

Avaliação de desempenho dos modelos de aprendizado de máquina para previsão de preços de ações do mercado financeiro

LUCAS EDUARDO PEREIRA TELES

Universidade do Estado do Amazonas (UEA)

Manaus – AM – Brazil

Email: lept.eng21@uea.edu.br

CARLOS M. S. FIGUEIREDO

Universidade do Estado do Amazonas (UEA)

Manaus – AM – Brazil cfigueiredo@uea.edu.br

Abstract

This paper presents a comparative analysis of machine learning models in order to predict the closing price of six stocks chosen from different sectors in different forecast scenarios. In addition, we used five metrics to measure how well they performed, and we used comparative charts of actual and predicted price over several days for the prediction. The development of the project consisted of both raw data and transformed data for the tests and used two recent models proposed for time series forecasting.

Resumo

Este artigo apresenta uma análise comparativa de modelos de aprendizado de máquina com o intuito de prever o preço de fechamento de seis ações escolhidas de setores distintos em diferentes cenários de previsão. Além disso, foram usadas cinco métricas para medir o desempenho que tiveram e utilizamos gráficos comparativos de preço real e previsto ao longo de vários dias para a predição. O desenvolvimento do projeto consistiu tanto dados no formato bruto quanto dados transformados para os testes e utilizados dois modelos recentes propostos para previsão de séries temporais.

INTRODUÇÃO

O Mercado de Ações é um segmento de Mercado de Capitais que está inserido no Mercado Financeiro, que por sua vez desempenha um papel essencial na economia global permitindo que empresas, governos e indivíduos levantem capitais, gerenciem riscos e alcancem seus objetivos financeiros, e são realizadas operações financeiras de prazos que variam entre médio e longo, possibilitando que as empresas de capital aberto financiem seus investimentos [11]. As ações são títulos de uma empresa que estão negociadas no mercado de capitais e representam uma fração auferida ao terceiro, proveniente de aplicação de seu recurso naquela companhia com direito a recebimento de seus lucros distribuídos aos titulares de ações em forma de dividendos [9]. Alocar eficientemente o capital é um trabalho fundamental presente na economia e para isso é investido em setores com o intuito de obter retornos maiores [19].

O aprendizado de máquina (Machine Learning - ML) é uma área da inteligência artificial voltada para o desenvolvimento de algoritmos e modelos com capacidade de adquirir conhecimento e habilidades a partir de dados, sem depender de

programação explícita. Essa abordagem permite que os modelos aprendam e melhorem suas próprias capacidades com o decorrer do tempo, adaptando-se a diversos contextos e tarefas e realizando previsões sobre eventos futuros [7]. Um sistema de ML é apresentado com muitos exemplos significativos para uma tarefa, e neles se encontra estrutura estatística que eventualmente permite que o sistema venha com regras para automatizar a tarefa [10].

O termo série temporal é largamente utilizado na literatura em diversos campos de pesquisa por causa da sua flexibilidade na representação de diferentes tipos de dados. Apesar de não ser um termo recente, ainda é utilizado frequentemente em vários trabalhos atuais [6]. Uma série temporal é uma sequência de valores de uma variável observada ao longo do tempo. A ordem dos dados é essencial e os valores vizinhos são dependentes [18].

O objetivo deste projeto é usar métodos de séries temporais, modelos de ML, para realizar a predição de preço das ações utilizando datasets com preços de compra e venda desses ativos e realizar uma análise comparativa verificando quais foram os melhores parâmetros e identificar qual modelo se saiu melhor para cada cenário.

As principais contribuições deste trabalho envolve mostrar uma análise dos resultados obtidos através dos testes realizados considerando os cenários, pré-processamento dos dados e a diversidade dos modelos utilizados, dos simples para os mais complexos.

A estrutura que o projeto seguirá é a seguinte: Na seção 2 é mostrado exemplos de trabalhos que faz a predição dos preços das ações. Na seção 3 é mostrado os passos seguidos, além de quais modelos foram utilizados. Na seção 4 será exibido os dados que foram extraídos e a forma como foram utilizados, além dos resultados obtidos com os hiperparâmetros usados. Na seção 5 são as considerações finais e trabalhos futuros e na seção 6 é a parte dos agradecimentos.

TRABALHOS RELACIONADOS

Prever os preços das ações não é uma atividade recente, sempre existiu a necessidade de saber o que vai acontecer com o mercado no futuro ou até mesmo no dia seguinte. Esta seção trará alguns trabalhos relacionados aos problemas de predição de valores e suas resoluções propostas na literatura com o propósito de entender as respostas produzidas. No trabalho de [3] foram escolhidas 20 ações com maiores indicadores na B3 e NASDAQ-100 onde os indicadores técnicos foram usados como forma de previsão para essas ações, a análise do preditor consistiu na construção de dois modelos, o primeiro prevê o preço usando indicadores técnicos e o segundo utiliza uma rede neural artificial treinada com base nos indicadores. Os resultados obtidos mostraram que com o uso de redes neurais gerou uma diminuição das perdas e aumento nos ganhos, sem seu uso gera bons resultados em um mercado que não sofre com muita volatilidade.

O trabalho de [18] teve o propósito de fazer uma análise usando redes neurais e máquina de vetores de suporte para prever preços futuros e tendências. Utilizando dados de fechamento diários das ações, foi mostrado ser favorável o uso de redes neurais complexas combinada com a análise técnica e fundamentalista.

No trabalho do [8] foi usado dois tipos de redes neurais para as ações da PETR3, PETR4, ITUB4 e ABEV3 utilizando série temporal univariada e multivariada e utilizando validação cruzada, treino-teste-validação e teste da razão da variância para a hipótese de Random Walk para avaliar o poder de predição dos modelos.

Neste trabalho o [14] utilizou rede neural recorrente para prever o preço de cinco ações que possuem maiores transações diárias para alguns cenários propostos para cada uma com o objetivo de verificar se as duas redes escolhidas conseguem obter boas performances para os valores da série e para outras características.

O objetivo do projeto de [5] foi analisar o desempenho de seis modelos de aprendizado de máquina no qual possuem baixo consumo computacional para realizar previsões dos preços de cinco ações da bolsa de valores brasileira onde o período de extração dos dados coletados foi de 11 anos. Além disso, realizaram dois experimentos com os ativos analisando-os nos períodos de pré e durante a pandemia.

O trabalho do [15] trata da construção de uma carteira de ações americanas alocada com base em algoritmos e regressão. Escolhendo as ações que compõem o índice SP500 foi realizado uma análise exploratória e tratamento dos dados, logo em seguida a construção e avaliação da carteira. Os resultados mostraram que a utilização de ML permitiu a escolha de ações com maior potencial de valorização.

MÉTODOS PROPOSTOS

Nesta seção é descrita a maneira como foi realizada as análises da predição dos ativos coletados utilizando modelos clássicos (modelos estatísticos sobre previsão de séries temporais) e modelos híbridos (modelos com a utilização de redes neurais, além de modelos mais recentes).

Coleta dos Dados

Os dados foram coletados através da API Yahoo! Finances e buscados seis ativos de diferentes setores como é mostrado na Tabela 1. Além disso, o intervalo de tempo selecionado foi do dia 29/12/2019 até 01/08/2023 totalizando 891 dias, mas para a TAESA foi 890 dias por ter um registro NaN (Valor ausente).

Ação	Ticket	Setor
Banco do Brasil	BBAS3	Financeiro
TAESA	TAE11	Energia
Copasa	CSMG3	Saneamento Básico
Magazine	MGLU3	Varejo
Totvs	TOTS3	Tecnologia
Índice Ibovespa	BOVA11	índice

Tabela 1: Informações das ações coletadas

Pré-processamento

Antes de iniciar o treinamento, é necessário realizar o pré-processamento dos dados. A primeira parte é criar dois tipos de dados para serem passados para os modelos: Dados absolutos e dados diferenciados. Os dados absolutos não sofrerão qualquer tipo de transformação diferentemente dos dados diferenciados onde será pegado o valor atual e subtraído pelo valor anterior. O processo de recuperação é feito pegando o valor de referência, salvo antes da transformação, e somado com cada valor subtraído. A segunda parte foi separá-los em janelas temporais de x dias e a previsão é o próximo registro da sequência, o valor que foi previsto vira um preditor para a próxima previsão sendo essa a abordagem de previsão iterativa. A terceira parte foi preparar os dados para treinamento e teste, dependendo do modelo foi preparado um conjunto de dados de

validação, A divisão consistiu em 80% dos dados em treinamento e 20% para teste e do treinamento foi separado novamente 80% em treino e 20% para validação.

Modelos

- **ARIMA:** São modelos autoregressivos integrados com média móvel. Identificam os valores da variável dependente relacionados com seus próprios valores passados tendo como base a ideia de “deixar os dados falarem por si mesmos”, como o [12] ressalta.
- **SARIMA:** SARIMA é uma extensão do ARIMA que lida com a sazonalidade dos dados [2].
- **Holt-Winters:** A suavização exponencial de Holt-Winters é usada quando os dados exibem tendência e sazonalidade. Os dois principais modelos são Modelo aditivo exibindo sazonalidade aditiva e Modelo multiplicativo exibindo sazonalidade multiplicativa [13].
- **MLPs:** São Perceptrons de Multicamadas. Rede neural artificial com uma ou mais camadas ocultas contendo um número indeterminado de neurônios.
- **LSTM:** Unidades de Memória de Longo Prazo, ou simplesmente LSTM, são um tipo especial de Rede Neural Recorrente capaz de aprender dependências de longo prazo. Além disso, tem a capacidade de remover ou adicionar informações ao estado da célula, cuidadosamente regulado por estruturas chamadas de portões, que são uma forma opcional de permitir a passagem de informações [17].
- **GRU:** Semelhante a LSTM, a Gated Unit Recurrent (GRU) surgiu como uma tentativa para resolver o problema de dissipação de gradientes das Redes Neurais Recorrentes [14].
- **Transformers:** É uma arquitetura codificador-decodificador que depende exclusivamente do mecanismo de atenção e estão no domínio do Processamento Natural de Linguagem, com transformadores generativos pré-treinados sendo aplicados a muitos dos desafios do domínio [4].
- **TimeGPT:** Modelo pré-treinado para previsão de séries temporais que pode produzir previsões precisas em uma gama diversificada de domínios e aplicações sem formação adicional [1].
- **AutoGluon:** É um AutoML de serviço aberto para previsão probabilística de séries temporais. Somado a isso, pode gerar previsões pontuais e estocásticas para coleções de séries temporais univariadas [16].

Teste e Hiper-parâmetros

Os ajustes feitos são essenciais para alcançar melhores resultados de acordo com os dados passados e seu processamento. Por serem diferentes o seu refinamento levou tempo para melhorar a resposta dos modelos. Modelos clássicos são mais simples e fácil de treinar em comparação aos híbridos que dependendo da arquitetura utilizada e dos hiper-parâmetros selecionandos podem ser demorados.

Teste 1

Os modelos utilizados nesse teste foram o ARIMA, SARIMA, Holt-Winters, MLP, LSTM, GRU e o Transformers. Na tabela 2 é mostrado os hiper-parâmetros do modelo ARIMA onde a Ordem significa os parâmetros (p, d, q) que são parâmetros de, respectivamente, auto-regressivo, diferenciação e média móvel. O segundo hiper-parâmetro é a Tendência que vai adicionar tendência determinística, o termo “t”

significa que o modelo terá tendência linear e “-” significa que não tem tendência. Impor estacionariedade exige, quando é verdadeiro, que os parâmetros auto-regressivos correspondam ao processo de estacionariedade e a Escala de concentrado vai concentrar (Verdadeiro) ou não (Falso) a escala da variação do termo de erro fora da probabilidade.

Ação	Ordem	Tendência	Impor estacionariedade	Escala de concentrado
Banco do Brasil	(2, 3, 5)	-	Verdadeiro	Falso
TAESA	(0, 0, 5)	-	Falso	Verdadeiro
Copasa	(1, 0, 5)	t	Falso	Verdadeiro
Magazine	(3, 1, 5)	-	Falso	Verdadeiro
Totvs	(0, 0, 5)	-	Verdadeiro	Falso
Índice Ibovespa	(0, 1, 5)	-	Verdadeiro	Falso

Tabela 2: Hiper-parâmetro do modelo ARIMA

Para o Sarima, os hiper-parâmetros testados foram a Ordem e Tendência onde a Ordem agora vai mudar para (p, d, q, s), os três primeiros termos seguem a explicação do ARIMA e o termo “s” significa a aplicação da sazonalidade. Para as ações testadas os ajustes foram o seguinte: Banco do Brasil com a Ordem (2, 0, 5, 12) sem tendência; TAESA com a Ordem (1, 0, 5, 4) com tendência “c” que inclui termo constante; Copasa com a Ordem (0, 0, 5, 12) com tendência “t”; Magazine com a Ordem (3, 0, 5, 4); Totvs com a Ordem (4, 0, 5, 6); Ibovespa com a Ordem (4, 0, 5, 6) com a tendência “c”.

Na tabela 3 é exibido os hiper-parâmetros do Holt-Winters. O termo Sazonalidade indica os tipos de sazonalidades que será trabalhado onde “mul” indica que os padrões sazonais variam em proporção ao nível da série e “add” tem os padrões sazonais constantes. O Período Sazonal é a quantidade de períodos em um ciclo sazonal completo, para a Tendência o termo “-” significa que não há tendência, “mul” é a taxa de crescimento/declínio ser proporcional ao nível atual da série temporal e “add” e a taxa ser constante ao longo do tempo. A Tendência Amortecida quando é Verdadeira indica que o componente de tendência deve ser amortecido.

Ação	Sazonalidade	Período Sazonal	Tendência	Tendência Amortecida
Banco do Brasil	mul	100	-	Falso
TAESA	mul	100	-	Falso
Copasa	mul	100	-	Falso
Magazine	mul	100	mul	Verdadeiro
Totvs	mul	50	mul	Verdadeiro
Índice Ibovespa	add	20	add	Verdadeiro

Tabela 3: Hiper-parâmetro do modelo Holt-Winters

Nas tabelas 4, 5, 6 e 7 a explicação dos hiper-parâmetros é a mesma. Os Neurônios refere-se a quantidade de neurônios em cada camada, se tiver (60, 60) quer dizer que foram duas camadas contendo 60 neurônios cada. A Função de Ativação é uma transformação não-linear que vai ativar a saída do neurônio com base nas entradas recebidas e as funções utilizadas no projeto foram: relu, linear, selu e elu. Quando tiver, por exemplo, (linear, relu) com 2 camadas de neurônios significa que a primeira camada tem a função linear e a segunda, relu. Quando tiver N camadas e ter uma função de ativação quer dizer que todas as camadas tem aquela ativação.

Se tiver, por exemplo, uma camadas a mais de função de ativação em comparação ao número de camadas de neurônios que dizer que a camada de saída terá a última função especificada da camada de ativação, por padrão a camada de saída tem a

função relu. O número de Épocas quer dizer a quantidade de execuções que o modelo treinou (no projeto não teve callbacks). Tamanho do Lote é a quantidade de amostras coletadas dos dados de treino que são processadas pelo modelo antes de atualizar os pesos durante o treinamento. O Dropout vai eliminar aleatoriamente uma porcentagem de conexões entre neurônios, isso ajuda a diminuir o problema do overfitting que faz o modelo "aprender demais" sobre os dados e não pegar tão bem as características dos dados de treinamento.

Ação	Neurônios	Ativação	Épocas	Tamanho do Lote	Dropout
Banco do Brasil	(60, 20)	relu	90	20	-
TAESA	(65, 40)	relu	110	28	2,5%
Copasa	(70, 60, 50)	relu	110	36	-
Magazine	(70, 70, 50)	(selu, linear, linear)	110	36	-
Totvs	(50, 50)	relu	110	28	-
Índice Ibovespa	(70, 70, 70)	relu	120	28	-

Tabela 4: Hiper-parâmetro do modelo MLPs

Ação	Neurônios	Ativação	Épocas	Tamanho do Lote
Banco do Brasil	(60, 12)	relu	100	28
TAESA	(65, 25)	(linear, relu)	110	28
Copasa	(40, 40, 40)	relu	130	40
Magazine	(60, 60, 60)	relu	120	28
Totvs	(70, 70, 70)	selu	130	28
Índice Ibovespa	(60, 60, 60)	(relu, relu, elu)	110	36

Tabela 5: Hiper-parâmetro do modelo LSTM

Ação	Neurônios	Ativação	Épocas	Tamanho do Lote
Banco do Brasil	55	relu	80	20
TAESA	65	linear	90	24
Copasa	50	relu	110	28
Magazine	(60, 60)	relu	130	24
Totvs	(70, 70)	selu	120	24
Índice Ibovespa	(70, 50)	linear	120	24

Tabela 6: Hiper-parâmetro do modelo GRU

Ação	Neurônios	Ativação	Épocas	Tamanho do Lote	Dropout
Banco do Brasil	200	relu	65	28	20%
TAESA	(200, 200)	linear	70	24	10%
Copasa	(200, 200)	relu	80	28	2%
Magazine	(200, 200)	relu	80	24	1%
Totvs	200	(selu, elu)	80	24	1%
Índice Ibovespa	200	elu	80	24	1%

Tabela 7: Hiper-parâmetro do modelo Transformers

Teste 2

Para este teste, os modelos utilizados são SARIMA, Holt-Winters, MLP, GRU e Transformers e foi testada para a ação do Banco do Brasil para prever 5 dias, mas para os modelos híbridos a janela temporal usada foi de 5 dias. Para o SARIMA foi usada

uma biblioteca diferente e a melhor Ordem foi (3, 1, 2)(0, 0, 0) para os dados absolutos e (3, 0, 2)(0, 0, 0) para os dados diferenciados.

Para o Holt-Winters, os hiper-parâmetros considerados foram: Tendência, Sazonalidade e Período sazonal onde o primeiro e último foram explicados no teste 1. A Sazonalidade aplicada tem os mesmos valores "add" e "mul" onde o primeiro a amplitude sazonal é constante ao longo do tempo e o segundo varia proporcionalmente ao nível da série. Para os dados absolutos os ajustes foram "add", "mul" e 200, respectivamente, considerando os hiper-parâmetros mencionados acima. Para os dados diferenciados os ajustes foram "add", "add" e 220, respectivamente.

Para os modelo usando redes neurais, os hiper-parâmetros considerados foram: Quantidade de Neurônios, Função de Ativação, Quantidade de Epocas e Tamanho do Lote. Considerando o MLP para os dados absolutos foram usadas 3 camadas ocultas contendo 60 neurônios cada onde as duas primeiras função de ativação foi elu e a terceira, relu. Além disso a quantidade de épocas foi 100 e o tamanho do lote foi 30. Para os dados diferenciados foram usadas 4 camadas ocultas contendo 100 neurônios cada onde em todas a função de ativação é a relu. Somado a isso, a quantidade de épocas foi 150 com um tamanho do lote sendo 32.

Considerando o GRU para os dados absolutos foi usada 1 camada oculta contendo 60 neurônios com função de ativação relu e treinado com tamanho do lote igual a 20 e épocas sendo 80. Para os dados diferenciados foram usadas 3 camadas ocultas contendo 100 neurônios cada onde a primeira contém a função de ativação elu e as outras duas, linear. Além disso, o tamanho do lote é 32 e a quantidade de épocas usadas foi 130.

Considerando o Transformers para os dados absolutos foi usada uma camada oculta contendo 200 neurônios com função de ativação selu e com tamanho do lote sendo 24 épocas sendo 65. Para os dados diferenciados foi uma camada oculta contendo 150 neurônios com função de ativação selu e com tamanho do lote e épocas sendo 32 e 70, respectivamente.

Teste 3

Após o teste 2, foi escolhido os dois melhores modelos (GRU e Transformers, mas a utilizada nesse teste foi o GRU) e agora considerando os indicadores preço do Dólar e Índice Ibovespa para a previsão do Banco do Brasil, TAESA, Copasa, Magazine e Totvs. O número de dias previstos foram de 178 dias, também considerando os dados sendo absolutos e diferenciados, como é mostrado na tabela 8.

Ação	Neurônios	Ativação	Épocas	Tamanho do Lote
Banco do Brasil	(130, 130, 130)	selu	130	32
TAESA	120	relu	150	32
Copasa	(120, 120)	(relu, elu)	150	64
Magazine	130	linear	140	32
Totvs	130	linear	130	64
Ações	Neurônios	Ativação	Épocas	Tamanho do Lote
Banco do Brasil	130	linear	130	32
TAESA	120	linear	130	64
Copasa	130	elu	140	32
Magazine	(130, 130)	(selu, linear)	140	32
Totvs	120	linear	140	64

Tabela 8: Hiper-parâmetro do modelo GRU considerando indicadores. Tabela de cima são os dados absolutos e a debaixo os dados diferenciados

Teste 4

Nesse experimento foi utilizado os modelos GRU, Transformers, TimeGPT e AutoGluon, esses dois últimos são modelos recentemente criados. A janela temporal usada é de 15 dias para prever 5, 15, 25 e 35 dias. Na tabela 9 é referente ao modelo GRU na ação do Banco do Brasil.

Na tabela 10 é usado o modelo GRU na ação da TAESA.

Na tabela 11 é usado o modelo GRU na ação da Copasa.

Na tabela 12 é usado o modelo GRU na ação da Magazine Luiza.

Na tabela 13 é usado o modelo GRU na ação da Totvs.

Na tabela 14 é usado o modelo GRU no Índice Ibovespa.

Na tabela 15 é usado o modelo Transformers na ação do Banco do Brasil.

Na tabela 16 é usado o modelo Transformers na ação da TAESA.

Na tabela 17 é usado o modelo Transformers na ação da Copasa.

Na tabela 18 é usado o modelo Transformers na ação da Magazine Luiza.

Na tabela 19 é usado o modelo Transformers na ação da Totvs.

Na tabela 20 é usado o modelo Transformers no Índice Ibovespa.

Para o modelo Transformers, a função de ativação na última camada (saída) foi testada com outras funções.

Dias previstos	Neurônios	Ativação	Épocas	Tamanho do Lote
5 dias	100	selu	120	64
15 dias	(100, 100, 100)	(relu, relu, elu)	140	64
25 dias	(120, 120)	elu	150	64
35 dias	(140, 140)	(linear, relu)	160	32
Dias previstos	Neurônios	Ativação	Épocas	Tamanho do Lote
5 dias	(120, 120, 120)	relu	130	32
15 dias	(130, 120)	(relu, selu)	130	32
25 dias	(120, 120)	(relu, linear)	160	32
35 dias	(120, 120)	(relu, relu)	140	32

Tabela 9: Hiper-parâmetro do modelo GRU para o Banco do Brasil para os dados absoluto (tabela de cima) e diferenciados (tabela de baixo)

Dias previstos	Neurônios	Ativação	Épocas	Tamanho do Lote
5 dias	(110, 110)	elu	130	64
15 dias	(120, 120)	selu	130	32
25 dias	(120, 120)	selu	130	64
35 dias	(120, 120, 120)	elu	150	32
Dias previstos	Neurônios	Ativação	Épocas	Tamanho do Lote
5 dias	(120, 120)	selu	120	64
15 dias	(120, 120, 120)	selu	120	64
25 dias	(120, 120, 120)	elu	120	64
35 dias	(120, 120, 120)	linear	130	64

Tabela 10: Hiper-parâmetro do modelo GRU para a TAESA para os dados absoluto (tabela de cima) e diferenciados (tabela de baixo)

Dias previstos	Neurônios	Ativação	Épocas	Tamanho do Lote
5 dias	(110, 110)	selu	130	32
15 dias	(120, 120)	elu	130	64
25 dias	(120, 120, 120)	elu	130	64
35 dias	(120, 120)	selu	130	64
Dias previstos	Neurônios	Ativação	Épocas	Tamanho do Lote
5 dias	(120, 120, 120)	relu	150	32
15 dias	(120, 120)	(relu, linear)	140	64
25 dias	(130, 130)	(relu, elu)	140	64
35 dias	(120, 120, 120)	relu	140	64

Tabela 11: Hiper-parâmetro do modelo GRU para a Copasa para os dados absoluto (tabela de cima) e diferenciados (tabela de baixo)

Dias previstos	Neurônios	Ativação	Épocas	Tamanho do Lote	
5 dias	(120, 120)	(relu, linear)	150	64	
15 dias	(120, 120, 100)	linear	140	64	
25 dias	(120, 120)	(relu, linear)	150	64	
35 dias	120	linear	150	64	
Dias previstos	Neurônios	Ativação	Épocas	Tamanho do Lote	Dropout
5 dias	130	selu	160	32	5%
15 dias	(120, 120)	(relu, elu)	140	64	-
25 dias	(130, 130)	(selu, elu)	140	32	-
35 dias	(120, 120)	selu	140	32	-

Tabela 12: Hiper-parâmetro do modelo GRU para a Magazine para os dados absoluto (tabela de cima) e diferenciados (tabela de baixo)

Dias previstos	Neurônios	Ativação	Épocas	Tamanho do Lote
5 dias	(120, 120)	(relu, selu)	140	64
15 dias	(120, 120)	(relu, selu)	140	32
25 dias	(120, 120)	(relu, selu)	130	64
35 dias	(120, 120)	(elu, selu)	140	64
Dias previstos	Neurônios	Ativação	Épocas	Tamanho do Lote
5 dias	120	elu	140	32
15 dias	(110, 120)	relu	110	64
25 dias	120	elu	140	64
35 dias	(120, 120)	elu	150	64

Tabela 13: Hiper-parâmetro do modelo GRU para a Totvs para os dados absoluto (tabela de cima) e diferenciados (tabela de baixo)

Dias previstos	Neurônios	Ativação	Épocas	Tamanho do Lote
5 dias	(130, 130, 130)	(relu, relu, selu)	140	32
15 dias	(120, 120)	(relu, linear)	140	32
25 dias	(120, 120)	(relu, linear)	140	32
35 dias	(120, 120)	(elu, selu)	140	64
Dias previstos	Neurônios	Ativação	Épocas	Tamanho do Lote
5 dias	120	elu	140	32
15 dias	(110, 120)	relu	110	64
25 dias	120	elu	140	64
35 dias	(120, 120)	elu	150	64

Tabela 14: Hiper-parâmetro do modelo GRU para o Índice Ibovespa para os dados absoluto (tabela de cima) e diferenciados (tabela de baixo)

Dias previstos	Neurônios	Ativação	Épocas	Tamanho do Lote	Dropout
5 dias	200	(linear, relu)	95	32	10%
15 dias	(200, 200)	(relu, relu, elu)	100	32	(10%, 10%)
25 dias	(200, 200)	(relu, linear, linear)	130	64	(10%, 10%)
35 dias	(200, 200)	(relu, relu, linear)	110	64	(10%, 10%)

Dias previstos	Neurônios	Ativação	Épocas	Tamanho do Lote	Dropout
5 dias	200	(elu, relu)	110	32	10%
15 dias	220	linear	100	64	10%
25 dias	220	(linear, selu)	100	32	10%
35 dias	230	linear	100	32	10%

Tabela 15: Hiper-parâmetro do modelo Transformers para o Banco do Brasil para os dados absoluto (tabela de cima) e diferenciados (tabela de baixo)

Dias previstos	Neurônios	Ativação	Épocas	Tamanho do Lote	Dropout
5 dias	200	(linear, selu)	70	32	5%
15 dias	200	(elu, linear)	80	32	5%
25 dias	200	(elu, selu)	70	32	5%
35 dias	200	(selu, selu)	70	64	5%
Dias previstos	Neurônios	Ativação	Épocas	Tamanho do Lote	Dropout
5 dias	200	(linear, linear)	70	64	10%
15 dias	200	(linear, linear)	80	64	10%
25 dias	200	(elu, linear)	70	64	5%
35 dias	200	(linear, selu)	70	64	5%

Tabela 16: Hiper-parâmetro do modelo Transformers para a TAESA para os dados absoluto (tabela de cima) e diferenciados (tabela de baixo)

Dias previstos	Neurônios	Ativação	Épocas	Tamanho do Lote	Dropout
5 dias	200	(relu, selu)	70	32	5%
15 dias	200	(selu, elu)	70	64	5%
25 dias	200	(relu, selu)	70	32	5%
35 dias	200	(elu, relu)	70	64	5%
Dias previstos	Neurônios	Ativação	Épocas	Tamanho do Lote	Dropout
5 dias	200	(linear, elu)	70	64	10%
15 dias	200	(elu, elu)	80	64	10%
25 dias	200	(elu, selu)	70	64	5%
35 dias	200	(selu, selu)	70	32	5%

Tabela 17: Hiper-parâmetro do modelo Transformers para a Copasa para os dados absoluto (tabela de cima) e diferenciados (tabela de baixo)

Dias previstos	Neurônios	Ativação	Épocas	Tamanho do Lote	Dropout
5 dias	200	(linear, relu)	70	32	5%
15 dias	200	(relu, elu)	70	32	5%
25 dias	200	(linear, elu)	70	64	5%
35 dias	200	(selu, elu)	70	64	5%
Dias previstos	Neurônios	Ativação	Épocas	Tamanho do Lote	Dropout
5 dias	200	(selu, elu)	80	64	5%
15 dias	200	(selu, selu)	70	64	5%
25 dias	200	(linear, elu)	80	32	5%
35 dias	200	(linear, selu)	70	64	5%

Tabela 18: Hiper-parâmetro do modelo Transformers para a Magazine para os dados absoluto (tabela de cima) e diferenciados (tabela de baixo)

Dias previstos	Neurônios	Ativação	Épocas	Tamanho do Lote	Dropout
5 dias	200	(elu, relu)	70	64	5%
15 dias	200	(selu, relu)	70	32	5%
25 dias	200	(elu, relu)	70	64	5%
35 dias	200	(selu, relu)	70	32	5%
Dias previstos	Neurônios	Ativação	Épocas	Tamanho do Lote	Dropout
5 dias	200	(linear, elu)	70	32	5%
15 dias	200	(linear, elu)	70	32	5%
25 dias	200	(elu, elu)	80	32	5%
35 dias	200	(elu, elu)	80	32	5%

Tabela 19: Hiper-parâmetro do modelo Transformers para a Totvs para os dados absoluto (tabela de cima) e diferenciados (tabela de baixo)

Dias previstos	Neurônios	Ativação	Épocas	Tamanho do Lote	Dropout
5 dias	200	(linear, selu)	70	64	5%
15 dias	200	(elu, elu)	50	64	10%
25 dias	210	(relu, selu)	70	32	10%
35 dias	200	(elu, linear)	70	32	10%
Dias previstos	Neurônios	Ativação	Épocas	Tamanho do Lote	Dropout
5 dias	210	(linear, linear)	60	64	10%
15 dias	220	(linear, selu)	60	64	5%
25 dias	220	(linear, selu)	80	32	5%
35 dias	200	(selu, linear)	70	64	10%

Tabela 20: Hiper-parâmetro do modelo Transformers para a Índice Ibovespa para os dados absoluto (tabela de cima) e diferenciados (tabela de baixo)

Para o TimeGPT não foi necessário criar janelas temporais, apenas o ajuste dos hiper-parâmetros. Os ajustes feitos foram o tipo de modelo especificado (se é "timegpt" ou "timegpt-1-long-horizon") e Passos de ajustes que é a etapa do modelo pré-treinado para se adequar aos dados específicos, melhorando a capacidade de previsão, ele representa um valor inteiro que foi variado de 1 a 40, tanto para os dados absolutos quanto para os dados diferenciados.

Para o AutoGluon também não foi necessário criar janelas temporais e sim um processamento para transformar em um dataframe contendo as colunas Data, ticket da Ação e o valor de fechamento. Além disso, foi especificado a frequência das datas, que são diárias, a variável alvo que é o valor de fechamento e quantos dias para serem previstos.

MÉTRICAS

As métricas utilizadas para avaliar os resultados obtidos foram:

MSE (Mean Squared Error): Métrica amplamente utilizada para avaliar a precisão de modelos preditivos, especialmente em contextos de regressão.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (1)$$

RMSE (Root Mean Squared Error): Mede a diferença média entre os valores preditos do modelo estatístico e valores reais. Matematicamente, é o desvio padrão da distância dos valores para a linha de regressão.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2)$$

MAE (Mean Absolute Error): Calcula a média da diferença absoluta entre os valores previstos pelo modelo e os valores observados.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3)$$

MAPE (Mean Absolute Percentual Error): Bastante utilizada em previsão de séries temporais, mostra o erro percentual em relação aos valores reais.

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (4)$$

Coefficiente de Determinação (R²): Demonstra o nível de confiança que o modelo possui.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (5)$$

EXPERIMENTOS

Esta seção visa mostrar os resultados dos experimentos realizados nos modelos em cada ação e foram separados em quatro tipos de testes. Os resultados serão a média de cada resultado das métricas dos modelos para cada cenário, menos para o Experimento 2 que é considerado o RMSE e MAPE diretamente porque só foi usada uma ação.

Experimento 1

Neste experimento foi utilizados todas as 6 ações e previstos 178 dias. Como mencionado anteriormente, foi calculado a média de cada resposta das métricas de cada ação de acordo com o modelo utilizado, Os dois melhores modelos foram o ARIMA e MLP para o primeiro teste como é mostrado na tabela 21.

Modelos	MSE	RMSE	MAE	MAPE	R2
ARIMA	8.065	2.316	1.868	7.115%	-3.401%
SARIMA	25.856	3.001	2.396	7.753%	-46.110%
Holt-Winters	16.547	3.664	2.988	10.650%	-197.614%
MLP	9.197	2.552	2.139	7.045%	-41.382%
LSTM	11.217	2.929	2.302	8.459%	-71.675%
GRU	9.647	2.623	2.192	7.585%	-45.531%
Transformers	13.109	3.135	2.556	8.084%	-201.503%

Tabela 21: Média dos resultados dos modelos testados nas ações.

Experimento 2

Neste experimento o teste foi apenas na ação do Banco do Brasil e desta vez apenas utilizando 5 modelos sendo o SARIMA, Holt-Winters, MLP, GRU e Transformers para uma janela temporal de 5 dias para prever 5 dias. Os dois melhores modelos foram o MLP e o GRU tanto para os dados absolutos quanto para os dados diferenciados como é mostrado na tabela 22, mas o modelo Transformers estava conseguindo pegar características, mesmo que as métricas não sejam tão favoráveis. Por causa disso para

os próximos testes foi apenas utilizados o GRU e Transformers e a adição de dois novos modelos.

Experimento 3

Neste experimento foi testado o modelo GRU para todas as ações, mas para o teste foi utilizado técnicas como pesquisa em grade (grid search) para ajudar a testar várias configurações de hiper-parâmetros. Como é mostrado na tabela 23, em média os melhores resultados foram utilizando os dados absolutos, mas para o próximo teste não foi mais utilizado o grid search.

Modelos	RMSE sem diff	RMSE com diff	MAPE sem diff	MAPE com diff
SARIMA	1.301	1.301	2.664%	2.662%
Holt-Winters	0.573	0.596	0.966%	0.987%
MLPs	0.371	0.662	0.745%	1.239%
GRU	0.403	0.445	0.783%	0.848%
Transformers	0.400	1.148	0.614%	2.332%

Tabela 22: Somente para a Ação do Banco do Brasil

Métricas	Dados Absolutos	Dados Diferenciados
MSE	0.253	29.631
RMSE	0.412	4.029
MAE	0.310	3.386
MAPE	1.568%	12.343%
R2	74.102%	-403.662%

Tabela 23: Resultados do modelo GRU para cada métrica considerando os dados absolutos e diferenciados

Experimento 4

Neste experimento foi testado os modelos GRU, Transformers, TimeGPT e AutoGluon em todas as ações para prever 5, 15, 25 e 35 dias tanto para os dados absolutos quanto para os dados diferenciados e consideranco a métrica R2 como determinante para dizer o melhor, visto que para as outras métricas quanto maior for a quantidade de dias previstos mais erro é acumulado. Na tabela 24 é mostrado os resultados para o modelo GRU onde o cenário de 25 dias foi melhor em comparação aos outros cenários do modelo tanto para os dados absolutos quanto para os dados diferenciados. Além disso, o R2 foi maior para os dados diferenciados mostrando que o processamento foi adequado para melhorar os resultados.

Métricas	Prever 5 dias	Prever 15 dias	Prever 25 dias	Prever 35 dias
MSE	0.219	0.374	0.620	1.053
RMSE	0.334	0.481	0.624	0.757
MAE	0.287	0.400	0.518	0.624
MAPE	0.668%	1.229%	1.714%	1.868%
R2	-104.847%	-31.474%	19.622%	-10.913%
Métricas	Prever 5 dias	Prever 15 dias	Prever 25 dias	Prever 35 dias
MSE	0.200	0.973	0.415	0.816
RMSE	0.308	0.779	0.477	0.790
MAE	0.239	0.697	0.389	0.661
MAPE	0.580%	2.007%	1.276%	2.526%
R2	-39.444%	-379.939%	36.182%	-38.533%

Tabela 24: Resultados do modelo GRU considerando os dados absolutos (tabela de cima) e diferenciados (tabela de baixo)

Na tabela 25 é mostrado os resultados para o modelo Transformers onde os melhores resultados foram para os cenários de 35 dias para os dados absolutos e 25 dias para os dados diferenciados.

Métricas	Prever 5 dias	Prever 15 dias	Prever 25 dias	Prever 35 dias
MSE	0.375	0.645	1.024	1.380
RMSE	0.391	0.618	0.836	0.893
MAE	0.326	0.542	0.685	0.772
MAPE	0.766%	1.671%	3.346%	2.164%
R2	-43.164%	-120.615%	-142.007%	-32.125%
Métricas	Prever 5 dias	Prever 15 dias	Prever 25 dias	Prever 35 dias
MSE	0.222	0.741	0.916	1.521
RMSE	0.369	0.708	0.752	1.009
MAE	0.329	0.772	0.656	0.851
MAPE	0.836%	1.675%	1.994%	3.116%
R2	-347.851%	-237.983%	-66.202%	-357.172%

Tabela 25: Resultados do modelo Transformers considerando os dados absolutos (tabela de cima) e diferenciados (tabela de baixo)

Na tabela 26 é mostrado os resultados para o modelo TimeGPT onde o melhor resultado foi para o cenário de 35 dias para os dados absolutos e 15 dias para os dados diferenciados.

Métricas	Prever 5 dias	Prever 15 dias	Prever 25 dias	Prever 35 dias
MSE	0.420	0.861	1.122	0.744
RMSE	0.404	0.719	0.905	0.786
MAE	0.337	0.614	0.794	0.661
MAPE	1.590%	2.080%	3.786%	4.843%
R2	-146.690%	-334.250%	-255.235%	-119.567%
Métricas	Prever 5 dias	Prever 15 dias	Prever 25 dias	Prever 35 dias
MSE	0.225	0.354	3.589	16.885
RMSE	0.362	0.494	1.377	3.180
MAE	0.313	0.426	1.197	2.764
MAPE	1.320%	1.670%	4.160%	11.01%
R2	-643.430%	-187.900%	-883.200%	-3134.950%

Tabela 26: Resultados do modelo TimeGPT considerando os dados absolutos (tabela de cima) e diferenciados (tabela de baixo)

Na tabela 27 é mostrado os resultados para o modelo AutoGluon onde o melhor resultado foi para o cenário de 35 dias para os dados absolutos e 15 dias para os dados diferenciados.

Métricas	Prever 5 dias	Prever 15 dias	Prever 25 dias	Prever 35 dias
MSE	0.800	1.323	2.927	8.842
RMSE	0.661	0.939	1.401	2.024
MAE	0.596	0.768	1.197	1.780
MAPE	1.607%	2.019%	3.696%	5.327%
R2	-796.979%	-312.937%	-230.232%	-206.907%
Métricas	Prever 5 dias	Prever 15 dias	Prever 25 dias	Prever 35 dias
MSE	1.303	1.090	2.684	8.563
RMSE	0.889	0.809	1.487	2.710
MAE	0.762	0.672	1.294	2.340
MAPE	2.340%	2.236%	6.439%	12.872%
R2	-4175.908%	-491.962%	-925.772%	-2225.861%

Tabela 27: Resultados do modelo AutoGluon considerando os dados absolutos (tabela de cima) e diferenciados (tabela de baixo)

CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS

A partir das análises comparativas, é possível concluir que os modelos sofrem bastante variações nos resultados considerando o tipo de cenário, a forma como os dados são tratados e quantos dias deverão ser previstos. Além disso, a escolha entre um ou outro modelo também afetará os resultados desejados, visto que dependendo se é um modelo clássico ou híbrido e da arquitetura configurada para uma situação específica as respostas obtidas podem variar de múltiplas formas.

Por fim, os trabalhos futuros envolvem a utilização de mais dias quando utilizar as janelas temporais podendo ser 25, 35 ou mais dias para passar como treinamento, bem como definir outras abordagens que não seja a previsão iterativa usada no presente artigo, usar janelas fixas como resposta pode ser uma alternativa. Também é interessante realizar uma análise exploratória detalhada sobre os dados para focar nas suas características e assim melhorar a arquitetura e hiper-parâmetros dos modelos.

Agradecimentos

Agradeço à FAPEAM (Fundação de Amparo à Pesquisa do Estado do Amazonas) pelo auxílio financeiro e material provido e ao LSI (Laboratório de Sistemas Inteligentes) por fornecer o espaço de trabalho e todo o apoio/dicas para desenvolvimento do projeto.

REFERÊNCIAS

1. Cristian Challu Azul Garza and Max Mergenthaler-Canseco. Timept-1. *arXiv*, 2024.
2. M Bilgili. Time series forecasting on cooling degree-days (cdd) using sarima model. *Nat Hazards*, 2023.
3. Wagner Igarashi e Deisy Cristina Corrêa Igarashi Charles Chacim Esteves Pereira. Análise comparativa de modelos de indicadores técnicos e de redes neurais para predição de ações. *International Journal of Image, Graphics and Signal Processing*, 2022.
4. Z. Chen. Stock price prediction with denoising autoencoder and transformers. *Highlights in Science, Engineering and Technology*, 2024.
5. Gabriel de Lima e Silva, Rennan de Aguiar Ramos, and Igor Barbosa da Costa. Avaliação de algoritmos de aprendizado de máquina para a predição da bolsa de valores do brasil. 2022.
6. Guilherme Albertini de Oliveira e Paulo Sérgio Silva Rodrigues. Predição do mercado financeiro com uma arquitetura de extração de contexto baseada em decomposição de series temporal. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022.
7. Eggleston Patricio de Oliveira Souza. Análise comparativa de modelos de machine learning e convolutional neural networks na detecção de falhas em máquinas rotativas. Dissertação (mestrado em engenharia de produção), Universidade Federal de Pernambuco, Caruaru, 2024.
8. Thiago José Pinheiro e Hermes Yukio Higachi. Modelos de deep learning e previsão de ações: Estudo de casos da bolsa brasileira. *International Journal of Image, Graphics and Signal Processing*, 2022.
9. Tito Francisco Ianda e Marcelo Scherer Perlin. Mercado de ações: um estudo da diferença de preços entre ações ordinárias e preferenciais no brasil. In *European Conference on Computer Vision (ECCV)*, 2014.
10. Paulo Cesar Pereira Alves e Sandra Cristina Costa Prado. Estudo comparativo entre algoritmos de machine learning aplicados à previsão de séries temporais do mercado financeiro. *IEEE Conference on Computer Vision and Pattern Recognition*, 2022.
11. Dante Augusto Couto Barone e Luiz Otávio Vilas Boas Oliveira Fabio Alberto Prochnow. Programação genética para predição de séries temporais aplicadas a mercados financeiros. *International Journal of Image, Graphics and Signal Processing*, 2013.
12. Porter D. Gujarati, D. Econometria básica. *Porto Alegre, AMGH Editora Ltda.*, 2011.
13. Prajakta S. Kalekar. Time series forecasting using holt-winters exponential smoothing. *Kanwal Rekhi school of information Technology*, 2004.
14. Matheus Morgado Correa; Flávia Cristina Bernardini e Jane Vieira Volotto. Marcos Vinicius De Oliveira Ribeiro. Previsão de preços de ações baseada em redes neurais recorrentes lstm e gru. *International Journal of Image, Graphics and Signal Processing*, 2022.

15. Giancarlo Lucca Gracaliz Dimuro Tiago Asmus Mario Gambim, Heloisa A. Camargo. Uma estratégia para alocação de carteira de ações usando algoritmos de aprendizado de máquina e regras fuzzy. *Porto Alegre: Sociedade Brasileira de Computação*, 2023.
16. Nick Erickson Huibin Shen Alexander Shirkov Tony Hu Bernie Wang Oleksandr Shchur, Ali Caner Turkmen. Autogluon–timeseries: Automl for probabilistic time series forecasting. *Proceedings of the Second International Conference on Automated Machine Learning, PMLR*, 2023.
17. Julia Naves Rodrigues and Dra. Nádia Giaretta Biase. Inflação no brasil: uma aplicação de séries temporais e redes neurais recorrentes. 2021.
18. Lucas Pimenta Vasco and Heloisa de Arruda Camargo. Um estudo de redes neurais recorrentes no contexto de previsões no mercado financeiro, 2020.
19. Jeffrey Wurgler. Financial markets and the allocation of capital. *Journal of Financial Economics*, page 2, 2000.