Detecção de Padrões Fraudulentos em Contratos Inteligentes do Ethereum Usando Aprendizado de Máquina para Classificação de Transações

Lorenzo Henrique Rebelo (Engenharia de Computação Poli-USP)
Paolla Queiroz (Engenharia de Computação Poli-USP)
Vinicio Mendes (Engenharia de Produção Poli-USP)

Datathon de Criptomoedas
2024

Agenda

1. Overview
2. Fundamentação Teórica
3. Metodologia
4. Resultados
5. Conclusões

Overview

Objetivos:

- Identificar padrões de transação fraudulentos em contratos inteligentes do Ethereum utilizando modelos de aprendizado de máquina, para classificar transações como fraudulentas ou legítimas.

Questão norteadora:

- Quais padrões de transação indicam um possível esquema de Ponzi em contratos inteligentes do Ethereum?”

Metodologia:

- Descrição dos Dados
- Processamento de Dados
- Modelos de Machine Learning Utilizados
- Métricas de Avaliação

Conclusão:

- Principais Aprendizados
- Limitações do Trabalho
- Próximos Passos

Fundamentação teórica

A Securities and Exchange Commission dos Estados Unidos (SEC) define **esquemas de pirâmide** como um investimento fraudulento que envolve o pagamento de retorno aos investidores antigos a partir de fundos de novos investidores (SEC, 2013)

Esquema Ponzi é uma antiga forma de fraude que muitas pessoas odeiam e buscam. O golpe é construído em traição e mentiras, pelas quais os organizadores e especuladores enganam conjuntamente investidores inocentes, fomentando crenças na obtenção dos benefícios esperados. Os esquemas Ponzi entraram nas finanças digitais a partir das finanças tradicionais ao longo do tempo e são chamados de **esquema Ponzi inteligente** (SPS) no mundo da moeda digital/virtual. (Bartoletti et al., 2020).

Risco é um fator relevante no que tange a investimentos, pois espera-se uma correlação positiva com o retorno. Dessa forma, é necessário um maior retorno para atrair o investidor médio quando o risco apresentado é maior. Sua definição em finanças se refere as chances do retorno total de um determinado ativo financeiro ser diferente do previsto, isto é, as chances potenciais de se perder parte ou todo o investimento original (Ricciardi, 2008).

Fundamentação teórica

De acordo com a **definição** da Comissão de Valores Mobiliários dos EUA e o item de **esquema Ponzi** na Wikipédia, um esquema Ponzi tem as seguintes características.

- É uma fraude de investimento.
- Seus retornos para investidores existentes vêm dos novos investidores.
- Seu organizador tem pouco ou nenhum lucro legítimo.
- Seu organizador promete altas recompensas para novos investidores.
- Ele pode manter um negócio sustentável até que não consiga pagar os saques exigidos pelos investidores porque não há novos patrícios suficientes.
- Um fundo normal se tornaria um esquema Ponzi enquanto suas fontes legais de renda são permanentemente cortadas.

Os esquemas SPS operam em três fases: **bootstrap, hiperoperação e colapso**, caracterizando o ciclo de vida do esquema desde seu início até o colapso inevitável quando novos investidores deixam de ingressar (Boshmaf et al., 2020).

- **Bootstrap**: É a frase inicial que o fundo é conhecido por uma quantidade muito pequena de investidores que estão esperando e vendo.
- **Hiperoperação**: Durante esta fase, mais e mais investidores participam do esquema, todos acreditam que poderiam obter altos retornos e convidar novos investidores sinceramente.
- **Colapso**: Esta fase começa com os investidores não podem retirar o seu investimento e perder a sua confiança no esquema, pois, naquela época, não havia novos investimentos suficientes para apoiar as recompensas aos primeiros investidores. Em geral, antes desta fase, o trapaceiro vai fugir, o que aceleraria esta fase a vir.

Dados e Metodologia

Base de dados

Dataset que contém transações conhecidas como fraudulentas e válidas realizadas na rede Ethereum

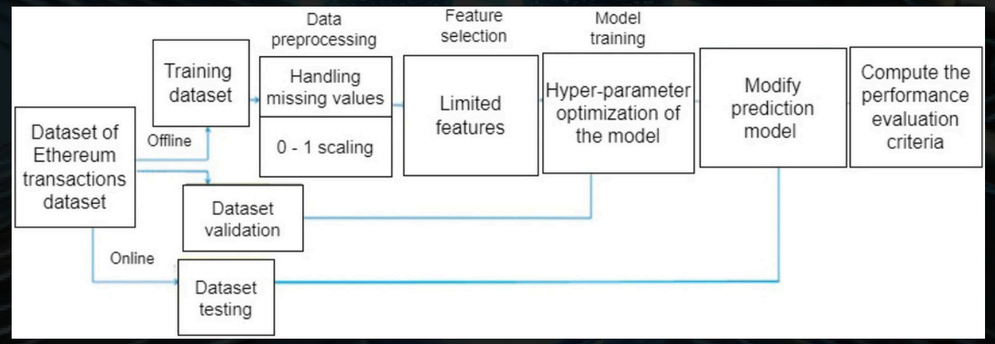
Conjunto de dados desbalanceado que contém:
7637 transações verdadeiras
2179 transações fraudulentas

Link dataset

	Address	FLAG	Avg min betwee...	Avg min betwee...	Time
	9816 unique values	0 1	0 430k	0 482k	0
	0x0009277775ac7d0d59eaad8fee3d10ac6c805e8	0	844.26	1093.71	704785
	0x0002b44ddb1476db43c868bd494422ee4c136fed	0	12709.07	2958.44	121821
	0x0002bda54cb772d040f779e88eb453cac0daa244	0	246194.54	2434.02	516729
	0x00038e6ba3fd5c09aedb96697c8d7b8fa6632e5e	0	10219.6	15785.09	397555
	0x00062d1dd1afb6fb02540ddad9cdebfe568e0d89	0	36.61	10707.77	382472
	0x000895ad78f4403ecd9468900e68d6ee506136fd	0	9900.12	375.48	20926

Processamento dos dados

1. Separação de Features e Target:
   - Features relevantes separadas da variável-alvo (FLAG).
2. Normalização:
   - Uso de MinMaxScaler para padronizar as features entre [0, 1].
3. Balanceamento de Classes:
   - Aplicação de SMOTE (Synthetic Minority Over-sampling Technique) para gerar exemplos sintéticos da classe minoritária.
4. Simulação e Integração de Dados:
   - Dados reais e simulados combinados:
     - 500k legítimos | 500k fraudulentos | 200k desafiadores.
5. Imputação de Dados Faltantes:
   - Tratamento de valores ausentes para evitar inconsistências.



Features

As features escolhidas para o modelo de detecção refletem características comportamentais e estruturais frequentemente associadas à fraude em análise

.

- Alta frequência de transações
  (Avg min between sent/received tnx).
- Curta duração de atividade
  (Time Diff between first and last (Mins)).
- Interações amplas com muitos participantes
  (Sent tnx e Received Tnx).
- Comportamento suspeito de criação de contratos
  (Number of Created Contracts).

- **Total de 1,2M de transações processadas ao final.**
  - 80% Treinamento
  - 20% Teste para evitar overfitting.

Dados e Metodologia

Explicando XGBoost

- **O que é**: É um modelo de **aprendizado de máquina** baseado em **árvores de decisão** que combina várias "árvores de decisão" que aprendem com os erros das anteriores para melhorar as predições. Ele analisa os padrões nos dados (como tempo entre transações ou número de contratos) e decide se uma transação é legítima ou fraudulenta.

- **Como usamos:** Primeiro, **ampliamos o conjunto de dados de contratos inteligentes** relacionados a esquemas Ponzi e eliminamos o desbalanceamento do dataset aplicando amostragem sintética adaptativa. Em seguida, **definimos quatro tipos de conjuntos de features** baseados nos códigos de operação (opcodes) dos contratos inteligentes.

F1 - Score

O F1-Score é usado como uma **métrica final para verificar se o modelo está equilibrado**, especialmente em conjuntos de dados desbalanceados, como fraudes (classe minoritária) versus transações legítimas.

- Valores mais próximos de 1 indicam um equilíbrio ideal entre precisão e recall.
- O F1-Score é útil quando o custo de falsos negativos ou falsos positivos é alto, como em fraudes.

$$F_1 = 2 \times \frac{Precision \times Recall}{(Precision + Recall)}$$

ROC AUC

- A curva ROC (Receiver Operating Characteristic) r**elaciona a taxa de verdadeiros positivos (TPR) com a taxa de falsos positivos (FPR)** em diferentes limiares de decisão.
- A AUC (Area Under the Curve) é a área sob essa curva e indica o quão bem o modelo separa as classes.
  - AUC = 1.0: O modelo classifica todas as amostras corretamente (perfeito).
  - AUC = 0.5: O modelo não é melhor do que um classificador aleatório.
  - AUC > 0.7: Indica um bom desempenho na maioria dos casos.

$$AUC = \int_0^1 TPR(FPR) d(FPR)$$

XGBoost realiza classificação binária

- Input: Conjunto de features {X} e labels {y}.
  - y={0,1} representando classes (legítima ou fraudulenta).
  - X = {x1, x2, ...xm} features de cada amostra.
- Output: Probabilidade de uma transação ser fraudulenta P(y=1 | X)).

Após o treinamento, o XGBoost calcula a importância de cada feature com base na redução de erro (ganho) que a feature proporcionou em cada divisão da árvore

função de perda

$$L = \sum_{i=1}^{m} l\left(y_i, \hat{y_i}\right) + \sum_{k=1}^{K} \Omega\left(f_k\right)$$

$l\left(y_i, y_i\right)$: Função de perda

$\Omega\left(f_k\right)$: Termo de regularização para evitar overfitting

Resultados

```
Cenários Desafiadores:
Predições Fraudulentas: 987
Predições Legítimas: 13
Probabilidade Média para Fraude: 0.99
Desvio Padrão das Probabilidades: 0.11

Dados Legítimos:

              precision    recall  f1-score   support

           0       1.00      1.00      1.00       500
           1       0.00      0.00      0.00         0

    accuracy                           1.00       500
   macro avg       0.50      0.50      0.50       500
weighted avg       1.00      1.00      1.00       500


Dados Fraudulentos:

              precision    recall  f1-score   support

           1       1.00      1.00      1.00       500

    accuracy                           1.00       500
   macro avg       1.00      1.00      1.00       500
weighted avg       1.00      1.00      1.00       500
```

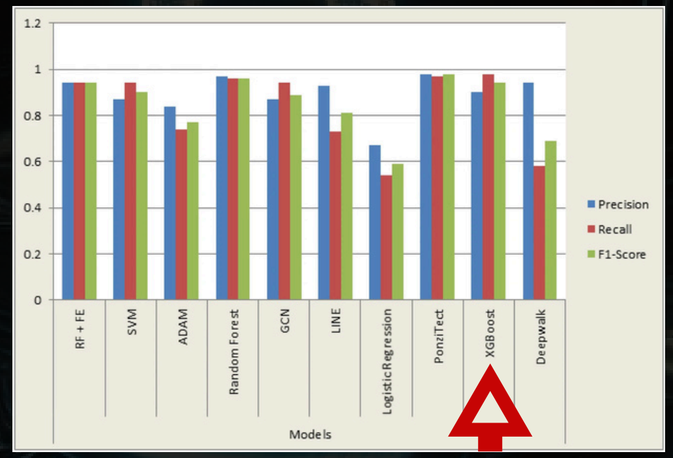
Desempenho nos Dados de Teste Gerais

- Precision: 99%-100%
- Recall: 100%
- F1-Score: 100%
- Acurácia: 100%
- ROC AUC: 1.00 (perfeito)

Avaliação em Cenários Desafiadores

- Predições Fraudulentas: 987 de 1000
- Predições Legítimas: 13 de 1000
- Probabilidade Média para Fraude: 0.99
- Desvio Padrão das Probabilidades: 0.11

Acurácia de 98.7%



Eficiência de modelos de aprendizado de máquina para esquemas de Ponzi

Como as métricas de validação foram usadas:

- **Acurácia**: Avalia o desempenho geral do modelo, verificando a proporção de predições corretas em relação ao total de casos.
- **Precision**: Mede o quanto das transações classificadas como fraudulentas pelo modelo realmente eram fraudes, ajudando a reduzir falsos positivos.
- **Recall**: Verifica a capacidade do modelo de detectar todas as fraudes reais, garantindo baixa taxa de falsos negativos.
- **F1-Score**: Média harmônica entre Precision e Recall, garantindo que o modelo fosse eficiente tanto na identificação de fraudes quanto em minimizar erros.
- **ROC AUC:** Avalia a capacidade do modelo de separar corretamente fraudes de transações legítimas, independentemente do limiar de decisão. Quanto mais próximo de 1, melhor a separação entre as classes.

Considerações Finais

Aprendizados:

A transparência inerente à blockchain, onde todas as transações são públicas e imutáveis, possibilita uma **análise contínua e detalhada de atividades fraudulentas por meio de diversas features**. Isso permite identificar padrões associados a fraudes, especialmente em **esquemas de Ponzi**, com **elevada taxa de detecção e precisão**.
Nosso projeto demonstra o potencial de transformar a forma como fraudes são detectadas e prevenidas em redes blockchain. Se expandido para incluir mais dados reais, incorporar novas features, e ser adaptado a diversas redes, como Bitcoin e Binance Smart Chain, pode se tornar um divisor de águas na segurança do ecossistema cripto. Além de reduzir riscos, reforça a confiança dos usuários e investidores.

Detecção de Esquemas Ponzi seguindo as suas fases bem definidas:

As fases já citadas criam um padrão previsível e consistente que pode ser capturado com alta precisão por modelos de aprendizado de máquina. Ao identificar comportamentos anômalos, como frequência elevada de transações e ciclos curtos de atividade, o nosso modelo é capaz de distinguir atividades fraudulentas de transações legítimas.

A implementação criada utilizou modelos de aprendizado de máquina para capturar padrões comportamentais e transacionais que caracterizam fraudes, como frequência elevada de transações e ciclos de vida curtos. Por meio da análise detalhada dessas características, o modelo mostrou alta precisão ao distinguir atividades legítimas de fraudulentas, criando uma base sólida para detecção de fraudes de forma eficiente.

Próximos Passos

- Expandir a Base de Dados
- Avaliar Diversos Modelos de Machine Learning e comparar suas acurácias
- Aprimorar as Features
- Reduzir Ruído nos Dados (como isolamento de outliers por meio de algoritmos como Isolation Forest)
- Integrar Dados Externos (como listas negras de endereços suspeitos)
- Expandir para Outras Redes Blockchain

Limitações

- Uma parte significativa do dataset é gerada sinteticamente. Embora válidos, esses dados podem não capturar todas as nuances das transações reais.
- Apesar das features escolhidas serem as mais relevantes, elas podem não capturar aspectos mais profundos, como comportamentos maliciosos avançados.

5º DESAFIO DE DADOS (DATATHON) DE MOEDAS DIGITAIS FGV EESP

Obrigado!

Lorenzo H. Rebelo
Engenharia de Computação
Poli -USP

Paolla C. Queiroz
Engenharia de Computação
Poli -USP

Vinicio Mendes
Engenharia de Produção
Poli -USP

FGV

USP

POLICHAIN.XYZ