**Relatório João Gabriel Faus & João Delomo**

## Resumo

Este projeto explora a aplicação de predição linear para a previsão de valores futuros em séries temporais financeiras. Utilizamos dados coletados de séries temporais de preços de ativos financeiros para treinar um modelo de predição linear, testando-o em diferentes escalas temporais, como diária, semanal e mensal. Os resultados obtidos foram comparados com os encontrados em um artigo científico, que utilizou modelos baseados em redes neurais dinâmicas para previsão financeira. Este relatório apresenta uma análise dos resultados e discute a viabilidade da predição linear em comparação com métodos mais avançados.

## Introdução

A previsão de séries temporais é um elemento central na análise de dados financeiros. Antecipar os movimentos de preço pode auxiliar em decisões de investimento e operações de trade, permitindo uma resposta mais eficiente às variações do mercado. Neste projeto, utilizamos um modelo de predição linear para prever valores de séries temporais financeiras, considerando uma abordagem que avalia diferentes escalas de tempo (diária, semanal e mensal). Esta análise busca avaliar se a predição linear é eficaz para capturar as tendências dos dados financeiros e se pode ser utilizada como uma ferramenta preditiva viável.

## Metodologia

### Coleta de Dados

Os dados foram obtidos da API **Yahoo Finance**, utilizando a biblioteca *yfinance* no Python. As séries temporais dos preços de fechamento de ativos financeiros foram coletadas para diferentes períodos, abrangendo escalas de 1, 5 e 10 anos. Esses dados foram divididos em conjunto de **treinamento** (70%) e **teste** (30%).

### Divisão dos Dados

Os dados coletados foram segmentados em períodos de treinamento e teste, seguindo uma divisão padrão de 70% dos dados para treino e 30% para teste. Essa divisão foi aplicada em cada uma das escalas temporais (diária, semanal e mensal) para observar a consistência do modelo em capturar tendências de curto, médio e longo prazo.

### Modelo de Predição Linear

O modelo preditivo utiliza uma abordagem de regressão linear para prever o próximo valor da série temporal com base nos valores anteriores. A fórmula geral da predição é , onde X representa uma matriz das entradas passadas e w é o vetor de pesos ajustado pelo modelo. A predição foi realizada em uma janela de variáveis anteriores, com uma otimização via descida de gradiente.

### Escalas de Análise

Para capturar variações em diferentes horizontes temporais, analisamos a série temporal em três escalas: diária, semanal e mensal. Isso permitiu observar como o modelo linear responde a dados com frequências diferentes, ajudando a avaliar a aplicabilidade do modelo para predições de curto, médio e longo prazo.

## Implementação

### Descida de Gradiente

A otimização do modelo foi realizada por meio de uma **descida de gradiente** manual, utilizando a função *autograd* para calcular os gradientes. A função de erro quadrático médio (MSE) foi escolhida como métrica de avaliação, e os pesos foram ajustados em cada iteração com base na minimização do MSE.

### Normalização dos Dados

Para garantir uma escala consistente e melhorar a estabilidade do modelo, os dados foram normalizados entre 0 e 1 antes de serem passados para o modelo de treinamento e teste.

### Métrica de Avaliação

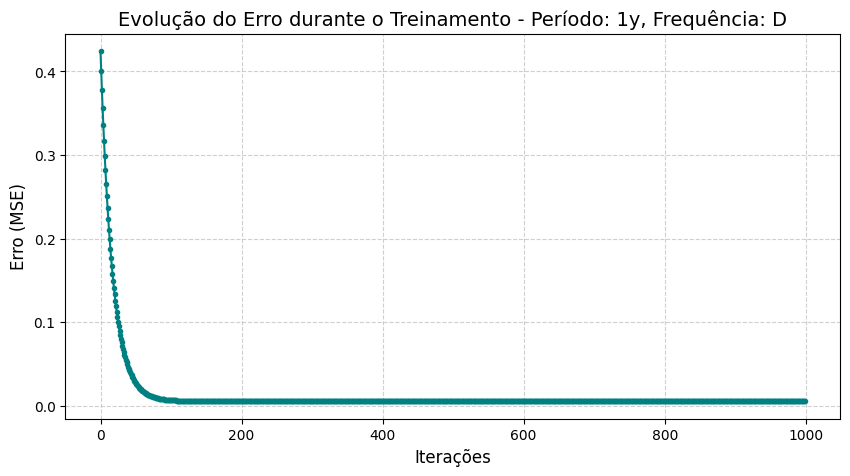
A **métrica MSE** foi utilizada para avaliar o desempenho do modelo nas predições de teste. Essa métrica indica a precisão média da predição em relação aos valores reais, ajudando a comparar o desempenho nas diferentes escalas de tempo.

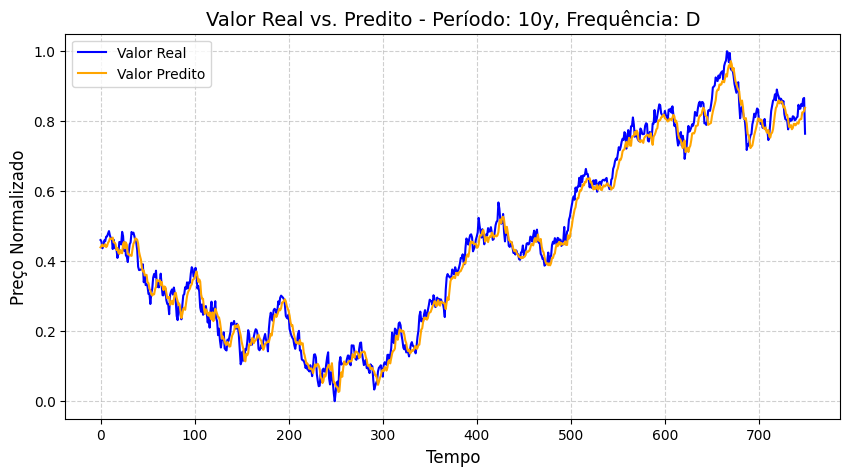
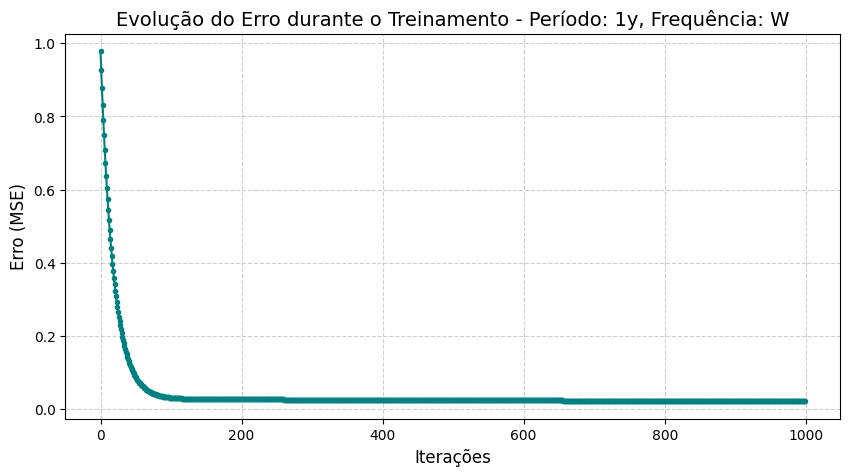
## Resultados

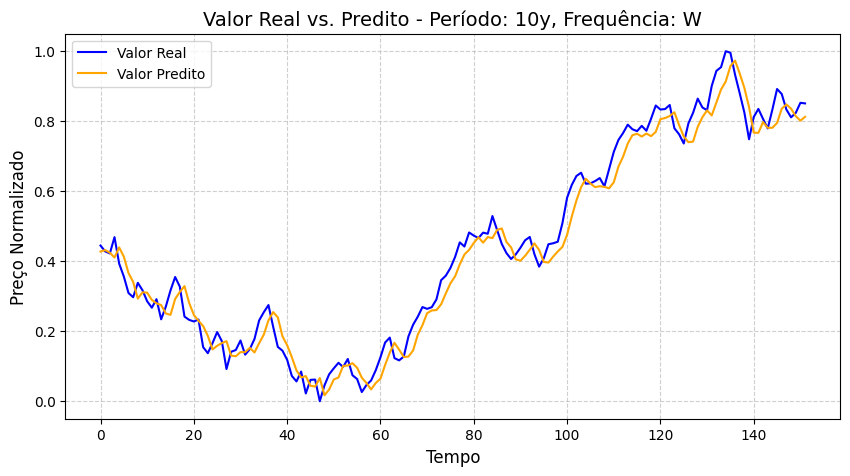
### Visualização dos Gráficos

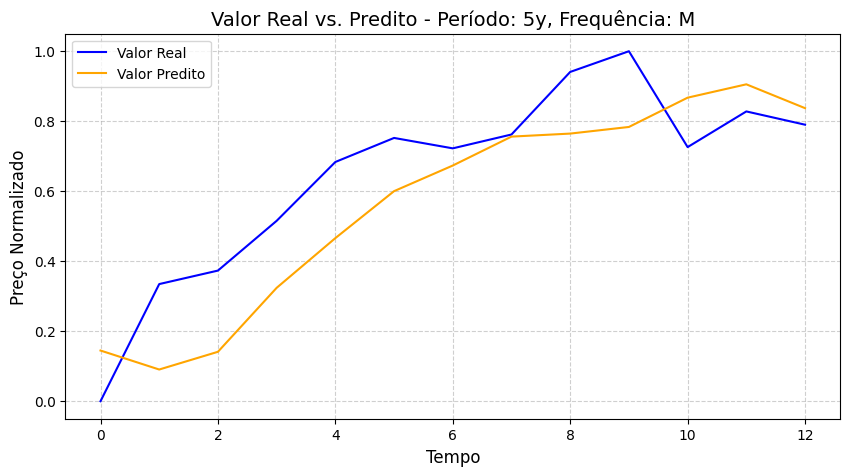
Para cada escala temporal (diária, semanal e mensal), geramos gráficos que comparam os valores reais com os valores preditos pelo modelo. Cada gráfico contém:

* A **linha dos valores reais** da série temporal no conjunto de teste.
* A **linha dos valores preditos** pelo modelo linear.
* Gráficos de **evolução do erro** durante o treinamento, que mostram a convergência do modelo em cada escala temporal.









### Análise dos Resultados

Observamos que o modelo linear apresentou diferentes níveis de precisão dependendo da escala temporal:

* Na escala **diária**, o modelo capturou flutuações menores e variações rápidas, porém com um aumento de erro.
* Na escala **semanal**, houve uma suavização das variações, revelando uma tendência mais clara em comparação com os dados diários.
* Na escala **mensal**, o modelo obteve uma linha de tendência mais estável e apresentou a menor taxa de erro, indicando melhor desempenho para previsões de longo prazo.

Essas observações sugerem que a predição linear pode ser mais eficaz em capturar tendências de médio a longo prazo, enquanto apresenta limitações em capturar flutuações de curto prazo.

## Referências

* Costa, A. F. (Ano). Previsão de séries temporais financeiras por meio de redes neurais dinâmicas e processos de transformação de dados: uma abordagem empírico-comparativa. Disponível em: <https://repositorio.unesp.br/items/b6056e26-d360-458b-8bf4-0e788022ef86/full>
* Documentação da biblioteca *yfinance*: <https://pypi.org/project/yfinance/>
* Documentação do *autograd*: <https://github.com/HIPS/autograd>