Model for Detecting Fraudulent Transactions in Credit Card Data

Farhan Quadri, Jiayi Ding, Monsoo Kim and Mahesh Nidhruva

W207 – December 10th, 2024

Berkeley
UNIVERSITY OF CALIFORNIA

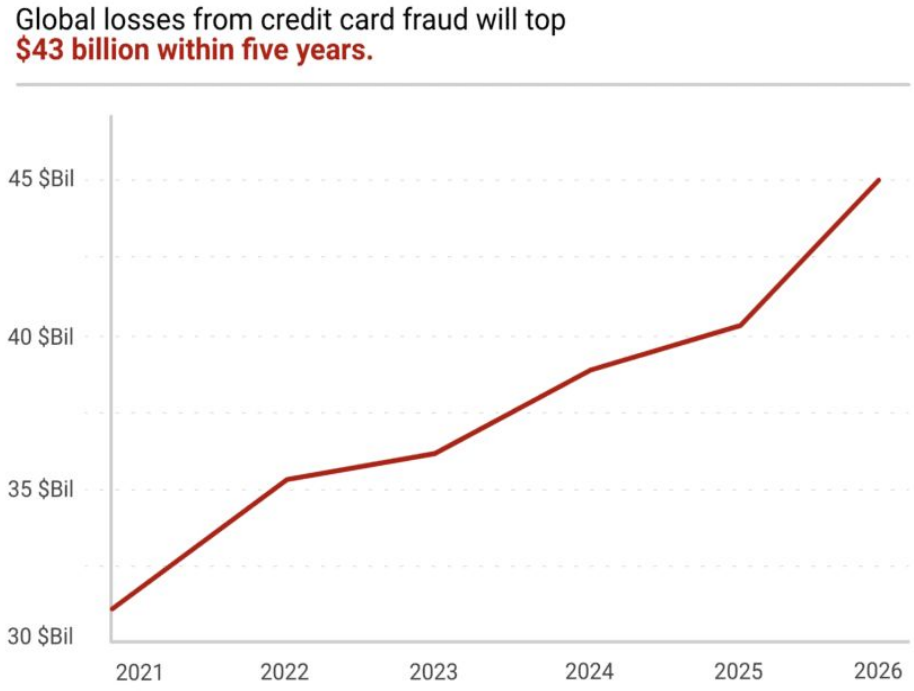
Contents

Background

- **What is Credit Card Fraud?**
  - Unauthorized transactions that result in financial losses for both cardholders and institutions.
  - The rise of online transactions has amplified the need for robust fraud detection systems.

- **Challenges in Detecting Fraud**
  - Imbalanced Datasets: Fraudulent transactions are rare, making them harder to detect.
  - Dynamic Fraud Patterns: Constantly evolving tactics require adaptable detection systems.
  - Real-Time Processing: Systems must operate quickly and accurately to prevent financial losses.

Growing Challenge of Credit Card Fraud



Source:
https://merchantcostconsulting.com/lower-credit-card-processing-fees/credit-card-fraud-statistics/

- 46% of global credit card fraud happens in the US
- Credit card fraud worldwide will reach **$43 billion by 2026**

Motivation

a. Growing need for advanced fraud detection systems to protect financial resources and consumer trust.
b. Enhance the reliability of digital payment platforms.

Introduction

- **Objective**
  - To build a machine learning model that accurately detects fraudulent transactions in real-time
  - Reduce false positives while maintaining high detection rates for fraudulent activity

- **Approach**
  - Exploratory Data Analysis (EDA) to understand patterns in fraudulent transactions.
  - Feature engineering and selection to optimize the model's performance.
  - Build and compare multiple machine learning models, including Logistic Regression, Random Forest, and Neural Networks.

Dataset Overview and Cleaning

- **Source**
  Kaggle Credit Card Fraud Detection Dataset (2023)

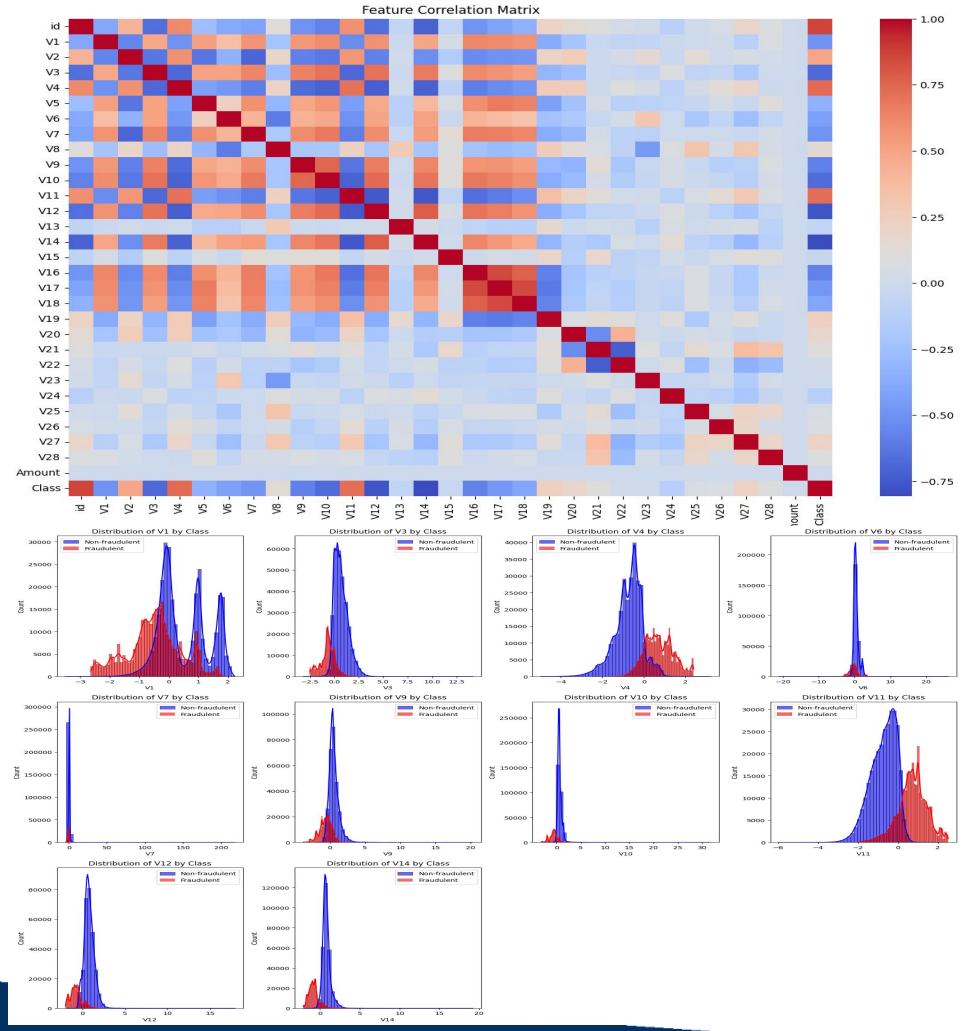
- **Size**
  568,630 transactions with 31 features

- **Key Features**
  - V1–V28: Anonymized transaction attributes
  - Amount: Transaction value
  - Class: Target variable indicating fraud (1) or legitimate (0)

- **Data Cleaning**
  - Checked for duplicates and missing values – did not find any

Exploratory Data Analysis (EDA) Summary

- **Correlation Analysis**
  - Heatmap plot allows us to analyze correlation of each variable
  - The features V1, V3, V4, V6, V7, V9, V10, V11, V12, and V14 exhibit significant correlations with the Class variable.
  - Features like V12 and V14 show strong negative correlation with fraud class
- **Distribution Insights**
  - Features like V1, V3, and V11 demonstrate distinct patterns for fraudulent cases.
  - V10 and V7 show overlaps, reducing their utility.



Berkeley
UNIVERSITY OF CALIFORNIA

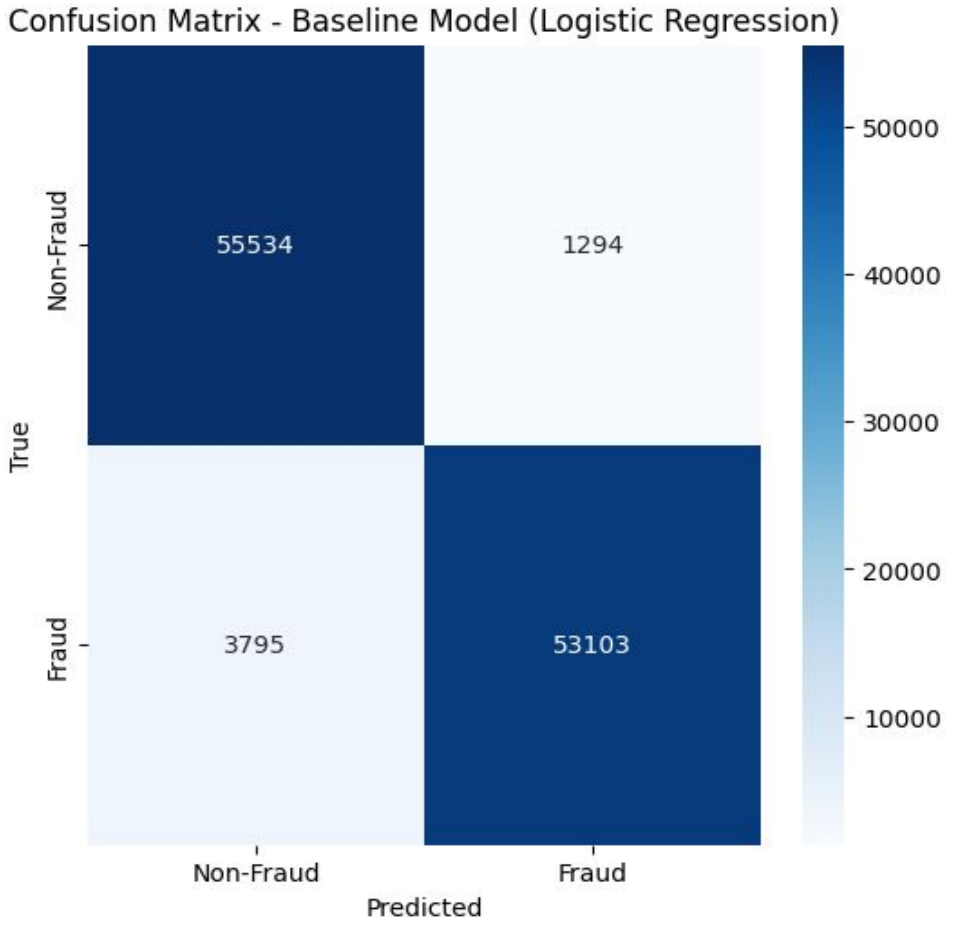
Feature Selection and Data Standardization

- **Selected Features**
  - Features V1, V3, V4, and V11 were particularly found to be important for identifying fraudulent activity, as the the fraudulent transactions (in red) often have different distributions or distinct regions within the histograms
  - Features like V12 and V14 show strong negative correlation with fraud class
- **Data Standardization**
  - Features were not uniformly scaled – Some features exhibit a wider range of values, while others show more concentrated distributions, which can lead to inconsistencies in model performance
  - Therefore, standardization was applied to ensure all features are on the same scale
- **Data Splitting**
  - The dataset was partitioned into three subsets: training (60%), validation (20%), and test sets (20%)

Model Experimentation - Logistic Regression

Logistic regression model was built using features V1, V3, V4, V11, V12, and V14
This model achieves a classification accuracy of 95%, demonstrating strong performance.

Data Subset	Accuracy
Train	0.95598
Test	0.95642
Validation	0.95525

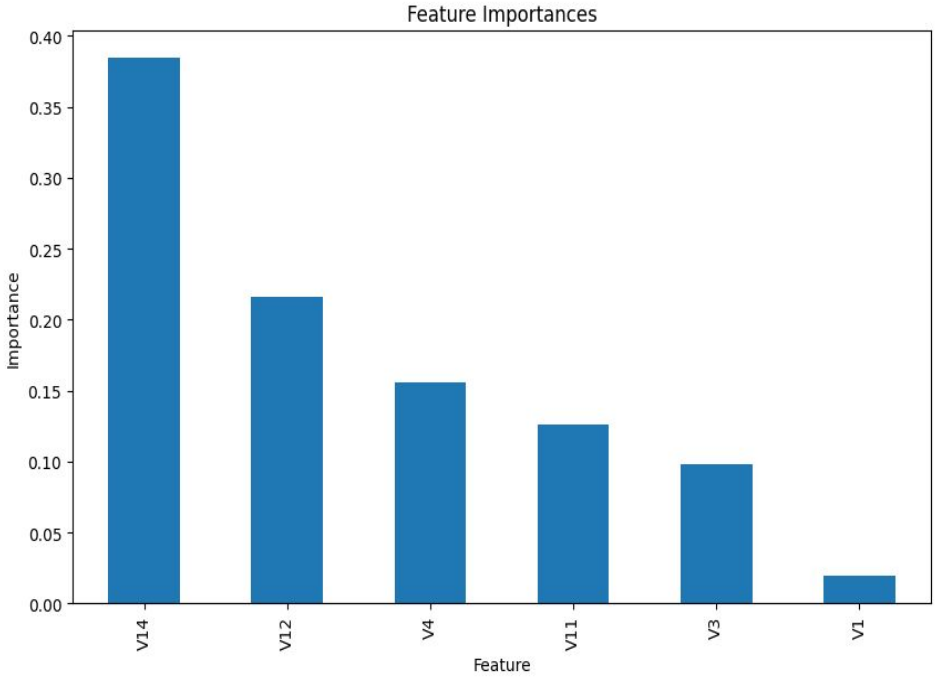


Precision: 0.98 | Recall of 0.93 | F1-score: 0.95

Berkeley
UNIVERSITY OF CALIFORNIA

Model Experimentation - Random Forest

- Model achieves an impressive accuracy of 98%
- V14 is the most impactful features where importance value nearly 0.37
- Model demonstrates excellent performance on the validation data, achieving near-perfect precision, recall, and F1-scores for both classes. It effectively identifies most instances and performs consistently well across both classes.

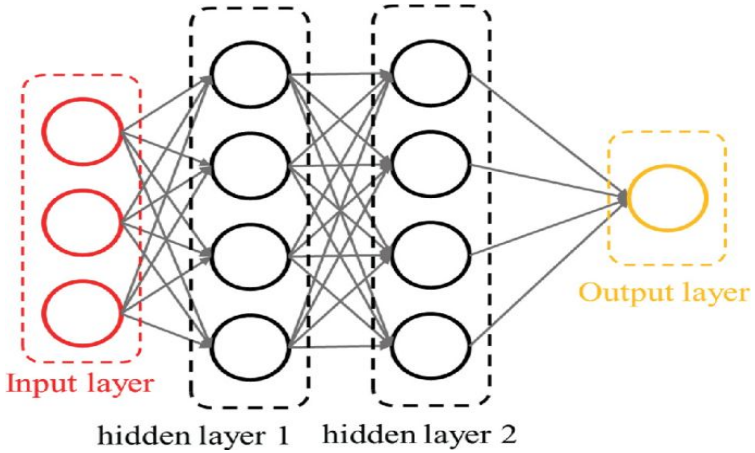


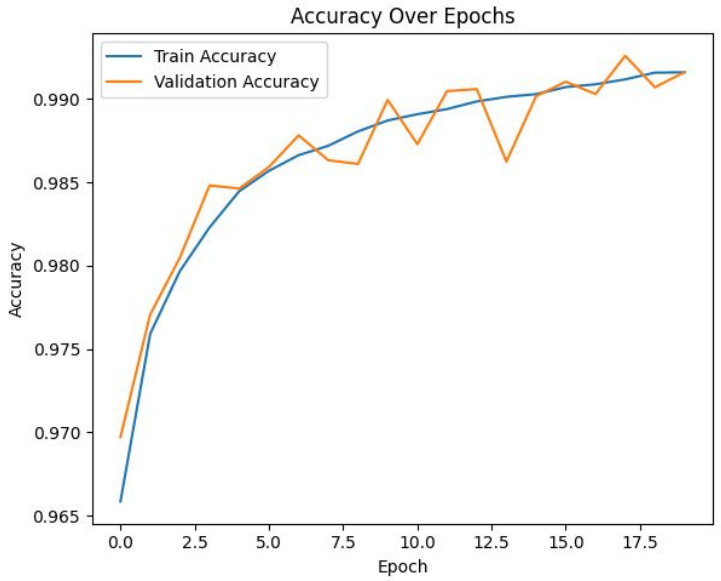
Data Subset	Accuracy
Train	0.98128
Test	0.98068
Validation	0.98012

Classification Report (Validation Data):

	precision	recall	f1-score	support
0	0.99	0.99	0.99	56828
1	0.99	0.99	0.99	56898
accuracy			0.99	113726
macro avg	0.99	0.99	0.99	113726
weighted avg	0.99	0.99	0.99	113726

Model Experimentation - Multi-Layer Neural Network

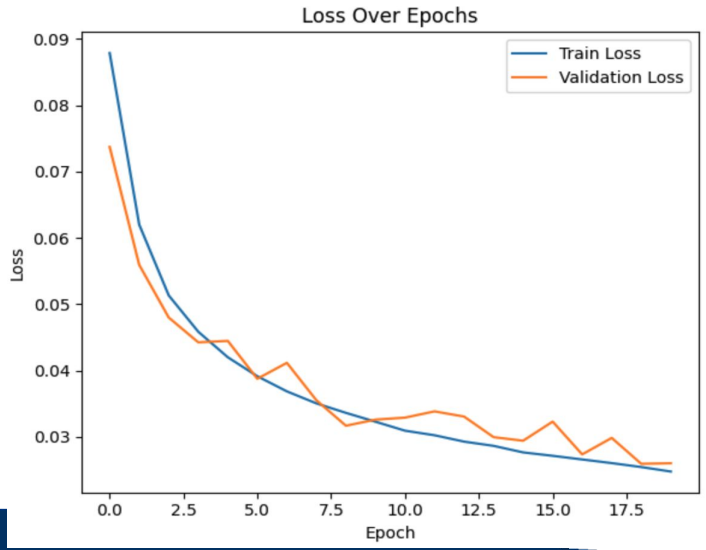




Hidden Layer 1: A fully connected layer with 128 neurons and ReLU activation that processes the input features to extract higher-level representations

Hidden Layer 2: Another dense layer with 48 neurons and ReLU activation that further refines the features extracted by the first hidden layer.

Learning rate: 0.001



Results

Model	Training Accuracy	Validation Accuracy	F1-Score	Comments
Logistic Regression (Baseline)	95.6%	95.5%	0.95	Strong performance but limited in capturing non-linear relationships.
Random Forest	98.1%	98.0%	0.98	Balanced across fraud and non-fraud classes
Neural Network (TensorFlow)	99.3%	99.3%	0.99	Indicates high precision and recall

Conclusions

- **Effectiveness of Machine Learning**
  - Advanced models like Random Forest and Neural Networks outperform traditional rule-based systems and baseline logistic regression for fraud detection

- **Real-Time Application**
  - The models demonstrate potential for real-time fraud detection with high accuracy and minimal latency

- **Impact**
  - These models can significantly reduce financial losses and operational costs for financial institutions while improving customer trust

Considerations

- **Fairness and Bias**
  - Bias mitigation achieved through feature scaling and class balancing
  - Models must be evaluated for potential biases in predictions that could disproportionately affect certain users
  - Anonymization of features reduces ethical concerns, but care is needed to ensure fairness
- **Real-World Deployment**
  - Continuous monitoring and updating of models are required to adapt to evolving fraud patterns
  - Consider trade-offs between explainability and performance, especially for high-stakes decisions
- **False Positives vs. False Negatives**
  - Reducing false positives is essential to minimize disruptions for legitimate transactions
  - False negatives, though rare, could lead to significant financial losses

Thank You!

Appendix

References & Resources

- References
  - Nidula Elgiriyewithana. (2023). Credit Card Fraud Detection Dataset 2023 [Data set]. Kaggle. https://doi.org/10.34740/KAGGLE/DSV/6492730
  - https://merchantcostconsulting.com/lower-credit-card-processing-fees/credit-card-fraud-statistics/
- Resources
  - Ran on Google Colab Environment:
    - CPU and RAM:
      - model name: Intel(R) Xeon(R) CPU @ 2.20GHz
      - MemTotal: 13290460 kB
    - No TPU used

Contribution

Team Member	Model Development	Presentation
Farhan Quadri	1. Conducted exploratory data analysis on dataset to find key features for model development. Performed data cleaning and pre-processing on dataset. 2. Developed linear regression model and random forest classifier model. Analyzed model accuracy and performed tests to ensure no over-fitting.	1. Worked on dataset overview, EDA, and feature selection slides.
Jiayi Ding	1. Develop models, including a simple neural network, a multi-layer neural network, and perform hyperparameter tuning to optimize performance. 2. Add detailed descriptions for each section of the codebase and finalize the Colab notebook for clarity and presentation.	1. Fill the model experimentation part of slides
Moonsoo Kim	1. Proposed initial guidance on the project's objectives and model selection, emphasizing the need for a machine learning model that could accurately identify fraudulent transactions based on historical data. 2. Imported and preprocessed the raw dataset, addressing missing values, outliers, and inconsistent formatting to ensure data integrity.	1. Background, Introduction, Results, Conclusions, and Considerations
Mahesh Nidhruva	1. Independently analyzed data to identify features and ran the regression, random and hypertuning 2. Integrated specifically the sections on correlation analysis and data distribution for the key features in the final deliverable notebook	1. EDA, Feature Selection and Data Standardization section 2. Included the NeurIPS checklist and the associated information in relevant sections

NeurIPS Checklist

Checklist Question	Response	Reference
Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?	Yes	Covered under the Objective. Motivation and the Conclusions & Considerations
Have you read the ethics review guidelines and ensured that your paper conforms to them?	Yes	NeurIPS Checklist
Did you discuss any potential negative societal impacts of your work?	Yes	Covered under the “Conclusions and Considerations” section - potential misuse/errors due to the false positives affecting the user experience
Did you describe the limitations of your work?	Yes	Covered under the “Conclusions and Considerations” section

NeurIPS Checklist

Checklist Question	Response	Reference
Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)?	Yes	Code, inference and instructions in the Jupyter notebook submitted
Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)?	Yes	Refer to the EDA, Feature Selection, Data Standardization sections in the notebook and presentation
Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)?	Yes	Error analysis included standard deviation over multiple runs. Plotted confidence intervals for key metrics.
Did you include the amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)?	Yes	Included in the References and Resources section

NeurIPS Checklist

Checklist Question	Response	Reference
If your work uses existing assets, did you cite the creators?	Yes	Included the References section in Appendix
Did you mention the license of the assets?	Yes	Included the References section in Appendix
Did you include any new assets either in the supplemental material or as a URL?	N/A	Kaggle dataset is referred using the URL
Did you discuss whether and how consent was obtained from people whose data you're using/curating?	N/A	The Kaggle dataset leveraged for the modeling was already anonymized
Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content?	N/A	The Kaggle dataset leveraged for the modeling was already anonymized