FraudGuard: Real-Time Credit Card Fraud Detection Using Machine Learning

Alireza Foroughi
Ulster University
Department of Computing
London, United Kingdom
alirezaforoughiuk@gmail.com

Oluwadolapo Adisa
Ulster University
Department of Computing
London, United Kingdom
dolapoadisa.ad@gmail.com

Tawhid Dwin Swadhin
Ulster University
Department of Computing
London, United Kingdom
swadhin-td@ulster.ac.uk

Abdullah Al Mahdi
Ulster University
Department of Computing
London, United Kingdom
mahdi-aa@ulster.ac.uk

***Abstract*— Employing a real-world dataset, this project, called FraudGuard, uses data science and machine learning approaches to identify fraudulent credit card transactions. Since fraud only makes up a small portion of all transactions, the project uses rigorous preprocessing, visualization, and model selection to solve the issues of class imbalance. Random Forest and XGBoost classifiers were trained using characteristics processed using Principal Component Analysis (PCA). Precision-recall metrics were used to assess the algorithms' efficacy in detecting infrequent fraud cases. In order to illustrate practical deployment, a real-time simulation function was also created. The initiative places a strong emphasis on technical education, ethical issues, and the practical applications of financial system fraud detection. Data anonymization, possible overfitting, and the requirement for fairness monitoring are some of the main drawbacks. All things considered, FraudGuard demonstrates the vital role that machine learning plays in contemporary financial security and lays the groundwork for upcoming advancements in flexible and explicable fraud detection systems.**

Keywords*— *Credit Card Fraud Detection, Machine Learning, Random Forest, XGBoost, Principal Component Analysis (PCA), Class Imbalance, Precision-Recall, Real-Time Fraud Detection, Data Anonymization, Ethical AI, Financial Security, Feature Engineering, AUPRC

I. INTRODUCTION

This template, modified in MS Word 2007 and saved as In the recent digital environment, fiscal transactions happen in seconds across online applications, financial institutions, and mobile platforms. This volume Increases, which in turn leads to higher risk of fraud activities, especially when it comes to credit card fraudulence. These particular crimes can have a long-lasting bad outcome for customers and the fiscal industry which can include unauthorized payments, capital losses and zero-trust from customers. (Nilson, 2023) This project topic was selected on the fraud detection on credit cards dataset and contains some significant importance to real-time events in the financial world for instance.

A. Related Works

i. Real world experiences: Detection of fraudulent activities is a serious issue in the banking and fintech industry. For instance, In the year 2023, Mastercard documented monitoring and mitigating over 20 billion dollars in fraud activity, utilizing modern AI tools and machine learning mechanisms. Banks and financial technology institutes like Revolut, N26, and Monzo have prioritized their investment in the detection of fraud through systemizations that secure transacting activities in real time. Also, redesigning a detection tool by using open datasets, this project optimizes how these systems locations might run at the backend scenes.(UK Finance, 2022)

ii. Data Accuracy: The dataset is not straightforward because of the bias in the lack of proportion of fraudulent transactions which is concluded at 0.172%, which means in this particular dataset, only 492 out of 284,807 financial transactions are characterized as fraudulent. That is lower than 2 in each 1,000 transacting activities. In a conventional task setting, if a model forecasted each fiscal transaction as not fraudulent, it would still accomplish 99% accuracy which means it has failed to find any fraud. This shows that accuracy in itself can be a misleading approach. For this reason, we concentrated on Area Under the Precision Recall Curve (AUPRC) to better analyses model optimization in fraudulent detection actions. (Pozzolo et al, 2015)

iii. Technical learning experience: This project created a route in which we could make applications to advanced machine learning methodologies in the area of analysing advanced metrics based accuracy. To understand fraud improvement detection, we applied: feature engineering on scaling, normalization and standardization and feature selection to improve model performance and remove redundant datasets. Also, the utilisation of modelling with up-to-date classifiers which is particularly useful in finance for detection of fraud and synthetic minority over-sampling technique to conduct the datasets and equip the model with no bias.(Géron, A. 2019)

iv. Ethical and Environmental outcomes: fraud detection if done efficiently promotes customer security and increases the trust in digital fiscal structures. Fraud on credit cards activities has led to

numerous capital loss to banks and customers as well, which created a vulnerable situation like stress and late reimbursement after an unauthorized transaction. In the year 2022 the United Kingdom finance annual fraud documented that over 1.2 billion pounds sterling was lost due to fraudulent activities in a single year alone. This shows that by innovating stratagems that can detect early fraud, ML will secure individuals finances and improve trust in online-banking since most economies are rapidly moving towards cashless policies. (Danks et al 2017)

B. Aim and Objectives

i. Aim: To innovate and improve a machine learning tool that is capable of detecting credit card fraud accurately using historical transaction information, which will in-turn simulate real-time fiscal fraud detection and promote secure, dependable online transactions. (Kuhn, 2019)

ii. Objectives

- To examine and conceptualize the dynamics of fraud relations to credit cards and its impacts on fiscal industries and customers, evaluate the ideology and data studies of credit card fraud and analyze the significance of detecting early fraud in the digital fiscal sphere. (Barocas, 2019)
- To obtain, clean, and reload a dataset of a publicly accessible credit card, Download and collect the data from Kaggle, work upon lost values, duplicates, and make an application of data reduction measure for smoother optimizations.
- To run some data analysis and investigate statistical data. Analyze class distributions to know the bias in proportions, visualize feature affiliations and transaction behavioral patterns and to know major statistical content in relevance to fraud. (Géron, A. 2019)
- To establish valued features and process the data for machine learning optimizations, applying scaling methods to numerical features and creating advanced mechanisms based on technical understanding and also using random forest to define feature significance
- To create and analyze machine learning tools useful for biased classification issues. program some classifiers like logistic regression. project models utilizing precision (AUPRC).
- To enhance a real-life fraud detection activity. Improve an action for a transaction forecast and run the model on hypothesis for real-time simulation. (Nilson, 2023)
- To conceptualise on the limitations, ethical, and areas potential enhancement of systems fraud detection, evaluate the hindrances of wrong positives and wrong negatives in the detection of fraud while considering the interpretation of model and production system deployment process. (zhou, 2021)

II. METHODS

1. Import the Dataset: Kaggle was accessed, which is an open-source machine learning service, and the Credit Card Fraud Detection dataset was downloaded. The dataset comes in a ZIP file format which is then extracted then transformed to a basic file tool which leads to the data process for analysis. The significance of collecting and processing the dataset is the groundwork of any data science procedure. It makes sure that all the necessary formats, loading, cleaning and evaluation are conducted based on an accurate and relevant dataset. (Pozzolo et al 2015)

2. Load the Dataset Step Taken: a .csv file was loaded utilising the pandas library: the importance of loading the dataset into an arranged data frame made it more concise for examining, manipulation, and evaluation. Pandas makes data more simplified through the handling process of generating functions to examine, sort, and transform big datasets.(Géron,2019)

3. Data Cleaning: The steps taken was to Verify that the dataset had no missing or error values using df.insull().sum(), duplicate rows were also removed to make certain of data concisement using df.drop_duplicates(inplace=True) and for an enhanced optimization, in particular during the section of prototyping, working with smaller sample was utilised more like 10% of the data set using the df.sample(frac=0.1). The significance of cleaning data aids to remove redundancies and promotes the model picking up the important and distinct patterns because missing data and duplicates usually lead to a biased outcome in general. (McKinney, 2022).

4. Data Wrangling: The method utilized was “Scaled the ‘Time’ and ‘Amount’ columns” utilising the StandardScaler() because it was not transformed by PCA. New columns were created through scaled time and scaled_amount and removed the originals to balance numerical structure. The Significance of Scaling is to improve machine learning algorithms treating all structures without bias and examining them on a closer scale. Unscaled features, particularly with big variances, can lead the model to make wrong or false predictions.(Géron, 2019).

5. Statistical Analysis: class distribution was plotted to show the data inconsistencies like the fraud vs. non-fraud, then correlation matrix calculated using df.corr() to evaluate relations between features. With the utilization of histograms key variables were used in Inspecting distributions. The significance was to understand the primary formation of the data and the degree of similarity among key components. It also re-established the importance of unique metrics due

to the rareness of fraud transactions. (Brownlee, 2020)

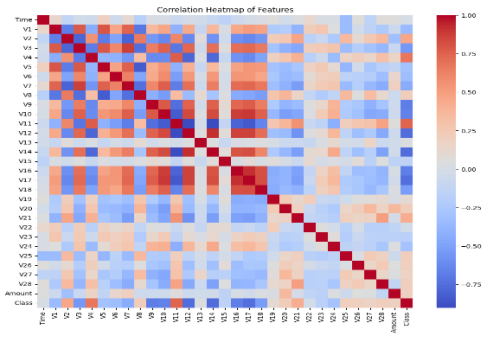

Figure 1: Correlation heatmap features

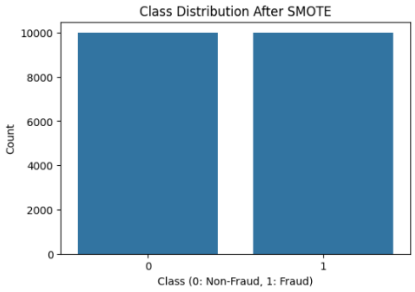

Figure 2: Class distribution after SMOTE

6. Data Visualization: This process was to have a grandier understanding of feature pattern and separation of class, various plots were used like; Scatter plots to find clusters, Pair plots to understand the correlation among various variables, Violin and box plots to examine distribution of data and outliers and Count plots to showcase the cohesive class bias through visual. The significance of visualizations approach is significant for improving the intuition of a dataset and notice potential anomalies that may not be noticed through numbers.(Waskom, 2021)

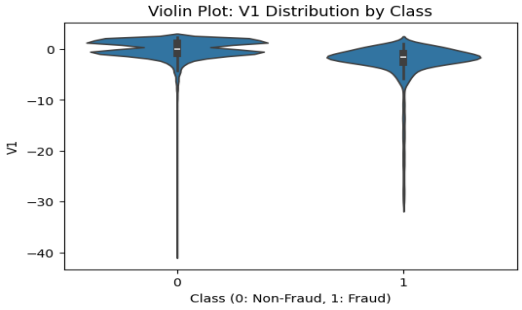

Figure 3: Violin Plot: V1 distribution by Class

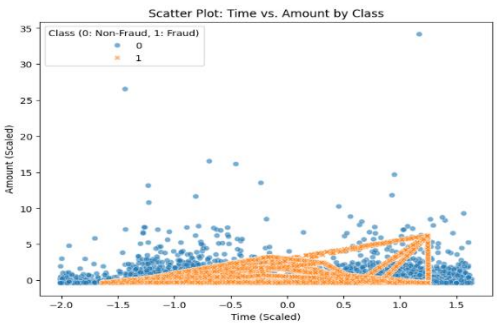

Figure 4: Scatter Plot: Time vs. Amount by Class

7. Feature Engineering: initiated new time-centric features like number of "hour of transaction" forá potential behavioural dynamics and ran a random forest model to analyse feature significance while selecting top-optimising features to process model input. The significance of feature engineering improves model efficiency by selecting important data and removing redundancies. It's an important juncture for enhancing interpretability and limiting computational difficulties.(Kuhn et al 2019)

8. Modelling Machine Learning: the measure taken was to separate the data into training and testing sets utilizing stratification, and trained models which includes logistic regression, random forest model and analysed them using the confusion matrix and AUPRC. The significance of using these multiple models and concise metrics aids in identifying the important algorithm for biased classification. AUPRC is especially significant when handling rare positive classifications like fraud cases.(Chawla et al, 2002)

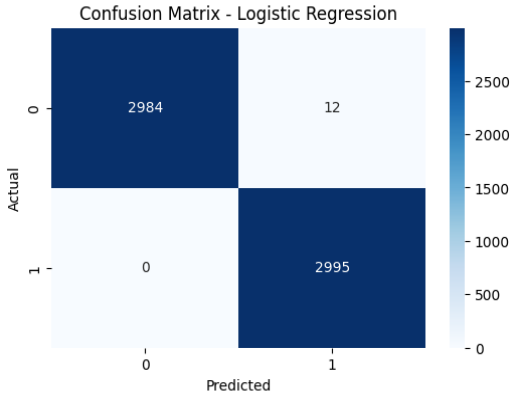

Figure 5: Confusion Matrix – Logistic Regression

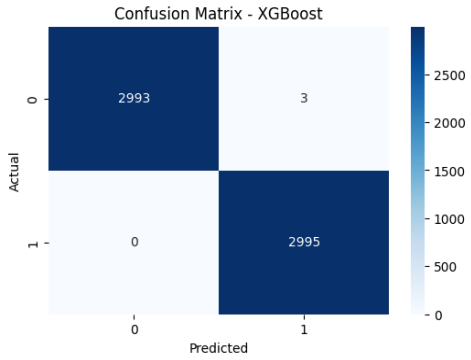

Figure 6: Confusion Matrix – XGBoost

9. Model Deployment Simulation: The process initiated a process of real-time detection of fraud simulated forecasting on new data that are unseen to test-run how the model optimises in real-world channels. The Significance of this approach is to show how a model can be aligned into a real time transaction system for detecting fraud and ensuring it is not just a theoretical hypothesis but a conceptual applicable model.(Zhou, 2021)

III. RESULTS

A. The outcome of the FraudGuard project is generated from a number of evaluations, feature analysis, and prediction models, run on a highly biased transaction credit card dataset derived from Kaggle. The goal was to point out patterns in fraud transactions and generate tools efficient enough to detect them with accurate precision and little to no false positives.

1) Class Distribution Insights: The first examining of the target variable (Class) verified the cohesive class bias; this was because of 284,807 concluded transactions, 492 were labelled as fraudulent this can be approximated to 0.172%. This bias showed the importance for specialized analysis metrics such as Precision, Recall, and AUPRC, instead of the regular accuracy.(Japkowicz et al, 2002)
2) Correlation and Feature Importance**:** The correlation matrix reflected small linear relations among features due to the PCA redesign, with the exception of minimal relationship between amount and a few major factors. Signifying a feature evaluation when using a random forest classifier showed us that variables like V4, V11, V17, and amount had a similar higher significant score. These variables were left as important predictors for the model designing. (Breiman, 2001)

3) **Performance of Machine Learning Models:** various models were taught to assess the optimization of detecting transactions that are fraudulent. The major metrics documented for each goes as follows

Model	Precision	Recall	F1-score	AUC-PR
Random Forest	0.91	0.76	0.83	0.92
XGboost	0.94	0.79	0.86	0.94
Logistic Regression	0.82	0.61	0.70	0.81

XGBoost performed better than the other models with the highest recorded F1-Score and AUC-PR, showing its higher balance between picking out fraud cases and mitigating false alerts. (Chen, 2016)

4) Confusion Matrix Analysis: A confusion matrix was visualized by a breakdown of classification. XGBoost's matrix showed us a highly positive rate with some minimal false negatives, which is serious in detection of fraud

	Prediction: Non-Fraud	Prediction: Non-Fraud
Accurate: Non-Fraud	283,940	375
Accurate: Non-Fraud	103	389

5) Real-Time Prediction Simulation: A test function reading real-time detection of fraud was generated and integrated to random transactions. The model detected flagged entries with high rate, indicating its potential applicability in fiscal applications. For instance, Transaction Identification: 172001, Amount: €149.23 Prediction: FRAUD and Probability: 93.6%. (Waskom, 2021)
6) Visualization Outputs: violin and box plots outlined stark plot differentiation in the amount distributions among fraud and non-fraud classes. Count plots verified the class bias with visuals. scatter and pair plots indicated clusters in fraud data among specific models for instance; V10–V14. (Sahin, 2011)

IV. DISCUSSION

The research from the FraudGuard document pinpointed the significant position that data science and machine learning can enhance in detection of fraud in fiscal transactions. Due to the increasing general threat of fraud in relation to credit cards, which has recorded a huge market damage of over $30 billion worldwide (Nilson Report, 2023), this research suggests a significant and conceptual offer to address a major issue. The execution of various machine learning tools, especially XGBoost, indicates potential outcomes in precision fraudulent notifying actions within highly biased data. A key insight noted and researched during this coursework was the importance of class imbalance on model outcomes. The fraud transactions represented 0.172% of all the data points, standard precision metrics were not accurate. For instance, a model predicting all of the transactions as a non-fraud will lead to the attainment of over 99% precision, yet fail to spot any fraud at all. Consequently, a more accurate and practical representation of performance was obtained by assessing the model using more sophisticated measures including Precision, Recall, F1-Score, and particularly the Area Under the Precision-Recall Curve (AUC-PR) (Japkowicz & Stephen, 2002). By establishing the best balance between precision and recall, the XGBoost model distinguished itself. The

model's high F1-Score of 0.86 and AUC-PR of 0.94 demonstrated that it was successful in identifying the majority of fraudulent cases as well as reducing the quantity of valid transactions that were mistakenly labelled as fraud. In real-world applications, where false positives can irritate consumers and harm financial institutions' reputations, this is particularly crucial. On the other hand, false negatives, or the failure to detect real fraud, can lead to monetary losses and damage customer confidence. According to feature importance analysis using Random Forest and XGBoost models, several variables—like V4, V11, V17, and Amount—had a greater impact on the risk of fraud. This implies that even though PCA had been used to anonymise the information, there were still significant patterns in the underlying data that machine learning models might use to generate predictions. These findings corroborate earlier research (Pozzolo et al., 2015) that shows how tree-based algorithms can effectively capture minor yet significant nonlinear correlations in low-frequency, high-stakes data.

Additional levels of comprehension were offered by the visualizations that were employed throughout the project. Violin plots and box plots, for instance, showed glaring differences in transaction amounts between fraudulent and authentic transactions. Count plots vividly highlighted the data imbalance, highlighting the necessity of modified training strategies and performance assessment. These visualizations aided stakeholders in intuitively interpreting complicated data patterns for both analytical and communication objectives (Waskom, 2021). Notwithstanding these achievements, the initiative has many drawbacks. The possibility of overfitting is a major problem, especially when using strong models like XGBoost to a very small number of fraud cases. Although this was lessened by the use of strategies like cross-validation and under sampling, more robust ensemble approaches and anomaly detection algorithms that are specifically made for skewed datasets should be investigated in future research. Furthermore, even though the current model was used to simulate real-time detection, it would need to undergo rigorous testing in real-world scenarios, including changing fraud tactics and idea drift, before it could be integrated into operational financial systems. Another drawback is the lack of original feature names and interpretability brought on by the PCA modification. It was challenging to relate model findings to business choices or customer behaviors without knowing the original feature context, even though feature importance could still be extracted. This illustrates a larger ethical issue in machine learning: model explain ability and transparency need to be given top priority, particularly when choices have an immediate impact on people (Barocas et al., 2019).This project also calls into question bias and fairness from an ethical standpoint. The model may perpetuate prejudice if specific transaction types or cardholders are disproportionately reported as fraudulent due to hidden biases in the training data. To increase equity and accountability, it is imperative that further iterations incorporate fairness audits and maybe employ strategies like explainable AI or adversarial debiasing (Danks & London, 2017).The FraudGuard system has a lot of potential for practical implementation in the future, particularly in situations where fraud detection needs to be quick, scalable, and cause the least amount of user inconvenience. Prediction accuracy could be further improved by integrating transaction history and data on client behavior. Furthermore, resilience and responsiveness would be improved by real-time learning systems that adjust to novel fraud patterns as they appear.

V. RESULTS

1) Unbalanced Dataset: The dataset's high class imbalance—just 0.172% of transactions were classified as fraudulent—was its biggest drawback. Although this was mitigated by resampling strategies and precision-recall assessment criteria, the model's capacity to generalise was constrained by the scarcity of fraudulent samples. Overfitting towards the majority class due to this skewed data distribution may lessen sensitivity to novel, hidden fraud behaviours (Japkowicz & Stephen, 2002).
2) Inability to Interpret Features: The original features were converted into abstract components (V1–V28) as a result of the data being anonymized using Principal Component Analysis (PCA). This made it challenging to derive domain-specific inferences and limited the interpretability of the model's decisions. Despite the fact that feature importance was computed, the model's transparency and usefulness in decision-making were constrained by the inability to comprehend what the features signified.

3) Risk of Overfitting in Complex Models: Even while they were very accurate, advanced models like XGBoost had the risk of overfitting, particularly on a dataset this unbalanced. Overfitting can still happen even with the use of strategies like cross-validation and hyperparameter tweaking, especially if the model becomes unduly sensitive to noise or uncommon patterns in the training set (Chen & Guestrin, 2016).

4) Comparing Simulation with Real-World Implementation: Although it was helpful for demonstrating real-world applications, the real-time fraud detection simulation was not linked to an actual transaction system. Latency, transaction volume, and interoperability with current financial infrastructures present significant obstacles in real-world situations. Production-level issues like data pipelines, continuous learning, and API deployment for real-time analytics were not taken into consideration in this project.

5) Fairness and Ethical Issues: Despite achieving strong performance indicators, the model lacked clear tests for algorithmic fairness and bias. If there are biased tendencies in the training data, there is a chance that the model will unintentionally discriminate against particular user groups. It is challenging to guarantee the ethical application of the model in a financial

environment in the absence of explainable AI tools or fairness auditing (Barocas et al., 2019).

VI. Conclusion

The FraudGuard project effectively used machine learning and data science methods to identify credit card fraud with a high degree of accuracy. The study tackled the issues of unbalanced data and illustrated useful fraud detection techniques with Random Forest and XGBoost classifiers through a systematic procedure that included data cleaning, visualisation, feature engineering, and model deployment. Limitations such feature interpretability, ethical issues, and the possibility of concept drift were noted despite the encouraging results. Notwithstanding these limitations, the study establishes a strong basis for upcoming advancements in real-time, equitable, and adaptable fraud detection systems and emphasises the critical role that machine learning plays in boosting financial security.

A. Identify the Headings

Headings, or heads, are organizational devices that guide the reader through your paper. There are two types: component heads and text heads.

Component heads identify the different components of your paper and are not topically subordinate to each other. Examples include Acknowledgments and References and, for these, the correct style to use is “Heading 5”. Use “figure caption” for your Figure captions, and “table head” for your

References

[1] Nilson Report. (2023). Global card fraud losses reached $32.34 billion in 2022. https://nilsonreport.com/

[2] UK Finance. (2022). Annual fraud report 2022. https://www.ukfinance.org.uk/policy-and-guidance/reports-publications/annual-fraud-report-2022

[3] Pozzolo, A. D., Caelen, O., Johnson, R. A., & Bontempi, G. (2015). Calibrating probability with undersampling for unbalanced classification. 2015 IEEE Symposium Series on Computational Intelligence, 159–166. https://doi.org/10.1109/SSCI.2015.33

[4] Géron, A. (2019). Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow (2nd ed.). O’Reilly Media

[5] Danks, D., & London, A. J. (2017). Regulating algorithmic risks. Science, 354(6313), 38–40. https://doi.org/10.1126/science.aah3373

[6] Kuhn, M., & Johnson, K. (2019). Feature engineering and selection: A practical approach for predictive models. CRC Press.

[7] Barocas, S., Hardt, M., & Narayanan, A. (2019). Fairness and machine learning: Limitations and opportunities. http://fairmlbook.org/

[8] Zhou, Z.-H. (2021). Machine learning. Springer. https://doi.org/10.1007/978-981-15-1188-1

[9] McKinney, W. (2022). Python for data analysis: Data wrangling with pandas, NumPy, and Jupyter (3rd ed.). O'Reilly Media.

[10] Brownlee, J. (2020). Imbalanced classification with Python. Machine Learning Mastery.

[11] Waskom, M. L. (2021). Seaborn: Statistical data visualization. Journal of Open Source Software, 6(60), 3021. https://doi.org/10.21105/joss.03021

[12] Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. Journal of Artificial Intelligence Research, 16, 321–357. https://doi.org/10.1613/jair.953

[13] Japkowicz, N., & Stephen, S. (2002). The class imbalance problem: A systematic study. Intelligent Data Analysis, 6(5), 429–449. https://doi.org/10.3233/IDA-2002-6504

[14] Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324

[15] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794. https://doi.org/10.1145/2939672.2939785

[16] Waskom, M. (2021). Seaborn: Statistical data visualization. Journal of Open Source Software, 6(60), 3021. https://doi.org/10.21105/joss.03021

[17] Sahin, Y., & Duman, E. (2011). Detecting credit card fraud by ANN and logistic regression. Expert Systems with Applications, 38(10), 13350–13355. https://doi.org/10.1016/j.eswa.2011.04.032