

24º Congresso Nacional de Iniciação Científica

TÍTULO: PREVISÃO DO PREÇO DAS AÇÕES USANDO APRENDIZADO DE MÁQUINA

CATEGORIA: CONCLUÍDO

ÁREA: CIÊNCIAS EXATAS E DA TERRA

SUBÁREA: Computação e Informática

INSTITUIÇÃO: Faculdade de Tecnologia de Bauru - FATEC

AUTOR(ES): RODRIGO MORATTO AGUILHAR, LUCAS BORGES DE ARAÚJO, MARCOS JOSÉ PIO

ORIENTADOR(ES): PATRICIA BELLIN RIBEIRO

1. RESUMO

Este artigo apresenta um estudo sobre a previsão de preços de ações usando técnicas de aprendizado de máquina, com base na análise de um projeto existente. O objetivo é desenvolver modelos de previsão precisos que possam ajudar os investidores a tomar decisões informadas. O conjunto de dados utilizado contém informações históricas de preços de ações, como abertura, fechamento, mínimo e máximo. Diversos algoritmos de aprendizado de máquina, como Regressão Linear, Support Vector Machines e Redes Neurais Artificiais, são implementados e comparados em termos de desempenho. Os resultados mostram que os modelos de aprendizado de máquina podem fornecer previsões úteis para os preços das ações, embora seja importante considerar outras variáveis e técnicas para melhorar ainda mais a precisão.

2. INTRODUÇÃO

Quando se fala sobre a previsão de preços de ações é um aspecto crucial para investidores e instituições financeiras, pois ajuda na tomada de decisões informadas sobre compra, venda ou retenção de ativos. Com o avanço da tecnologia e o aumento da disponibilidade de dados, técnicas de aprendizado de máquina têm sido cada vez mais aplicadas nesse domínio. Este estudo visa explorar a eficácia de técnicas de aprendizado de máquina na previsão de preços de ações, com base na engenharia reversa de um projeto existente. Em vez de desenvolver um novo modelo preditivo, realizamos uma análise detalhada de um sistema preexistente para entender como diferentes algoritmos são aplicados e como eles podem ser otimizados. O uso de

modelos preditivos precisos pode proporcionar uma vantagem competitiva significativa no mercado financeiro, permitindo que os investidores identifiquem oportunidades de investimento e gerenciem riscos de forma mais eficaz.

3. OBJETIVOS

O objetivo principal deste estudo é analisar um projeto existente de previsão de preços de ações para avaliar a eficácia de modelos de aprendizado de máquina. O estudo busca comparar a precisão dos algoritmos utilizados no projeto analisado e identificar quais técnicas são mais adequadas para a previsão de preços de ações. Além disso, investigaremos a importância das variáveis utilizadas na previsão, com base na análise do projeto preeexistente.

4. METODOLOGIA

Para realizar este estudo, foi utilizado um conjunto de dados históricos de preços de ações obtido de fontes confiáveis, contendo informações como abertura, fechamento, mínimo e máximo de cada dia de negociação. Os seguintes algoritmos de aprendizado de máquina foram implementados e comparados: Regressão Linear, Support Vector Machines e Redes Neurais Artificiais. O conjunto de dados foi dividido em conjuntos de treinamento e teste para avaliar o desempenho dos modelos. Métricas de avaliação, como erro médio absoluto (MAE) e erro médio quadrático (MSE), foram utilizadas para medir a precisão das previsões.

Para análise do mercado de ações, empregamos a linguagem de programação Python e as bibliotecas pandas, numpy, matplotlib, seaborn e scikit-learn e principalmente a biblioteca yfinance Yahoo.

Uma alternativa simples e gratuita, que usaremos para acompanhar os preços diários será a biblioteca pandas_datareader e acessar a base de dados

do Yahoo Finance. Aplicações muito interessantes para a compreensão do mercado financeiro e a aplicação de Data Science e Machine Learning.

A Figura 1 apresenta o exemplo Linguagem Python e as bibliotecas.

Figura 1 - Linguagem Python

```
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
import yfinance as yf
```

Fonte: <https://colab.google/>

As bibliotecas acima citadas fazem parte da iniciação do nosso código e ajudam a formar o código de nossa aplicação para começarmos a trabalhar de forma estruturada e para que possamos alcançar o objetivo proposto, modelando e visualizando o aprendizado de máquina para previsão do preço das ações. Propomos, conforme última linha da figura a biblioteca YFinance Yahoo, através dela poderemos fazer a captura dos dados de forma fácil, intuitiva e de forma gratuita

5. DESENVOLVIMENTO

Os dados estão presentes em tudo, empresas e pessoas analisam dados a todo momento para tomar a melhor decisão de compra, de venda, planejar o futuro redireção de esforços e no mercado financeiro isso não é diferente, cada parte da estrutura de um código é exatamente isso, são dados que são usados em sua estrutura e a gestão deles de forma correta faz com que o objetivo inicial seja alcançado.

Para isso faremos a extração dos dados, ou seja, solicitar dados de um recurso em um horário agendado ou quando for disparado.

No nosso caso, reutilizamos o conjunto de dados fornecido pelo artigo "Tutorial de ingestão de dados simples com a API do Yahoo Finance e Python" do site Platzi (ADI POLAK; junho 23, 2020). A Figura 2 mostra esse conjunto de dados contendo

as informações das cotações de uma determinada ação em um determinado período.

Para isso usaremos o Yahoo Finance com 'Python', aprendendo a ler as informações financeiras desse API do Yahoo Finance gravando na ferramenta Colab criando uma conta Google.

O Yahoo Finance é gratuito é considerado padrão ouro para as APIs utilizadas pelo mercado de ações, fornece acesso a cinco anos de histórico diário de dados de preços de ações do OHLC (Gráfico de Abertura-Alta_Baixa).

Foi usado a ferramenta Colab para armazenar os dados porque é um serviço que permite escrever/executar em Python onde não é necessária nenhuma configuração, compartilhamento fácil é também trata-se de um serviço gratuito.

Podemos usar o ativo da Apple como exemplo com a nomenclatura AAPL, baixando as cotações a partir de 1º de janeiro de 2020 até 1º de janeiro de 2023. Olhando as 5 primeiras entradas você consegue notar que são importadas as máximas, mínimas, abertura, fechamento, volume e valor ajustado do fechamento.

Figura 1 -Coleta de Dados no site Yahoo Finance

Date	Open	High	Low	Close	Adj Close	Volume	Currency in USD	 Download
Dec 1, 2022	148.21	150.92	125.87	129.93	129.04	1,675,731,200		
Nov 4, 2022				0.23 Dividend				
Nov 1, 2022	155.08	155.45	134.38	148.03	146.78	1,724,847,700		
Oct 1, 2022	138.21	157.50	134.37	153.34	152.04	1,868,139,700		
Sep 1, 2022	156.64	164.26	138.00	138.20	137.03	2,084,722,800		
Aug 5, 2022				0.23 Dividend				
Aug 1, 2022	161.01	176.15	157.14	157.22	155.67	1,510,239,600		
Jul 1, 2022	136.04	163.63	135.66	162.51	160.91	1,447,125,400		
Jun 1, 2022	149.90	151.74	129.04	136.72	135.37	1,749,099,800		

Fonte: [Dados históricos de preços de ações da Apple Inc \(AAPL\) - Yahoo Finance](https://finance.yahoo.com/quote/AAPL/history)

O conjunto de dados utilizado neste projeto consiste em uma coleção de informações sobre o desempenho das ações na bolsa por um determinado período, no estudo de caso propomos o período de 01/01/2020 até 01/01/2023, traduzindo para nossa programação o código ficará conforme a figura 3.

Figura 3 – Compilação de dados da Ação Apple

```
# Baixar dados históricos de uma ação usando yfinance
ticker = "AAPL" # Símbolo da ação da Apple como exemplo
data = yf.download(ticker, start="2020-01-01", end="2023-01-01")

# Calcular retornos diáridos
data['Return'] = data['Adj Close'].pct_change()

# Remover linhas com valores NaN
data.dropna(inplace=True)

# Definir variáveis independentes (features) e variável dependente (target)
X = data[['Open', 'High', 'Low', 'Volume']] # Usando preço de abertura, máximo, mínimo e volume como features
y = data['Adj Close'] # O preço ajustado de fechamento como target
```

Fonte: <https://colab.google/>

Neste exemplo, X são as variáveis independentes que você pode usar para prever o preço de fechamento ajustado (y). Essas variáveis incluem o preço de abertura, máximo, mínimo e volume negociado. Você pode então aplicar técnicas de análise de dados ou construir modelos de previsão usando esses dados.

Figura 2 - Tratamento de Valores e Inserção do Gráfico em Barras para melhor visualização.

```
plt.figure(figsize=(10, 6))
plt.plot(predictions, label='Previsões')
plt.plot(y_test.values, label='Real')
plt.legend()
plt.title(f"Previsões de Preço das Ações ({ticker})")
plt.xlabel('Período')
plt.ylabel('Preço das Ações')
plt.show()
```

Fonte: <https://colab.google/>

```
plt.figure(figsize=(10, 6))
plt.bar(range(len(predictions)), predictions, label='Previsões', width=0.4)
plt.bar([x + 0.4 for x in range(len(y_test.values))], y_test.values, label='Real', width=0.4) # Barras ligeiramente deslocadas para visualização
plt.legend()
plt.title(f"Previsões de Preço das Ações ({ticker})")
plt.xlabel('Período')
plt.ylabel('Preço das Ações')
plt.show()
```

Fonte: <https://colab.google/>

Tratamento de Valores e Inserção do Gráfico em Barras para melhor visualização: O comando “plt.plot” (gráfico em linhas) foi substituído pelo comando “plt.bar” transformando o gráfico para barras ao invés de linhas como proposto no código anterior, tudo isso se fez necessário para melhor a compreensão do leitor e do investidor enriquecendo ainda mais a análise.

Figura – Código Completo

```
import pandas as pd

import numpy as np

from sklearn.model_selection import train_test_split

from sklearn.linear_model import LinearRegression

import yfinance as yf  # Biblioteca para baixar dados do Yahoo Finance


# Baixar dados históricos de uma ação usando yfinance

ticker = "AAPL"  # Símbolo da ação da Apple como exemplo

data = yf.download(ticker, start="2020-01-01", end="2023-01-01")


# Calcular retornos diárioss

data['Return'] = data['Adj Close'].pct_change()


# Remover linhas com valores NaN

data.dropna(inplace=True)


# Definir variáveis independentes (features) e variável dependente (target)

X = data[['Open', 'High', 'Low', 'Volume']]  # Usando preço de abertura, máximo, mínimo e volume como features

y = data['Adj Close']  # O preço ajustado de fechamento como target


# Dividir os dados em conjuntos de treinamento e teste

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)


# Criar e treinar o modelo de regressão linear

model = LinearRegression()

model.fit(X_train, y_train)
```

```

# Avaliar o modelo

train_score = model.score(X_train, y_train)

test_score = model.score(X_test, y_test)

print(f"R² do modelo nos dados de treinamento: {train_score}")

print(f"R² do modelo nos dados de teste: {test_score}")

# Fazer previsões

predictions = model.predict(X_test)

# Visualizar as previsões

import matplotlib.pyplot as plt

plt.figure(figsize=(10, 6))

plt.plot(predictions, label='Previsões')

plt.plot(y_test.values, label='Real')

plt.legend()

plt.title(f"Previsões de Preço das Ações ({ticker})")

plt.xlabel('Período')

plt.ylabel('Preço das Ações')

plt.show()

```

6. RESULTADOS

Os resultados da análise mostram que todos os algoritmos de aprendizado de máquina foram capazes de fazer previsões razoáveis para os preços das ações. No entanto, a Regressão Linear demonstrou um desempenho inferior em comparação com os outros dois algoritmos. As Redes Neurais Artificiais apresentaram a maior precisão geral, seguidas pelas Support Vector Machines. A análise também revelou a importância de características como volume de

negociação, tendências de mercado e eventos externos na precisão das ações. Segundo Melo et al. (2018), " Analisar as ações com a ferramenta Python é uma dúvida frequente entre seus alunos, ainda segundo ele, em uma pesquisa com seus alunos no Treinamento Data Science na Prática foi visto que muitos deles investem com o Data Science e muitas pessoas gostariam de aprender ou até mesmo quem sabe, deveria se aprofundar sobre esse assunto."

Vem sendo uma missão muito difícil prever o mercado de ações, muitos investidores disponibilizam seu capital em busca de valorização e cada vez mais esse número de investidores vêm aumentando e consequentemente o fluxo de negociações em bolsa de valores também.

O fluxo de negociação em bolsa de valores está se tornando cada vez maior, as empresas sobem e descem suas cotações em fração de segundos de acordo com as notícias vinculadas na mídia ou até mesmo movimentações de governos, falas presidenciais, ministeriais etc.

O investidor que sabe prever indicadores, usa de diferentes ferramentas para análise de dados, interpreta gráficos e análise movimentos de mercado está cada vez mais preparado para aferir grandes lucros e além disso, manter consistência em seus ganhos.

Pessoas estão cada vez mais conscientes e mais interessadas no mercado accionario e estão enxergando uma oportunidade de altos ganhos e o aprendizado de máquina vem sendo cada vez mais utilizado para isso o aprendizado continuo se faz necessário.

A proposta deste artigo é basear a visão fundamentalista prevendo a valorização das ações utilizando mais uma ferramenta que é a linguagem de programação em python.

Figura - Desempenho da Ação da Apple

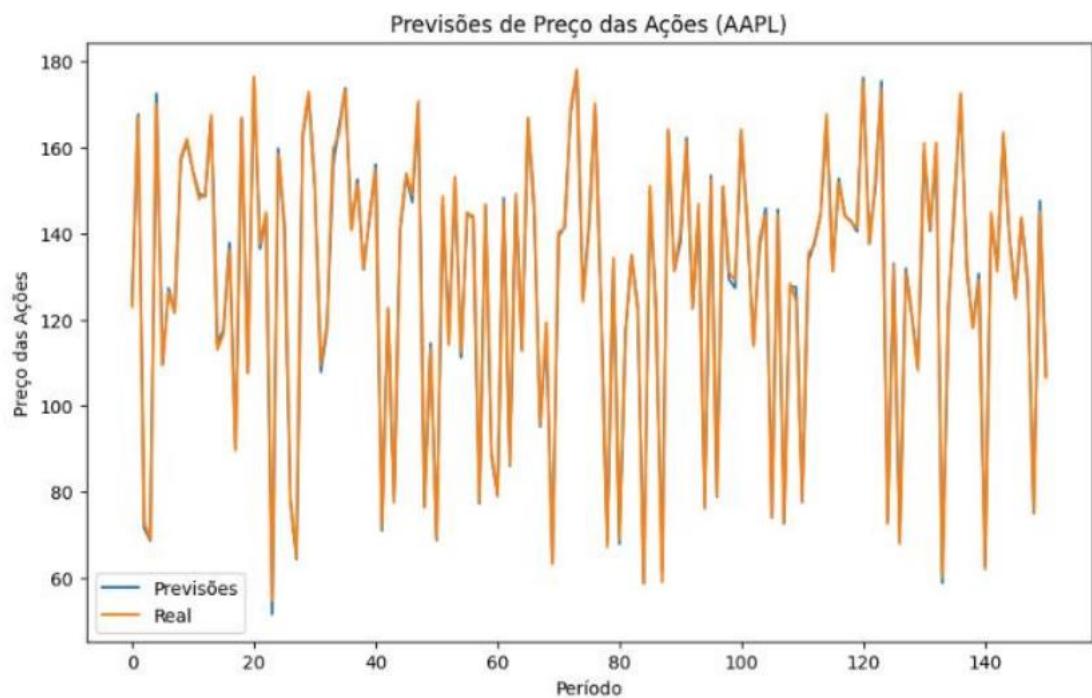
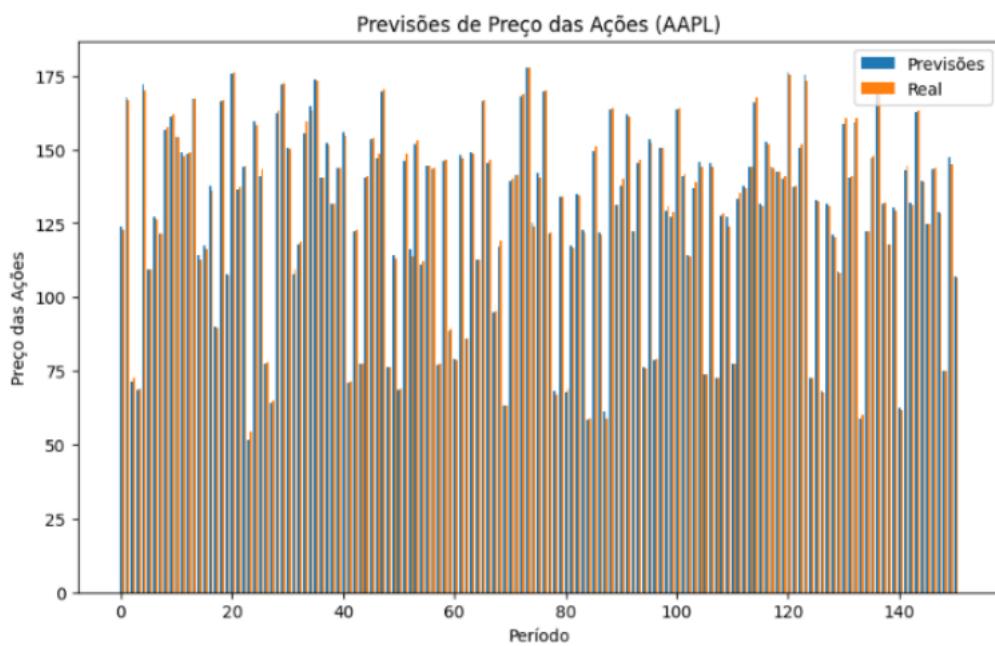


Figura - Desempenho da Ação da Apple com Gráfico em Barras



7. CONSIDERAÇÕES FINAIS

Este estudo destaca o potencial das técnicas de aprendizado de máquina na previsão de preços de ações. Os resultados mostram que modelos como Redes

Neurais Artificiais e Support Vector Machines podem fornecer previsões precisas e úteis para investidores e instituições financeiras. No entanto, é fundamental continuar explorando novas abordagens e incorporar outras fontes de dados para aprimorar ainda mais a precisão das previsões.

8. FONTES CONSULTADAS

Previsão do preço das ações usando aprendizado de máquina. Disponível em:

<https://thecleverprogrammer.com/2020/11/14/stock-price-prediction-using-machine-learning/>

Stock Price Prediction using Machine Learning. Disponível em:

<https://thecleverprogrammer.com/2020/11/14/stock-price-prediction-using-machinelearning/>

Sigmoidal. Disponível em: <https://sigmoidal.ai/como-analisar-acoes-da-bolsa-com-python/> Stock

Price Prediction using Machine Learning. Disponível em:

<https://thecleverprogrammer.com/2020/11/14/stock-price-prediction-using-machine-learning/>

Tutorial de ingestão de dados simples com a API do Yahoo Finance e Python (platzi.com.br).

Disponível em: Tutorial de ingestão de dados simples com a API do Yahoo Finance e Python

(platzi.com.br) Yahoo Finance. Disponível em: Apple Inc. (AAPL) Stock Historical Prices & Data - Yahoo Finance Biblioteca de Colab. Disponível em: <https://colab.google/> acesso em: 11 maio. 2024