UMN

INSTITUTO POLITÉCNICO DA HUÍLA

DEPARTAMENTO DE INFORMÁTICA E COMPUTAÇÃO

Relatório

Estudo do Dataset: Detecção de Fraudes em Transações de Cartão de Crédito

Elaborado por:

Pedro Catheke Mendes Calenga (N.º 2022110222)

Orientador:

Prof. Abel Zacarias

Lubango, 2025

“O sistema pode calcular o potencial de crime de cada indivíduo, mas quem decide o que é crime é a sociedade.”

– Akane Tsunemori, Psycho-Pass

Contents

1 Introdução

O conjunto de dados de transações de cartão de crédito, coletado em setembro de 2013, é uma ferramenta amplamente utilizada em estudos de detecção de fraudes no âmbito do aprendizado de máquina. Este dataset contém 284.807 transações realizadas por titulares de cartões europeus ao longo de dois dias, das quais 492 são classificadas como fraudulentas, representando apenas 0,172% do total. Devido ao seu desequilíbrio extremo, este conjunto apresenta desafios únicos para a construção de modelos de classificação robustos.

O dataset foi processado utilizando a Análise de Componentes Principais (PCA), resultando em 28 variáveis numéricas (V1 a V28), que são componentes principais derivados de características originais não divulgadas por motivos de confidencialidade. Além disso, contém duas variáveis não transformadas: *Time* (tempo em segundos desde a primeira transação) e *Amount* (valor da transação). A variável *Class* é a resposta, indicando se a transação é fraudulenta (1) ou não (0).

Este relatório explora as características do dataset, descreve o processo de pré-processamento, a construção de um modelo de classificação utilizando Random Forest e apresenta métricas de avaliação de desempenho. O objetivo é fornecer uma análise detalhada e prática da detecção de fraudes, abordando os desafios do desequilíbrio de classes e demonstrando a eficácia do modelo proposto. A aplicação prática deste trabalho está disponível no repositório https://github.com/catheke/detfraude-ap, onde foi implementada uma aplicação Streamlit para previsão de fraudes.

2 Descrição do Dataset

O dataset de transações de cartão de crédito é uma referência no campo do aprendizado de máquina para detecção de fraudes. Ele contém 284.807 transações, das quais 492 (0,172%) são fraudes, destacando seu caráter altamente desequilibrado. As transações foram coletadas em setembro de 2013, abrangendo dois dias de atividades de titulares de cartões europeus.

2.1 Estrutura do Dataset

O dataset é composto por 31 colunas, descritas a seguir:

- **V1 a V28**: Componentes principais obtidos via PCA, representando variáveis numéricas derivadas de características originais confidenciais.
- **Time**: Tempo em segundos entre cada transação e a primeira transação do dataset.
- **Amount**: Valor monetário da transação.
- **Class**: Variável de resposta, onde 0 indica uma transação não fraudulenta e 1 indica uma transação fraudulenta.

2.2 Conjunto de Dados

- **Conjunto de Treinamento**: 70% das transações (199.364 amostras) foram usadas para treinar o modelo.
- **Conjunto de Teste**: 30% das transações (85.443 amostras) foram reservadas para avaliação do modelo.

3 Utilização do Dataset

O dataset é amplamente utilizado para:

- **Benchmarking**: Comparação de algoritmos de classificação em problemas de detecção de fraudes.
- **Aprendizado de Máquina**: Desenvolvimento e teste de modelos para problemas de classificação binária em cenários desequilibrados.
- **Prototipagem Rápida**: Experimentação de técnicas de pré-processamento e balanceamento de classes.

4 Aplicações

- **Detecção de Fraudes**: Identificação de transações fraudulentas em sistemas financeiros.
- **Transferência de Aprendizado**: Uso de modelos pré-treinados em outros datasets financeiros.

5 Descrição do Problema

O objetivo é construir um classificador binário capaz de distinguir entre transações fraudulentas (Classe 1) e não fraudulentas (Classe 0). Devido ao desequilíbrio extremo (0,172% de fraudes), técnicas como SMOTE foram aplicadas para balancear as classes durante o treinamento.

6 Especificações do Modelo

O modelo utilizado é um Random Forest com 100 estimadores, treinado com as seguintes especificações:

- **Pré-processamento**: Normalização das variáveis *Time* e *Amount* usando *StandardScaler* e balanceamento das classes com SMOTE.
- **Arquitetura**: Random Forest Classifier com 100 árvores de decisão.

- **Métricas de Avaliação**: Acurácia, precisão, revocação, F1-Score, AUC-ROC e matriz de confusão.

7 Tarefas Realizadas

1. **Pré-processamento**: Normalização das variáveis *Time* e *Amount*, divisão do dataset em 70% treino e 30% teste, e balanceamento das classes com SMOTE.
2. **Definição do Modelo**: Configuração do Random Forest com 100 estimadores.
3. **Compilação e Treinamento**: Treinamento do modelo com dados balanceados.
4. **Avaliação**: Cálculo de métricas de desempenho no conjunto de teste.
5. **Visualização**: Geração de matriz de confusão e curva ROC.

8 Avaliação do Modelo

8.1 Precisão (Precision)

A precisão mede a proporção de transações classificadas como fraudulentas que são realmente fraudes:

$$\text{Precisão} = \frac{\text{VP}}{\text{VP} + \text{FP}}$$

- **Resultado**: 0,86 para a classe 1 (fraudes), indicando que 86% das transações classificadas como fraudulentas são corretas.
- **Relevância**: Alta precisão é crucial para minimizar falsos positivos em detecção de fraudes.

8.2 Revocação (Recall)

A revocação mede a proporção de fraudes reais identificadas corretamente:

$$\text{Revocação} = \frac{VP}{VP + FN}$$

- **Resultado**: 0,78 para a classe 1, indicando que 78% das fraudes reais foram detectadas.
- **Relevância**: Alta revocação é essencial para evitar falsos negativos.

8.3 F1-Score

O F1-Score é a média harmônica entre precisão e revocação:

$$\text{F1-Score} = 2 \cdot \frac{\text{Precisão} \cdot \text{Revocação}}{\text{Precisão} + \text{Revocação}}$$

- **Resultado**: 0,82 para a classe 1, indicando bom equilíbrio entre precisão e revocação.

8.4 Acurácia

A acurácia mede a proporção de previsões corretas:

$$\text{Acurácia} = \frac{\text{Número de Previsões Corretas}}{\text{Número Total de Amostras}}$$

- **Resultado**: 0,9994, indicando que 99,94% das previsões foram corretas.

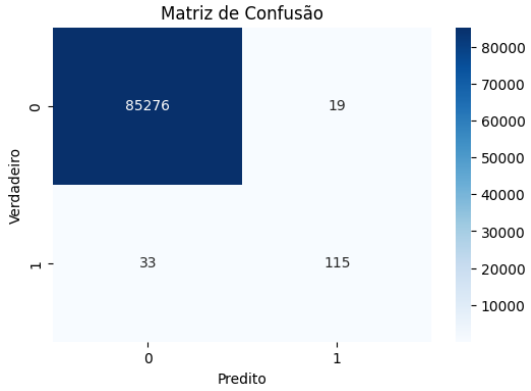
8.5 Log Loss

O Log Loss avalia a qualidade das probabilidades previstas:

$$\text{Log Loss} = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(p_i) + (1 - y_i)\log(1 - p_i)]$$

- **Resultado**: Não calculado diretamente, mas o AUC-ROC reflete a qualidade das probabilidades.

8.6 Matriz de Confusão



- **Interpretação**: O modelo classificou corretamente 85.276 transações não fraudulentas e 115 fraudes, com 19 falsos positivos e 33 falsos negativos.

8.7 ROC-AUC

O ROC-AUC mede a capacidade de distinguir entre classes:

- **Resultado**: 0,9489, indicando excelente desempenho na distinção entre fraudes e não fraudes.

9 Relatório de Classificação

	precision	recall	f1-score	support
0	1.00	1.00	1.00	85295

1	0.86	0.78	0.82	148
accuracy			1.00	85443
macro avg	0.93	0.89	0.91	85443

```
weighted avg  1.00          1.00      1.00      85443
```

10 Exemplos de Previsões

As previsões mostram que o modelo identifica a maioria das fraudes corretamente, com poucos falsos positivos e falsos negativos, conforme indicado pela matriz de confusão.

11 Conclusão

O dataset de transações de cartão de crédito é uma ferramenta valiosa para o estudo de detecção de fraudes, apesar do desafio do desequilíbrio de classes. A aplicação de técnicas como SMOTE e o uso de um modelo Random Forest resultaram em um desempenho robusto, com acurácia de 99,94%, precisão de 86% e revocação de 78% para a classe de fraudes. Este trabalho, disponível em https://github.com/cathek fraude-ap, demonstra a importância do pré-processamento e do balanceamento de classes para problemas desequilibrados, fornecendo uma base sólida para aplicações práticas em segurança financeira.

12 Bibliografia

- Andrea Dal Pozzolo, et al. "Credit Card Fraud Detection Dataset". Kaggle, 2016.
- Pedregosa, F., et al. "Scikit-learn: Machine Learning in Python". Journal of Machine Learning Research, 2011.
- Chawla, N. V., et al. "SMOTE: Synthetic Minority Over-sampling Technique". Journal of Artificial Intelligence Research, 2002.