

# Utilização de Redes Neurais Recorrentes do tipo LSTM na previsão de preços de ações

Gabriel Gobbo A. Valera, Engenharia de Computação – Universidade Federal de Santa  
Catarina (UFSC)

gobbogabriel@grad.ufsc.br

## **Resumo**

*Este artigo apresenta uma investigação sobre a aplicação de redes neurais recorrentes, especificamente Long Short-Term Memory (LSTM), na previsão de preços de ações no mercado financeiro. A revisão da literatura identificou estudos relevantes que abordam o uso de LSTMs nesse contexto. Os resultados englobam a aplicação das LSTMs para capturar dependências temporais complexas nos dados históricos, possibilitando uma análise do comportamento dos preços das ações. Utilizando dados do Banco do Brasil como exemplo, observou-se que as previsões da LSTM conseguiram capturar o aumento no preço das ações, levando a resultados financeiros favoráveis. No entanto, é importante considerar que a previsão de preços de ações é influenciada por fatores externos e incertezas, portanto, o uso das LSTMs deve ser complementado por outras abordagens. Em conclusão, este estudo ressalta o potencial das LSTMs como uma ferramenta valiosa para auxiliar investidores na tomada de decisões no mercado financeiro, fornecendo previsões mais precisas e embasadas.*

*Palavras chave: Inteligência Artificial, LSTM, mercado financeiro.*

## **Abstract**

*This article presents an investigation into the application of recurrent neural networks, specifically Long Short-Term Memory (LSTM), in predicting stock prices in the financial market. The literature review identified relevant studies that address the use of LSTMs in this context. The results encompass the application of LSTMs to capture complex temporal dependencies in historical data, enabling an analysis of stock price behavior. Using Banco do Brasil data as an example, it was observed that LSTM predictions were able to capture the increase in stock prices, leading to favorable financial outcomes. However, it is important to consider that stock price prediction is influenced by external factors and uncertainties; therefore, the use of LSTMs should be complemented with other approaches. In conclusion, this study highlights the potential of LSTMs as a valuable tool to assist investors in decision-making in the financial market, providing more accurate and informed predictions.*

*Keywords: Artificial Intelligence, LSTM, finance market.*

## **1 Introdução**

O mercado financeiro é um dos ambientes mais dinâmicos e fascinantes do mundo, despertando interesse global. Nele, uma ampla gama de operações envolvendo fluxo de capitais é realizada, dentre elas destaca-se a compra e venda de ações. Essa atividade atrai cada vez mais investidores em busca de oportunidades lucrativas, dispostos a assumir riscos ao adquirir ações de empresas com potencial de sucesso.

Nesse contexto, vale destacar que seu capital é colocado em jogo, movido pela convicção de que escolhas inteligentes e estratégicas podem gerar retornos significativos. Entretanto, é de extrema importância compreender as tendências atuais do mercado de ações e fundamentar as previsões temporais diretamente a essas informações, os investidores estarão mais aptos a aproveitar oportunidades lucrativas e minimizar riscos.

Com o avanço da inteligência artificial, os algoritmos têm se tornado cada vez mais sofisticados e gerado técnicas avançadas de aprendizado de máquina. Esses algoritmos de IA têm a capacidade de processar grandes volumes de dados históricos e em tempo real, permitindo a identificação de padrões complexos e a extração de informações precisas. Nesse contexto, a inteligência artificial tem sido uma ferramenta essencial para investidores, proporcionando maior eficiência na previsão de ações e contribuindo para aprimorar a tomada de decisões no mercado financeiro.

No âmbito da Inteligência artificial, possui a área chamada redes neurais recorrentes, elas são um tipo de rede neural artificial projetada para reconhecer padrões em sequências de dados, como texto, genomas, caligrafia, palavra falada ou dados de séries numéricas que emanam de sensores, bolsas de valores e agências governamentais. Esses algoritmos consideram tempo e sequência, eles têm uma dimensão temporal [1].

Dentro das redes neurais recorrentes, destaca-se o tipo específico chamado Long-Short Term Memory (LSTM). Projetadas para evitar o problema de dependência de longo prazo, as LSTMs possuem a capacidade natural de recolher informações por períodos extensos [2]. Essa característica faz dela uma excelente ferramenta para prever o custo de ações, permitindo uma análise mais precisa de dependências temporais complexas nos dados históricos.

## 2 Metodologia

A metodologia empregada neste artigo é baseada em um processo de revisão sistemático da literatura. Iniciamos definindo as melhores palavras chaves para buscar artigos semelhantes ao tema proposto. Nesse contexto, utilizamos diversas técnicas para escolhermos os artigos referências, utilizando critérios de inclusão (CI), critérios de exclusão (CE) e questões de pesquisas (QP) listadas a seguir e respondidas neste artigo:

- QP1: Qual a fundamentação teórica deste trabalho?
- QP2: Quais são os dados utilizados neste trabalho?
- QP3: Como funciona a implementação do algoritmo?
- QP4: Quais foram os resultados obtidos?
- Para a seleção dos artigos, foram escolhidos os seguintes critérios de inclusão (CI)
- CI1: Documentos publicados entre os anos de 2017 e 2023;
- CI2: Estar relacionado com uso de LSTM para o mercado financeiro para prever custos de ações;
- E os critérios de exclusão aplicados na revisão sistemática deste artigo foram:
- CE1: Documentos que estão ligados a mercado de estoque (Stock market);
- CE2: Documentos que fujam do tema LSTM aplicado ao mercado financeiro;
- CE3: Documentos com menos de 5 páginas. Neste item o objetivo é remover os resumos, deixando apenas documentos completos;
- CE4: Documentos com línguas diferentes do português ou inglês.

Com o objetivo de encontrar os documentos acadêmicos próximos, foram utilizadas as bases de dados Google Scholar e IeeeXplore. Nesse contexto, foram encontrados 373 documentos no Google Scholar, 51 no IeeeXplore, totalizando 424 documentos acadêmicos, desses 424 documentos encontrados, utilizando os métodos de inclusão e exclusão, em sequência classificação de Base de dados acadêmica (Google Scholar e IeeeXplore), Título, Autor(es), Ano da publicação, Instituição, Tipo (Artigo, TCC, Dissertação ou Tese), Tipo do artigo (Conferência ou revista), Estado e Nível Educacional. Resultou em 36 documentos no google scholar e 9 documentos no IeeeXplore, totalizando 45 documentos. Os Anais de Documentos 2017 a 2023 do Google Scholar e IeeeXplore estão disponíveis em seu site. Para o Google

Scholar foi utilizada a string de busca “Inteligência Artificial AND LSTM AND Mercado Financeiro”. Enquanto no IeeeXplore, foi utilizada a string de busca: "Artificial Intelligence"AND LSTM AND Finance Market”.

As respostas das questões de Q1 a Q4 foram consolidadas após a segunda fase e são apresentadas na próxima seção.

### **3. Desenvolvimento**

#### **3.1 Fundamentação teórica**

##### **3.1.1 Contextualização Histórica**

Com a introdução das redes neurais recorrentes, um avanço significativo foi alcançado no tratamento de dados. No entanto, esse progresso revelou uma limitação na capacidade de armazenamento de memória de longo prazo, resultando em perda de informações importantes. Diante dessa questão, evidenciou a necessidade de desenvolver uma variante que pudesse lidar com esse problema, e assim surgiu a LSTM (Long Short-Term Memory).

As redes neurais recorrentes tradicionais enfrentavam dificuldades em capturar dependências de longo prazo em sequências de dados. Embora fossem capazes de processar informações sequenciais, as RNNs sofriam de um problema conhecido como "desvanecimento do gradiente". Esse problema ocorria quando o gradiente usado para atualizar os pesos da rede diminuía exponencialmente à medida que a informação percorria as camadas da RNN, resultando em uma perda de memória de eventos distantes no tempo.

Hochreiter e Schmidhuber (1997) resolveram esse problema introduzindo a arquitetura LSTM. A ideia central por trás da LSTM era criar unidades de memória especiais que permitissem que informações relevantes fossem retidas por longos períodos de tempo. O próprio nome da LSTM já indica sua função principal: a capacidade de armazenar informações a curto e longo prazo. Ao combinar uma célula de memória de longo prazo com uma de curto prazo, a LSTM conseguiu superar a limitação anterior e resolver o problema de perda de dados.

Nesse contexto, o surgimento da arquitetura LSTM foi um marco importante, pois possibilitou a captura e retenção de dados com dependências de longo prazo, abrindo caminho para a realização de previsões precisas em diversas áreas. Com as células de memória e os mecanismos de controle de fluxo de informação, as LSTMs superaram o problema do desvanecimento do gradiente, permitindo que informações relevantes fossem mantidas ao longo do tempo. Essa inovação representou um avanço significativo no campo das redes neurais recorrentes e teve um impacto profundo no desenvolvimento de modelos de aprendizado de máquina.

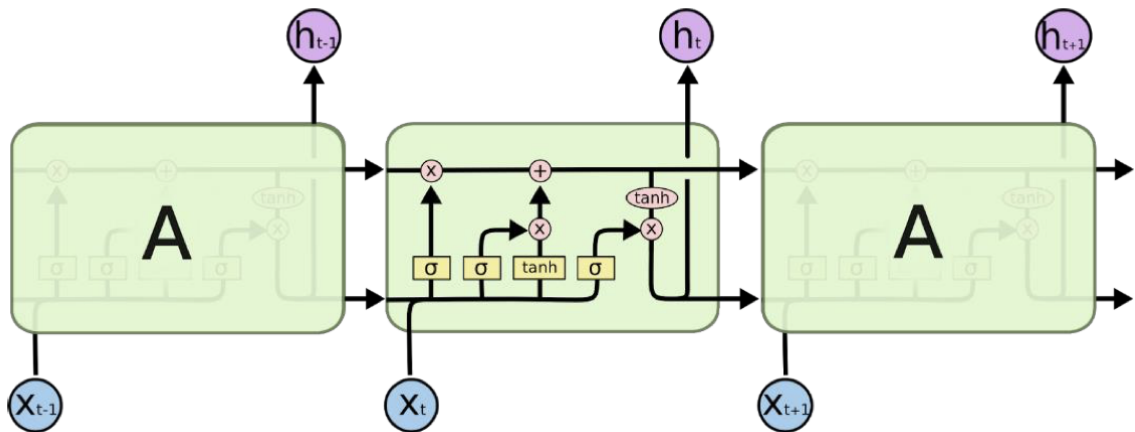
### **3.1.2 LSTM**

A arquitetura LSTM tem sido amplamente utilizada no campo de aprendizado de máquina para lidar com a modelagem de sequências, devido às suas capacidades avançadas. Ela supera as limitações das RNNs tradicionais, como o desvanecimento do gradiente, e é capaz de capturar e reter dependências de longo prazo nos dados (Géron, 2019).

A célula LSTM é composta por um estado de célula que armazena informações de longo prazo e um estado oculto que retém informações de curto prazo. Essa estrutura complexa permite um melhor desempenho, treinamento mais rápido e uma maior capacidade de detectar padrões e relações temporais complexas nos dados sequenciais.

Nesse contexto, as LSTM possuem uma estrutura em cadeia, mas com quatro camadas que interagem de forma especializada, diferentemente das redes neurais convencionais que possuem apenas uma camada [5]. Segue abaixo uma figura com sua estruturação:

**Figura 1** - Estrutura das células LSTM.



Fonte: Olah 2015

A partir desta figura, é importante ressaltar que a LSTM possui 3 portões (*gates*) para proteger e controlar o estado da célula, sendo eles:

- Forget gate: é responsável por decidir quais informações antigas serão esquecidas ou mantidas na memória de longo prazo.
- Input gate: determina quais informações novas serão adicionadas à memória de longo prazo.
- Output gate: regula a saída das informações processadas pela célula LSTM.

Esses portões garantem que a LSTM possa armazenar informações relevantes, descartar informações irrelevantes e fornecer saídas precisas com base nos dados de entrada.

Além disso, há outro componente fundamental da LSTM que é a célula de memória, responsável por armazenar informações de longo prazo. Ela é projetada para reter informações relevantes por um período prolongado, permitindo que a LSTM capture dependências de longo prazo nos dados de entrada. A célula de memória é atualizada por meio dos portões (*gates*) da LSTM, que controlam a adição, remoção e atualização das informações armazenadas.

### 3.2 Dados utilizados

Para avaliar a eficácia da LSTM na previsão de valores futuros, utilizei dados do site Yahoo Finance referentes às ações BBSA3.SA (Banco do Brasil) no período de 01/01/2018 a 30/12/2021, totalizando 3 anos de informações. Em seguida, executei a previsão para o primeiro mês de 2022, visto que não havia informações deste mês disponíveis no arquivo. Os dados obtidos das ações BBSA3.SA apresentavam os seguintes referenciais:

- "Date": a data em que a ação foi registrada;
- "Open": o preço de abertura da ação;
- "High": o preço máximo atingido pela ação durante o dia;
- "Low": o preço mínimo que a ação alcançou durante o dia;
- "Close": o preço de fechamento da ação;
- "Adj Close": o preço ajustado considerando eventos e ajustes relevantes;
- "Volume": a quantidade de ações negociadas durante o dia.

**Figura 2** - Dataset presente no código.

	Date	Open	High	Low	Close	Adj Close	Volume
0	2018-01-02	32.090000	32.930000	32.090000	32.930000	23.474606	9178100
1	2018-01-03	33.099998	33.419998	32.820000	33.349998	23.774012	8180700
2	2018-01-04	33.520000	33.959999	33.450001	33.669998	24.002119	9625200
3	2018-01-05	33.700001	33.900002	33.290001	33.669998	24.002119	5267800
4	2018-01-08	33.639999	33.750000	33.340000	33.750000	24.059156	5130600
...	...	...	...	...	...	...	...
984	2021-12-23	29.299999	29.480000	29.160000	29.219999	25.943962	7337800
985	2021-12-27	29.250000	29.490000	29.200001	29.280001	25.997232	5746200
986	2021-12-28	29.340000	29.389999	29.010000	29.110001	25.846296	8640700
987	2021-12-29	29.180000	29.309999	28.940001	28.959999	25.713108	7666400
988	2021-12-30	29.129999	29.190001	28.770000	28.850000	25.615444	12582800

989 rows x 7 columns

**Fonte:** disponível em:

<https://colab.research.google.com/drive/1eN9blhzPAmHX6aO8AX01udWx7lnrXPnD?usp=sharing>

A escolha de um conjunto abrangente de dados foi fundamental para permitir que a rede neural trabalhasse com uma quantidade significativa de informações. Além disso, optei por realizar a previsão para uma data já existente no conjunto de dados, a fim de avaliar a precisão do resultado obtido. Essa abordagem proporcionou uma maneira confiável de verificar a capacidade da LSTM em fornecer previsões satisfatórias.

Essa estratégia de utilizar dados históricos e verificar a acurácia da previsão em uma data conhecida é essencial para estabelecer a confiança na utilização da LSTM como uma ferramenta eficaz de previsão. Essa análise crítica dos resultados permite determinar se a rede neural é capaz de fornecer insights relevantes e confiáveis para previsões futuras no contexto específico das ações do Banco do Brasil.

### 3.3 Algoritmo utilizado

#### 3.3.1 Bibliotecas utilizadas

Neste artigo, a programação foi realizada na plataforma Google Colab, utilizando como principal biblioteca o TensorFlow (versão 2.12). O TensorFlow foi responsável por construir e treinar o modelo LSTM utilizado na previsão de valores de ações. Além do TensorFlow, também foram utilizadas as bibliotecas NumPy, Pandas e Matplotlib. O NumPy e o Pandas desempenharam um papel essencial no processamento e manipulação dos dados, enquanto o Matplotlib possibilitou a visualização clara e concisa dos resultados obtidos.

Dentro da biblioteca TensorFlow, foram importadas as seguintes classes de camadas específicas para construir modelos de aprendizado profundo: Input, LSTM, GRU, SimpleRNN, Dense e GlobalMaxPooling1D. Cada classe desempenha uma função específica no processamento de sequências:

- **Input:** é responsável pela camada de entrada do modelo.
- **LSTM:** é uma camada de memória de longo prazo recorrente que processa sequências.
- **GRU:** é uma camada de unidade recorrente de portão que também processa sequências.
- **SimpleRNN:** é uma camada de memória recorrente simples para processamento de sequências.
- **Dense:** é uma camada totalmente conectada que mapeia características para saídas desejadas.
- **GlobalMaxPooling1D:** é uma camada de pooling global que extrai o valor máximo ao longo de uma dimensão temporal em sequências unidimensionais.



Além disso, importei a classe Model do TensorFlow Keras, que fornece as funcionalidades necessárias para criar modelos personalizados e realizar tarefas de treinamento e previsão.

Também importei as classes SGD e Adam da biblioteca TensorFlow Keras, que são otimizadores utilizados durante o treinamento de modelos de aprendizado profundo para ajustar os pesos do modelo e reduzir a função de perda. O SGD é um otimizador clássico que atualiza os pesos com base no gradiente médio de um subconjunto aleatório dos dados de treinamento. O Adam, por sua vez, é um otimizador popular que combina ideias de outros algoritmos e se adapta dinamicamente às taxas de aprendizado.

A combinação dessas bibliotecas e classes proporcionou as funcionalidades necessárias para implementar o algoritmo de previsão de valores de ações com a LSTM, bem como para realizar o treinamento e ajuste dos modelos de aprendizado profundo no contexto deste estudo.

### **3.3.2 Funcionamento do algoritmo**

O algoritmo utiliza os dados históricos de preço de fechamento das ações do Banco do Brasil, selecionando a coluna "Close". Essa escolha foi feita com base em experimentos anteriores, que mostraram um melhor desempenho da LSTM com essa variável.

Os dados são normalizados usando o escalonamento padrão (StandardScaler), garantindo que todas as variáveis estejam na mesma escala. Isso permite que a LSTM capture padrões e tendências com mais precisão, contribuindo para a eficácia do modelo de previsão.

Em seguida, são definidos dois parâmetros importantes: T, que representa o tamanho da janela de sequência, ou seja, o número de dias anteriores considerados para prever o próximo valor, e D, que indica a dimensão dos dados de entrada.

As sequências de entrada e seus valores correspondentes são criados a partir dos dados normalizados. Cada sequência é composta por T dias anteriores e o valor correspondente é o próximo dia. Essas sequências são armazenadas em uma lista X e os valores correspondentes em uma lista Y.

A arquitetura do modelo LSTM é definida, onde a camada de entrada é seguida por uma camada LSTM com um número específico de unidades e, em seguida, uma camada densa que mapeia as características extraídas para a saída desejada.

O modelo é compilado com uma função de perda (mean squared error - MSE) e um otimizador (Adam) para ajustar os pesos do modelo durante o treinamento.

Após o treinamento do modelo utilizando os dados de treinamento e validação, os gráficos de perda e validação são plotados para avaliar o desempenho do modelo ao longo das épocas.

Em seguida, são feitas previsões para o próximo dia e para o próximo mês. A última sequência de dados disponíveis é usada para prever o próximo dia, e a sequência é atualizada para prever o próximo valor em cada iteração, gerando as previsões para cada dia do próximo mês.

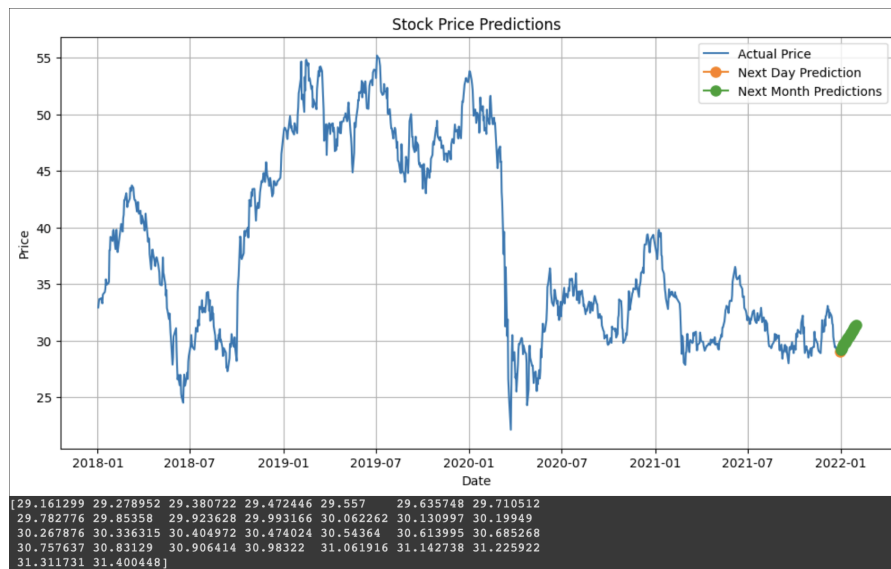
Os valores previstos são desnormalizados para obter os preços reais das ações. Essas previsões são então plotadas em um gráfico juntamente com os valores reais das ações, permitindo visualizar a performance do modelo em relação aos dados reais.

O algoritmo estará disponível no seguinte link:  
<https://colab.research.google.com/drive/1eN9blhzPAmHX6aO8AX01udWx7lnrXPnD?usp=sharing>

### **3.4. Resultados**

Após a execução do código, foram obtidos resultados promissores na previsão de valores de ações do Banco do Brasil. As previsões para o próximo dia e para o próximo mês foram realizadas com sucesso e plotadas em um gráfico. Abaixo, é apresentada a imagem do gráfico com os valores previstos das ações.

**Figura 3 - Resultados obtidos da ação BBSA3.SA**



**Fonte:** acervo pessoal do autor.

Ao observar o gráfico, pode-se notar que houve uma valorização significativa na ação, com um aumento de preço de R\$29,16 para R\$31,40, resultando em um lucro total de R\$2,24. Nesse contexto, um investidor que optasse por investir nesse momento teria obtido um lucro considerável, tornando essa decisão altamente vantajosa em termos de investimentos em ações.

É importante ressaltar que a LSTM não teve acesso aos dados reais das ações durante o processo de previsão, ou seja, ela fez a predição sem conhecimento dos preços desses dias específicos. Em seguida, será apresentada uma imagem que mostra os valores reais das ações do Banco do Brasil.

**Figura 4 - Valor real da ação BBSA3.SA**



**Fonte:** Retirado do site google finanças.

Nos primeiros 30 dias do ano de 2022, observou-se um aumento nos preços das ações do Banco do Brasil. A seguir, estão listados os preços de fechamento das ações nesses dias, considerando que a bolsa de valores não opera nos finais de semana:

- 03/01/2022: R\$ 28,82
- 04/01/2022: R\$ 28,85
- 05/01/2022: R\$ 28,37
- 06/01/2022: R\$ 28,60
- 07/01/2022: R\$ 28,63
- 10/01/2022: R\$ 28,78
- 11/01/2022: R\$ 28,90
- 12/01/2022: R\$ 29,18
- 13/01/2022: R\$ 29,65
- 14/01/2022: R\$ 30,41
- 17/01/2022: R\$ 30,47
- 18/01/2022: R\$ 30,74
- 19/01/2022: R\$ 31,01
- 20/01/2022: R\$ 31,20
- 21/01/2022: R\$ 31,20
- 24/01/2022: R\$ 31,35
- 25/01/2022: R\$ 32,16
- 26/01/2022: R\$ 31,85
- 27/01/2022: R\$ 32,29
- 28/01/2022: R\$ 32,49

Ao comparar os resultados reais com as previsões da LSTM, observa-se que a ação também apresentou um aumento considerável, conforme previsto pela LSTM. No entanto, a ação teve um aumento ainda maior do que o previsto pela LSTM, resultando em um lucro adicional de R\$1,09 em relação à previsão. Apesar dessa diferença, ambos os resultados foram positivos, o que evidencia uma previsão precisa por parte da LSTM. Se um investidor tivesse utilizado a LSTM como base para suas decisões de investimento, teria obtido um bom resultado financeiro, uma vez que a LSTM previu corretamente a valorização da ação.

## **4 Conclusão**

Conclui-se que a LSTM é uma poderosa rede neural quando se trata de dados sequenciais. No entanto, ao lidar com previsões em mercados altamente voláteis, como o mercado de ações, é arriscado confiar exclusivamente nessa abordagem. A LSTM tem sua eficácia limitada, uma vez que se concentra apenas no processamento dos dados sequenciais, enquanto o mercado de ações é influenciado por uma infinidade de fatores externos, como pronunciamentos de autoridades, lançamentos de novas tecnologias, flutuações cambiais e eventos noticiosos relevantes.

Nesse contexto, é importante destacar que confiar integralmente em uma LSTM para a previsão de valores de ações não é viável. No entanto, se utilizada com o propósito específico de analisar e prever tendências em dados sequenciais, a LSTM demonstra uma eficácia notável.

É fundamental reconhecer que a previsão de valores de ações requer uma abordagem mais abrangente, considerando-se uma variedade de fatores externos e ações humanas que influenciam o mercado. Para obter previsões mais precisas, é recomendado combinar a LSTM com outras técnicas de análise e levar em conta dados não sequenciais, como pronunciamentos políticos, mudanças econômicas e eventos significativos. Dessa forma, será possível obter uma visão mais completa e fundamentada sobre as flutuações do mercado de ações.

Em suma, embora a LSTM seja altamente eficaz no processamento de dados sequenciais, é crucial considerar sua aplicação com cautela no contexto da previsão de valores de ações. A combinação de abordagens mais abrangentes e a análise de fatores externos podem melhorar significativamente a precisão das previsões e fornecer uma base mais sólida para a tomada de decisões no mercado financeiro.

## **5 Referências**

1. Deep Learning (Ian J. Goodfellow, Yoshua Bengio and Aaron Courville), MIT Press, 2016.

2. OLAH, C. Understanding LSTM Networks. 2015. Disponível em: <<https://colah.github.io/posts/2015-08-Understanding-LSTMs/>>.
3. Hochreiter, S, and J Schmidhuber. "Long short-term memory." Neural computation vol. 9,8 (1997): 1735-80. doi:10.1162/neco.1997.9.8.1735
4. GÉRON, Aurélien. Hands-On Machine Learning with Scikit-Learn, Keras & TensorFlow. 2. ed. Sebastopol: O'reilly, 2019. 286 p.
5. <https://repositorio.pucgoias.edu.br/jspui/handle/123456789/3537>