

## 1) Título

Predição de Preços de Criptomoedas Utilizando Modelos de Séries Temporais Lineares, Não Lineares e Híbridos

## 2) Integrantes

- **Nome:** Marcello Linard Teixeira
- **RA:** 10419338
- **E-mail:** [Seu e-mail aqui]

## 3) Resumo

Este projeto visa desenvolver e avaliar modelos de Inteligência Artificial para a predição de preços de criptomoedas. O escopo foca na análise comparativa de três famílias de modelos de séries temporais: lineares (ARIMA), não lineares (LSTM) e híbridos (ARIMA-LSTM). Os dados históricos do Bitcoin (BTC) foram obtidos em formato **.csv** a partir do portal **investing.com**. A metodologia emprega a técnica de *sliding windows* para a criação das janelas de dados de entrada. O desempenho de cada modelo será comparado quantitativamente utilizando a métrica *Root Mean Squared Error* (RMSE), permitindo a identificação da abordagem preditiva mais eficaz para a série temporal analisada.

## 4) Introdução

**a. Contextualização** O mercado de criptomoedas é conhecido por sua alta volatilidade e pelo crescente interesse de investidores e pesquisadores. A natureza dinâmica e complexa dos preços desses ativos digitais representa um desafio significativo para a previsão, tornando-se um campo fértil para a aplicação de técnicas avançadas de Inteligência Artificial.

**b. Justificativa** A capacidade de prever, mesmo que com margem de erro, a direção dos preços de criptomoedas pode oferecer insights valiosos. Este projeto se justifica pela relevância prática de aplicar modelos de IA a um problema financeiro real e complexo, explorando a capacidade de algoritmos de aprendizado de máquina em identificar padrões em séries temporais financeiras ruidosas.

**c. Objetivo** O objetivo principal é construir, comparar e avaliar modelos preditivos de IA. Os objetivos específicos são: obter um dataset histórico do Bitcoin, realizar seu pré-processamento e análise exploratória, implementar e treinar modelos representativos das famílias linear, não linear e híbrida, e por fim, avaliar comparativamente o desempenho dos modelos para identificar o mais eficaz.

**d. Opção do projeto** ML/DL – Aplicação de frameworks de Machine Learning/Deep Learning para problemas de regressão/predição.

## 5) Fundamentação Teórica (Resumida)

Este projeto se baseia em conceitos de séries temporais e aprendizado de máquina. Uma série temporal é uma sequência de pontos de dados indexados no tempo. A técnica de *sliding windows* (janelas deslizantes) é utilizada para transformar a série em um problema de aprendizado supervisionado. Serão exploradas três famílias de modelos: **Modelos Lineares** (como o ARIMA, que captura relações lineares), **Modelos Não Lineares** (como a rede neural LSTM, eficaz em aprender dependências de longo prazo) e **Modelos Híbridos** (que combinam as forças de ambos). A avaliação será feita pela métrica **RMSE (Root Mean Squared Error)**, que penaliza erros maiores.

## 6) Descrição do Problema

O problema a ser resolvido é o de predição de séries temporais univariadas. Utilizando a técnica de *sliding windows*, dado uma sequência de preços de fechamento diário de uma criptomoeda nos últimos N dias, o objetivo é desenvolver um modelo  $f$  que preveja o preço de fechamento no dia seguinte, minimizando o erro entre o valor real e o previsto.

## 7) Aspectos Éticos do uso da IA e sua Responsabilidade

O desenvolvimento de uma solução de IA para predição financeira carrega responsabilidades éticas. É crucial destacar que os resultados não são aconselhamento financeiro, dada a volatilidade do mercado. A solução deve ser transparente quanto às suas limitações, documentando claramente as métricas de erro para evitar uma falsa sensação de precisão. Além disso, a metodologia deve incluir a separação rigorosa de dados de treino e teste para mitigar o risco de sobreajuste (*overfitting*) e fornecer uma avaliação honesta do desempenho do modelo.

## 8) Dataset

O dataset foi obtido manualmente do portal [investing.com](https://investing.com), contendo o histórico diário de preços (Abertura, Máxima, Mínima, Fechamento) e Volume do Bitcoin (BTC) de julho de 2014 a setembro de 2025, totalizando 4099 registros. A análise exploratória e a preparação dos dados foram realizadas em Python, utilizando a biblioteca Pandas, e estão documentadas no notebook [01\\_Exploratory\\_Data\\_Analysis.ipynb](#). As etapas incluíram a conversão de tipos de dados, limpeza de valores numéricos e textuais, e ordenação cronológica. Para a modelagem, os dados de preço foram normalizados na escala de 0 a 1 e transformados em janelas de 60 dias para servirem de entrada aos modelos.

## 9) Metodologia e Resultados Esperados

A metodologia para a segunda fase do projeto seguirá os seguintes passos:

1. Carregamento dos dados pré-processados, salvos na primeira etapa.

2. Implementação e treinamento de um modelo ARIMA sobre os dados não normalizados.
3. Implementação e treinamento de uma rede neural LSTM sobre os dados normalizados e janelados.
4. Implementação de um modelo híbrido que combina as duas abordagens.
5. Avaliação de cada modelo sobre o conjunto de teste utilizando a métrica RMSE.

Espera-se obter uma tabela comparativa com o desempenho dos três modelos. A hipótese é que o modelo híbrido apresentará o menor erro (RMSE), seguido pelo modelo LSTM, e por fim o ARIMA, devido à capacidade dos modelos mais complexos em capturar padrões não lineares presentes na série histórica do Bitcoin.

## **10) Referências**

Irei detalhar o conteúdo e em breve subo este pdf com essa secção atualizada.

## **11) Bibliografia**

Irei detalhar o conteúdo e em breve subo este pdf com essa secção atualizada.