

Programa de Pós-graduação em Ciência da Computação  
Faculdade de Computação

# MACHINE LEARNING APLICADO A FORENSE DE CRIPTOMOEDAS

**Aluno:** Pedro Henrique Resende Ribeiro

**Orientador:** Rodrigo Sanches Miani

**Coorientador:** Ivan da Silva Sendin

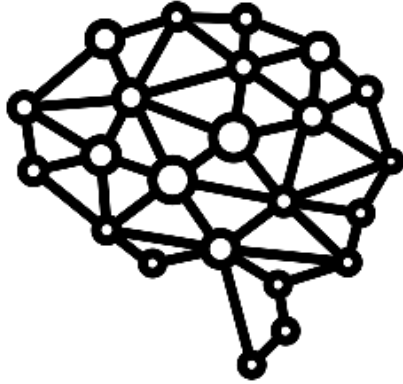
# Conteúdo

1. Introdução ao machine learning
  - 1.1. Conceitos básicos
  - 1.2. Justificativa do uso de machine learning
2. Algoritmos de machine learning
  - 2.1. Visão geral
  - 2.2. Escolha do algoritmo para o exemplo
3. Preparação dos dados
  - 3.1. Coleta e pré-processamento
  - 3.2. Técnicas de limpeza e transformação de dados
  - 3.3. Feature engineering
  - 3.4. Qualidade dos dados
4. Exemplo prático: árvore de decisão
  - 4.1. Descrição do problema
  - 4.2. Construção do modelo
  - 4.3. Métricas de avaliação
  - 4.4. Resultados e discussão
5. Desafios e considerações
  - 5.1. Desafios técnicos e limitações
  - 5.2. Empresas que realizam investigações
6. Trabalhos futuros
  - 6.1. Linhas de pesquisa



# INTRODUÇÃO

## 1.1. Conceitos básicos

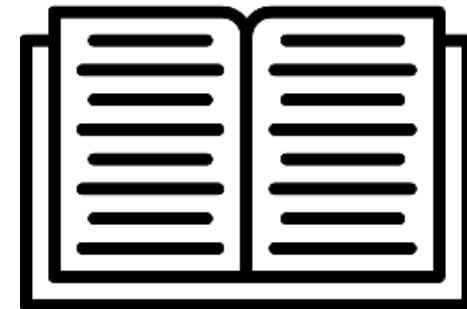


Machine Learning (ML) se concentra no desenvolvimento de algoritmos que **permitem que computadores aprendam a partir de dados** e façam previsões ou decisões sem serem explicitamente programados para isso. Em essência, o machine learning capacita sistemas a **identificar padrões em grandes volumes de dados**.

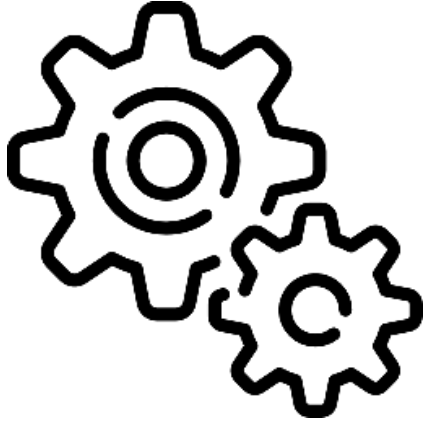
**Aprendizado supervisionado:** o algoritmo é treinado com um conjunto de dados rotulados.

**Aprendizado não supervisionado:** os dados de treinamento não têm rótulos.

**Aprendizado por reforço:** o algoritmo aprende a tomar decisões por meio de um processo de tentativa e erro, recebendo recompensas ou punições com base nas ações tomadas.



## 1.1. Conceitos básicos

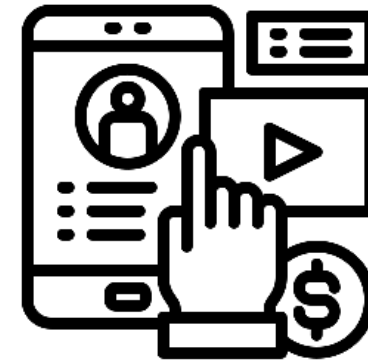


Os componentes de um modelo de machine learning podem ser descritos como:

Dados
Modelo
Treinamento
Avaliação
Hiperparâmetros

Os algoritmos de machine learning são amplamente utilizados em diversas áreas, como reconhecimento de fala, visão computacional, sistemas de recomendação, e na [forense de criptomoedas](#), como veremos ao longo desta apresentação.

[Recomendação](#): ver os trabalhos da FACOM...



## 1.2. Justificativa do uso de machine learning



**Volume e complexidade dos dados:** As redes de criptomoedas geram um grande volume de transações a cada segundo, muitas vezes com padrões complexos e interconexões difíceis de identificar manualmente.

**Detecção de atividades ilícitas:** as criptomoedas são usadas em atividades ilícitas, como lavagem de dinheiro, financiamento ao terrorismo e fraude. Algoritmos de ML podem ser treinados para identificar comportamentos suspeitos em transações, sinalizando-as para investigações mais aprofundadas.



## 1.2. Justificativa do uso de machine learning



**Evolução contínua:** O ambiente de criptomoedas está em constante evolução, com novas ameaças e técnicas surgindo regularmente.

**Integração com outras ferramentas:** As técnicas de machine learning pode ser integrado com outras ferramentas de análise forense digital para criar um sistema de monitoramento mais robusto e abrangente.

**Recomendação:** pesquisar sobre OSINT...

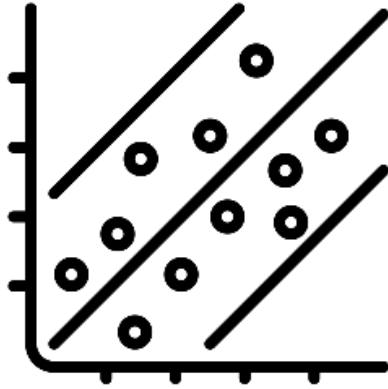




# **ALGORITMOS DE MACHINE LEARNING**



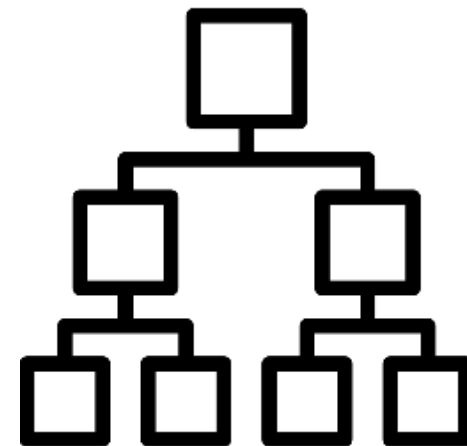
## 2.1. Visão geral



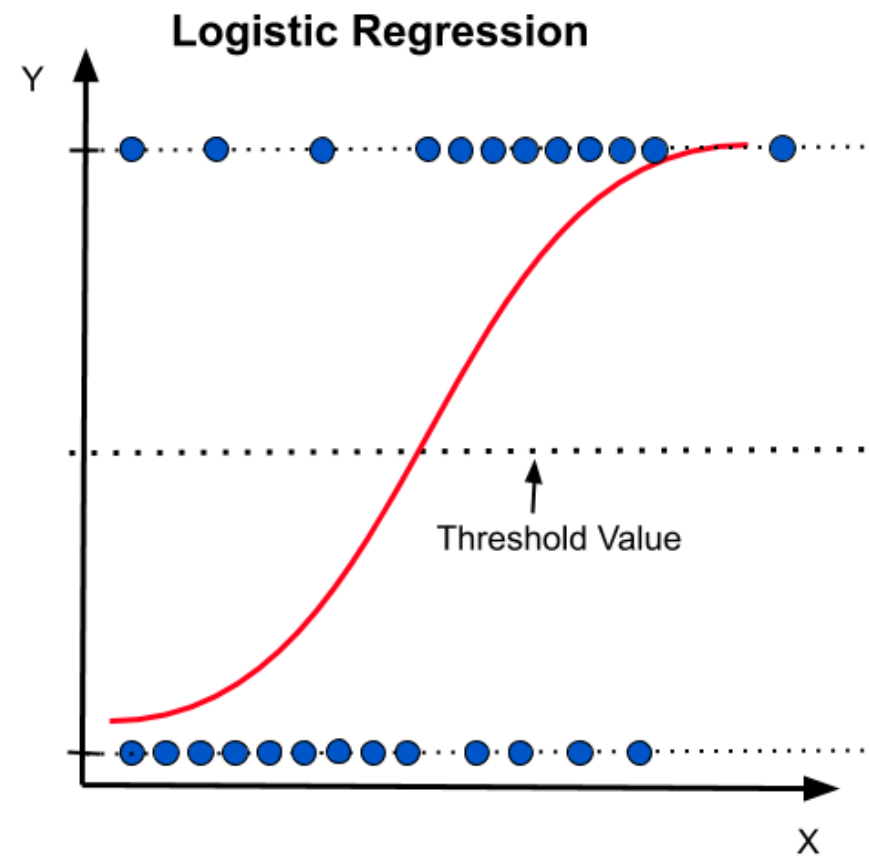
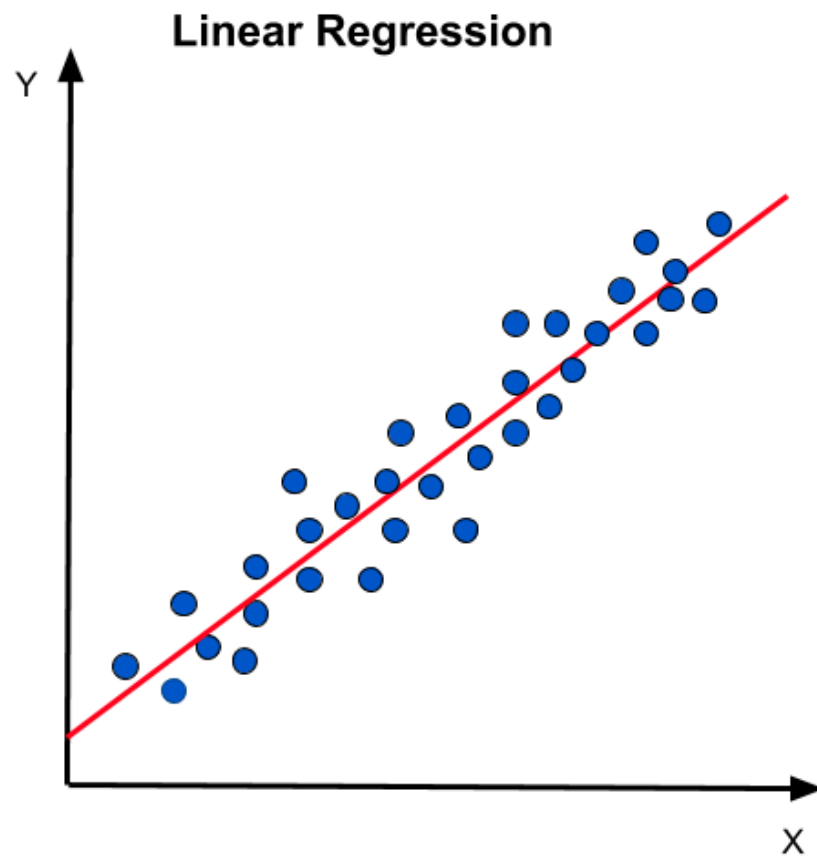
**Regressão linear:** Busca a melhor linha reta que minimiza o erro entre as previsões e os valores reais. Embora simples, pode ser muito eficaz em casos onde há uma relação linear entre as variáveis.

**Regressão logística:** É usado para problemas de classificação binária. Ele estima a probabilidade de pertencer a uma determinada classe, baseando-se em uma combinação linear de características.

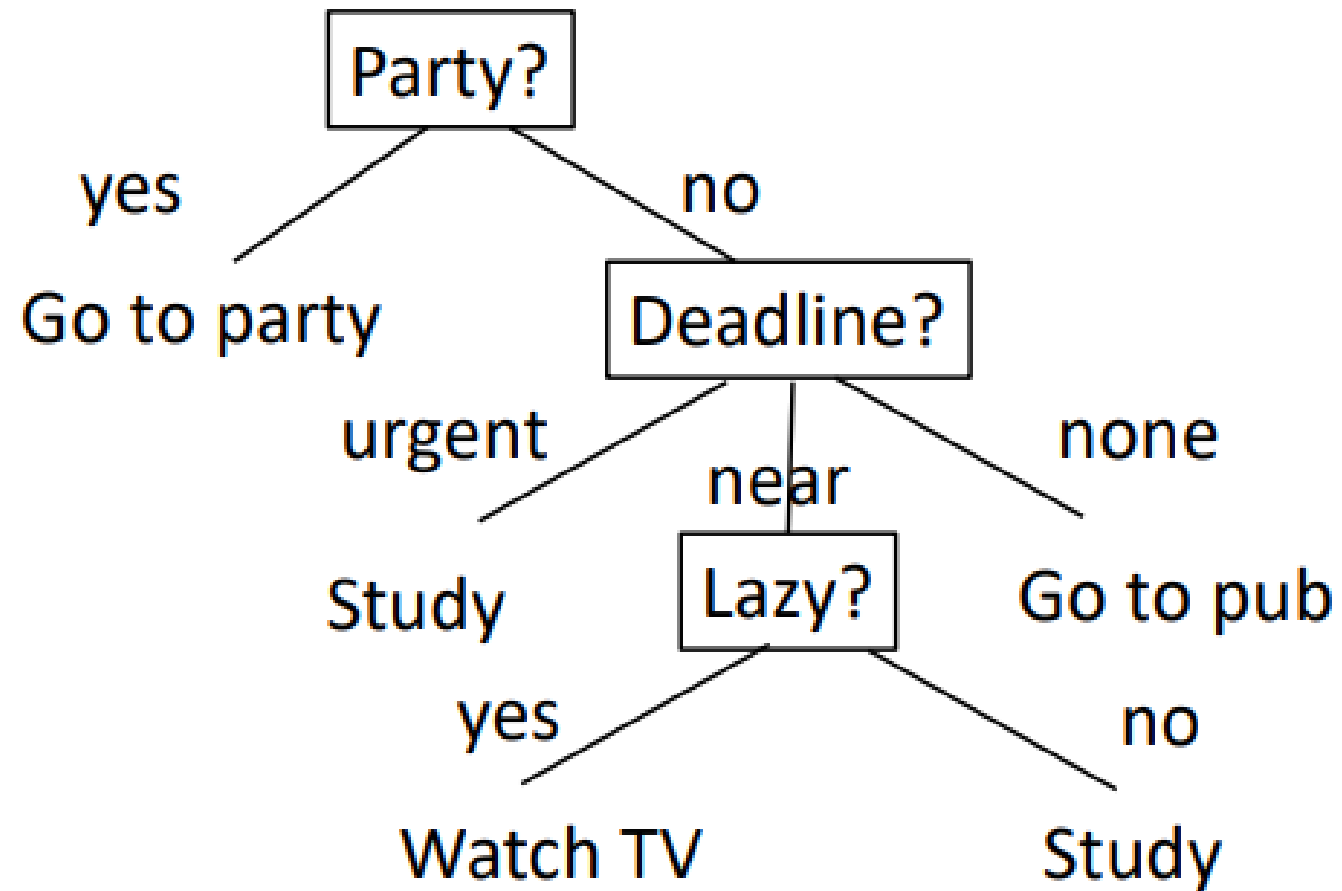
**Árvores de decisão:** Um método de classificação que divide os dados em subconjuntos baseados em perguntas sequenciais, levando a uma árvore de decisões. Cada nó da árvore representa uma decisão com base em uma característica, e as folhas representam as previsões finais.



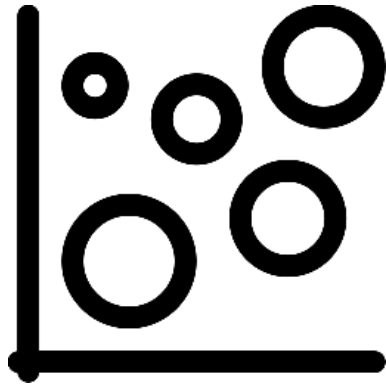
## 2.1. Visão geral



## 2.1. Visão geral

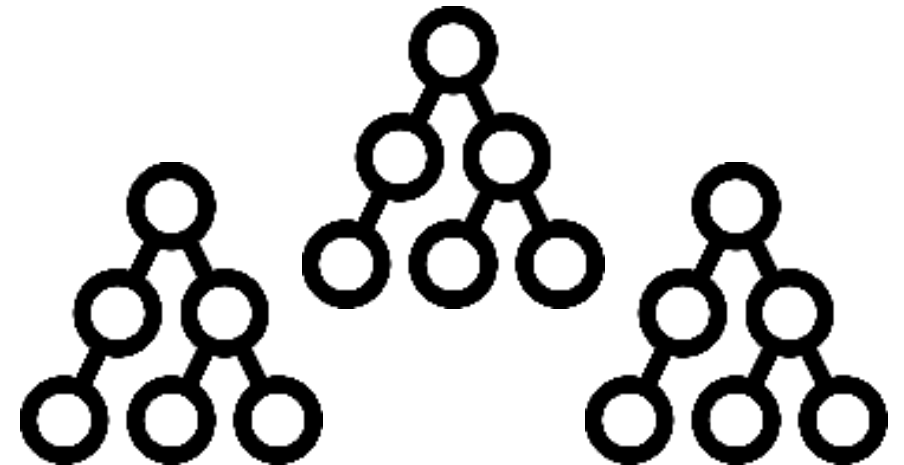


## 2.1. Visão geral

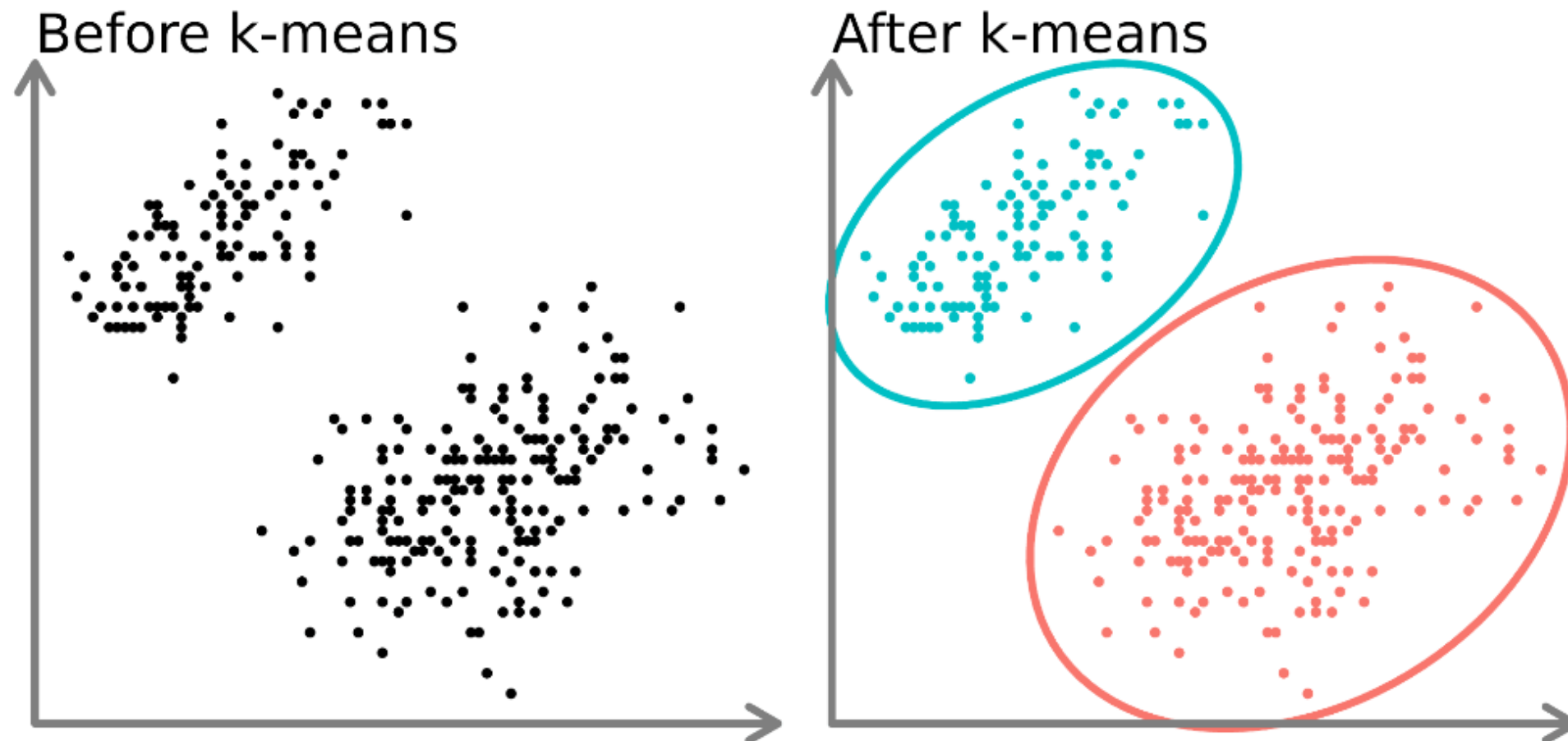


**K-Means:** Um dos algoritmos de clustering mais conhecidos, o K-Means agrupa dados em K clusters, onde cada dado é associado ao cluster com o centroide mais próximo. É útil para descobrir grupos naturais dentro dos dados sem a necessidade de rótulos.

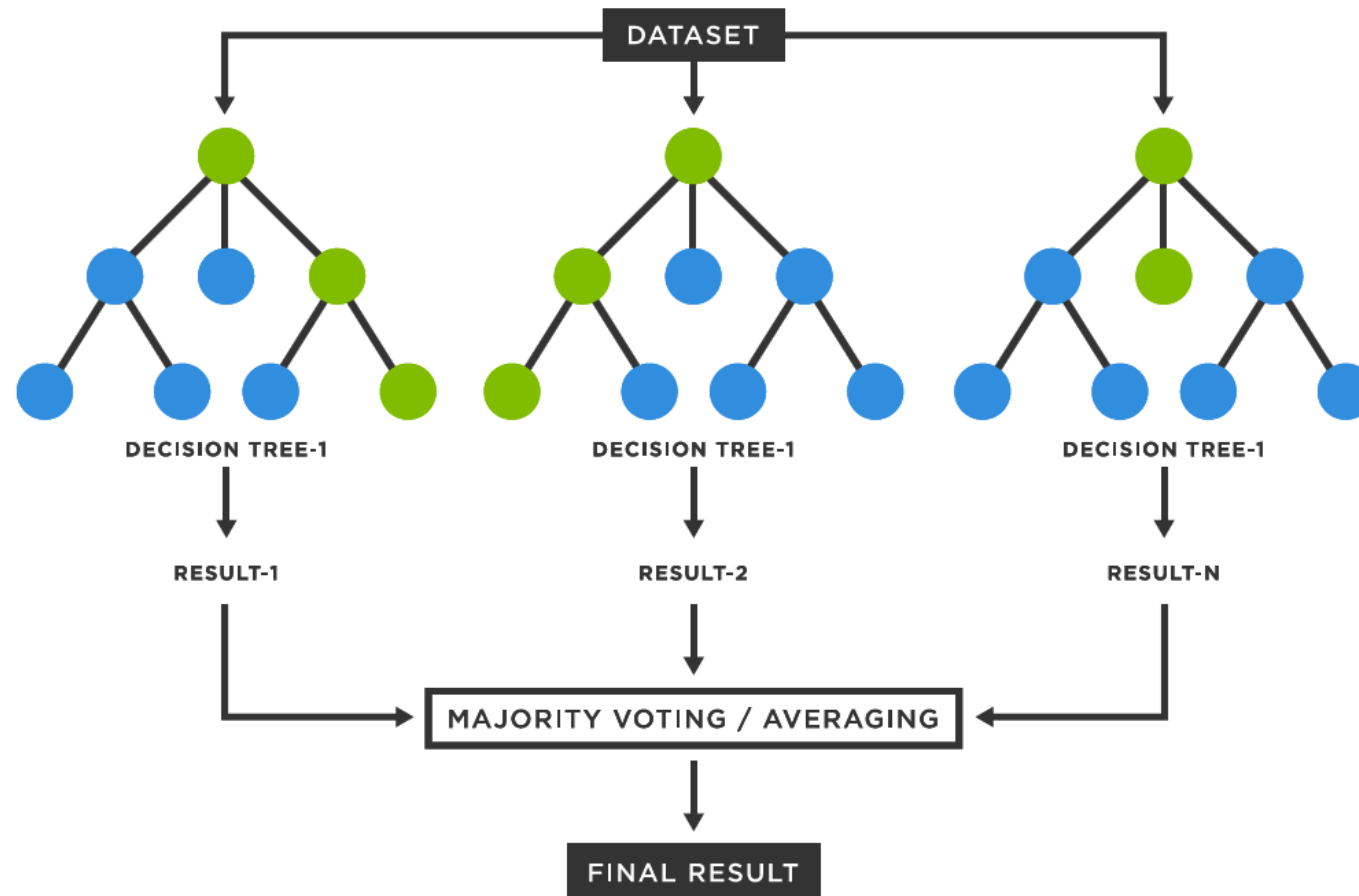
**Random forest:** Uma extensão das árvores de decisão, o Random Forest cria uma "floresta" de árvores de decisão, onde cada árvore é treinada em diferentes subconjuntos dos dados. As previsões de todas as árvores são combinadas, resultando em um modelo robusto e menos propenso ao overfitting.



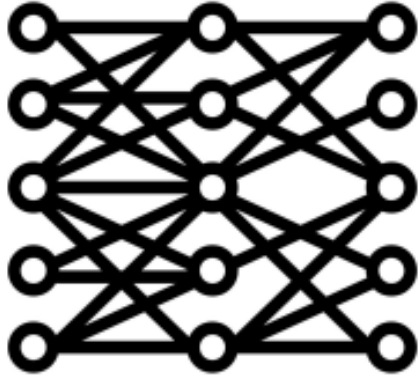
## 2.1. Visão geral



## 2.1. Visão geral



## 2.1. Visão geral



**Redes neurais artificiais:** São compostas por camadas de nós interconectados. Cada nó processa uma entrada e transmite um sinal modificado para a próxima camada. Redes neurais são extremamente flexíveis e podem modelar relações não lineares complexas, sendo a base para técnicas mais avançadas, como deep learning.

Cada um desses algoritmos tem suas próprias vantagens e desvantagens, e a escolha do algoritmo certo **depende de fatores como o tipo de problema, a natureza dos dados**, e a necessidade de interpretabilidade versus performance.



## 2.2. Escolha do algoritmo para o exemplo



Após conhecer alguns algoritmos, como decidir qual deles utilizar? Existem alguns fatores que ajudam a decidir...

- o Artigos científicos
- o Características dos algoritmos
- o Simplicidade de implementação

### Detecting Illicit Entities in Bitcoin using Supervised Learning of Ensemble Decision Trees

Pranav Nerurkar  
Dept of CE & IT, VJTI  
Dept of Data Science, NMIMS  
Mumbai, Maharashtra  
+91 9619997797

panerurkar\_p16@ce.vjti.ac.in

Yann Busnel  
SRCD Department  
IMT Atlantique  
Rennes, France  
+33 299127000

yann.busnel@imt-atlantique.fr

Romarc Ludinard  
SRCD Department  
IMT Atlantique  
Rennes, France  
+33 299127000

romarc.ludinard@imt-atlantique.fr

Kunjal Shah  
Dept of CE & IT, VJTI  
Mumbai, Maharashtra  
+91 9819027140  
kunjal1999@gmail.com

Sunil Bhirud  
Dept of CE & IT, VJTI  
Mumbai, Maharashtra  
+91 2224198101  
sgbhirud@ce.vjti.ac.in

Dhiren Patel  
Dept of CE & IT, VJTI  
Mumbai, Maharashtra  
+91 2224198101  
dhiren29p@gmail.com

#### ABSTRACT

Since its inception in 2009, Bitcoin has been mired in controversies for providing a haven for illegal activities. Several types of illicit users hide behind the blanket of anonymity. Uncovering these entities is key for forensic investigations. Current methods utilize machine learning for identifying these illicit entities. However, the existing approaches only focus on a limited category of illicit users. The current paper proposes to address the issue by implementing an ensemble of decision trees for supervised learning. More parameters allow the ensemble model to learn discriminating features that can categorize multiple groups of illicit users from licit users. To evaluate the model, a dataset of 2059 real-life entities on Bitcoin was extracted from the Blockchain. Nine features were engineered to train the model for segregating 28 different licit-illicit categories of users. The proposed model provided a reliable tool for forensic study. Empirical evaluation of the proposed model vis-a-vis three existing benchmark models was performed to highlight its efficacy. Experiments showed that the specificity and sensitivity of the proposed model were comparable to other models.

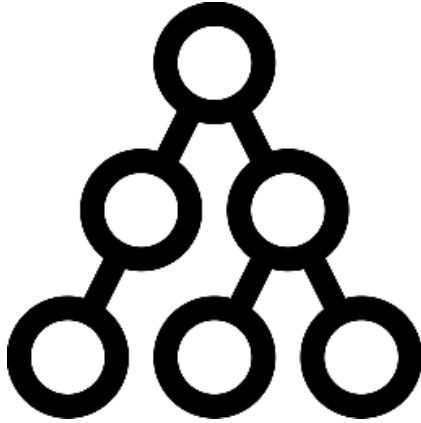
#### 1. INTRODUCTION

Bitcoin<sup>1</sup> platform has attracted anti-social elements [1] as it creates hurdles for law enforcement to trace suspicious transactions due to the anonymity and privacy [2]. As bitcoin became financially significant, the emergence of Ponzi schemes, money laundering, frauds, embezzlements, extortion, and tax evasion [3] practices were seen. These businesses used the blanket of secrecy afforded by Bitcoin to mislead the audit trail. It was speculated that in 2017, BTCs worth \$770 million were traded for illicit activities [4], a quarter of bitcoin users were malicious and 46% of all bitcoin activity was illegal [5].

Due to voluminous data generated about bitcoin transactions on the Blockchain, machine learning became a popular technique for tracking and scrutinizing illicit users or transactions. Existing literature surveyed on detecting illegal activities using Machine Learning (ML) had focused on illegal transactions, identifying suspicious bitcoin users (extortionists, ponzi scams, darknet markets, ransomware, human traffickers, frauds), detecting money laundering, identifying mixing services, identifying bitcoin exchanges, identifying illegal transactions, identifying bitcoin

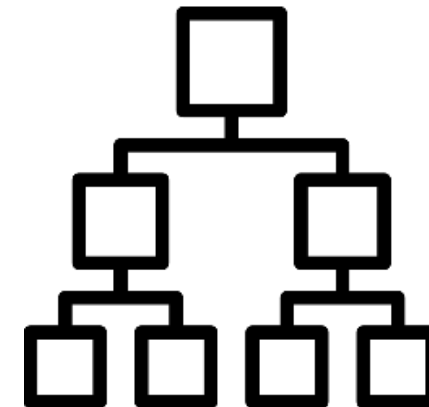


## 2.2. Escolha do algoritmo para o exemplo



Árvores de decisão são relativamente fáceis de treinar e implementar. Isso é particularmente importante em ambientes de forense digital, onde há uma necessidade constante de desenvolver e ajustar rapidamente modelos para responder a novas ameaças ou mudanças no comportamento das transações.

Na forense de criptomoedas, pode haver uma diferença significativa entre o número de transações legais e ilegais. Árvores de decisão, especialmente quando combinadas com técnicas como o "random forest", podem lidar bem com esse desbalanceamento, ajustando o peso das classes e garantindo que o modelo permaneça eficaz.



**3**

## **PREPARAÇÃO DOS DADOS**

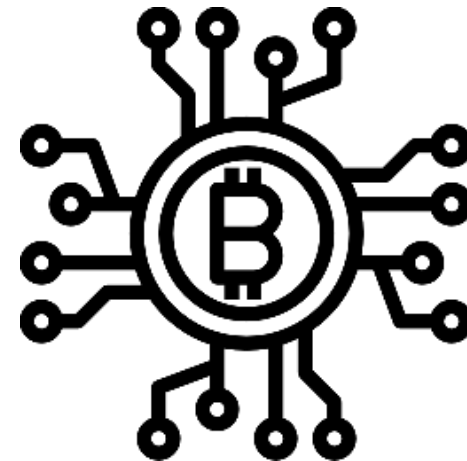
## 3.1. Coleta e pré-processamento



A primeira etapa na aplicação de machine learning é a **coleta e o pré-processamento** dos dados. A qualidade dos dados afeta diretamente a eficácia dos modelos de machine learning, tornando essa fase fundamental para o sucesso de qualquer análise.

Existem algumas fontes de dados disponíveis para obter informações sobre criptomoedas, como:

- o Blockchain
- o Exploradores de blockchain
- o Exchanges



## 3.1. Coleta e pré-processamento



The screenshot shows a web browser window with the address bar displaying "walletexplorer.com". The website has a dark blue header with the text "WalletExplorer.com: smart Bitcoin block explorer" in yellow and white. To the right of the header is a search bar with the placeholder text "Search address/txid/wallet id/xpub/firstbits". Below the header, the main content area has a large heading "Bitcoin block explorer with address grouping and wallet labeling" in green. Underneath this heading is a text input field with the placeholder text "Enter address, txid, [firstbits](#) (first address characters), first txid characters, XPUB/YPUB/ZPUB, internal wallet id, or service name:". Below the input field is a blue "Search" button. At the bottom of the page, there is a paragraph of text: "New: Searching by XPUB is much improved! Now it supports all XPUB formats, it scans all derivation paths, and all address types, it is much faster and it works even for very large wallets. 'Transaction view' for an XPUB is also more usable."

WalletExplorer.com: smart Bitcoin block explorer

Search address/txid/wallet id/xpub/firstbits

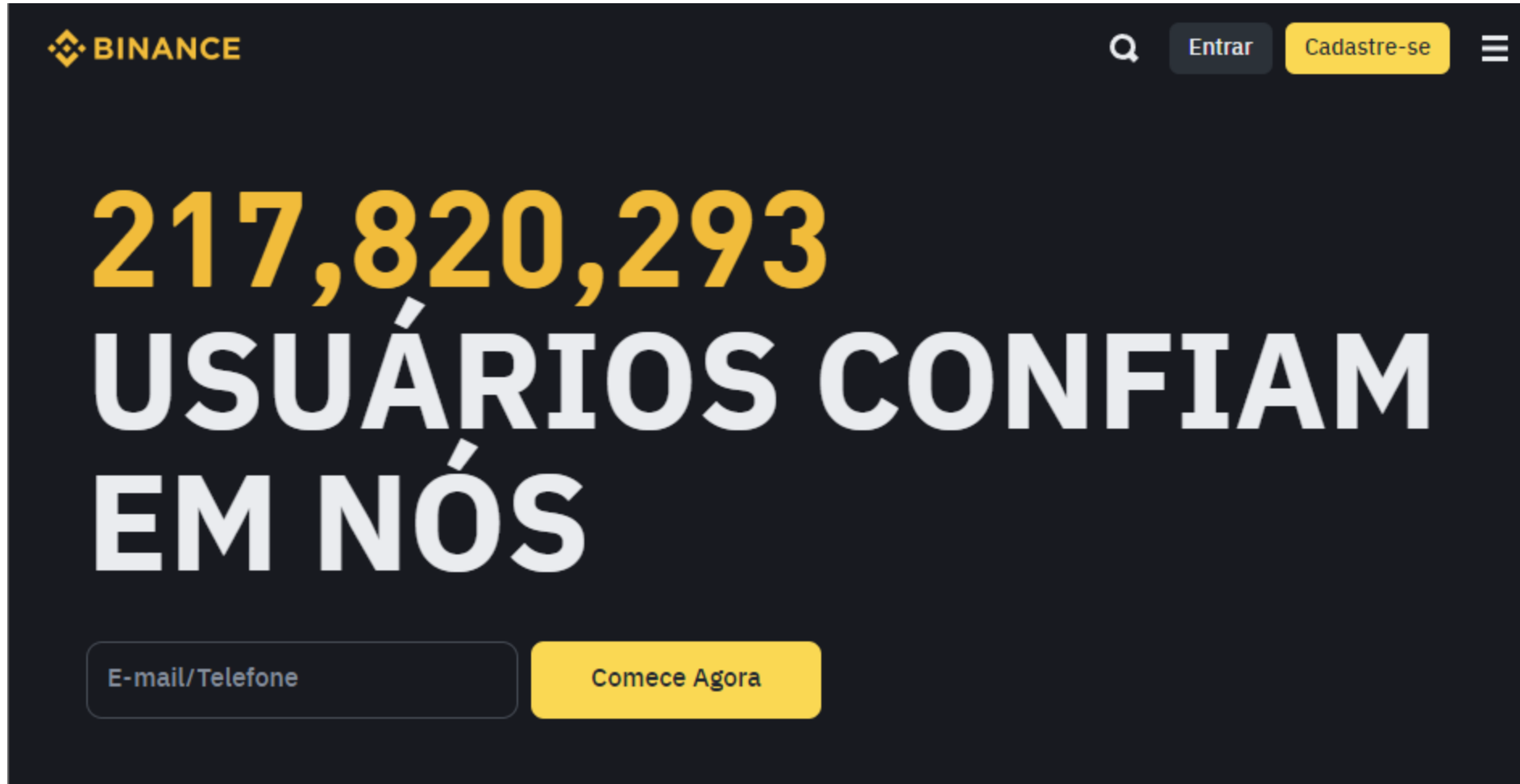
### Bitcoin block explorer with address grouping and wallet labeling

Enter address, txid, [firstbits](#) (first address characters), first txid characters, XPUB/YPUB/ZPUB, internal wallet id, or service name:

Search

New: Searching by XPUB is much improved! Now it supports all XPUB formats, it scans all derivation paths, and all address types, it is much faster and it works even for very large wallets. "Transaction view" for an XPUB is also more usable.

## 3.1. Coleta e pré-processamento



## 3.1. Coleta e pré-processamento

 **Blockchain.com**

 Início


 Preços


 Gráficos

 NFTs

 DeFi

 Academia

 Notícias

 Programadores


 Wallet

 Exchange






 Sign In





Sponsored:  
**WIN BIG**  
with  
**Leicester city!** [Play Now at BC.GAME](#)

**Win 8.88 BTC** 

**1000x BTC BONUS** 

**Play Slots & Win!** 

USD

 **Bitcoin BTC**  
\$59.157,97 -0.64% -382,18 



**10.965**  
Transações • 0.13 TPs

**850.073**  
Blocos • Último 4m17s

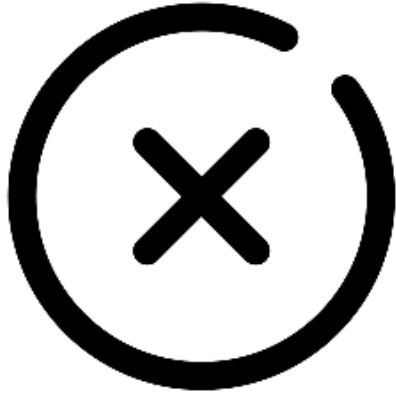
**\$7.241.529**  
Enviado hoje

**8.32 EH/s**  
Hashrate de rede

“The attacker isn't adding blocks to the end. He has to go back and redo the block his transaction is in and all the blocks after it, as well as any ne...” [\(Read More\)](#)

*Satoshi Nakamoto*  
Email • nov. de 2008

## 3.1. Coleta e pré-processamento



Mesmo com diversas fontes de dados disponíveis, encontrar uma base que possua exemplos de transações lícitas e ilícitas classificadas é um grande desafio.

Como resolver esse problema?

Elliptic++ Dataset

<https://github.com/git-disl/EllipticPlusPlus>

### Elliptic++ Dataset: A Graph Network of Bitcoin Blockchain Transactions and Wallet Addresses

The Elliptic++ dataset consists of 203k Bitcoin transactions and 822k wallet addresses to enable both the detection of fraudulent transactions and the detection of illicit addresses (actors) in the Bitcoin network by leveraging graph data.

If you have any questions or create something with this dataset, please let us know by email: [yelmougy3@gatech.edu](mailto:yelmougy3@gatech.edu).

DATASET CAN BE FOUND HERE: [Google Drive](#)

#### Dataset Summary

The Elliptic++ dataset contains a transactions dataset and an actors (wallet addresses) dataset.

Elliptic++ Transactions Dataset:

# Nodes (transactions)	203,769
# Edges (money flow)	234,355
# Time steps	49
# Illicit (class-1)	4,545
# Licit (class-2)	42,019
# Unknown (class-3)	157,205
# Features	183

## 3.1. Coleta e pré-processamento



Desenvolvido pela [Elliptic](#), uma empresa especializada em análise e segurança de blockchain. O objetivo principal do Elliptic++ dataset é [facilitar a pesquisa e o desenvolvimento de métodos para identificar e classificar transações ilícitas na blockchain do Bitcoin](#).

### Elliptic++ Dataset: A Graph Network of Bitcoin Blockchain Transactions and Wallet Addresses

The Elliptic++ dataset consists of 203k Bitcoin transactions and 822k wallet addresses to enable both the detection of fraudulent transactions and the detection of illicit addresses (actors) in the Bitcoin network by leveraging graph data.

If you have any questions or create something with this dataset, please let us know by email: [yelmougy3@gatech.edu](mailto:yelmougy3@gatech.edu).

DATASET CAN BE FOUND HERE: [Google Drive](#)

#### Dataset Summary

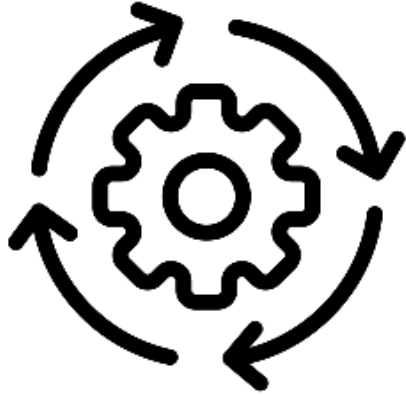
The Elliptic++ dataset contains a transactions dataset and an actors (wallet addresses) dataset.

Elliptic++ Transactions Dataset:

# Nodes (transactions)	203,769
# Edges (money flow)	234,355
# Time steps	49
# Illicit (class-1)	4,545
# Licit (class-2)	42,019
# Unknown (class-3)	157,205
# Features	183



## 3.1. Coleta e pré-processamento



As etapas de **pré-processamento** envolvem:

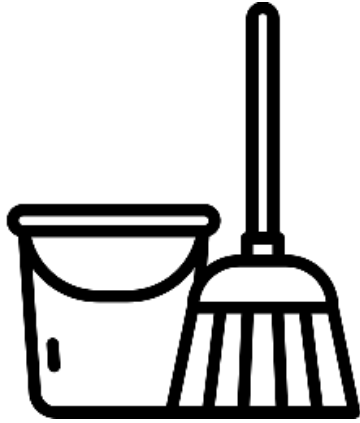
- o Limpeza dos dados
- o Normalização e padronização
- o Transformação dos dados
- o Anonimização e privacidade

No exemplo prático não foi necessário pré-processar os dados, pois o dataset Elliptic++ já se encontra pronto para uso.

**Observação:** Mesmo que o dataset seja de alta qualidade, é importante verificar se existe alguma inconsistência.



## 3.2. Limpeza e transformação dos dados

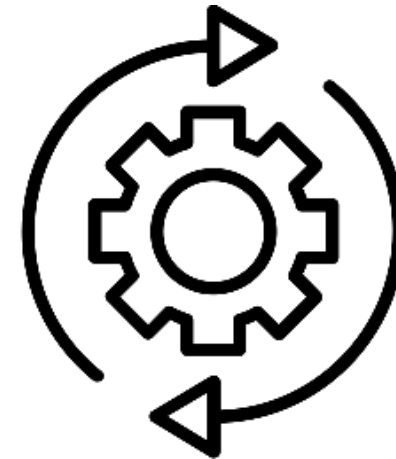


As técnicas de **limpeza** dos dados envolvem:

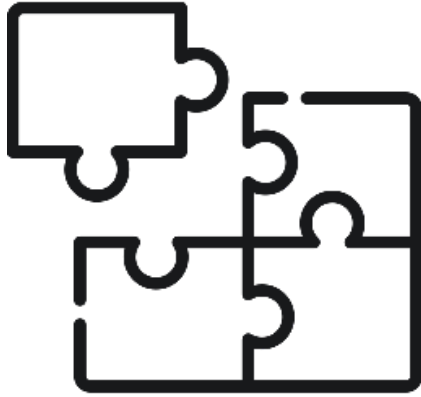
- o Tratamento de dados faltantes
- o Remoção de duplicatas
- o Correção de erros e inconsistências
- o Tratamento de outliers

As técnicas de **transformação** dos dados envolvem:

- o Normalização e padronização
- o Codificação de dados categóricos
- o Redução de dimensionalidade
- o Criação de novas features



### 3.3. Feature engineering



Feature engineering é a prática de usar **conhecimento do domínio do problema** para selecionar, criar ou transformar dados brutos em features que **melhorem a performance dos modelos** de machine learning. As features são os atributos ou características que o modelo usa para aprender padrões e fazer previsões.

A qualidade das features pode ser mais **determinante para o sucesso** de um modelo de machine learning do que a escolha do algoritmo em si. Isso ocorre porque algoritmos de machine learning **baseiam-se nas features para identificar padrões nos dados**.



### 3.3. Feature engineering



Além de criar e transformar features, a feature engineering também envolve a [seleção das features mais relevantes](#) para o problema. Algumas técnicas comuns incluem:

- o Análise de correlação
- o Seleção de features
- o Eliminação de features irrelevantes

Feature Engineering não é uma tarefa trivial e envolve vários desafios, como:

- o [Conhecimento do domínio](#)
- o Overfitting
- o Tempo e custo computacional



## 3.4. Qualidade dos dados



A qualidade dos dados refere-se à medida em que os dados atendem aos **critérios de precisão, consistência, integridade e relevância**. Dados de alta qualidade são aqueles que são corretos, completos, atualizados e relevantes para o propósito da análise.

A qualidade dos dados **impacta diretamente a eficácia das análises e a precisão dos modelos** de machine learning. Dados de baixa qualidade podem levar a uma série de problemas, incluindo:

- o Decisões erradas
- o Modelos incorretos
- o Aumento de custos
- o Impacto na reputação





## **EXEMPLO PRÁTICO: ÁRVORE DE DECISÃO**

## 4.1. Descrição do problema



O principal desafio é [distinguir transações legítimas das ilegítimas ou suspeitas](#). Dada a natureza pseudônima do Bitcoin, essa tarefa torna-se complexa. As transações ilegais são cuidadosamente ocultas entre milhões de transações legítimas, dificultando a identificação através de métodos tradicionais de análise.

Para atingir o objetivo, será construído um [modelo simplificado de árvore de decisão](#) utilizando o [dataset Elliptic++](#). Como visto anteriormente, o dataset possui rótulos (ilegal, legal, desconhecido) e a árvore de decisão é fácil de implementar.



## 4.2. Construção do modelo

### Google Colab

[https://colab.research.google.com/  
drive/1hF0rIYm\\_dkUPPNOcKOnEt4I3DNwWp3UQ?authuser=1  
#scrollTo=iTlqEpj3j9Hw](https://colab.research.google.com/drive/1hF0rIYm_dkUPPNOcKOnEt4I3DNwWp3UQ?authuser=1#scrollTo=iTlqEpj3j9Hw)



## 4.3. Métricas de avaliação

### Acurácia

A acurácia é uma das métricas mais básicas e amplamente utilizadas para avaliar modelos de classificação. Ela é definida como a proporção de previsões corretas (tanto positivas quanto negativas) em relação ao total de previsões feitas pelo modelo. Matematicamente, a acurácia é expressa como:

$$Acuracia = \frac{Total\ de\ previsoes\ corretas}{Total\ de\ previsoes}$$

Embora a acurácia seja útil para ter uma visão geral do desempenho do modelo, ela [pode ser enganosa em conjuntos de dados desbalanceados](#), onde uma classe pode ser muito mais frequente que as outras. Nesse caso, o modelo pode ter uma alta acurácia simplesmente por prever a classe majoritária na maioria das vezes, sem ser realmente eficaz na detecção de classes minoritárias, como transações ilegais.

## 4.3. Métricas de avaliação

### Matriz de confusão

A matriz de confusão é uma ferramenta que fornece uma visão detalhada das previsões do modelo, mostrando a contagem de verdadeiros positivos, verdadeiros negativos, falsos positivos e falsos negativos. Ela é particularmente útil para identificar os tipos de erros que o modelo está cometendo e pode guiar ajustes no modelo ou na escolha de métricas.

- o Verdadeiros Positivos (VP): Transações ilegais corretamente classificadas como ilegais.
- o Verdadeiros Negativos (VN): Transações legais corretamente classificadas como legais.
- o Falsos Positivos (FP): Transações legais incorretamente classificadas como ilegais.
- o Falsos Negativos (FN): Transações ilegais incorretamente classificadas como legais.

## 4.3. Métricas de avaliação

### Precisão

A precisão mede a proporção de previsões positivas corretas em relação ao total de previsões positivas feitas pelo modelo. Ela responde à pergunta: "Dentre todas as transações que o modelo classificou como ilegais, quantas realmente eram ilegais?"

$$Precisao = \frac{Verdadeiros\ positivos}{Verdadeiros\ positivos + Falsos\ positivos}$$

Alta precisão significa que há poucos falsos positivos, ou seja, poucas transações classificadas erroneamente como ilegais.

## 4.3. Métricas de avaliação

### Recall

O recall mede a proporção de verdadeiros positivos em relação ao total de casos reais da classe positiva. Ela responde à pergunta: "Dentre todas as transações realmente ilegais, quantas o modelo conseguiu identificar?"

$$\text{Precisao} = \frac{\text{Verdadeiros positivos}}{\text{Verdadeiros positivos} + \text{Falsos negativos}}$$

Alto recall significa que o modelo **é eficaz em capturar a maioria das transações ilegais**, mesmo que isso resulte em alguns falsos positivos.

## 4.3. Métricas de avaliação

### F1-score

O F1-score é a média harmônica da precisão e do recall. Ele é útil quando há um trade-off entre precisão e recall e é necessário um equilíbrio entre os dois. O F1-score é especialmente relevante em casos de dados desbalanceados, onde ambas as métricas são importantes.

$$F1 - score = 2 \times \frac{Precisao \times Recall}{Precisao + Recall}$$

Um F1-score alto indica que o modelo **tem um bom equilíbrio entre precisão e recall**, sendo capaz de identificar transações ilegais com precisão e abrangência.

# 4.4. Resultados e discussões

Relatório de classificação:				
		precision	recall	f1-score
Classes:				
1 - Ilícita	1	0.55	0.58	0.57
2 - Lícita	2	0.66	0.68	0.67
3 - Desconhecida	3	0.91	0.90	0.90



## **DESAFIOS E CONSIDERAÇÕES**

## 5.1. Desafios técnicos e limitações

1

Encontrar uma base de dados que contenha uma quantidade de rótulos balanceada (anonimidade das criptomoedas).

2

A construção de um modelo que seja robusto o suficiente também é desafio considerável. Lidar com a diversidade de usos das criptomoedas não é uma tarefa fácil.



## 5.2. Empresas que realizam investigações



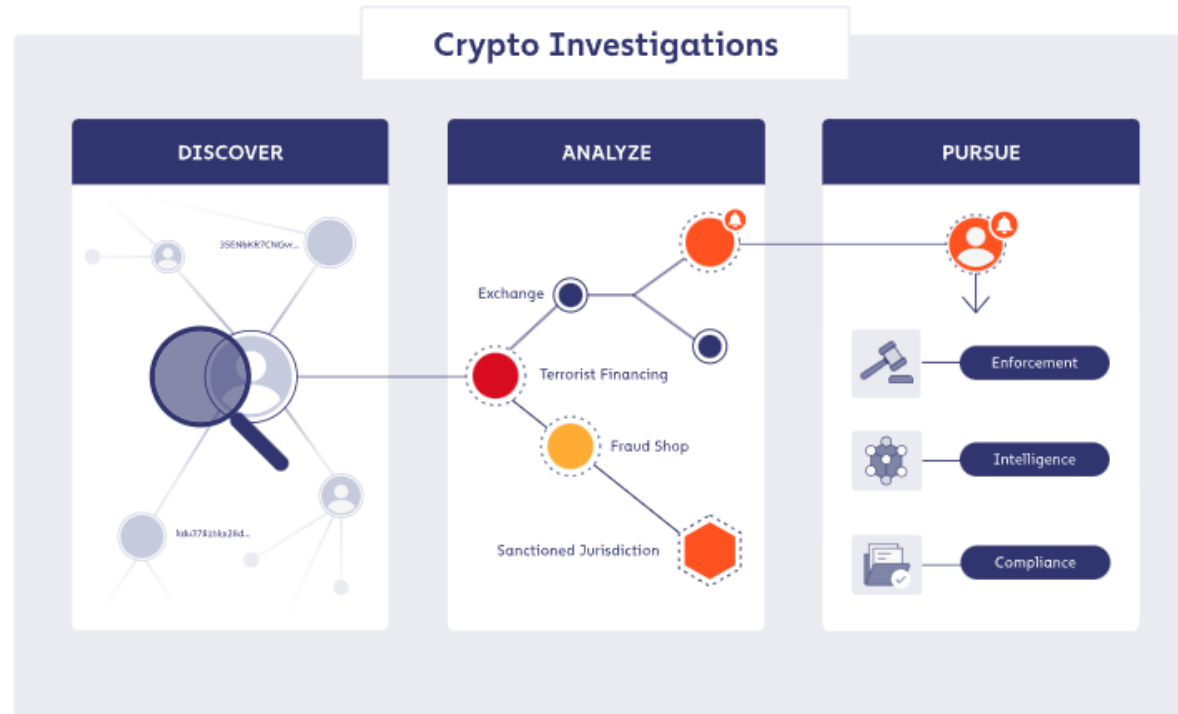
=

Crypto Investigations

Discover

Analyze

Pursue



## 5.2. Empresas que realizam investigações

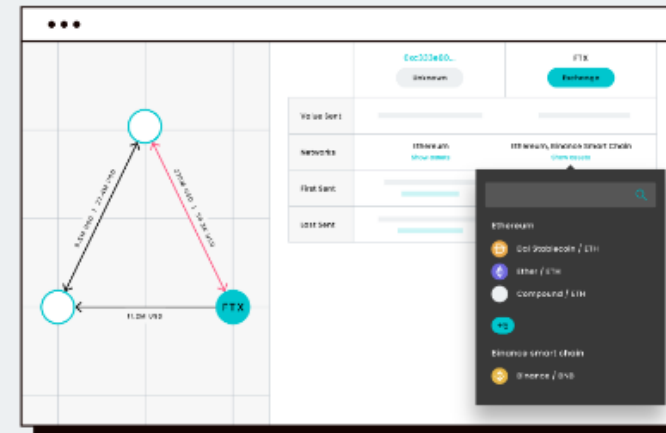
ELLIPTIC



[Home](#) / [Crypto Investigations](#)

### Elliptic Investigator

Conduct single-click investigations across blockchains and assets with ease. Instantly visualize the flow of crypto funds through wallets, entities and transactions to find meaningful evidence quickly and reduce the time and resources needed to close cases.



**Cross-chain Crypto Investigations You Can Trust**

## 5.2. Empresas que realizam investigações



Crypto Investigators

<https://cryptoinvestigators.com> · Traduzir esta página ·

### Crypto Investigators - Blockchain, Crypto Assets, CBDC, Web3 ...

Crypto Investigators is a world leader in **cryptocurrency investigations** for the purpose of crypto assets recovery, civil litigation, criminal complaints, and ...



Kroll

<https://www.kroll.com/services> · Traduzir esta página ·

### Cryptocurrency Risk, Investigation and Compliance Services

Kroll is the leading global provider of **crypto** compliance, risk, and **investigative** services. Since the introduction of the first virtual asset in 2009, ...



CipherBlade

<https://cipherblade.com> · Traduzir esta página ·

### CipherBlade: Blockchain Investigation and Expert Agency

CipherBlade conducts **cryptocurrency-forensic investigations**, assists with **cryptocurrency** recovery, and provides **cryptocurrency** expert witness services.



crystalintelligence.com

<https://crystalintelligence.com> · Traduzir esta página ·

### Crystal Investigations - Crystal Intelligence Analytics for Crypto ...

We work with private and public sectors to trace suspicious **crypto** transactions and uncover real-world identities. We also prepare a range of reports for audits ...



**TRABALHOS FUTUROS**

## 6.1. Linhas de pesquisa

### Detecting anomalous **cryptocurrency** transactions: An AML/CFT application of **machine learning**-based **forensics**

[N Pocher](#), [M Zichichi](#), [F Merizzi](#), [MZ Shafiq](#), [S Ferretti](#) - Electronic Markets, 2023 - Springer

... While we do not aim to offer a review of the techniques of **cryptocurrency forensics**, in this section, we describe a few works that provided an application to the concepts introduced above...

☆ Salvar Citar Citado por 23 Artigos relacionados Todas as 9 versões

### Anti-money laundering in **bitcoin**: Experimenting with graph convolutional networks for financial **forensics**

[M Weber](#), [G Domeniconi](#), [J Chen](#), [DKI Weidele](#)... - arXiv preprint arXiv ..., 2019 - arxiv.org

... While this approach shows the graph structure carries in the binary classification problem, and that this can be used with standard **machine learning** techniques, it is challenging to ...

☆ Salvar Citar Citado por 442 Artigos relacionados Todas as 7 versões

### **Crypto**-preserving investigation framework for deep **learning** based malware attack detection for network **forensics**

[S Bhardwaj](#), [M Dave](#) - Wireless Personal Communications, 2022 - Springer

... In this article, a **crypto**-evidence preservation and evidence collecting model is proposed. ... **learning** and **machine learning** classifiers. The various studies have shown that deep **learning** ...

☆ Salvar Citar Citado por 17 Artigos relacionados Todas as 3 versões

### Supervised **learning** model for identifying illegal activities in **Bitcoin**

[P Nerurkar](#), [S Bhirud](#), [D Patel](#), [R Ludinard](#), [Y Busnel](#)... - Applied ..., 2021 - Springer

... **forensic** investigation. However, issues that **machine learning** models face in **Bitcoin forensics**

... Lack of ground truth labeled information is seen in **Bitcoin**, which is not observed in other ...

☆ Salvar Citar Citado por 56 Artigos relacionados Todas as 7 versões

### Demystifying fraudulent transactions and illicit nodes in the **bitcoin** network for financial **forensics**

[Y Elmougy](#), [L Liu](#) - Proceedings of the 29th ACM SIGKDD Conference ..., 2023 - dl.acm.org

... transactions and **BTC** flow from a ... of **Bitcoin** addresses representing unique **Bitcoin** users. Second, we perform fraud detection tasks on all four graphs by using diverse **machine learning** ...

☆ Salvar Citar Citado por 6 Artigos relacionados Todas as 3 versões



**DÚVIDAS**

Programa de Pós-graduação em Ciência da Computação  
Faculdade de Computação

# MACHINE LEARNING APLICADO A FORENSE DE CRIPTOMOEDAS

**Aluno:** Pedro Henrique Resende Ribeiro

**Orientador:** Rodrigo Sanches Miani

**Coorientador:** Ivan da Silva Sendin