**Program: Artificial Intelligence Analysis, Design and Implementation**

**Course: AIDI-2005 - CAPSTONE TERM II**

**AI Application Name: Credit Card Fraud Detection**

**Group Number:** 10

**Group Member:**

1. **Harshit Morwal** – Student ID: 100870851
2. **Nasar Islam** – Student ID: 100829880
3. **Pushkal Suri** – Student ID: 100799314

**Introduction**

**Objective**: Credit card fraud is a major problem for banks, merchants, and customers alike. By creating a credit card fraud detection application, banks and merchants can better protect their customers’ information and ensure that their transactions are secure. By using advanced analytics and artificial intelligence, these applications can detect unusual or suspicious activity and alert the customer or merchant to potential fraud. This helps to reduce losses and improve customer satisfaction.

Chart, bar chart

Description automatically generated Bar chart

Description automatically generated

**Global non-cash transactions to reach 1.8 trillion by 2025.**

It is projected that the worldwide non-cash transactions will reach pre-pandemic levels by 2021 and will continue to rise beyond that. This is sparked by a surge in expenses and the accelerated acceptance of digital payments. According to the World Payments Report, the CAGR for global non-cash transactions is estimated to be 18.6% from 2020 to 2025, with a total volume of 1.8 trillion by the end of the period.

Now, while this might be exciting news, on the flip-side fraudulent transactions are on the rise as well. As per Nilson Report, which tracks the payments industry, predicts global fraud losses related to payment cards will reach $47.22 billion in 2031.

This is now becoming a serious problem since most of the time, a person who has become a victim of this fraud don't have any idea about what has happened until the very end. So, in this capstone project, what we have tried is to create a Web App for the detection of such type of frauds with the help of Machine Learning and Artificial Intelligence Technique. In the following sections, we will be explaining about the creation and importance of both a good Machine Learning model and the Web App.

|  |  |
| --- | --- |
| **Year** | **Amount (In billions)** |
| [Source: Card Fraud Worldwide](https://nilsonreport.com/publication_newsletter_archive_issue.php?issue=1232) | |
| 2021 | 32.34 |
| 2020 | 28.43 |
| 2019 | 28.65 |
| 2018 | 27.86 |
| 2017 | 23.97 |
| 2016 | 22.8 |
| 2015 | 21.84 |
| 2014 | 18.11 |

**Benefits of Credit Card Fraud Detection Application:**

1. **Improved Security:** Credit card fraud detection applications provide an additional layer of protection against fraud by monitoring credit card transactions and alerting users of suspicious activity. This helps to protect users’ confidential information, as well as their finances.
2. **Cost Savings:** Credit card fraud detection applications can save companies money by reducing the need to refund fraudulent purchases, as well as the cost of investigating and prosecuting fraudsters.
3. **Increased Customer Satisfaction:** Credit card fraud detection applications provide customers with peace of mind knowing that their financial information is secure.
4. **Improved Business Reputation:** By taking proactive steps to protect customers’ financial information, businesses can improve their reputation and gain the trust of potential customers.
5. **Enhanced Analytics:** Credit card fraud detection applications can provide businesses with valuable insights, such as customer spending patterns and trends. This can help businesses to better understand their customer base and make more informed decisions.

**Current drawbacks of Credit Card Fraud Detection Application:**

1. **False Positives and False Negatives:** Credit card fraud detection applications can produce false positives and false negatives, resulting in either valid transactions being blocked, or fraudulent transactions being allowed.
2. **Limited Data Sources:** Credit card fraud detection applications are typically limited to the data sources they integrate with. If the underlying data sources do not contain the necessary information, the application may be unable to detect fraud.
3. **Complex Rules and Algorithms:** Credit card fraud detection applications use complex rules and algorithms to detect fraudulent activity. This can make the application difficult to configure or maintain.
4. **Slow Processing Times:** Credit card fraud detection applications can be slow to process transactions, resulting in a poor user experience.
5. **Expensive to Implement and Maintain:** Credit card fraud detection applications can be expensive to implement and maintain, especially for smaller businesses.

**Our Proposal and Benefits of Our application:**

1. We will be using 2-3 Machine Learning model to find the best model and further we will be fine tuning the model with the help of hyperparameter tuning. We will be using Logistic Regression, Random Forest Classification, Support Vector Classification algorithms. These models will provide the output in the form of confusion matrix which has details related to false negative and false positive. Our algorithm will be aiming to reduce the probability false negative and false positive thereby improving the accuracy of fraud prediction.
2. Also, dataset is having more of non-fraud observation therefore we will be using statistical techniques to counter the imbalanced class problems.
3. Our application will also store new data entries into the database which will further help us to retrain and test our Machine Learning models.
4. Instead of implementing complex rule-based algorithm, dynamic and optimized machine learning techniques will use probabilistic approach to detect the fraud transactions and thereby reducing the processing time and maintenance.

**Dataset**

1. Link to the dataset: <https://www.kaggle.com/getting-started/186583>
2. Dataset has been taken from Kaggle Public data library. This dataset is synthetic or encrypted and features may have been reduced as result of a PCA Dimensionality reduction to protect user identities and sensitive features(v1-v28).
3. Dataset was last updated on 3-5-2021 and has 284807 instances, 31attributes.
4. This dataset represents transactions made within two days, with 492 frauds out of 284,807 transactions. The dataset is highly imbalanced, with the positive class (fraudulent) making up his 0.172% of all transactions. This includes only numeric input variables that are the result of PCA transformations.

**EDA**

**Chart, bar chart

Description automatically generated**

**Related Work**

1. **Reproducible ML for Credit Card Fraud Detection:** This book has covered areas related to :
   1. Simulator of synthetic transaction data
   2. Reproducibility: has Jupyter book that contains code.
   3. Evaluation and performance matrix
   4. Treatment of Imbalanced class
   5. Deep Learning
   6. Link: <https://fraud-detection-handbook.github.io/fraud-detection-handbook/Chapter_1_BookContent/BookContent.html>
2. **Credit Card Fraud Detection: A Systematic Review**

This review provides an overview of research conducted on Credit Card Fraud Detection (CCFD). It will emphasize the challenge of class imbalance and the different Machine Learning techniques. It will also discuss the most effective evaluation metrics for CCFD, since the data is sensitive and not easy to access. Additionally, web sources for available datasets and trending software tools used for deploying CCFD will be outlined.

Link: <https://link.springer.com/chapter/10.1007/978-3-030-38501-9_29>

1. **A Comprehensive Survey on Machine Learning Techniques and User Authentication Approaches for Credit Card Fraud Detection:** This report provides a comprehensive overview of current approaches and solutions to the issue of credit card fraud detection. The initial segment of the survey focuses on classical machine learning models, which rely on traditional transaction features for fraud prediction. Subsequently, the survey covers more advanced user authentication methods, which employ behavioral biometrics to identify individuals based on their unique behavior when interacting with electronic devices.

Link: <https://arxiv.org/abs/1912.02629>

1. **Credit card fraud detection using machine learning: A survey:** Has research on data-driven credit card fraud detection to identify any transactions which are not authorised by the legitimate cardholder. Various machine learning techniques are used to tackle the complexities of this process.

Link: <https://arxiv.org/abs/2010.06479>

**How Different Our Work is from above studies?**

We will be implementing the solution on web application unlike computer-based software systems that are stored locally on the Operating System of the device. Web applications are accessed by the user through a web browser with an active internet connection. These applications are programmed using a client–server modeled structure—the user is provided services through an off-site server that is hosted by a third-party. We will be using various parts of the above research and will try to combine and implement those techniques to increase the accuracy of the machine learning model. We will try and understand various short coming of the above research and will try and cover up those in our ML model and web-based application.

**Bibliography**

1. <https://www.paymentscardsandmobile.com/world-payment-report-non-cash-transaction-growth-hit-hard-by-covid/>
2. <https://www.bankrate.com/finance/credit-cards/credit-card-fraud-statistics/#fraud>
3. [https://fraud-detection-handbook.github.io/fraud-detection handbook/Chapter\_1\_BookContent/BookContent.html](https://fraud-detection-handbook.github.io/fraud-detection%20handbook/Chapter_1_BookContent/BookContent.html)
4. [Source: Card Fraud Worldwide](https://nilsonreport.com/publication_newsletter_archive_issue.php?issue=1232)