|  |  |
| --- | --- |
| CURSO DE CIÊNCIA DA COMPUTAÇÃO – TCC | |
| ( X ) PRÉ-PROJETO     (     ) PROJETO | ANO/SEMESTRE: 2023/1 |

ANÁLISE COMPARATIVA DE ALGORITMOS DE MACHINE LEARNING PARA A PREVISÃO DE VALOR DAS AÇÕES NA BOLSA DE VALORES

Jeferson Bonecher

Prof. Marcel Hugo – Orientador

# Introdução

A bolsa de valores é um ambiente centralizado e especializado voltado para negociações no mercado de capitais, onde empresas e empreendimentos ofertam seus títulos, ações e instrumentos financeiros com o objetivo de atrair investidores e agentes interessados (ALMEIDA, 2022). Essa plataforma de negociações permite às organizações captar recursos para financiar projetos e expandir atividades, enquanto investidores individuais e institucionais buscam diversificar seus portfólios e obter retornos financeiros através da compra e venda desses ativos. A bolsa de valores também promove a liquidez, eficiência, segurança e transparência do mercado, garantindo a integridade das operações realizadas entre os participantes.

Segundo Pinto (2020) uma das formas de fazer uma análise de um ativo é através da Análise Técnica. Ela é muito utilizada por investidores e profissionais do mercado financeiro para prever a direção dos preços de determinados ativos financeiros, com base na análise de gráficos que mostram a movimentação desses ativos ao longo do tempo. Por meio desta análise busca-se identificar padrões e tendências que possam indicar movimentos futuros e, assim, tomar decisões de "entrada" e "saída" com base nessas informações. A Análise Técnica é especialmente útil para operações de curto prazo na Bolsa de Valores, já que permite tomar decisões rápidas que podem resultar em lucro ou prejuízo em questão de minutos ou segundos, a depender da volatilidade do ativo em questão.

No campo da Análise Técnica, os algoritmos de machine learning, incluindo Redes Neurais Recorrentes (RNRs), têm ganhado cada vez mais espaço devido à sua capacidade de processar e analisar grandes volumes de dados em tempo real (XP Education, 2023). As RNRs são particularmente eficazes na análise de sequências temporais, como os dados de mercado financeiro, já que possuem a habilidade de capturar dependências temporais de longo prazo. Essa abordagem permite que investidores e *traders* identifiquem padrões e tendências nos mercados financeiros, auxiliando na tomada de decisões e potencialmente melhorando os resultados de suas operações (Matsumoto, 2019).

Nesse sentido, o presente trabalho tem como proposta realizar uma análise comparativa entre as redes neurais recorrentes Gated Recurrent Unit (GRU) e Long Short-Term Memory (LSTM) na previsão de valores de ações na bolsa de valores brasileira. Para tanto, serão utilizados dados históricos de empresas listadas na B3, buscando identificar qual dos algoritmos apresenta as melhores taxa de acerto para previsão dos preços das ações.

## OBJETIVOS

O objetivo deste trabalho é realizar uma comparação entre as redes neurais recorrentes LSTM e GRU para determinar qual delas apresenta o melhor desempenho na previsão dos preços das ações de cinco empresas listadas na bolsa de valores brasileira.

Os objetivos específicos são:

1. coletar e organizar os dados históricos de preços das ações das empresas Weg, Petrobras, Copel, Vale e Magazine Luiza na bolsa de valores brasileira;
2. implementar e treinar redes neurais recorrentes GRU e LSTM para prever os preços de fechamento das ações das empresas selecionadas;
3. avaliar o desempenho dos algoritmos GRU e LSTM utilizando métricas como acurácia, Mean Squared Error (MSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) e Root Mean Squared Error (RMSE).

# trabalhos correlatos

Nesta seção serão apresentados os trabalhos que correlacionam com o principal objetivo do presente trabalho. A subseção 2.1 traz o trabalho desenvolvido por Rossi (2019), em que foi criado um protótipo capaz de prever valores de ações através de Redes Neurais Artificiais. Na subseção 2.2 é apresentado um protótipo que prevê o preço das ações por meio de notícias (Bezerra,2022). Na subseção 2.3 está uma aplicação com o objetivo de prever valores de ações através de RNRs (Ribeiro *et al*. 2021).

## PROTÓTIPO PARA PREVISÃO DE AÇÕES UTILIZANDO REDES NEURAIS ARTIFICIAIS

O trabalho desenvolvido por Rossi (2019) tem como objetivo criar um protótipo capaz de prever tendências de preços de ações no mercado financeiro, utilizando Redes Neurais Artificiais (RNA) com foco na arquitetura LSTM. O intuito é fornecer uma ferramenta que auxilie investidores na tomada de decisões mais informadas ao realizar investimentos em ações na Bolsa de Valores. Os resultados foram satisfatórios atingindo uma taxa de 70% de acerto para o movimento ocorrido na semana.

Segundo Rossi (2019) o processo envolve diversas etapas, começando pela coleta de dados históricos das ações, que incluem preço de fechamento, preço de abertura, preço máximo, preço mínimo e volume negociado, a partir de fontes públicas. Os dados são, então, normalizados para que a amplitude das variações não afete o processo de treinamento e previsão da RNA. O conjunto de dados é dividido em subconjuntos de treinamento e teste, sendo que o primeiro é usado para treinar o modelo e o segundo para avaliar seu desempenho. A técnica de validação cruzada é aplicada para garantir a generalização do modelo e evitar o sobreajuste. Durante esse processo, o conjunto de treinamento é dividido em várias partições, e o modelo é treinado em cada uma delas, excluindo uma partição que é utilizada como conjunto de validação. A média dos resultados obtidos em todas as partições é utilizada como estimativa final do desempenho do modelo.

De acordo com Rossi (2019) para avaliar a eficácia do modelo proposto em situações reais, foram selecionadas cinco ações de alta liquidez e frequentemente negociadas na bolsa de valores: Raia Drogasil (RADL3), Vale SA (VALE3), Bradesco (BBDC4), Petrobras (PETR4) e WEG SA (WEGE3). Utilizando um período diário, o modelo empregou os últimos 50 valores de fechamento para prever os próximos 5 preços. A avaliação do desempenho do modelo foi realizada durante o treinamento da LSTM, utilizando as métricas de erro MSE e MAE. Os resultados obtidos podem ser observados na Tabela 1.

Tabela 1- Cálculos de erro durante o treinamento (em R$)

Tabela

Descrição gerada automaticamente

Fonte: Rossi (2019).

Os resultados mostraram que as ações RADL3 e WEGE3 obtiveram os melhores desempenhos, conforme apresentado na Tabela 2. O modelo conseguiu prever corretamente a tendência de alta ou baixa nas cotações em 7 das 10 semanas avaliadas. Por outro lado, as ações VALE3 e BBDC4 apresentaram os piores resultados, com uma taxa de acerto de 30% e 40% na previsão, respectivamente. Essa diferença de desempenho pode ser atribuída à alta volatilidade desses ativos.

Tabela 2 – Taxa de acerto das movimentações

Tabela

Descrição gerada automaticamente

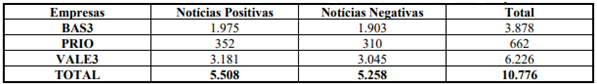
Fonte: Rossi (2019).

## REDES NEURAIS ARTIFICIAIS NA PREVISÃO DO PREÇO DAS AÇÕES NA BOLSA DE VALORES POR MEIO DE NOTÍCIAS

Bezerra (2022) apresenta um protótipo que utiliza redes neurais artificiais para sugerir a compra ou venda de ativos negociados na Bolsa de Valores brasileira, com base em notícias do mercado financeiro. O protótipo usa uma base de dados de 10.766 notícias coletadas através de um *Web Scraper*, com dados das empresas Banco do Brasil (BBAS3), PetroRIO (PRIO3) e Vale do Rio Doce (VALE3) no período de 01/01/2018 a 20/05/2022. A eficácia do protótipo foi avaliada com um percentual de acurácia de 52% nos testes.

De acordo com Bezerra (2022) o estudo examinou o efeito das notícias do mercado corporativo das três empresas nos preços das ações, observando os valores na abertura, no fechamento e no dia subsequente. O sentimento das notícias foi classificado como positivo quando o preço das ações aumentou e negativo quando diminuiu. Para analisar a relação entre as notícias e os preços das ações, empregou-se um algoritmo de classificação binária, designando 0 para reações negativas e 1 para reações positivas. Os achados indicaram que a previsão dos dados não é linear, e os gráficos evidenciaram a volatilidade do mercado, particularmente no período analisado (2018-2022). A Tabela 3 exibiu o total de sentimentos positivos e negativos relacionados às notícias e a quantidade total de notícias para cada empresa.

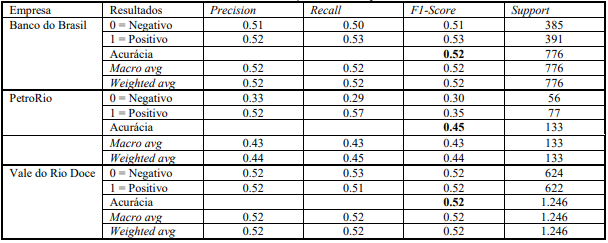
Tabela 3 – Sentimentos positivos e negativos em relação às notícias das empresas pesquisadas



Fonte: Bezerra (2022).

Segundo Bezerra (2022) os resultados obtidos apresentados na Tabela 4, destacam o recall das variáveis positivas para PetroRio (0,57) e Banco do Brasil (0,53), bem como a precision, que alcançou o mesmo valor (0,52) em todas as empresas. Entretanto, o F1-Score, que mede a efetividade geral do modelo, não ultrapassou 52% para BBAS3 e VALE3, e atingiu apenas 0,45 para PetroRio. Essa baixa efetividade pode estar relacionada ao baixo número de notícias da empresa no período analisado e sugere um foco investigativo futuro.

Tabela 4 – Relatório de Classificação do Protótipo na Matriz Confusão



Fonte: Bezerra (2022).

## Previsão de Preço de Ações Baseada em Redes Neurais Recorrentes LSTM e GRU

Ribeiro *et al*. (2021) conduziram uma pesquisa com o objetivo de prever os valores das ações no mercado brasileiro, empregando redes neurais recorrentes, tais como LSTM e GRU. Diversos cenários foram estabelecidos, combinando as referidas redes neurais a métricas de análise técnica, como médias móveis e bandas de Bollinger. Utilizaram-se dados do Yahoo Finance para criar cenários variáveis, buscando aprimorar as previsões dos valores futuros das ações. Métricas de erro, como MAPE, RMSE e U de Theil, foram empregadas na avaliação do desempenho dos modelos, auxiliando na tomada de decisões relativas aos investimentos. A pesquisa optou pela biblioteca Keras devido à sua eficiência e facilidade de uso, sendo desenvolvida em linguagem Python.

De acordo com Ribeiro *et al*. (2021) a rede LSTM possui 4 camadas com 100, 80 e 50 unidades nas três primeiras camadas, respectivamente, e uma camada de saída com 1 unidade, todas com uma taxa de Dropout de 0.2. O algoritmo de otimização utilizado é o Adam. A rede GRU é composta por 5 camadas, sendo 4 delas com 50 unidades e 0.2 de Dropout, e a camada de saída com 1 unidade. O algoritmo de otimização usado é o SGD. Ambas as redes utilizam o erro médio quadrático como métrica de perda.

Com isso Ribeiro et al. (2021) desenvolveu diferentes cenários para cada uma das ações selecionadas, tanto às redes LSTM quanto às redes GRU. O objetivo era verificar o desempenho das redes não apenas com os valores das séries temporais, mas também considerando outras variáveis (*features*), como volume de transações diárias e valor de fechamento do Índice Bovespa. Foram utilizadas variáveis conhecidas do mercado financeiro para análise técnica, como médias móveis e bandas de Bollinger. As variáveis escolhidas para os cenários estão apresentadas na Tabela 03.

Tabela 03 – Variáveis relacionadas aos cenários utilizados

Tabela

Descrição gerada automaticamente

Fonte: Ribeiro et al. (2021).

Para Ribeiro *et al*. (2021) os resultados indicaram que, de maneira geral, a inclusão de *features* adicionais aprimorou o desempenho das redes LSTM, resultando em médias mais baixas de erro percentual. Em relação às redes GRU, embora não tenha havido um cenário que se sobressaísse em todas as empresas analisadas, os resultados apontam que as *features* podem ser consideradas boas opções de métricas para previsões de valores de ações, devido ao baixo valor da média de erro percentual encontrado.

# proposta

Nesta seção será descrita a proposta deste trabalho, justificando o desenvolvimento, definindo os requisitos funcionais e não funcionais, as metodologias abordadas e por fim o cronograma.

## JUSTIFICATIVA

No Quadro 1 é apresentado um comparativo entre os trabalhos correlatos. As linhas representam as

características e as colunas os trabalhos.

Quadro - Comparativo dos trabalhos correlatos

|  |  |  |  |
| --- | --- | --- | --- |
| Trabalhos Correlatos  Características | Rossi (2019) | Bezerra (2022) | Ribeiro *et al*. (2021) |
| RNRs utilizadas | LSTM | LSTM | LSTM, GRU |
| Métricas utilizadas | MSE, MAE | Precision, Recall, f1-score | MAPE, RMSE U DE THEIL |
| Ações analisadas | RADL3,VALE3,BBDC4,  PETR4,WEGE3 | BBAS3, VALE3, PRIO3 | AMBEV, VALE3,  BBDC4, PETR4, ITUB4 |
| Histórico de preço | Sim | Não | Sim |
| Taxa de acerto | 70% | 52% | - |
| Disponibilizada interface de consulta | Sim | Sim | Não |

Fonte: elaborado pelo autor.

Conforme pode ser observado no Quadro 1, os trabalhos correlatos apresentados utilizam redes neurais recorrentes para analisar ações no mercado financeiro. Rossi (2019) e Bezerra (2022) usam LSTM como arquitetura de RNR, enquanto Ribeiro et al. (2021) empregam tanto LSTM quanto GRU. As métricas utilizadas para avaliar o desempenho são diferentes entre os estudos: Rossi (2019) utiliza MSE e MAE, Bezerra (2022) emprega Precision, Recall e f1-score, enquanto Ribeiro et al. (2021) usam MAPE, RMSE e U de Theil. Com estas métricas eles comparam as duas arquiteturas utilizadas junto com as quatro variáveis para definir qual rede teve o melhor desempenho. Os três estudos analisam diferentes conjuntos de ações, VALE3 sendo utilizada por todos. Apenas Rossi (2019) e Bezerra (2022) utilizam métricas mostrando sua taxa de acerto em porcentagem sendo 70% e 52% respectivamente. Ribeiro et al. (2021) utiliza outras métricas para medir seu desempenho e também é o único que não disponibiliza interface de consulta para seu modelo.

Esta proposta é relevante e significativa, pois aborda a complexidade da análise de ações no mercado financeiro e busca melhorar as técnicas existentes utilizando RNRs, como LSTMs e GRUs. Essas arquiteturas são capazes de capturar padrões temporais e extrair informações relevantes em séries temporais, tornando-as adequadas para o contexto do mercado financeiro. A investigação de diferentes métricas de desempenho, assim como a análise de conjuntos variados de ações, pode gerar percepções valiosas sobre a efetividade das RNRs na previsão do comportamento das ações e, consequentemente, auxiliar na otimização das estratégias de investimento.

As contribuições do estudo abrangem aspectos teóricos e práticos. Teoricamente, o estudo expande o conhecimento sobre a eficácia de diferentes arquiteturas de RNRs, como LSTMs e GRUs, na análise de ações e a relevância de distintas métricas de desempenho na avaliação de modelos de previsão no contexto do mercado financeiro. Em termos práticos, com base nas descobertas teóricas, o estudo oferece diretrizes para o desenvolvimento de modelos de análise de ações mais eficientes e precisos, resultando em melhores estratégias de investimento e auxiliando investidores e profissionais do mercado financeiro na tomada de decisões mais embasadas e potencialmente lucrativas.

## REQUISITOS PRINCIPAIS DO PROBLEMA A SER TRABALHADO

A ferramenta a ser desenvolvida deve:

1. gerar a previsão do valor de fechamento diário das ações (Requisito Funcional - RF);
2. gerar gráficos das previsões de valores das ações (RF);
3. gerar métricas definidas para comparar os dois algoritmos (RF);
4. ser implementada utilizando a linguagem Python (Requisito Não Funcional - RNF);
5. importar os dados históricos das ações através da biblioteca Yahoo Finance (RNF);
6. ser capaz de processar todo os históricos disponíveis dos preços das ações (RNF);
7. utilizar a biblioteca TensorFlow (RNF);
8. ser modelada seguindo os princípios das RNRs (RNF).

## METODOLOGIA

O trabalho será desenvolvido observando as seguintes etapas:

1. levantamento bibliográfico: buscar fontes bibliográficas com relação as RNRs, e bem como buscar trabalhos correlatos;
2. elicitação de requisitos da aplicação, baseando-se nas informações da etapa anterior, reavaliar os requisitos propostos para a aplicação;
3. processamento dos dados: importação e processamento dos dados utilizando a média dos preços de abertura, fechamento, preço mais alto e mais baixo em cada período de análise;
4. preparação dos dados para modelagem: os dados processados serão divididos em conjuntos de treinamento e teste, e serão normalizados para garantir que os algoritmos de machine learning possam aprender com eficiência;
5. implementação dos algoritmos: nesta etapa, serão implementadas as RNRs GRU e LSTM;
6. treinamento e validação dos modelos: os modelos GRU e LSTM serão treinados usando os conjuntos de treinamento e validação. Serão ajustados os hiperparâmetros para obter o melhor desempenho possível;
7. avaliação dos modelos: após o treinamento, os modelos serão avaliados no conjunto de teste com base em métricas como acurácia, MSE, MAE, MAPE e RMSE;
8. comparação dos algoritmos: nesta etapa, os resultados dos modelos GRU e LSTM serão comparados com base nas métricas obtidas.

As etapas serão realizadas nos períodos relacionados no Quadro 2.

Quadro - Cronograma

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 2023 | | | | | | | | | |
|  | ago. | | set. | | out. | | nov. | | dez. | |
| etapas / quinzenas | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 |
| levantamento bibliográfico |  |  |  |  |  |  |  |  |  |  |
| elicitação |  |  |  |  |  |  |  |  |  |  |
| processamento dos dados |  |  |  |  |  |  |  |  |  |  |
| preparação dos dados para modelagem |  |  |  |  |  |  |  |  |  |  |
| implementação dos algoritmos |  |  |  |  |  |  |  |  |  |  |
| treinamento e validação dos modelos |  |  |  |  |  |  |  |  |  |  |
| avaliação dos modelos |  |  |  |  |  |  |  |  |  |  |
| comparação dos algoritmos |  |  |  |  |  |  |  |  |  |  |

Fonte: elaborado pelo autor.

# REVISÃO BIBLIOGRÁFICA

Esta seção descreve brevemente sobre os assuntos que fundamentarão o estudo a ser realizado: Análise Técnica de ações e Redes Neurais Recorrentes.

Segundo Hochreiter e Schmidhuber (1997) a LSTM (Long Short-Term Memory) é uma arquitetura de aprendizado de máquina que aprimora as redes neurais recorrentes (RNRs) ao incorporar uma memória de longa duração, permitindo que o modelo compreenda a importância relativa das informações ao longo do tempo e aprenda e retenha informações relevantes de um passado mais remoto. Essa característica habilita as LSTMs a efetuar previsões mais precisas em diversas aplicações, como tradução automática, geração de texto e previsão de séries temporais.

De acordo com Cho et al. (2014) a GRU (Gated Recurrent Unit) é uma arquitetura de aprendizado de máquina que aperfeiçoa as redes neurais recorrentes (RNRs), incorporando mecanismos de atualização e redefinição para aprimorar a retenção de dependências temporais de longo alcance. Com uma estrutura simplificada e menor complexidade computacional em relação às LSTMs, as GRUs exibem desempenho semelhante em várias tarefas, incluindo análise de sentimentos, tradução automática e previsão de séries temporais.

Para avaliar o desempenho das RNRs serão utilizadas métricas de erro, como o MAE, que calcula a média dos valores absolutos das diferenças entre previsões e valores reais; o MSE, que dá mais peso aos maiores erros por meio da média das diferenças quadráticas; e o RMSE, que é a raiz quadrada do MSE e retorna o erro à unidade de medida do modelo. Além disso, o MAPE é uma métrica útil em previsão de demanda, mostrando a porcentagem média de erro em relação aos valores reais e calculando a soma dos valores absolutos das diferenças pelo valor absoluto dos valores reais, multiplicando o resultado por 100 (OLIVEIRA JÚNIOR, 2021).

Referências

ALMEIDA, Guilherme. **Bolsa de Valores: entenda tudo sobre esse mercado.** [2022]. Disponível em: https://www.certifiquei.com.br/bolsa-valores/. Acesso em: 20 abr. 2023.

Bezerra, Fabrício Oliveira. **Redes Neurais Artificiais na Previsão do Preço das Ações na Bolsa de Valores por Meio de Notícias**. [2022]. Trabalho de Conclusão de Curso (Bacharelado em Ciência da Computação) - Centro de Ciências Exatas e Naturais, Universidade Regional de Blumenau, Blumenau.

CHO, Kyunghyun; MERRIENBOER, Bart van; GULCEHRE, Caglar; BOUGARES, Fethi; SCHWENK, Holger; BENGIO, Yoshua. **Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation.** In: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). Doha, Qatar: Association for Computational Linguistics, 2014. p. 1724-1734. Disponível em: https://www.aclweb.org/anthology/D14-1179.pdf. Acesso em: 15 mar. 2023.

HOCHREITER, Sepp; SCHMIDHUBER, Jürgen. **Long short-term memory. Neural Computation, v. 9, n. 8, p. 1735-1780, 1997.** [1997]. Disponível em: https://www.bioinf.jku.at/publications/older/2604.pdf. Acesso em: 18 abr. 2023.

Matsumoto, **Daniel K.F. Estudo em Séries Temporais Financeiras utilizando Redes Neurais Recorrentes.** 2019. Disponível em: https://www.repositorio.ufal.br/bitstream/riufal/6813/3/Estudo%20em%20s%C3%A9ries%20temporais%20financeiras%20utilizando%20redes%20neurais%20recorrentes.pdf. Acesso em: 20 abr. 2023.

OLIVEIRA JÚNIOR, Clébio de. **Métricas para Regressão: Entendendo as métricas R², MAE, MAPE, MSE e RMSE.** [2021]. Disponível em: https://medium.com/data-hackers/prevendo-n%C3%BAmeros-entendendo-m%C3%A9tricas-de-regress%C3%A3o-35545e011e70. Acesso em: 19 abr. 2022.

PINTO, Leonardo. **O que é análise técnica: confira o guia completo.** [2020]. Disponível em: https://conteudos.xpi.com.br/aprenda-a-investir/relatorios/o-que-e-analise-tecnica/. Acesso em: 22 abr. 2023.

RIBEIRO, Marcos Vinicius de Oliveira; CORREA, Matheus Morgado. **Previsão de Preço de Ações Baseada em Redes Neurais Recorrentes LSTM e GRU.** [2021]. Disponível em: https://app.uff.br/riuff/bitstream/handle/1/26733/Matheus%20MorgadoTCC%20Marcos%20e%20%20final.pdf?sequence=1&isAllowed=y. Acesso em: 10 mar. 2023.

ROSSI, Rafael Bertoldi. **Protótipo para previsão de ações utilizando redes neurais artificiais**. [2019]. Trabalho de Conclusão de Curso (Bacharelado em Ciência da Computação) - Centro de Ciências Exatas e Naturais, Universidade Regional de Blumenau, Blumenau.

XP Education**. O machine learning pode ser utilizado no mercado financeiro? Vem que a gente explica!** [S.l.], [2023]. Disponível em: https://blog.xpeducacao.com.br/machine-learning-no-mercado-financeiro/. Acesso em: 22 abr. 2023.