**Projeto de Pesquisa e Planejamento de Atividades**

|  |  |  |
| --- | --- | --- |
| **Aluno**: Gustavo Menezes de Sirqueira | | **Data início curso**: \_\_/\_\_/\_\_\_\_ |
| **Orientador**: Daniel Alvarez Firmino | | **Defesa em:**  Dez/2024 |
| **Curso**: MBA Data Science e Analytics | **Modalidade**: Distância | Turma: 231 |

1. **Título do projeto** *(Inicial)*

**Comparação de Modelos de previsão de séries temporais aplicados a criptomoedas**

1. **Introdução**

Para se realizar transações entre 2 ou mais agentes no modelo financeiro tradicional, é necessária a existência de um intermediador, que por sua vez necessita de uma taxa para realizar a transação. Esse intermediário, usualmente bancos e instituições financeiras, mostra-se necessário para garantir a confiança, segurança e validade da transação. A moeda com a qual essa operação é realizada também precisa ser emitida por um órgão centralizador, como um banco central. Essa moeda muitas vezes tem seu lastro baseado em ouro ou confiança. Para abordar a questão mencionada acima, é preciso então a criação de um sistema que possa garantir transações seguras, confiáveis e válidas sem a necessidade de um intermediador e de um emissor centralizado(Patel et al., 2020; Vranken, 2017).

Em 2008 é proposto, no artigo “Bitcoin: A Peer-to-Peer Electronic Cash System” (Nakamoto, 2008), um conceito de moeda digital finita e transação online sem a necessidade de um intermediador, distribuído na rede e apoiado na tecnologia de “blockchain”, que por sua vez é baseada no artigo “How to time-stamp a digital document” (Haber e Stornetta, 1991). A tecnologia de transação atua como uma “caderneta pública” que é capaz de registrar movimentações em blocos, que por sua vez são encadeados a outros blocos gerados anteriormente, dando origem ao nome “blockchain” - cadeia de blocos. Devido ao caráter público da *blockchain*, existe transparência em suas transações. Um processo baseado em mecanismos de consenso e hash condicionados chamado de “mineração” garante que as transações válidas sejam adicionadas à corrente e, consequentemente, contabilizadas, além de garantir rastreabilidade e histórico. A mineração é realizada dentro dos nodos da rede e é um processo computacionalmente exigente e necessário para o funcionamento do sistema, gerando a necessidade de uma compensação aos agentes mineradores. Essa compensação é gerada pelo próprio sistema, e corresponde a uma quantia de bitcoin baseada na quantidade de bitcoins existentes na rede. Dessa forma, a cada 210000 blocos o sistema diminui a recompensa na metade, o que é chamado de “halving”, tornando a operação de mineração gradativamente menos rentável (Vranken, 2017).

Desde o surgimento do bitcoin, o interesse pelas criptomoedas cresceu consideravelmente e atraiu a atenção de investidores e empresas. Sua tecnologia deu origem a outras criptomoedas, com enfoques diversos, variando na criação de moedas focadas em transações dentro de redes específicas até a criação de redes especializadas em transações, seguro e “supply chain” no campo da agricultura (Shen et al., 2020).

O mercado de criptomoedas corresponde, hoje, março de 2024, a 2,6 trilhões de dólares, sendo aproximadamente metade desse valor, 1,28 trilhões de dólares, correspondente ao bitcoin (Forbes Digital Assets, 2024). Alguns dos fatores determinantes para essa relevância foi a possibilidade da utilização do bitcoin na diversificação em portfólios e o potencial de geração de lucro, desvencilhado da centralização governamental. Entretanto, um maior potencial de lucro vem acompanhado de maior risco, e consequentemente maior potencial de perdas, em especial com posições de longo prazo, indicando uma alta volatilidade por parte do criptoativo (Nedvd e Krisfoufek, 2023; Shen et al., 2020).

A previsão financeira desempenha um papel importante no mercado financeiro e no campo da economia, e é frequentemente utilizada na modelagem de ativos de alta volatilidade, na gestão de riscos e como ferramenta auxiliar na tomada de decisões. Alguns dos modelos mais comuns para esse fim são ARIMA para a modelagem e previsão da média e GARCH para lidar com a volatilidade, amplamente utilizados em artigos sobre a modelagem com retornos de bitcoin (Dyhrberg, 2015; Katsiampa, 2017; Phillip et al., 2017; Shen et al., 2020). Entretanto, a utilização de modelos de aprendizado de máquina não é tão amplamente difundida neste meio. Redes neurais recorrentes são capazes de aprender dependências de longo prazo em séries temporais, identificam características lineares e não lineares, apresentam maior sensibilidade a flutuações no tempo e podem preservar mais informação temporal durante o treino quando comparadas aos métodos econométricos tradicionais (Shen et al., 2020).

Levantados todos esses pontos, torna-se desejável a utilização de modelos capazes de lidar com a natureza imprevisível do bitcoin e que possuam características diferentes das já observadas em modelos comuns a literatura, tornando justificada a construção de um comparativo entre redes neurais recorrentes, em especial, GRU e LTSM, que possuem capacidade de lidar com o problema do gradiente desvanecente, e um modelo econométrico tradicional ARIMA. Dessa forma, o presente trabalho busca identificar qual dos modelos apresentará uma melhor performance de previsão de retornos do bitcoin.

1. **Objetivo**

O trabalho tem como objetivo determinar se os modelos de redes neurais recorrentes LSTM e GRU são capazes de apresentar melhor desempenho preditivo sobre os retornos de bitcoin quando comparados ao modelo tradicional ARIMA.

1. **Material e Métodos**

Os dados utilizados na pesquisa serão obtidos através do website CoinMarketCap.com (CoinMarketCap, 2024). Serão utilizados os valores de fechamento mensal do bitcoin, compreendendo o período de abril de 2013 até fevereiro de 2024. As primeiras 119 entradas serão utilizadas para o treino dos modelos, e as últimas 12 entradas para o teste. Será calculado o log-retorno entre os meses.

Os modelos utilizados serão ARIMA, sigla para “AutoRegressive Integrated Moving Average”, GRU, sigla para “Gated Recurrent Unit”, e LSTM, sigla para “Long Short-Term Memory”, que serão construídos com a linguagem python, por meio das bibliotecas Statsmodels (Statsmodels, 2024) e Keras (Keras, 2024). Para a determinação da ordem do modelo ARIMA-GARCH, será empregada a comparação das métricas AIC - Aikaike Information Criterion, BIC - Bayesian Information Criterion, significância dos seus parâmetros e a verificação da ACF - Autocorrelation Function e PACF - Partial Autocorrelation Function dos resíduos normalizados. Para a determinação dos modelos GRU e LTSM, será utilizada a técnica de “grid-search”na otimização dos hiperparâmetros.

A realização da comparação das previsões obtidas pelos modelos será realizada utilizando as métricas RMSE - Root Mean Square Error e sMAPE - Symmetric Mean Absolute Percentage Error. A utilização de critérios de mensuração de erros de categorias diferentes é recomendada a fim de se contabilizar pelas limitações que cada categoria apresenta (Shcherbakov et al., 2013).

1. **Resultados Esperados**

É esperada a construção de um comparativo entre os 3 modelos propostos, e a determinação da resposta à pergunta proposta nos objetivos acerca da performance dos modelos baseados em redes neurais recorrentes sobre o modelo econométrico tradicional ARIMA.

1. **Cronograma de Atividades**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Atividades planejadas** | **Mês** | | | | | | | | | |
| **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** |
| Construção do projeto de pesquisa | **x** |  |  |  |  |  |  |  |  |  |
| Construção do modelo econométrico |  | **x** | **x** |  |  |  |  |  |  |  |
| Construção dos modelos em deep learning |  |  | **x** | **x** |  |  |  |  |  |  |
| Entrega dos resultados preliminares |  |  |  | **x** | **x** |  |  |  |  |  |
| Desenvolvimento de resultados e conclusão |  |  |  |  |  | **x** | **x** |  |  |  |
| Correções referentes a revisão final do orientador |  |  |  |  |  |  | **x** | **x** |  |  |
| Entrega da versão escrita |  |  |  |  |  |  |  | **x** |  |  |
| Apresentação |  |  |  |  |  |  |  |  | **x** | **x** |
|  |  |  |  |  |  |  |  |  |  |  |

Projeto de Pesquisa; Resultados Preliminares; Entrega do Trabalho de Conclusão de Curso; Entrega da Apresentação da Defesa

1. **Referências Bibliográficas**

Arch. 2024. ARCH models in Python. Disponível em: <https://github.com/bashtage/arch>. Acesso em: 24 mar. 2024.

CoinMarketCap. 2024. Bitcoin price today, BTC to USD live price, marketcap and chart. Disponível em: <https://coinmarketcap.com/currencies/bitcoin/historical-data/>. Acesso em: 24 mar. 2024.

Dyhrberg, A.H. 2015. Bitcoin, gold and the dollar: A GARCH volatility analysis.

Finance Research Letters 1-8.

Forbes Digital Assets. 2024. Cryptocurrency Prices, Market Cap and Charts. Disponível em:

<[Cryptocurrency Prices, Market Cap and Charts | Forbes Digital Assets](https://www.forbes.com/digital-assets/crypto-prices/?sh=38fc4e422478)>. Acesso em: 23 mar. 2024.

Haber, S.; Stornetta, W.S. 1991. How to time-stamp a digital document. Journal of

Cryptology 3: 99-111.

Katsiampa, P. 2017. Volatility estimation for Bitcoin: A comparison of GARCH

models. Economics Letters.

Keras. 2024. Keras: Deep Learning for humans. Disponível em: <https://keras.io/>. Acesso em: 24 mar. 2024.

Nakamoto, S. 2008. Bitcoin: A Peer-to-Peer Electronic Cash System. Disponível em: https://bitcoin.org/bitcoin.pdf. Acesso em: 17 mar. 2024

Nedved, M.; Krisfoufek, L. 2023. Safe havens for Bitcoin. Finance Research Letters 51.

Patel, M.M.; Tanwar, S.; Gupta, R.; Kumar, N. 2020. A Deep Learning-based Cryptocurrency Price Prediction Scheme for Financial Institutions. Journal of Information Security and Applications 55.

Phillip, A.; Chan, J.; Peiris, S. 2017. A new look at Cryptocurrencies.

Economics Letters.

Shen, Z.; Qing, W.; Leatham, D.J. 2020. Bitcoin Return Volatility Forecasting: A Comparative Study between GARCH and RNN. Journal of Risk and Financial Management 14(7): 337.

Shcherbakov, M.; Brebels, A.; Shcherbakova, N.; Tyukov, A.; Janovsky, T.; Kamaev, V. 2013. A survey of forecast error measures. World Applied Sciences Journal 24: 171–176.

Vrankin, H. 2017. Sustainability of bitcoin and blockchains. Current Opinion in Environmental Sustainability 28: 1-9.