By Walter Andrés Paz Callizo

Banking and Fraud

What-If?

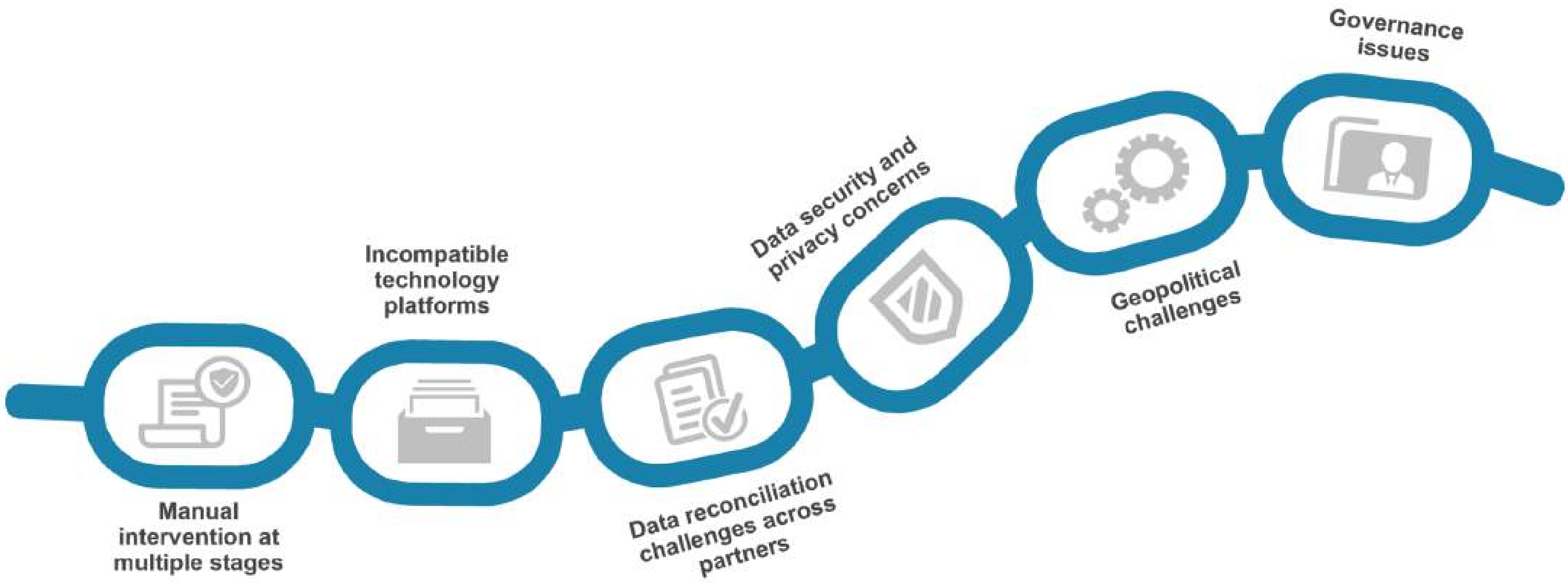
Executive Summary

This paper intends to assess the present security measures regarding credit accounts and manifest the prerequisite of reinforcing cybersecurity. These present protections, which have been established with effectiveness, are now being threatened by the emerging cyber threats that cause a potential risk for financial transactions.

In anticipation of these evils, the proposal of building and putting into practice solid cybersecurity frameworks to be accompanied by a contingency plan within financial institutions can considerably fill the vulnerabilities, making the customers to be more secure and minimizing possible breaches. The direct recommendations in this report promote strategic positioning of the financial ecosystem toward a more resilient and secure environment.

Table of Contents

Table of Contents



Introduction

At this time, the digital age is already ensuring that any financial transactions are kept secure. The credit accounts have their strong level of protection; they will always have to step up because of the nature of threats.

Complete cyber defense has to be resorted to for safeguarding customer transactions and sinking any potential breach. This document describes the reason for the combined upgrades of the cybersecurity system and also a contingency plan to cushion against risks within the six months time frame.

Background Analysis

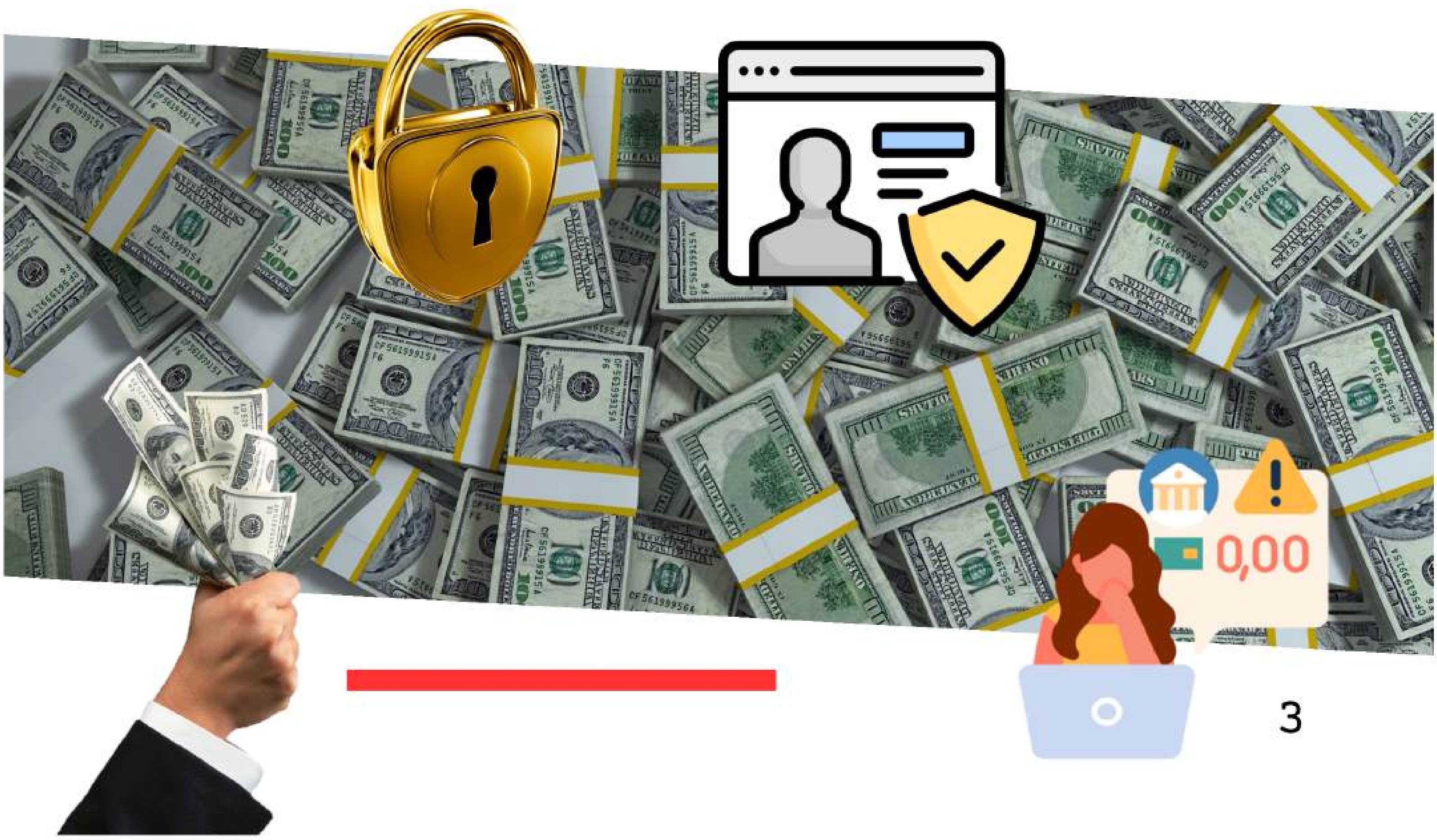
Huge banks globally whether it is in **Canada, United States of America, the European Union or even Central America (CA)** rely on various data formats which cover. Nonetheless, many actions taken for the Data's influence and security networks becomes crucial for the lead's loyalty programs and the integrity of their personal information (HYPR, 2024).

Fundamentally, the break through on **Credit Cards, Accounts, Online Services** and more is the basis for the benefits of any power and the liberty of every customer. Also, the right choices made as a company protecting the resources of every member of banks while maintain the insurance of a responsible institution (Western University (WU), 2024).

Background Analysis

Developing fraud analysis systems can help banking institutions prepare for potential crises by ensuring data integrity and supporting the modeling and expansion of products and services.

The current market in Banking and Cybersecurity is giving out the necessity for the projections of Growth in the market from 2024 to 2032 from **USD 193.73 billion in 2024 to USD 562.72** (Fortune Business Insights, 2024).



Problem Identification

First of all, we all seek problems coming for the sources of the income and funds registered for possible **Credit Frauds.**

No Correlation among the Variables as **Accounts** seem to be **Independent** one from the other.

Correlation will affect the selected models as it might be a Low Variance with no huge dispersion for data but hard for any proper grouping or ordering.

Data Cleaning Issues, understanding the Data Frame, Columns, Rows and even the outputs for the further analysis.

Problem Identification

Individual Customer's Cases which differ one from the other.

Data Verification and Money Source.

Data's Impact and Future's Certainty.

Relevance of Industry

The Banking Industry (BI) is the backbone of the nations with the integrity of the data delivering the construction of delicate systems which improve notably the understanding of any breaches with the strategic management for the national income management and the revenue agencies across the entire globe (World Bank (WB), 2025).

Furthermore, the constant inclusion of tariffs and sanctions by various countries can heavily change the policies for the data maintenance not only nationally but by continents.

Purpose of the Investigation

Not only discovering the changes among the various **Banking Accounts, Cheating and Fraudulent Data** inside of Data Frames but the best protection systems for the upgrading of Data Integrity and the considerations for any possible Fraudulent Credits in the Data for the customer's of the banks.

Objectives

Generate a analysis for the Account on Banking and Credit's Fraud through Data Modeling (DM).

Model the results of the analysis for the Credit's and Frauds inside of the Data Frame's Accounts as it is essential to seek for the Data Predictions (DP) based on the new Amount.

Formulate recommendations and conclusions to identify issues previous to a plan against any Crisis and Management of Security.

Scope

Benefits of working with such a prestigious source of information is helpful for the growth of the ideas and Cybersecurity for the many prevention programs which can be done before any potential **Fraud or Crisis Management.**

Empirical

Data Frame Construction & Key Variables:
Account Details: Account ID, user profile, credit history.

Transaction Features: Date, amount, location, transaction type.

Fraud Indicators: Frequency of high-value transactions, unusual geographic spending, card-not-present transactions.

Empirical

Labeling: Transactions categorized as fraudulent (1) or non-fraudulent (0) based on historical fraud reports.

Business

Fraud Trends: High-risk accounts often have rapid and frequent large withdrawals.

Fraudsters use synthetic identities to bypass security checks.

Digital payment fraud is increasing due to e-commerce growth.

Financial Loss Estimation: Fraudulent transactions often result in chargebacks, costing banks millions.

Business

Stolen credit card information is used for unauthorized purchases, leading to disputes.

Business Implications: Increased operational costs due to fraud investigations.
Compliance penalties for failing to detect fraud.

Loss of customer trust and reputational damage.

Methodology

Data Collection & Preprocessing

- Extract transaction data from banking records.
- Clean and normalize the data (handling missing values, duplicates, outliers).
- Feature engineering (creating fraud risk scores, transaction velocity metrics).

Model Development & Training

- Train fraud detection models using labeled transaction data.

Methodology

- Use techniques such as feature selection, hyperparameter tuning, and cross-validation to optimize model performance.

Fraud Detection & Real-time Monitoring

- Deploy models to monitor transactions in real-time.
- Implement an alert system for high-risk transactions.

Evaluation & Continuous Improvement

- Evaluate models using precision, recall, and F1-score to ensure accuracy.

Methodology

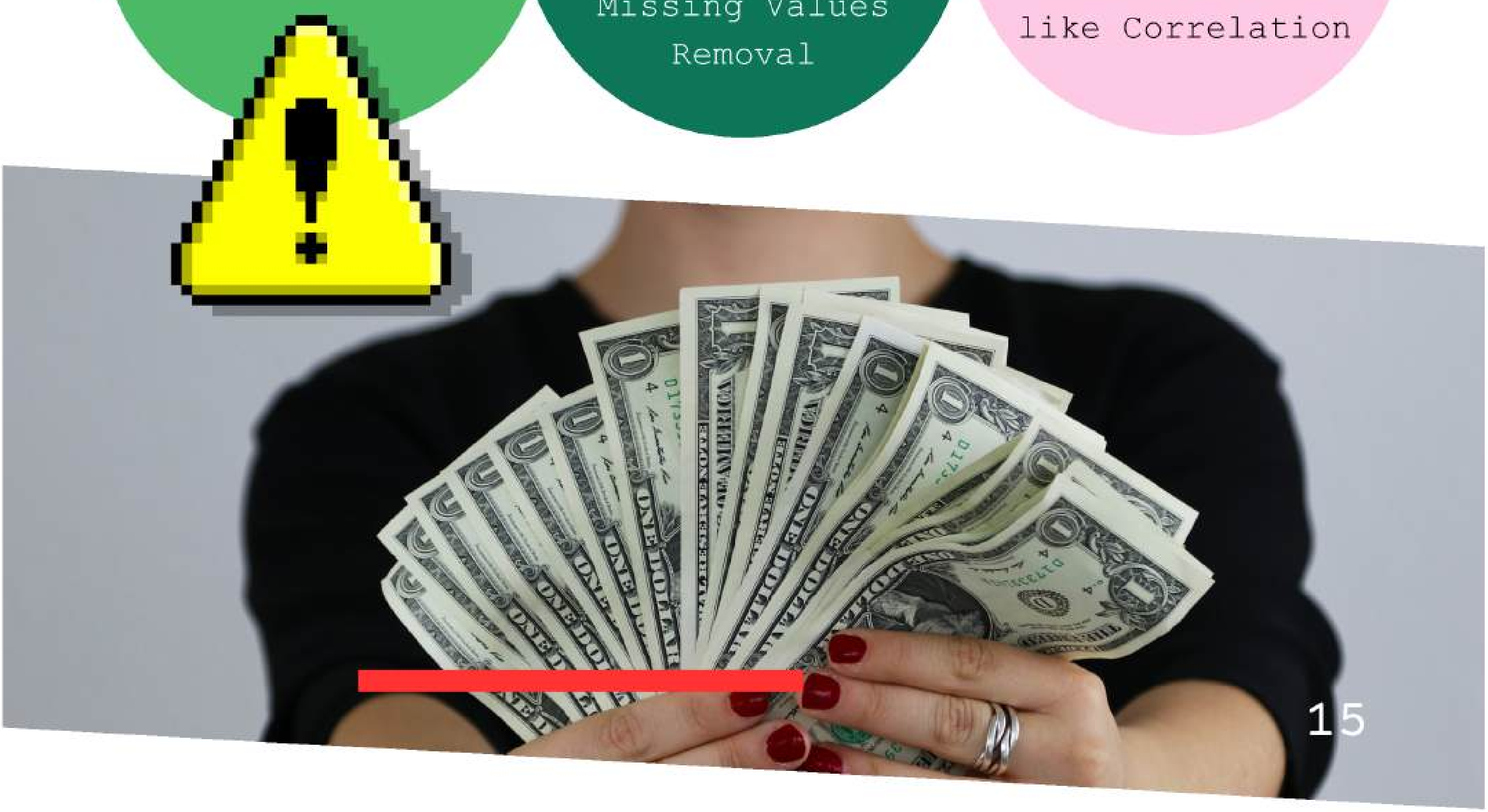
- Update models continuously as fraud patterns evolve.

Tools and Techniques used in analyzing the Data

Central Tendency and Dispersion Analysis

Mean, Mode, Median, Ranges, Quartiles, Shapes, Null Values and Missing Values Removal

Variance, Skewness, Kurtosis, BIAS among other considerations like Correlation



Tools and Techniques used in analyzing the Data

jupyter

ANACONDA

python™

SQL

PyTorch

Keras

Analysis

Correlation was too little inside of the **"Credit_Cards.csv"**, **"Credit_Accounts.csv"** **and "Fraud_Accounts.csv"** files meaning a different approach from dependency like **Linear Regression, K- Means Clustering** (Despite of it helping with Unsupervised Data) and other ideas had to be taken into account between **"Feature's Engineering"** or the creation of Subsets and the **"Neural Network's" Generation** for the dataset's predicted values of new values.

Central Tendencies Measurements

	V1	V2	V3	V4	V5 \
count	77338.000000	77338.000000	77338.000000	77338.000000	77338.000000
mean	-0.254918	-0.031977	0.678014	0.164544	-0.275915
std	1.883504	1.670498	1.395670	1.369866	1.384600
min	-56.407510	-72.715728	-33.680984	-5.172595	-42.147898
25%	-1.016589	-0.597243	0.187865	-0.726170	-0.893539
50%	-0.248318	0.070092	0.765732	0.183869	-0.308220
75%	1.153837	0.723977	1.396222	1.046651	0.261079
max	1.960497	18.902453	4.226108	16.715537	34.801666

	V6	V7	V8	V9	Amount
count	77338.000000	77338.000000	77338.000000	77338.000000	77337.000000
mean	0.096378	-0.114258	0.054293	-0.002388	97.617764
std	1.304456	1.250692	1.231153	1.147883	270.498883
min	-26.160506	-31.764946	-73.216718	-9.283925	0.000000
25%	-0.641995	-0.604823	-0.141463	-0.681631	7.680000
50%	-0.153551	-0.074168	0.068582	-0.084259	26.750000
75%	0.490920	0.417033	0.347917	0.634360	89.000000
max	22.529298	36.677268	20.007208	10.392889	19656.530000

Analysis

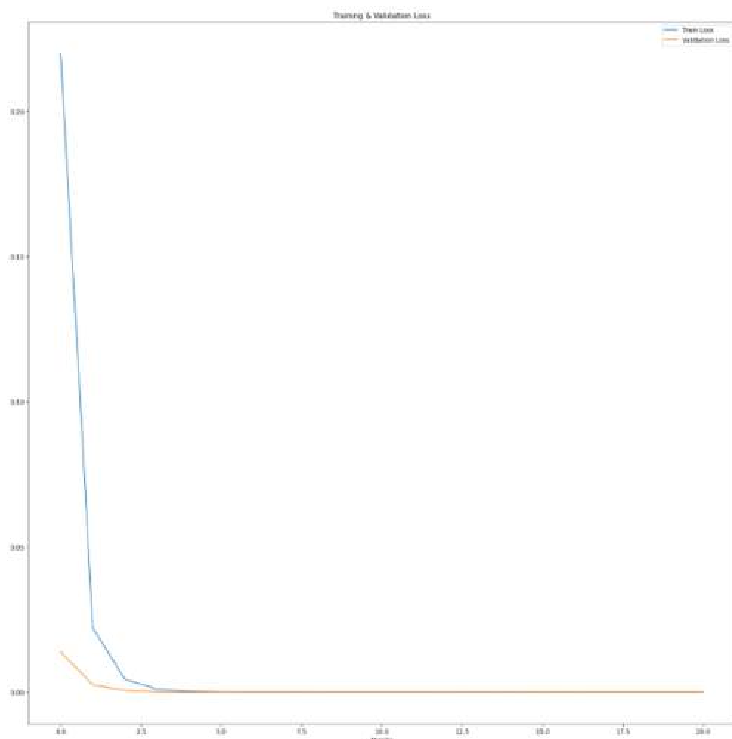
Neural Network Results

Test Loss: 2.370135007367935e-05

Test MAE: 0.0019733756780624 39

	V1	V2	V3	V4	V5	V6	V7 \
0	-1.359807	-0.072781	2.536347	1.378155	-0.338321	0.462388	0.239599
1	1.191857	0.266151	0.166480	0.448154	0.060018	-0.082361	-0.078803
2	-1.358354	-1.340163	1.773209	0.379780	-0.503198	1.800499	0.791461
3	-0.966272	-0.185226	1.792993	-0.863291	-0.010309	1.247203	0.237609
4	-1.158233	0.877737	1.548718	0.403034	-0.407193	0.095921	0.592941

	V8	V9	Amount	Predicted_Amount
0	0.098698	0.363787	149.62	81.938057
1	0.085102	-0.255425	2.69	-0.959219
2	0.247676	-1.514654	378.66	333.653839
3	0.377436	-1.387024	123.50	68.526367
4	-0.270533	0.817739	69.99	52.395428



Learning Rates for NN

Analysis

Neural Network Results

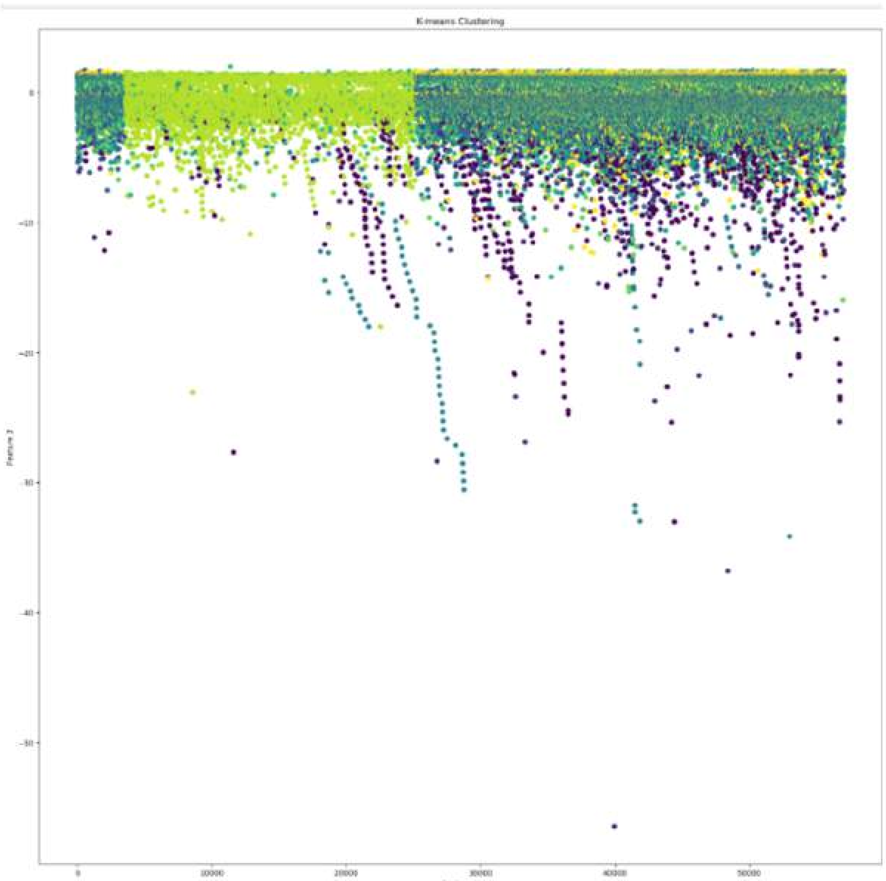
```
Test Loss: 0.00029631680080896139, Test MAE: 0.008996142074465752
2417/2417 ━━━━━━━━━━━━━━━━━━━━ 3s 1ms/step
         V1        V2        V3        V4        V5        V6        V7  \
0 -1.359807 -0.072781  2.536347  1.378155 -0.338321  0.462388  0.239599
1  1.191857  0.266151  0.166480  0.448154  0.060018 -0.082361 -0.078803
2 -1.358354 -1.340163  1.773209  0.379780 -0.503198  1.800499  0.791461
3 -0.966272 -0.185226  1.792993 -0.863291 -0.010309  1.247203  0.237609
4 -1.158233  0.877737  1.548718  0.403034 -0.407193  0.095921  0.592941

         V8        V9  Amount  Predicted_Amount  Predicted_V1
0  0.098698  0.363787  149.62         81.938057     -0.709048
1  0.085102 -0.255425    2.69         -0.959219      1.030538
2  0.247676 -1.514654  378.66        333.653839     -0.526372
3  0.377436 -1.387024  123.50         68.526367     -0.490198
4 -0.270533  0.817739   69.99         52.395428     -0.620062
```

The Neural Network (NN) based on results in some cases like the **"Amounts"** and **"Accounts from V1 to V9"** seem to lower due to the investigation on further accounts and the money being processed with the transactions and retrieval of funds by customers. Nonetheless, they are not correlated due to the accounts being different based on customers.

Analysis

K-Nearest Neighbor

Seeking manners to plotting into barplots for the visualization of the Amount's as an example of expected. First of all, let us remember our **Banking Accounts** have a "Uniform Distribution" which means **K-Nearest Neighbor or Clustering** might not work for this specific scenario as every case and bank account becomes different despite of similar Credit categories.



The Data despite of being clustered seems to have a **"uniformity"** among the data changes and **V's** in accounts for the amounts.

Analysis

Results

When using **Neural Networks (NN)** results were showcased with Data’s multivariate for the **Banking’s Accounts** requiring from the columns where a vast majority of accounts are **Non-Fraudulent** and protected.

```
  TransactionID AccountID  TransactionAmount      TransactionDate  \
0      TX000001   AC00128              14.09  2023-04-11 16:29:14
1      TX000002   AC00455             376.24  2023-06-27 16:44:19
2      TX000003   AC00019             126.29  2023-07-10 18:16:08
3      TX000004   AC00070             184.50  2023-05-05 16:32:11
4      TX000005   AC00411              13.45  2023-10-16 17:51:24

  TransactionType   Location DeviceID       IP Address MerchantID Channel  \
0           Debit  San Diego  D000380   162.198.218.92       M015     ATM
1           Debit    Houston  D000051      13.149.61.4       M052     ATM
2           Debit       Mesa  D000235   215.97.143.157       M009  Online
3           Debit    Raleigh  D000187   200.13.225.150       M002  Online
4          Credit    Atlanta  D000308     65.164.3.100       M091  Online

   CustomerAge CustomerOccupation  TransactionDuration  LoginAttempts  \
0           70             Doctor                   81              1
1           68             Doctor                  141              1
2           19            Student                   56              1
3           26            Student                   25              1
4           26            Student                  198              1

   AccountBalance PreviousTransactionDate  Fraudulent
0         5112.21     2024-11-04 08:08:08       False
1        13758.91     2024-11-04 08:09:35       False
2         1122.35     2024-11-04 08:07:04       False
3         8569.06     2024-11-04 08:09:06       False
4         7429.40     2024-11-04 08:06:39       False
```

```
print(df[df['Fraudulent'] == True])  # Shows only rows where Fraudulent is True
```

```
Empty DataFrame
Columns: [TransactionID, AccountID, TransactionAmount, TransactionDate, TransactionType, Location, DeviceID, IP Address, MerchantID, Channel, CustomerAg
e, CustomerOccupation, TransactionDuration, LoginAttempts, AccountBalance, PreviousTransactionDate, Fraudulent]
Index: []
```

Analysis

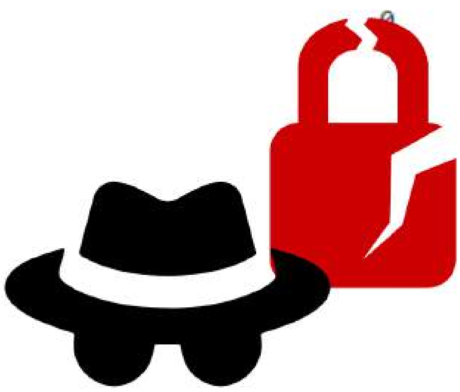
Results

With the same filter at least three Accounts include the **Fraudulent Status** for Credit Transactions over time. Apart from that most of the dataset is properly protected.

```
print(df_II[df_II['Fraudulent'] == True])  # Shows only rows where Fraudulent is True
```

```
        Time         V1         V2         V3         V4         V5        V6  \
46841  42951 -23.712839 -42.172688 -13.320825   9.925019 -13.945538  5.564891
54018  46253 -21.780665 -38.305310 -12.122469   9.752791 -12.880794  4.256017
58465  48401 -36.802320 -63.344698 -20.645794  16.715537 -20.672064  7.694002

              V7        V8        V9  ...        V22        V23       V24  \
46841  15.710644 -2.844253 -1.580725  ...  -6.320710 -11.310338  0.404175
54018  14.785051 -2.818253 -0.667338  ...  -5.619439 -10.547038  0.653249
58465  24.956587 -4.730111 -2.687312  ... -10.933144 -17.173665  1.180700

             V25       V26       V27       V28    Amount  Class  Fraudulent
46841  -4.547278 -1.577118 -2.357385  2.253662  12910.93    0.0        True
54018  -4.232409 -0.480459 -2.257913  2.082488  11898.09    0.0        True
58465  -7.025783 -2.534330 -3.602479  3.450224  1[illegible]    0.0        True

[3 rows x 32 columns]
```


```
       V24  \
.8  0.404175
038  0.653249
3665  1.180700

lass  Fraudulent
0.0        True
0.0        True
           True
```

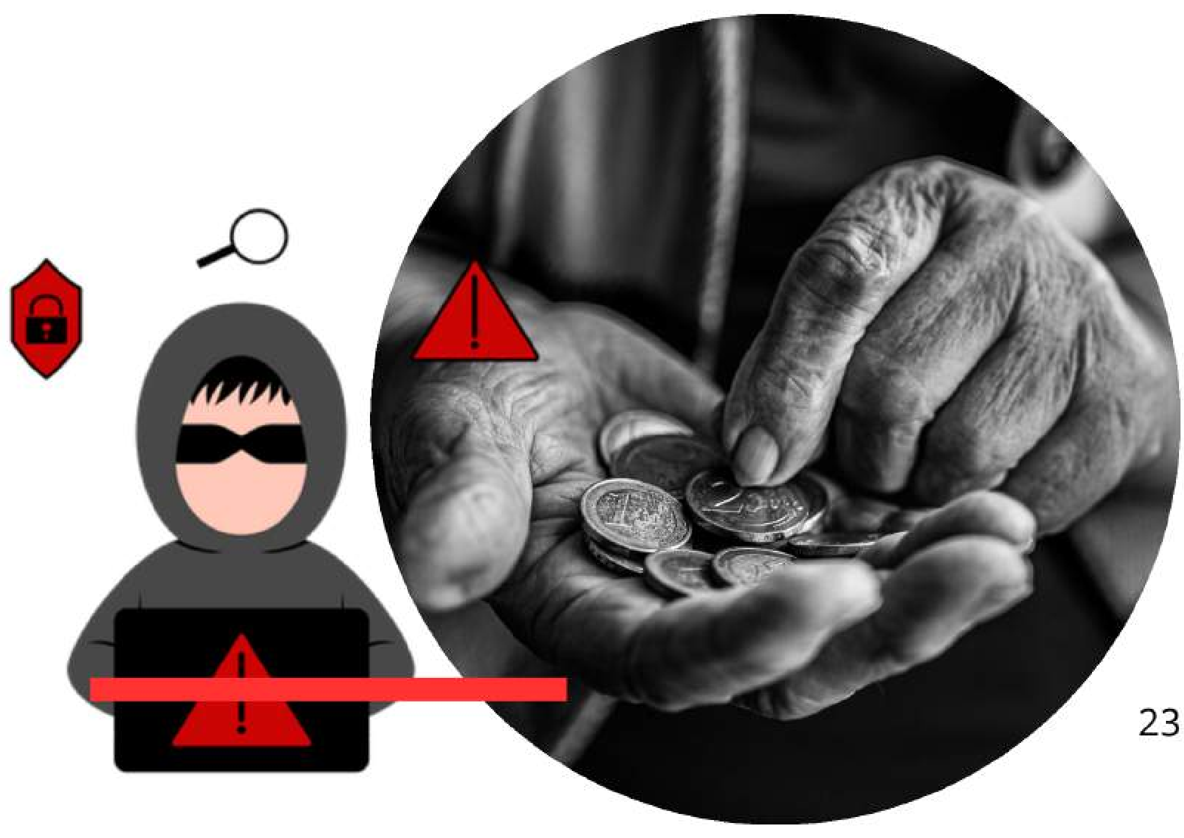



Analysis

Results

Any Fraud's means security was breached at a specific time for the **Credit Card's** transactions which stands for a tighter security methods as only three accounts today can mean millions stolen over the next months.

According to the website of Trans Union, any possible breach of data can significantly produce any trust issues as the **compliance, regulations, basis and the access to information within Data Governance (DG)** is under the analysis constantly costing millions to any financial institution (Trans Union, 2025).



Results

Results

After a deep analysis and segmentation for the needed **"Amounts"** within future changes as predicted and the notable decrease among the many transactions but little increase in the **Credit's** or Funds from the customer's databases. Additionally, there was found inside to the Credit Cards accounts the presence of three different **Fraudulent Accounts.**

Answers from Questions

Answers

Any Fraud's means security was breached at a specific time for the **Credit Card's** transactions which stands for a tighter security methods as only three accounts today can mean millions stolen over the next months.

According to the website of Trans Union, any possible breach of data can significantly produce any trust issues as the **compliance, regulations, basis and the access to information within Data Governance (DG)** is under the analysis constantly costing millions to any financial institution (Trans Union, 2025).

Furthermore, the optimization of strategies for the tactical and business changes according to the data's transparency inside of the bank's databases becomes crucial for the development of the **Credit's Services.**

Conclusion and Recommendations

Conclusion

Overall, the Credit's Accounts seem to be clear in means of protection, nonetheless, it is highly necessary to build-up a notable system for Cyber Security and the full contingency against issues within a time frame of less than 6 months to conduct less security issues and breaches checked on the common transactions done by customers.

Recommendations

- Creating a software for checking inside of the business and account's manager which helps deliver a transparent manipulation of data.
- Education of customers about any fraudulent transactions alongside preparation for the internal employees about what and what not to do about the account's security methods for any data integrity.
- Clean-Up any suspicious sources of information for the data leaks and branch of any crucial details about customers.

Conclusion and Recommendations

Recommendations

- Increase the Digital Protection and Information's Technology's budget for implementation of innovations among the security methods used by the bank currently.

- Increase the Digital Protection and Information's Technology's budget for implementation of innovations among the security methods used by the bank currently.

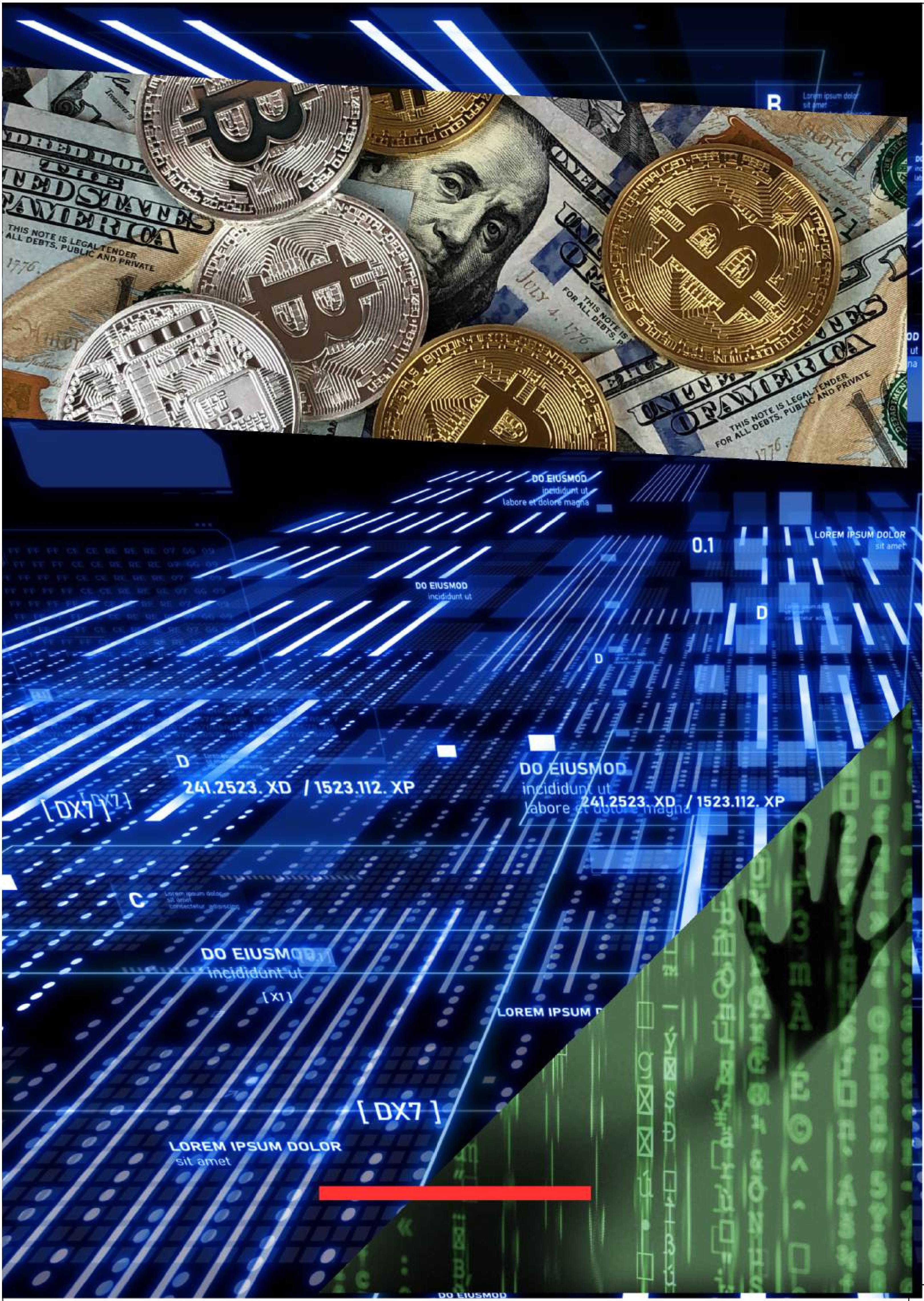
Bibliographical References

HYPR (2024). Cybersecurity Regulations for Financial Services in 2024 and Beyond. https://blog.hypr.com/top-financial-services-cybersecurity-regulations \

Trans Union (2025). What is Fraud Detection and Prevention in Banking? https://www.transunion.com/blog/what-is-fraud-detection-and-prevention-in-banking?atvy=%7B%22264995%22%3A%22Experience+B%22%7D

Western University (WU) (2024).Cyber Insurance. https://cybersmart.uwo.ca/for_it_support_providers/policies_compliance__risk_management/risk_management/cyber_insurance.html

World Bank (WB) (2025). World's GDP. https://data.worldbank.org/indicator/NY.GDP.MKTP.CD


THIS NOTE IS LEGAL TENDER
FOR ALL DEBTS, PUBLIC AND PRIVATE
DO EIUSMOD
incididunt ut
labore et dolore magna
LOREM IPSUM DOLOR
sit amet
0.1
DO EIUSMOD
incididunt ut
D
[DX7]
241.2523. XD / 1523.112. XP
DO EIUSMOD
incididunt ut
labore et dolore magna
C
DO EIUSMOD
incididunt ut
[X1]
[DX7]
LOREM IPSUM DOLOR
sit amet