

**MBA
USP
ESALQ**

Comparação de modelos de previsão de séries temporais aplicados a bitcoin

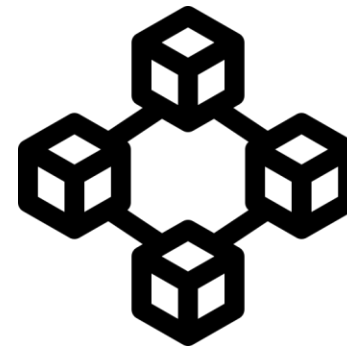
Gustavo Menezes de Sirqueira
Daniel Alvarez Firmino

Introdução

- O Bitcoin é uma criptomoeda baseada em blockchain, que por sua vez permite realização de transações sem intermediador. Desde seu surgimento, atraiu atenção de investidores e empresas devido ao seu potencial de geração de lucro. Em março de 2024 o mercado dessa criptomoeda correspondia a 1,28 trilhões de dólares.
- A volatilidade do Bitcoin torna desafiadora a previsão de seu preço, o que é necessário na tomada de decisões de investimento e na gestão de riscos.
- Modelos de previsão eficazes podem oferecer insights sobre o comportamento futuro do mercado, dessa forma, comparar o desempenho de diferentes modelos de previsão de séries temporais aplicados ao preço do Bitcoin torna-se justificado.

Objetivo:

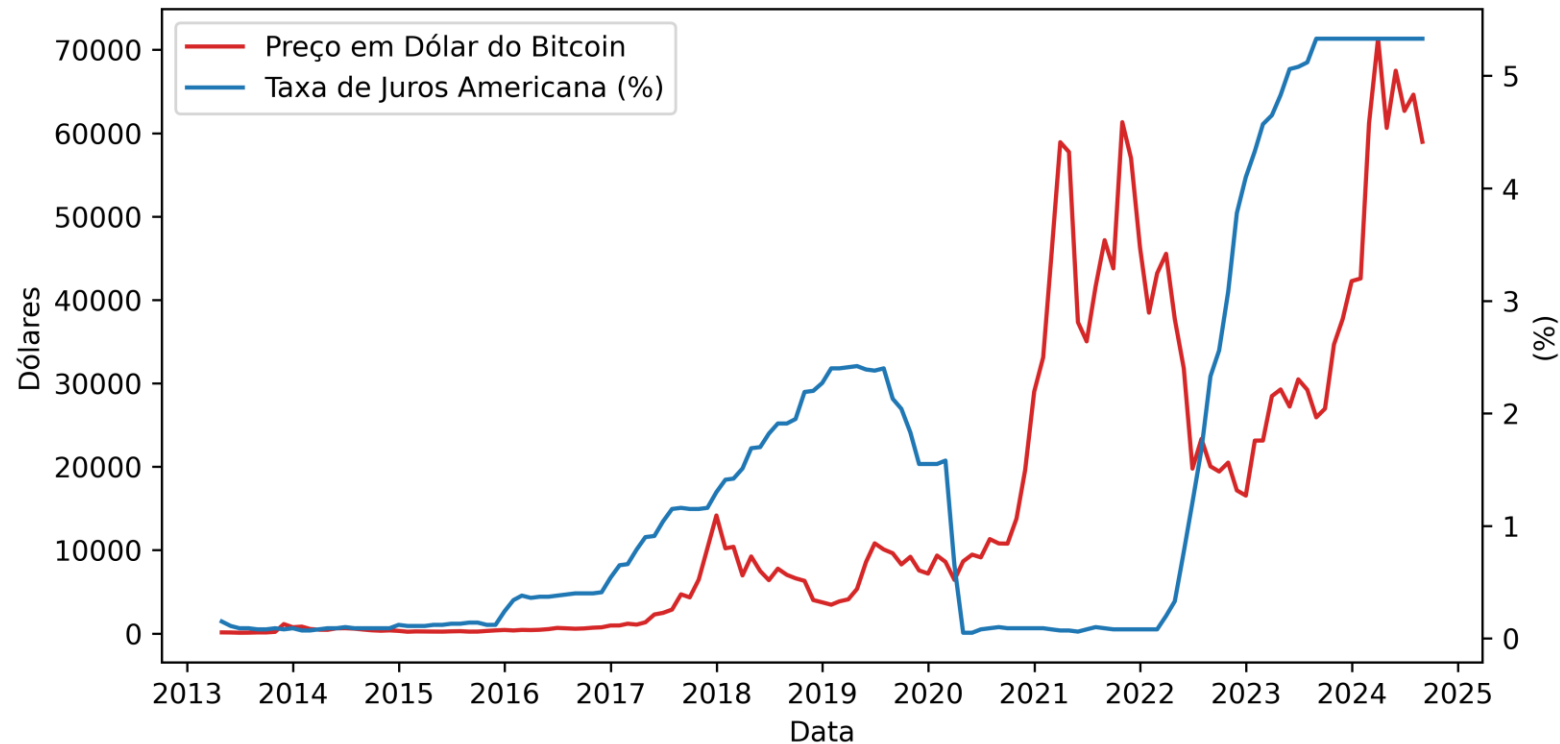
Analisar modelos tradicionais (ARIMA-GARCH) em comparação com modelos de redes neurais recorrentes (LSTM e GRU).



Materiais

Dados:

- Preço do bitcoin em dólar, frequência mensal, variável endógena, de Abril de 2013 até Agosto de 2024
- Taxa de juros americana, frequência mensal, variável exógena, de Abril de 2013 até Agosto de 2024



Materiais

Bibliotecas:

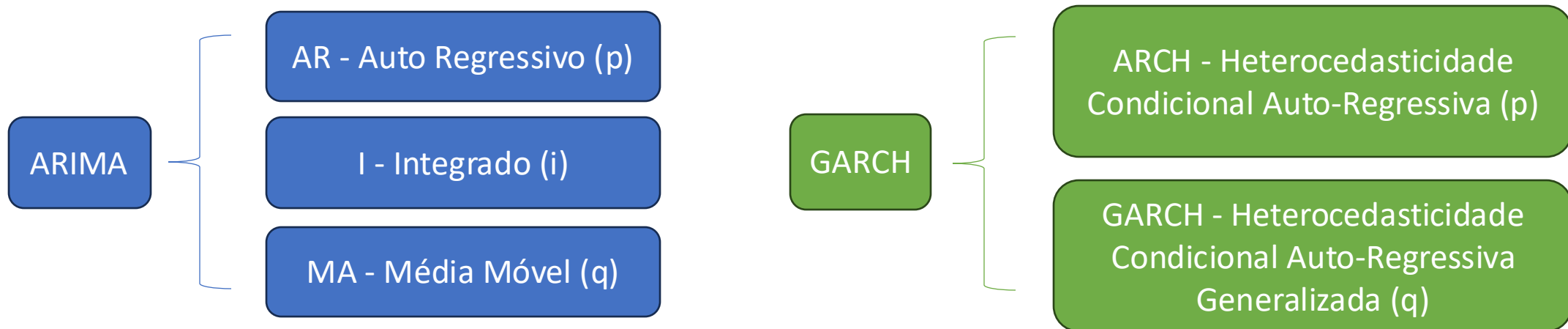
Para a construção dos modelos e manipulação dos dados, foi utilizada a linguagem Python, juntamente com as seguintes bibliotecas:

- Pandas
- Numpy
- Statsmodels
- Arch
- Scikit-learn
- Keras (Tensorflow)



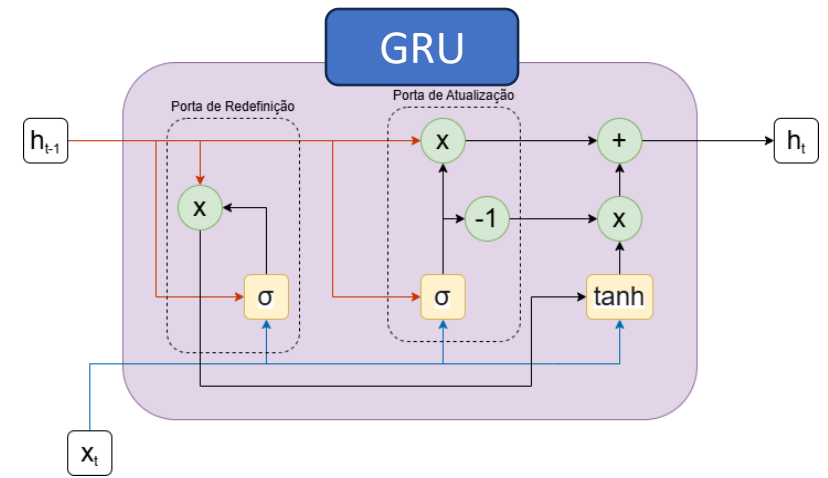
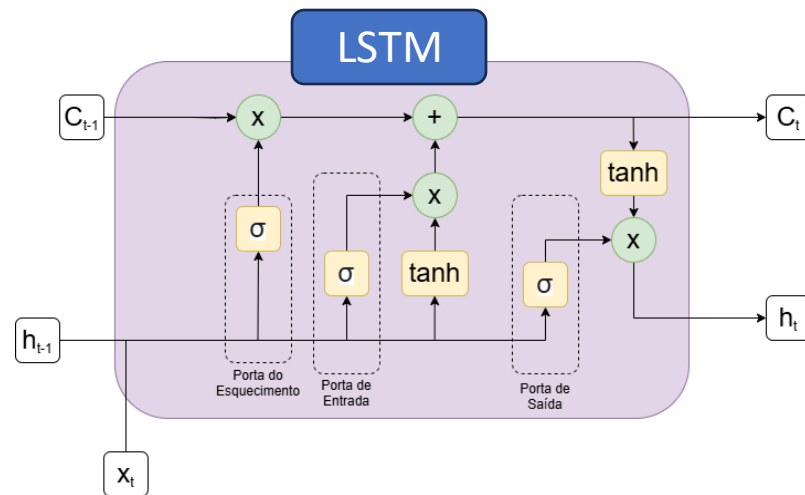
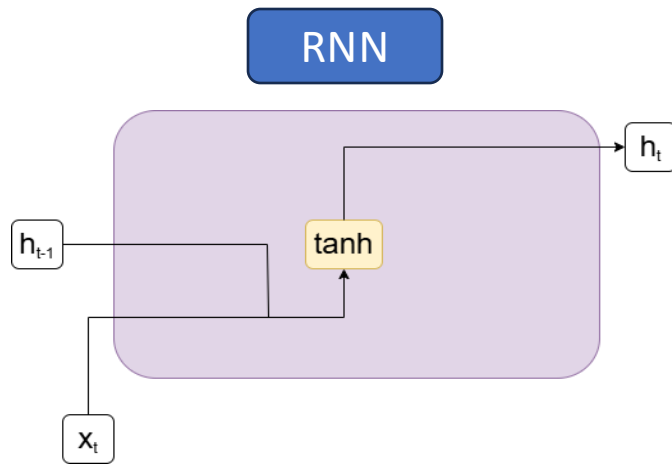
Métodos

- Auto-Regressive Integrated Moving Average [ARIMA]
- Generalized Auto-Regressive Conditional Heteroskedasticity [GARCH]



Métodos

- Recurrent Neural Network [RNN]
- Long Short-Term Memory [LSTM]
- Gated Recurrent Unit [GRU]



Estratégia Empírica

Passo 1 – Preço do bitcoin em dólar e Taxa de juros americana decompostas

Passo 2 – Separação de conjunto de treino com 127 observações e teste com 10

Passo 3 – Normalização das séries para entrada nos modelos LSTM e GRU

Passo 4 – Determinação das ordens do modelo ARIMA-GARCH (FAC, FACP, Dickey-Fuller Aumentado)

Passo 5 – Determinação das camadas do modelo LSTM (LSTM->Dropout->LSTM->Dropout->Densa)

Passo 6 – Determinação das camadas do modelo GRU (GRU->Dropout->GRU->Dropout->Densa)

Passo 7 – Determinação do otimizador e da função de custo (ADAM, RMSE)

Passo 8 – Random Search para hiperparâmetros

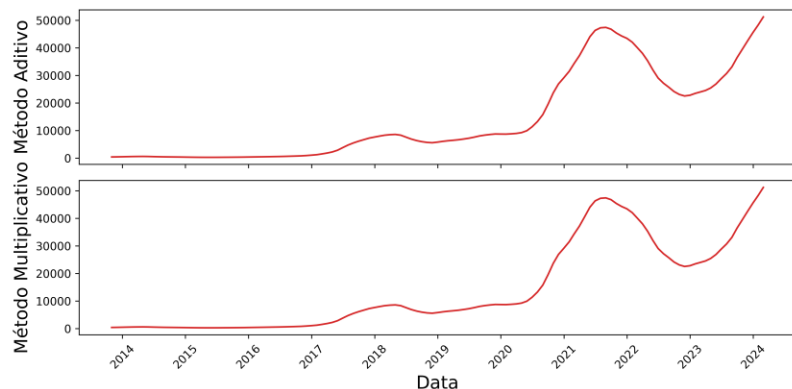
Passo 9 – Previsão

Passo 10 – Cálculo de sMAPE e RMSE

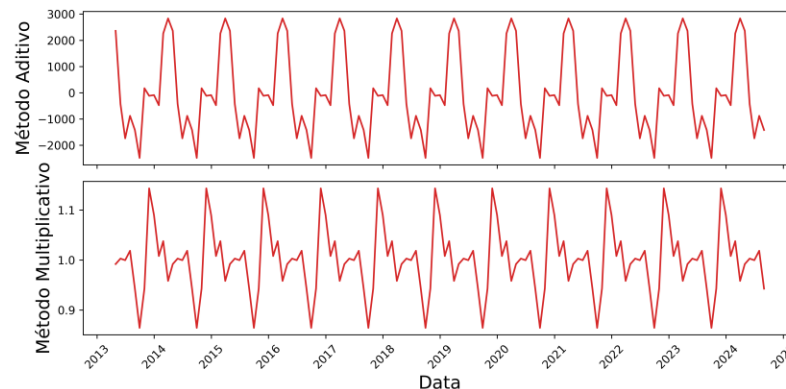
Passo 11 – Comparação das previsões e determinação do modelo com maior capacidade de previsão

Resultados e Discussões - Decomposição das Séries

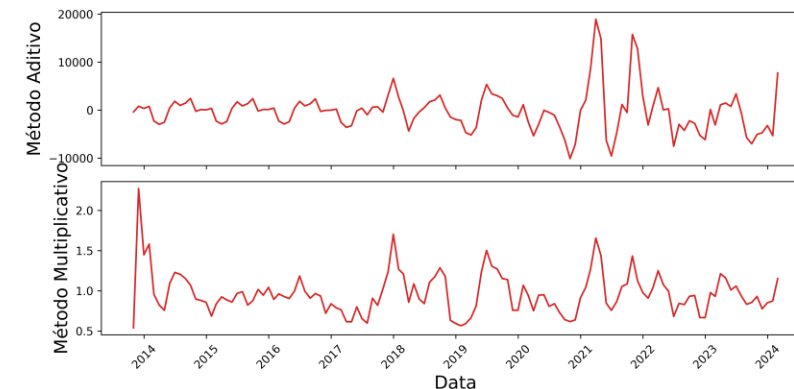
Preço em Dólar do Bitcoin - Tendência



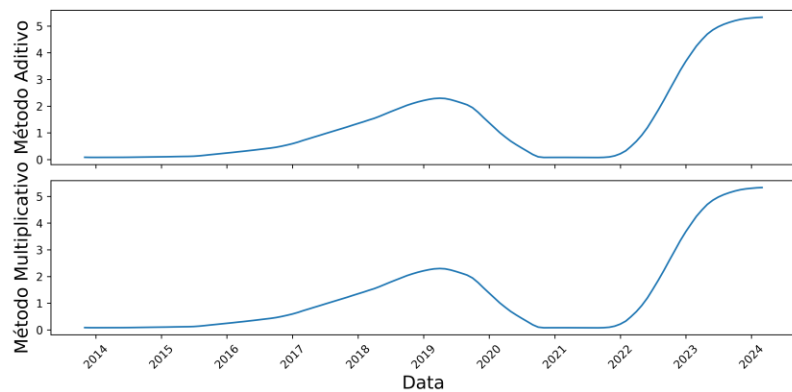
Preço em Dólar do Bitcoin - Sazonalidade



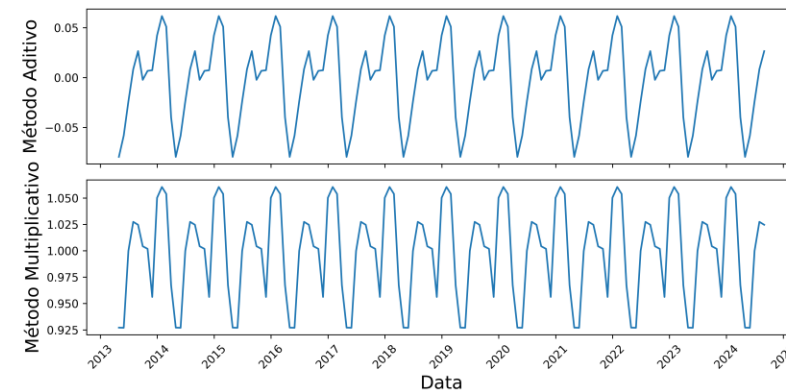
Preço em Dólar do Bitcoin - Resíduos



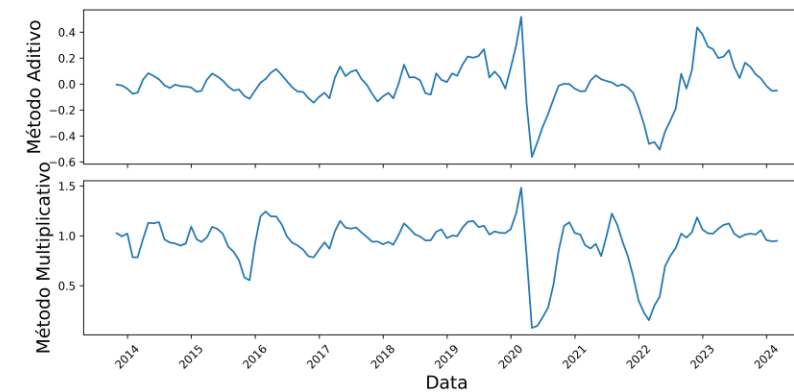
Taxa de Juros Americana - Tendência



Taxa de Juros Americana - Sazonalidade



Taxa de Juros Americana - Resíduos



Resultados e Discussões - Ordens do modelo ARIMA-GARCH

ARIMAX (2, 1, 2)			
Variável	Coeficiente	Erro Padrão	P Valor
Taxa de Juros Americana	-2278,3069	2305,498	0,323
AR Lag 1	-0,3668	0,12	0,002
AR Lag 2	-0,7519	0,099	0
MA Lag 1	0,5882	0,137	0
MA Lag 2	0,7096	0,135	0
σ^2	1,60E+07	0,099	0

GARCH (1,1)			
Variável	Coeficiente	Erro Padrão	P Valor
ω	2,85E+05	5,13E+04	2,60E-08
α_1	0,4219	6,40E-02	4,29E-11
β_1	0,5781	6,82E-02	2,43E-17

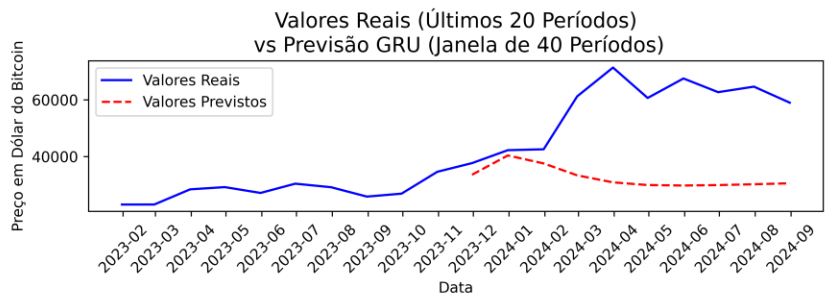
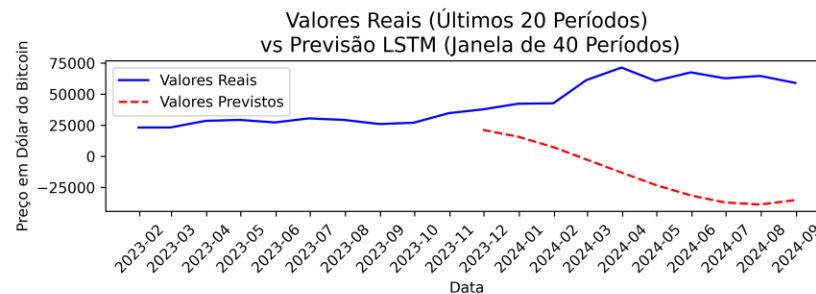
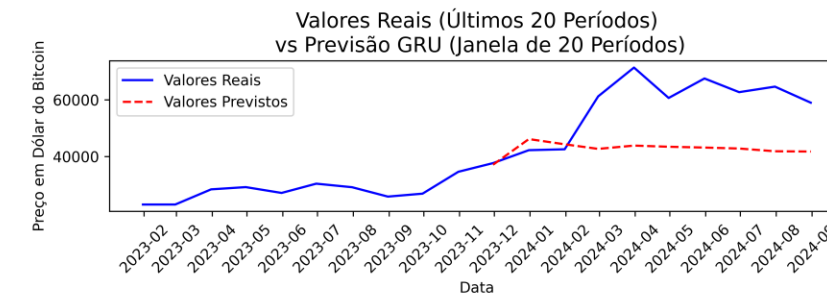
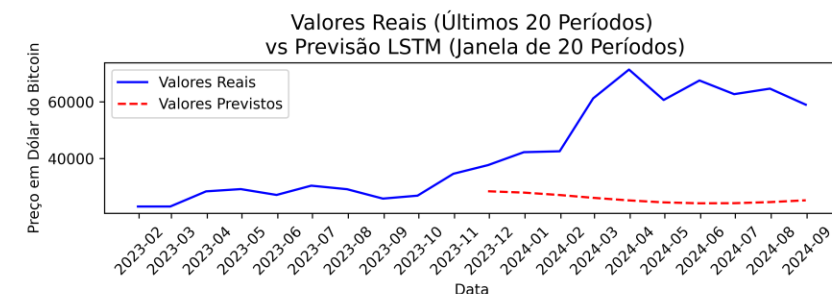
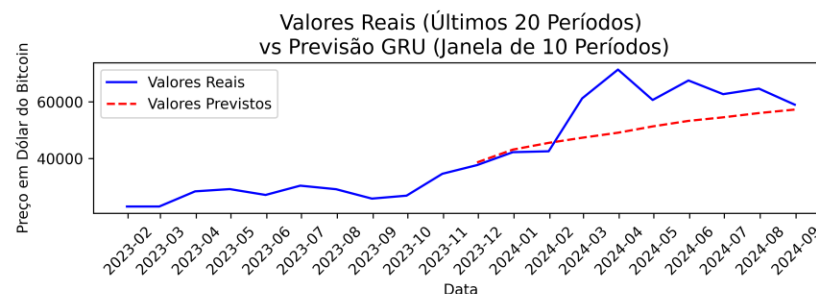
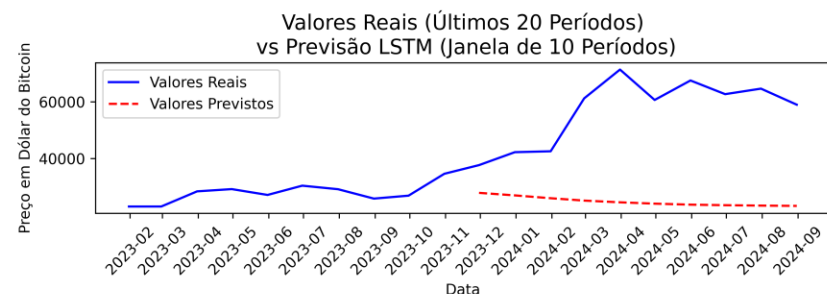
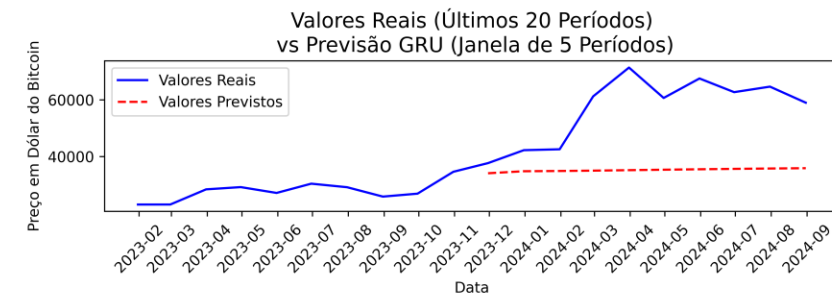
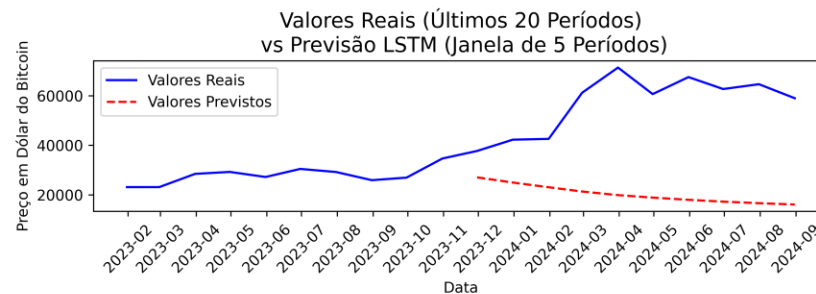
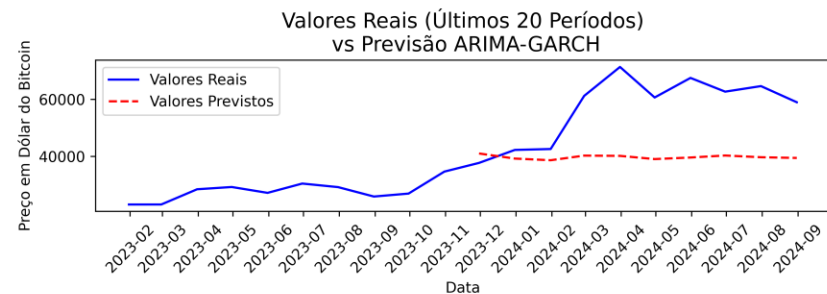
Resultados e Discussões - Hiperparâmetros

Hiperparâmetros e Valores possíveis	
Hiperparâmetro	Valores possíveis
Unidades	20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250
Taxa de “dropout”	0.2, 0.3, 0.4, 0.5
Taxa de aprendizagem	0.01, 0.001, 0.0001
Tamanho do lote	32, 64, 96, 128
Épocas	25, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300

Resultados e Discussões - Random Search

Hiperparâmetros otimizados						
Célula	Janela	Unidades	Taxa de Dropout	Taxa de Aprendizagem	Tamanho do Lote	Épocas
LSTM	5	100	0,4	0,01	128	125
LSTM	10	50	0,4	0,01	96	150
LSTM	20	100	0,4	0,001	128	125
LSTM	40	50	0,3	0,01	32	50
GRU	5	100	0,2	0,001	128	75
GRU	10	125	0,4	0,001	64	75
GRU	20	100	0,2	0,001	128	75
GRU	40	100	0,2	0,001	128	75

Resultados e Discussões - Previsões



Resultados e Discussões - Performance dos Modelos

Tabela de Performances (RMSE e sMAPE)

Modelo	RMSE	sMAPE
Previsão ARIMA-GARCH	20470,68272	34,89335347
Previsão LSTM (Janela de 5 Períodos)	39252,72054	91,91452952
Previsão LSTM (Janela de 10 Períodos)	34355,70298	75,45948516
Previsão LSTM (Janela de 20 Períodos)	33510,48607	72,38148643
Previsão LSTM (Janela de 40 Períodos)	77267,10151	168,7873815
Previsão GRU (Janela de 5 Períodos)	24206,055	44,62905862
Previsão GRU (Janela de 10 Períodos)	10622,66617	14,50398544
Previsão GRU (Janela de 20 Períodos)	17901,89902	28,89072889
Previsão GRU (Janela de 40 Períodos)	28034,02861	51,74115612

Conclusões

ARIMA-GARCH:

- Desempenho intermediário
- RMSE elevado, indicando previsões distantes dos valores reais
- sMAPE sugere maior controle na variação percentual média dos erros, refletindo certa estabilidade preditiva

Modelos LSTM:

- Desempenho inconsistente, com RMSE e sMAPE altos
- Piora com janelas de previsão maiores, janela de 40 períodos teve desempenho excepcionalmente ruim, sugerindo dificuldade em captar dinâmicas de longo prazo

Modelos GRU:

- Desempenho superior, especialmente nas janelas de 10 e 20 períodos
- GRU com janela de 10 períodos teve o menor RMSE (10.622,67) e menor sMAPE (14,50%)
- Desempenho deteriora em janelas maiores, especialmente na de 40 períodos

Possíveis Melhorias Futuras:

- Incorporar novas variáveis, como volume de transações e indicadores macroeconômicos
- Testar janelas intermediárias (15 ou 25 períodos) para identificar ponto ideal
- Ajustar a frequência dos dados para aumentar a amostra
- Utilizar outras sementes (SEED 42)