi, Gianluca. Scarff: a scalable framework for streaming credit card fraud detection with Spark, Information fusion,Elsevier, IEEE 2016
6. Carcillo, Fabrizio; Le Borgne, Yann-Aël; Caelen, Olivier; Bontempi, Gianluca. Streaming active learning strategies for real-life credit card fraud detection: assessment and visualization, International Journal of Data Science and Analytics, Springer International Publishing
7. Bertrand Lebichot, Yann-Aël Le Borgne, Liyun He, Frederic Oblé, Gianluca Bontempi Deep-Learning Domain Adaptation Techniques for Credit Cards Fraud Detection, INNSBDDL 2019: Recent Advances in Big Data and Deep Learning,
8. Fabrizio Carcillo, Yann-Aël Le Borgne, Olivier Caelen, Frederic Oblé, Gianluca Bontempi Combining Unsupervised and Supervised Learning in Credit Card Fraud Detection Information Sciences, 2019
9. Bertrand Lebichot, Gianmarco Paldino, Wissam Siblini, Liyun He, Frederic Oblé, Gianluca Bontempi Incremental learning strategies for credit cards fraud detection, IInternational Journal of Data Science and Analytics
10. Andrea Dal Pozzolo, Giacomo Boracchi, Olivier Caelen, Cesare Alippi, and Gianluca Bontempi. Credit card fraud detection: a realistic modeling and a novel learning strategy. *IEEE transactions on neural networks and learning systems*
11. *Bertrand Lebichot, Théo Verhelst, Yann-Aël Le Borgne, Liyun He-Guelton, Frédéric Oblé, and Gianluca Bontempi. Transfer learning strategies for credit card fraud detection. IEEE access*
12. *Sara Makki, Zainab Assaghir, Yehia Taher, Rafiqul Haque, Mohand-Said Hacid, and Hassan Zeineddine. An experimental study with imbalanced classification approaches for credit card fraud detection. IEEE Access*
13. *Rimpal R Popat and Jayesh Chaudhary. A survey on credit card fraud detection using machine learning. In 2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI) . IEEE, 2018.*
14. *Janvier Omar Sinayobye, Fred Kiwanuka, and Swaib Kaawaase Kyanda. A state-of-the-art review of machine learning techniques for fraud detection research. In 2018 IEEE/ACM Symposium on Software Engineering in Africa (SEiA), 11–19. IEEE, 2018.*
15. *Yu Sun, Ke Tang, Zexuan Zhu, and Xin Yao. Concept drift adaptation by exploiting historical knowledge. IEEE transactions on neural networks and learning systems*
16. *Kalyan Veeramachaneni, Ignacio Arnaldo, Vamsi Korrapati, Constantinos Bassias, and Ke Li. Aiˆ 2: training a big data machine to defend. In 2016 IEEE 2nd International Conference on Big Data Security on Cloud (BigDataSecurity), IEEE International Conference on High Performance and Smart Computing (HPSC), and IEEE International Conference on Intelligent Data and Security (IDS), 49–54. IEEE, 2016.*
17. *Ivan Tomek and others. Two modifications of cnn. IEEE Trans. Syst. Man Commun,*
18. Ehramikar S (2000) The enhancement of credit card fraud detection systems using machine learning methodology. Masters, University of Toronto, Canada