



UNIVERSIDADE FEDERAL DE SANTA CATARINA  
CENTRO DE CIÊNCIAS, TECNOLOGIAS E SAÚDE DO CAMPUS ARARANGUÁ  
CURSO DE GRADUAÇÃO EM ENGENHARIA DE COMPUTAÇÃO

Ramon Willian Tramontin

**Análise de Sentimentos e Aprendizado Profundo na Predição de Preços de  
Ativos listados em Bolsa de Valores**

Araranguá  
2024

Ramon Willian Tramontin

**Análise de Sentimentos e Aprendizado Profundo na Predição de Preços de  
Ativos listados em Bolsa de Valores**

Trabalho de Conclusão de Curso do Curso de Graduação em Engenharia de Computação submetido ao Centro de Ciências, Tecnologias e Saúde do Campus Araranguá da Universidade Federal de Santa Catarina para a obtenção do título de Bacharel em Engenharia de Computação.

Orientador: Prof. Alexandre Leopoldo Gonçalves, Dr.

Araranguá

2024

Ficha catalográfica gerada por meio de sistema automatizado gerenciado pela BU/UFSC.  
Dados inseridos pelo próprio autor.

Tramontin, Ramon Willian  
Análise de Sentimentos e Aprendizado Profundo na  
Predição de Preços de Ativos listados em Bolsa de Valores /  
Ramon Willian Tramontin ; orientador, Alexandre Leopoldo  
Gonçalves, 2024.  
40 p.

Trabalho de Conclusão de Curso (graduação) -  
Universidade Federal de Santa Catarina, Campus Araranguá,  
Graduação em Engenharia de Computação, Araranguá, 2024.

Inclui referências.

1. Engenharia de Computação. 2. Aprendizado profundo. 3.  
Análise de sentimentos. 4. Bolsa de valores. 5. Predição de  
mercado financeiro. I. Gonçalves, Alexandre Leopoldo. II.  
Universidade Federal de Santa Catarina. Graduação em  
Engenharia de Computação. III. Título.

Ramon Willian Tramontin

**Análise de Sentimentos e Aprendizado Profundo na Predição de Preços de  
Ativos listados em Bolsa de Valores**

Este Trabalho de Conclusão de Curso foi julgado adequado para obtenção do Título de Bacharel em Engenharia de Computação e aprovado em sua forma final pelo Curso de Graduação em Engenharia de Computação.

Araranguá, 18 de Junho de 2024.

---

Prof. Jim Lau, Dr.  
Coordenador do Curso

**Banca Examinadora:**

---

Prof. Alexandre Leopoldo Gonçalves, Dr.  
Orientador

---

Prof<sup>a</sup>. Andréa Sabedra Bordin, Dra.  
Avaliadora  
Universidade Federal de Santa Catarina

---

Prof. Martín Augusto Gagliotti Vigil, Dr.  
Avaliador  
Universidade Federal de Santa Catarina

# Análise de Sentimentos e Aprendizado Profundo na Predição de Preços de Ativos listados em Bolsa de Valores

## *Sentiment Analysis and Deep Learning for Predicting Stock Market Asset Prices*

Ramon Willian Tramontin \*      Alexandre Leopoldo Gonçalves †

2024, Junho

### Resumo

A bolsa de valores é um ambiente dinâmico e complexo onde investidores compram e vendem ativos visando retornos financeiros. A análise desse mercado é crucial, pois apoia a tomada de decisões, minimizando riscos. Apesar do uso de tecnologias computacionais ser empregado nas estimativas de tendências futuras há algum tempo, avanços nas áreas de Análise de Dados e Inteligência Artificial têm revolucionado este campo. Ao encontro dessas mudanças, o presente trabalho investiga a eficácia da combinação de técnicas de Análise de Sentimentos e Aprendizado Profundo para prever os preços de ativos financeiros negociados na bolsa de valores brasileira. O método proposto envolve a coleta e o pré-processamento de dados financeiros e publicações textuais, cujos sentimentos são extraídos. Após a normalização dos dados, esses são utilizados no treinamento e avaliação das Redes Neurais Artificiais modeladas para a predição dos preços das ações. Embora os resultados indiquem que a inclusão de sentimentos apenas durante parte do treinamento possa prejudicar os modelos preditivos, o método proposto mostrou eficácia em um experimento no qual os investimentos sugeridos alcançaram um lucro de 33,45% em 20 dias. Esses achados sugerem que a abordagem combinada de análise de sentimentos e aprendizado profundo tem potencial para aprimorar a previsão de preços de ativos, com perspectiva positiva para suportar a tomada de decisão no mercado financeiro.

**Palavras-chave:** Aprendizado profundo. Análise de sentimentos. Bolsa de valores. Predição de mercado financeiro.

---

\*ramon.tramontin@grad.ufsc.br

†a.l.goncalves@ufsc.br

# Análise de Sentimentos e Aprendizado Profundo na Predição de Preços de Ativos listados em Bolsa de Valores

## *Sentiment Analysis and Deep Learning for Predicting Stock Market Asset Prices*

Ramon Willian Tramontin \*      Alexandre Leopoldo Gonçalves †

2024, Junho

### Abstract

The stock exchange is a dynamic and complex environment where investors buy and sell assets aiming at financial gain. An analysis of this market is crucial because it supports decision-making and minimizes risks. Despite the fact that computational technologies have been used in estimating future trends for some time now, advances in the areas of Data Analysis and Artificial Intelligence have revolutionized this field. In light of these changes, the present work investigates the efficacy of the combination of Sentiment Analysis and Deep Learning techniques to predict the prices of financial assets traded in the Brazilian stock exchange. The proposed method encompasses the collection and preprocessing of financial data and written publications, whose sentiments are extracted. After the normalization of this data, it is then utilized in the training and validation of the Artificial Neural Networks modelled to predict share prices. Although the results indicate that the inclusion of sentiments only during part of the training can hinder the forecasting models, the proposed method has demonstrated efficacy in an experiment in which the suggested investments turned a profit of 33.45% in 20 days. These findings suggest that the combined approach of sentiment analysis and deep learning has the potential to enhance asset price prediction, with a positive outlook to support decision-making in the financial market.

**Keywords:** Deep learning. Sentiment analysis. Stock market. Stock market prediction.

---

\*ramon.tramontin@grad.ufsc.br

†a.l.goncalves@ufsc.br

# 1 Introdução

O mercado de ações é um sistema complexo onde títulos são negociados, muitas vezes com alta volatilidade, devido a inúmeros fatores. Negociar neste mercado envolve prever movimentos de ações para maximizar retornos e minimizar riscos. Classicamente, essas previsões dependiam de dados históricos, de indicadores de mercado e do desempenho das empresas. No entanto, o advento da era digital e o desenvolvimento de novas tecnologias começaram a transformar as abordagens tradicionais, introduzindo ferramentas analíticas mais sofisticadas que aumentam a precisão das previsões e melhoram as oportunidades de investimentos. Essa evolução abre caminho para a incorporação de técnicas avançadas de análise de dados nas avaliações tradicionais do mercado de ações, proporcionando um entendimento mais detalhado das dinâmicas do mercado e do comportamento dos investidores (REN; WU; LIU, 2019).

A compra e venda de ativos, comumente chamada de *trading* de ações, é uma atividade com a finalidade de capitalizar sobre suas variações de preço. Esta é uma atividade que depende fortemente da habilidade dos operadores em avaliar uma variedade de informações para tomar decisões rápidas e eficazes. À medida que as tecnologias de informação avançaram, houve um aumento significativo no volume de dados disponível em tempo real, permitindo aos participantes do mercado acessar informações cruciais para o embasamento de suas decisões. Com essa quantidade expressiva de dados, surge a necessidade de interpretá-los de forma eficiente para que uma alta lucratividade seja alcançada, o que motiva muitos estudos nesse sentido (JOSHI; GOEL, 2023; CHONG; HAN; PARK, 2017).

Para além dos dados técnicos, cada vez mais sistemas vêm utilizando dados de opinião pública. Isso tem se mostrado promissor no contexto de predição em mercados de ações. Neste contexto, tem-se a Análise de Sentimentos, do inglês, *Sentiment Analysis* (SA) que aproveita, por exemplo, os sentimentos públicos expressos em plataformas de mídia social como o X<sup>®</sup> (antigo Twitter<sup>®</sup>) para prever movimentos do mercado (PAGOLU *et al.*, 2016). Bollen, Mao e Zeng (2011) mostraram que pode ser observada correlação entre o humor público e o índice Média Industrial *Dow Jones*, do inglês *Dow Jones Industrial Average* (DJIA).

A partir disso, a evolução da Inteligência Artificial, do inglês, *Artificial Intelligence* (AI) e do Aprendizado Profundo, do inglês, *Deep Learning* (DL) têm revolucionado as metodologias de predição no mercado de ações, tornando os processos mais rápidos e precisos. Estas tecnologias têm capacidade para analisar grandes volumes de dados, capturando nuances que analistas humanos e métodos tradicionais podem ignorar. Dentre as ferramentas mais utilizadas nesse cenário estão as Redes Neurais Artificiais, do inglês, *Artificial Neural Networks* (ANNs) e suas variantes, que simulam o funcionamento do cérebro humano para identificar padrões e, com isso, realizar previsões. Duas das principais subdivisões das ANNs são a Redes Neurais Convolucionais, do inglês, *Convolutional Neural Networks* (CNNs) e a Redes Neurais Recorrentes, do inglês, *Recurrent Neural Networks* (RNNs) (NTI; ADEKOYA; WEYORI, 2020; BUSTOS; POMARES-QUIMBAYA, 2020).

A Rede Neural Convolutiva Temporal, do inglês, *Temporal Convolutional Neural Network* (TCN) é um tipo de CNN adaptada para a análise de dados temporais devido, principalmente, a uma estrutura de convolução causal. Essa estrutura incorpora convoluções prolongadas que permitem que a rede aproveite informações de intervalos de tempo mais amplos sem um aumento proporcional na complexidade computacional. Essa configuração preserva a sequência temporal dos dados e melhora a capacidade na detecção de padrões, por consequência, melhorando a previsão (ZUO, 2023).

Adicionalmente, a arquitetura de Memória de Curto Prazo Prolongada, do inglês, *Long Short-Term Memory* (LSTM), é uma sofisticada variante da RNN, sendo mais amplamente utilizada em estudos para a previsão na bolsa de valores devido à sua capacidade de reter informações dinamicamente e assegurar uma retropropagação mais eficiente (KHAT-TAK *et al.*, 2023). Essa capacidade é adquirida por meio de um mecanismo de aprendizado que as torna aptas a lidar com longas sequências de entrada. Ao lembrar seletivamente movimentos de preços históricos importantes, enquanto esquece informações não essenciais, as LSTMs podem fornecer uma estrutura robusta para prever tendências futuras do mercado (CHEN; ZHOU, 2021).

O objetivo geral deste trabalho consiste na proposição e desenvolvimento de um método de previsão de preços que combina Análise de Sentimentos com dados de cotações de ativos do mercado acionário brasileiro, empregando arquiteturas de Redes Neurais Profundas, do inglês, *Deep Neural Networks* (DNNs) tais como TCN e LSTM. Adicionalmente, pensando em uma avaliação prática, realizou-se diariamente a execução destes modelos, previamente treinados, para prever os valores das ações no dia subsequente, avaliando assim o potencial real da proposta.

O presente documento está organizado, além desta introdução, por outras cinco seções. A segunda seção apresenta o embasamento teórico necessário para a compreensão do estudo, introduzindo conceitos essenciais. Na terceira seção, são debatidos estudos anteriores que exploram a previsão de tendências através de ANNs. A quarta seção descreve o método proposto, seguido pela análise e discussão dos resultados, na quinta seção. A conclusão e as sugestões de futuras pesquisas são apresentadas na sexta seção.

## 2 Fundamentação Teórica

### 2.1 Mercado de Ações

O Mercado de Ações é um ambiente onde investidores compram e vendem ações, que são frações de propriedade de empresas. As bolsas de valores facilitam essas negociações e são reguladas para garantir a transparência e a justiça no mercado. Essas transações proporcionam aos investidores a oportunidade de ganhar dinheiro com o sucesso financeiro das empresas, seja pelo aumento de seu valor de mercado, seja pelos lucros que venham a ser distribuídos (GOMES, 2007).

A história do mercado financeiro foi marcada por diversos eventos que moldaram sua evolução. Desde a fundação do Banco Medici<sup>®</sup> em 1397, que introduziu conceitos essenciais para a área, até a criação da Companhia das Índias Orientais em 1599, a qual começou a comercializar suas ações na bolsa de valores de Amsterdã em 1602, marcando efetivamente o início do comércio de ações de empresas, o caminho estava

pavimentado para o desenvolvimento dos mercados de ações. A mania das tulipas em 1637 e a Bolha da South Sea em 1720, são exemplos emblemáticos dos riscos e volatilidades associados a investimentos especulativos e sublinham a necessidade de regulamentação. O colapso de Wall Street® em 1929, demonstrou profundamente a interconexão entre os mercados de ações e a economia mais ampla, levando à implementação de regulamentações financeiras rigorosas. Esses desenvolvimentos, ao lado de avanços tecnológicos e políticas governamentais, têm continuamente influenciado a estrutura e a função dos mercados de ações, destacando sua complexidade crescente e papel vital para a economia global (KISHTAINY; KINDERSLEY, 2013).

No Brasil, a história das bolsas de valores iniciou-se com a inauguração da primeira instituição de negociação de mercadorias e títulos em Salvador, em 1817. Seguiu-se a abertura de uma bolsa no Rio de Janeiro em 1820. Desde então, o setor financeiro do tem se adaptado continuamente às mudanças econômicas e políticas. Em particular, a Bolsa Livre de São Paulo, estabelecida em 1890, permanecendo ativa por 14 meses e, sendo sucedida pela Bolsa de Fundos Públicos de São Paulo em 1895, que foi um marco no desenvolvimento do mercado de ações do país. Em 1967 esta última passou a ser chamada de Bolsa de Valores de São Paulo, popularmente chamada de Bovespa. Menos de um ano depois, em 2 de janeiro de 1968, o IBOVESPA, principal índice do mercado de ações brasileiro, começou a ser calculado e pontuou um avanço significativo na estruturação financeira do Brasil (CASTRO *et al.*, 2019).

Após o estabelecimento da IBOVESPA, ocorreram a inauguração da BM&F em 1986 e a unificação das bolsas em 2000. A adoção do pregão eletrônico em 2005 e os IPOs da Bovespa e BM&F em 2007 simbolizaram etapas importantes. A fusão dessas entidades em 2008 criou o BM&FBOVESPA, destacando-se globalmente (MUB3, n.d.). Finalmente, em 2017, a fusão com a Cetip formou a bolsa de valores brasileira, Brasil, Bolsa, Balcão (B3), que apresenta um ambiente de pesquisa mais limitado em comparação com mercados mais desenvolvidos, como os dos EUA, por exemplo. A escassez de estudos representa um desafio para uma análise mais completa do mercado brasileiro, visto que algumas características únicas do mesmo podem nunca ter sido estudadas (ARAÚJO; MACHADO, 2018).

### 2.1.1 Análise de Tendência em Mercados de Ações

A habilidade de prever com precisão o preço de ações ou analisar suas tendências, empregando somente informações disponíveis no momento, representa uma vantagem competitiva significativa. Uma previsão acurada, além de ser valiosa por si só, também pode ser convertida em altos lucros ao detentor da informação. Essa lucratividade vem acompanhada de riscos elevados frente à volatilidade dos mercados de ações (MAQSOOD *et al.*, 2020)

Predizer tendências do mercado de ações é uma tarefa difícil e incerta devido à quantidade de variáveis que podem estar envolvidas e à dificuldade que se pode enfrentar para acessá-las e identificar a influência de cada uma delas. Métodos que visam atingir tal objetivo podem ser divididos em 2 principais classes: i) a análise técnica, que se baseia na análise do padrão comportamental de preços anteriores e afirma que esse padrão deve se repetir, e ii) a análise fundamentalista, que explora indicadores que não estejam exclusivamente relacionados ao histórico de preços, assim como interpreta dados textuais,

como relatórios de resultados, notícias financeiras e publicações em redes sociais (KUMAR *et al.*, 2022).

No Século XX, teorias econômicas inovadoras remodelaram a compreensão e a gestão dos mercados de ações, como a Hipótese do Mercado Eficiente, do inglês, *Efficient Market Hypothesis* (EMH) de Malkiel e Fama (1970), que postula que as informações referentes aos mercados de capitais são disseminadas de forma perfeita, fazendo com que os preços das ações reflitam instantânea e precisamente todas as informações disponíveis. Como consequência, investidores tomam decisões com base na suposição de que os preços dos títulos contemplam todas as informações pertinentes. No mercado brasileiro, o estudo de Castro e Marques (2013) aplicou a técnica de Estudo de Eventos para avaliar a relação entre a divulgação de informações contábeis e os preços das ações das empresas que compõem o Índice Bovespa, entre os anos de 2007 e 2009. Este estudo revela que, na maior parte do período analisado, as informações contábeis divulgadas não exerceram um impacto significativo nos preços dos títulos. Contudo, uma exceção notável ocorreu em 2009, momento em que efeitos significativos foram identificados nos retornos acumulados dos títulos subsequentemente à publicação das demonstrações, indicando uma interação dinâmica entre a divulgação de informações financeiras e a resposta do mercado de ações brasileiro.

A EMH é adotada por investidores em suas estratégias há muito tempo. Porém, estudos mais recentes desafiam esta hipótese questionando suas premissas e defendendo que o mercado de ações é parcialmente previsível. Grossman e Stiglitz (1980) já apontavam que algumas informações têm um custo para serem obtidas e concluem que se elas não puderem ser utilizadas para realizar previsões no mercado, trazendo lucro a quem as comprou, não há sentido em elas continuarem a ter esse custo. A influência que emoções e sentimentos extraídos de fontes de informações *online* têm sobre as decisões de investidores indicam que as movimentações do mercado são potencialmente previsíveis, caso baseiem-se nos sentimentos de notícias ou em assuntos em alta em redes sociais, como afirmam Carosia, Coelho e da Silva (2021). Mintarya *et al.* (2023) corroboram concluindo que nos dias de hoje a previsão do mercado de ações pode ser mais eficaz devido à capacidade tecnológica de processar com precisão grandes volumes de dados.

## 2.2 Análise de Sentimentos

A Análise de Sentimentos (SA) é um conceito que consiste na aplicação de técnicas de Processamento de Linguagem Natural, do inglês, *Natural Language Processing* (NLP) para extrair ou minerar opiniões a partir de textos. A SA permite avaliar a polaridade (positiva, negativa ou neutra) das informações contidas em um texto. Yi *et al.* (2003), entre os primeiros autores da área, introduziram essa forma de analisar textos utilizando uma metodologia que se baseava em dois recursos principais: um léxico (dicionário de termos) de sentimentos e uma base de dados de padrões de sentimentos.

Este ramo do conhecimento experimentou um rápido crescimento, passando a contar com a aplicação de técnicas inovadoras que permitiram uma compreensão mais aprofundada das emoções humanas. Com o DL, modelos tornaram-se mais aptos a identificar nuances emocionais com precisão. Além disso, a possibilidade de realizar análises em tempo real e adaptadas ao contexto, abriu caminhos para aplicações variadas, incluindo

o monitoramento de interações *online*, aprimoramento de sistemas de recomendação e o desenvolvimento de interfaces que respondem de maneira aparentemente consciente. Entretanto, a adaptação a diferentes contextos culturais e a interpretação dos aspectos mais sutis, como a ironia, ainda são desafios. Atualmente, estratégias de aprendizado auto-supervisionado buscam transcender essas limitações e enriquecer o entendimento dos sentimentos humanos de maneira mais abrangente (GANDHI *et al.*, 2023).

A SA pode ser aplicada a diferentes áreas, como por exemplo, na saúde, automatizando a compreensão de *feedbacks* e sentimentos dos pacientes em relação aos tratamentos, facilitando a melhoria dos serviços de atendimento. No contexto empresarial, a análise da satisfação e das preferências dos consumidores permite ajustes estratégicos decisivos. No ramo do entretenimento, os filmes, séries, jogos e outras mídias podem ter sua aceitação avaliada, direcionando o *marketing* e a criação de conteúdo. No âmbito educacional, facilita o monitoramento das percepções dos estudantes sobre os cursos e métodos de ensino, podendo levantar pontos de melhoria. Para as mídias sociais, é possível realizar direcionamento de conteúdo por perfil de usuário e extrair os tópicos em alta, por exemplo. Nos mercados financeiros, as técnicas de SA são particularmente interessantes, pois a compreensão das opiniões públicas sobre ativos e suas condições econômicas pode direcionar as estratégias de investimento (JIM *et al.*, 2024).

Além das opiniões básicas como positiva, neutra e negativa, é possível ainda que emoções sejam extraídas dos textos, possibilitando uma compreensão mais profunda dos sentimentos analisados. As informações que estão em redes sociais permitem desvendar nuances do comportamento humano *online*, identificando tendências e padrões de opinião que influenciam decisões de compra, políticas públicas e dinâmicas sociais. Através dessas análises, organizações podem monitorar a reputação de uma marca e responder adequadamente às expectativas e necessidades do público. Com as mídias sociais sendo responsáveis pela maior produção de texto escrito por humanos, a capacidade de captar e analisar esses sentimentos em grande escala oferece uma ferramenta poderosa para prever mudanças de mercado e antecipar crises (WANG *et al.*, 2016).

A rapidez com que plataformas de *microblogs* como o X<sup>®</sup> recebem dados na era da Web 2.0, estabelece a SA em tempo real como uma ferramenta preponderante para organizar e maximizar conteúdos em diversos domínios. Com o processamento adequado, é possível estruturar as informações de forma que sejam acessíveis poucos instantes após a ocorrência de um evento (NOLASCO; OLIVEIRA, 2019). Essa capacidade de processamento e análise em tempo real é especialmente relevante no contexto financeiro, onde publicações de personalidades podem desencadear flutuações no mercado. De modo geral, a reação do mercado a determinadas publicações pode ser atribuída não apenas ao conteúdo literal da mensagem, mas também à sua interpretação pelos investidores (ORTIZ, 2023).

## 2.3 Aprendizado Profundo

O surgimento do Aprendizado de Máquina, do inglês, *Machine Learning* (ML), fundamental para a evolução da AI contemporânea, começou na década de 1950 com a criação do *Perceptron*. Esse marco inicial proporcionou a base para a construção das redes neurais. Os modelos iniciais, no entanto, só tiveram suas limitações superadas após a

introdução do algoritmo de retropropagação, nos anos 80, o que permitiu o treinamento e aprendizado por meio de *Perceptron* Multicamada, do inglês, *Multilayer Perceptron* (MLP). Mesmo com esse avanço, as redes neurais ainda enfrentavam desafios de treinamento até 2006, quando foram introduzidos métodos para treinamento eficaz de arquiteturas profundas, incluindo Rede de Crença Profunda, do inglês, *Deep Belief Network* (DBN) e o uso de Máquinas de Boltzmann Restritas, do inglês, *Restricted Boltzmann Machines* (RBM) para pré-treinamento não supervisionado. Esses acontecimentos representaram um ponto de virada, elevando a popularidade das redes neurais e desencadeando o desenvolvimento de modelos sofisticados, como a CNN e a RNN (BENGIO, 2009; HASSAIRI; EJBALI; ZAIED, 2015).

A partir disso, surge o DL, entendido como um subconjunto do ML, se destacando como estrutura central na fundação teórica da inteligência computacional contemporânea e caracterizando-se pela existência de camadas ocultas em sua arquitetura MLP. Isto amplia o desempenho no entendimento e aprendizado de padrões complexos dos dados quando aplicado em tarefas com dados não estruturados, possibilitando aplicações em diversos campos. A adaptabilidade e eficiência do DL na aprendizagem a partir de vastas quantidades de dados enfatizam seu importante papel no avanço da AI, oferecendo assim uma ferramenta robusta para lidar com problemas computacionais complexos (SEZER; GUDELEK; OZBAYOGLU, 2020).

No núcleo da fundamentação teórica do DL está o conceito de ANN que foi projetada para imitar de forma bastante simplificada as células neurais interconectadas do cérebro humano. Essas redes consistem em camadas de nós, chamados de “neurônios”, cada um capaz de realizar cálculos simples. À medida que os dados atravessam essas camadas, a rede identifica padrões através da atribuição de pesos para cada neurônio e suas conexões, um processo que lembra a aprendizagem humana (KARIRI *et al.*, 2023).

As camadas de neurônios são categorizadas em camada de entrada, camadas ocultas (ou camadas intermediárias) e camada de saída. As camadas de entrada e de saída configuram o formato dos dados nas respectivas fases, enquanto as camadas ocultas manipulam os dados internamente através de operações matemáticas que possibilitam o discernimento de características importantes para a função que a rede visa executar. O processamento da rede acontece por meio da propagação direta, do inglês *forward propagation*, que envolve passar os dados de entrada pela rede neural para inferir um resultado. Neste processo, cada neurônio obtém a soma ponderada das suas entradas e a processa em uma função de ativação, em que o valor obtido deve ser enviado para os neurônios da próxima camada da rede seguindo a mesma lógica até a camada de saída (YANG *et al.*, 2024).

As funções de ativação desempenham um papel importante nas ANNs definindo a saída dos neurônios. Recebem este nome pois podem zerar o valor da saída de um neurônio, o equivalente a desativá-lo. Existem vários tipos de funções de ativação com características bastante distintas. As *sigmoid* e Tangente Hiperbólica (*tanh*) são funções tradicionalmente usadas devido à sua natureza diferenciável. Por outro lado, a função Unidade Linear Retificada, do inglês, *Rectified Linear Unit* (ReLU) e suas variantes oferecem soluções mais interessantes para a etapa de treinamento em redes profundas. Além disso, funções adaptativas como a *maxout* ou a Unidade Linear Exponencial, do inglês, *Exponential Linear Unit* (ELU) ajustam dinamicamente suas formas durante o treinamento, o que pode

levar a uma aprendizagem mais eficaz e adaptada às características específicas dos dados. Essas funções têm sido exploradas para melhorar a capacidade de generalização das redes (KALOEV; KRASSTEV, 2021; LAU; LIM, 2018).

Uma vez definida a arquitetura de uma ANN, assim como seus parâmetros de configuração, ocorre o processo de treinamento. Este envolve a utilização de dados, permitindo que uma rede neural ajuste e melhore seus pesos e vieses, do inglês, *biases*. Esta aprende características e padrões dos dados de treinamento à medida que itera sobre os dados (BATOOL; KHAN, 2022). Durante esse processo, uma função de perda é utilizada para avaliar o desempenho do modelo a cada iteração (época), avaliando a discrepância entre as saídas obtidas através da propagação direta e os valores esperados. O objetivo dessa etapa é minimizar esse valor de perda. Para tal, é realizada a retropropagação, do inglês, *backpropagation*, um método que calcula da camada de saída para a de entrada o gradiente da função de perda em relação a cada peso e viés, direcionando o ajuste dos mesmos de forma que a perda seja reduzida (SIREGAR; WANTO, 2017).

A disponibilidade de grandes conjuntos de dados sendo armazenados todos os dias e o aumento do poder computacional impulsionam a gama de atuação do DL. No reconhecimento de imagem, o DL pode identificar e classificar objetos e realizar reconhecimento facial com alta precisão. No campo do NLP, já revolucionou a maneira com que as máquinas interpretam e geram texto. Em neurociência, tem acelerado a compreensão das estruturas neurais. Na biologia molecular, está ajudando na análise de dados genéticos complexos, melhorando a capacidade de diagnosticar e tratar distúrbios. Na saúde, promoveu transformações significativas, com avanços na precisão dos diagnósticos por imagens. Nas ciências ambientais, contribui para a melhoria das previsões meteorológicas e das respostas a desastres naturais. Em robótica, permite a adoção de comportamentos adaptativos para ambientes complexos e não estruturados. Estende também seus benefícios à melhoria da acessibilidade para pessoas com deficiências, oferecendo ferramentas educacionais personalizadas. Estes e outros usos, ilustram a vasta capacidade desta tecnologia em resolver problemas nos mais diversos domínios (DEAN, 2022).

Apesar de um vasto volume de dados disponível, é um desafio encontrar dados informativos, com anotações e em volume suficiente para treinar modelos robustos. A variabilidade dos dados utilizados no treinamento é importante para que o modelo aprenda de forma mais geral e dinâmica. Durante o treinamento, tanto os dados utilizados, quanto a arquitetura e os parâmetros da rede podem levar ao *overfitting*, com o modelo se ajustando em demasia aos dados de treinamento e falhando em generalizar para novos dados ou ao *underfitting*, quando a rede é muito simples para aprender a complexidade dos dados, apresentando um mau desempenho tanto nos dados de treinamento quanto nos testes. (SAMANT *et al.*, 2022; MAHMUD *et al.*, 2021).

### 2.3.1 Rede Neural Convolutiva Temporal

Uma série temporal é um conjunto de observações feitas sequencialmente ao longo do tempo, geralmente em intervalos regulares. A habilidade de prever com precisão a tendência de uma série temporal possibilita uma projeção confiável do comportamento do que está sendo monitorado. Nesse contexto, o avanço das técnicas de DL tem desempenhado um papel crucial na melhoria da precisão dessas previsões. A CNN, conhecida por suas aplicações em processamento de imagens, possui variantes adaptadas para outras tarefas,

como a TCN, projetada para a predição de séries temporais. Essa adaptação permite que a rede trate a dependência temporal dos dados de forma mais eficiente (SAMAL; BABU; DAS, 2021).

A CNN emprega uma estrutura hierárquica com três tipos distintos de camadas, convolucionais, de agrupamento, do inglês *pooling*, e totalmente conectadas ou densas. Camadas convolucionais desempenham o papel de extratoras de características em uma rede através de filtros que convolvem sobre os dados a fim de detectar padrões específicos, como texturas ou bordas em imagens de duas dimensões. Isso preserva as relações espaciais entre as informações, diferente do que aconteceria em uma rede totalmente conectada. As camadas de agrupamento visam reduzir a dimensionalidade dos mapas de características, reduzindo o custo computacional e conferindo robustez contra variações de posição nos dados de entrada (KIM *et al.*, 2022).

A principal diferença entre a TCN e a CNN é a finalidade de cada uma delas, enquanto a CNN processa muito bem dados espaciais, como a análise de imagens e vídeos, a TCN é voltada para o processamentos de dados sequenciais, como séries temporais, sinais de áudios e outros dados onde a dinâmica temporal tem maior importância. Para tal, uma das diferenças na implementação dessas redes é a utilização de convoluções dilatadas, que permitem um aumento no tamanho das entradas dos filtros sem causar um aumento na complexidade computacional (DOGGART *et al.*, 2023).

Com a aplicação de operações convolucionais na dimensão temporal dos dados as TCN têm um aumento na probabilidade de sucesso na aprendizagem dos padrões temporais (KIM; REITER, 2017). Outro recurso chave é o uso de conexões residuais, onde as saídas de camadas anteriores são adicionadas às posteriores, combatendo o problema do desaparecimento do gradiente e apoiando a construção de redes mais profundas. A maior liberdade para processamento em paralelo faz esse tipo de ANN superar em velocidade de treinamento e em economia de memória as RNN, cujos cálculos dependem de resultados anteriores, como será melhor detalhado na [subseção 2.3.2](#). Essas características tornam a TCN uma escolha atraente para detectar padrões temporais complexos (LEA *et al.*, 2017).

Essas redes demonstram sua eficácia em uma ampla gama de aplicações aproveitando a habilidade de capturar dependências de longo prazo em dados sequenciais. São muito aplicadas para a previsão de demandas de produtos ou inventário, analisando históricos de vendas para identificar padrões sazonais e ciclos de demanda, evitando problemas de excessos ou faltas. No planejamento de produção, entregam previsões precisas das necessidades, melhorando a programação e a alocação de recursos materiais e humanos alinhados com a demanda esperada. Além disso, otimizam processos ao prever atrasos e gargalos, permitindo a gestão proativa de projetos. O suporte a análises em tempo real possibilita decisões rápidas e informadas, aumentando a agilidade e a eficiência operacional em diversos setores (JAHIN *et al.*, 2024).

O estudo de [Angaitkar, Janghel e Sahu \(2023\)](#) aponta um modelo que integra Rede Neural Convolucional Unidimensional, do inglês, *One-Dimensional Convolutional Neural Network* (1DCNN) com convoluções causais dilatadas, melhorando sua capacidade de processar dados sequenciais em campos receptivos extensos e destaca o desempenho superior do modelo em comparação com métodos existentes, alcançando uma precisão de 94,44% no conjunto de dados de treinamento. Esse avanço demonstra o potencial

significativo da utilização de TCNs em uma variedade de aplicações práticas que exijam uma análise eficiente e precisa de dados sequenciais.

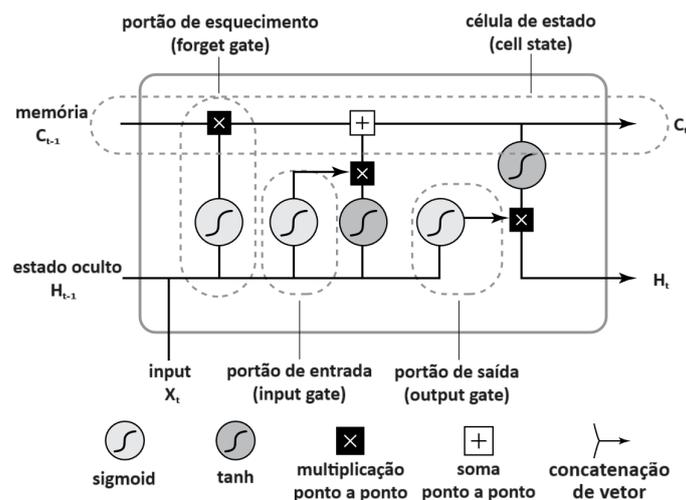
Embora existam avanços significativos, em determinados domínios, a aplicabilidade das TCN é comprometida por alguns desafios. Um desses desafios, é a complexidade no treinamento, pois para alcançar o desempenho ótimo, os parâmetros da rede devem ser devidamente ajustados. Além disso, a aplicação em tempo real implica em demandas de processamento contínuo que podem ser extremamente intensas em termos de recursos, limitando seu uso em ambientes com restrições de capacidade de processamento. Tais limitações sublinham a necessidade de pesquisa e desenvolvimento contínuos para aprimorar a eficiência e a aplicabilidade em cenários com recursos escassos (VANAMBATHINA *et al.*, 2023).

### 2.3.2 Memória de Curto Prazo Prolongada

Uma RNN se caracteriza por conexões recorrentes entre os neurônios, o que retém informações de uma entrada para as próximas. Cada neurônio das camadas ocultas recebe como entrada, além do valor que um neurônio de uma ANN comum receberia, os valores da saída da iteração anterior dos neurônios desta mesma camada (MUELLER; MASSARON, 2019).

Hochreiter e Schmidhuber (1997) introduzem a ideia de LSTM como uma solução para o problema do desaparecimento e da explosão de gradientes que as RNN enfrentam durante a retropropagação em redes muito profundas e com longos períodos temporais. A LSTM adiciona aos neurônios o conceito de blocos de memória, que a torna capaz de considerar dependências tanto de curto quanto de longo prazo nos valores de forma seletiva. Os blocos de memória são responsáveis por resolver o problema de desaparecimento de gradientes (ARAYA; VALLE; ALLENDE, 2020).

Figura 1 – Estrutura de um neurônio LSTM



Fonte: Adaptado de Mueller e Massaron (2019).

A Figura 1 apresenta a estrutura de um neurônio LSTM. Nessa estrutura, a entrada atual  $X_t$  e o estado oculto anterior  $H_{t-1}$  são concatenados e então processados pelo portão

de esquecimento, pelo portão de entrada e pelo portão de saída, que através de funções de ativação *sigmoid* e *tanh* regulam a retenção e o esquecimento de informações. O portão de esquecimento decide quanto da memória anterior é descartada, enquanto o portão de entrada determina quais novos dados serão adicionados ao estado da célula. Finalmente, o portão de saída influencia a informação que será passada ao estado oculto. As operações de multiplicação e a soma ponto a ponto entre os vetores são essenciais para atualizar o estado da célula  $C_t$  e produzir o estado oculto  $H_t$ , que juntos garantem que o neurônio LSTM mantenha uma memória eficaz de informações relevantes ao longo do tempo. Este mecanismo permite que as redes que o implementam superem os problemas relacionados à retropropagação, tornando-as excelentes para lidar com dados com dependências sequenciais (ZAINI *et al.*, 2022).

Perrot *et al.* (2022) afirmam que as RNNs do tipo LSTM podem ser consideradas as de maior sucesso, devido ao seu desempenho notável em diversas aplicações práticas. Zaytar e Amrani (2016) destacaram sua alta eficácia na previsão meteorológica, enquanto Harrou *et al.* (2021) evidenciaram seu uso em múltiplas aplicações, como reconhecimento de escrita à mão, modelos de linguagem, tradução e sintetização de voz, destacando a versatilidade e a potencial capacidade dessa tecnologia em superar desafios complexos em áreas variadas.

Apesar de apresentar ótimos resultados, as redes LSTM ainda enfrentam alguns desafios, que motivam os estudos em melhorias. Complexidade não linear e atrasos temporais nos dados de entrada dificultam a capacidade das redes de gerar previsões precisas. A exigência por grandes volumes de dados para um treinamento adequado, a complexidade inerente do modelo e sua natureza recorrente contribuem para o alto custo computacional e o tempo extenso necessário para o treinamento, limitando sua aplicabilidade em sequências de dados muito longas. Por fim, a capacidade de generalização das LSTMs podem ser comprometida em cenários não familiares ou quando os dados de treinamento não refletem a diversidade esperada na aplicação real, destacando a necessidade de uma cuidadosa seleção e preparação de dados (HUANG; MENG; HUANG, 2023).

### 3 Trabalhos Correlatos

Por meio de pesquisas em publicações científicas, selecionaram-se estudos para promover um levantamento de métodos e técnicas de predição de preço de negociação de ativos em bolsa de valores. O levantamento foi realizado nas plataformas ACM Digital Library<sup>®</sup>, ScienceDirect<sup>®</sup>, Scopus<sup>®</sup>, Web of Science<sup>®</sup>, IEEEExplore<sup>®</sup>, Periódicos Capes<sup>®</sup> e Springer Link<sup>®</sup>, buscando artigos científicos em língua inglesa publicados entre 2018 e 2023 cujos seguintes termos estivessem presentes: “*Stock Market*” e “*Sentiment Analysis*” e (“*Deep Learning*” ou “*Machine Learning*”) e (*tweet* ou *news* ou *microblog* ou “*social media*”). A escolha do período de cinco anos se justifica pelo grande volume de artigos encontrados, permitindo assim uma análise mais atualizada das técnicas emergentes e das tendências recentes nas áreas de pesquisa que suportam este trabalho.

A pesquisa em cada uma das bases foi realizada de acordo com os seguintes critérios. Nas bases ACM Digital Library<sup>®</sup>, Web of Science<sup>®</sup>, IEEEExplore<sup>®</sup>, ScienceDirect<sup>®</sup>, Periódicos Capes<sup>®</sup> e Scopus<sup>®</sup>, foram considerados o título, resumo e palavras-chave. Na Scopus<sup>®</sup> e Web of Science<sup>®</sup>, a seleção se restringiu a artigos de periódicos, desconsiderando-

se contribuições de congressos. Na Springer Link<sup>®</sup>, devido a limitações da ferramenta, a busca focou exclusivamente no título, abrangendo apenas artigos com acesso livre através do Portal de Periódicos da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

Tabela 1 – Resultados da revisão de literatura de 2018 à 2023

Base de Artigos Acadêmicos	Quantidade de artigos resultantes
ACM Digital Library <sup>®</sup>	4
ScienceDirect <sup>®</sup>	19
Scopus <sup>®</sup>	84
Web of Science <sup>®</sup>	76
IEEEExplore <sup>®</sup>	85
Springer Link <sup>®</sup>	27
Periódicos Capes <sup>®</sup>	29
Total de resultados	324

Fonte: Elaborado pelo autor (2023).

A [Tabela 1](#) apresenta o número de publicações recuperadas em cada base de dados. Inicialmente foram identificados 324 trabalhos. Após a remoção de duplicatas e análise dos títulos, 102 artigos apresentaram títulos aderentes ao trabalho. A análise dos resumos resultou em 58 artigos selecionados. Ao avaliar as introduções, destacaram-se 20 artigos de maior relevância. Destes, o texto completo foi lido e os 10 artigos mais correlatos com esta pesquisa foram selecionados e serão descritos a seguir.

No trabalho de [Carosia, Coelho e da Silva \(2021\)](#) uma metodologia foi proposta para avaliar a relação entre os sentimentos presentes em notícias publicadas em língua portuguesa pelos portais G1, Estadão e Folha de São Paulo e as variações no mercado financeiro brasileiro. O estudo foi estruturado em três etapas: (i) seleção da arquitetura de ANN e suas configurações mais adequadas à análise de sentimentos das notícias em português; (ii) aplicação do teste de causalidade de Granger para identificar a correlação entre os sentimentos analisados e as oscilações no mercado; e (iii) proposição de estratégias de investimento baseadas na SA aplicada para o mercado financeiro brasileiro. No decorrer do estudo, as arquiteturas de ANNs testadas foram MLP, LSTM e CNN. Os resultados mostraram que os sentimentos predominantes diários das notícias financeiras influenciam as compras e vendas das ações na bolsa e que considerar os sentimentos analisados como parte de uma estratégia de investimento pode resultar em lucros em curto e longo prazo.

O estudo de [Yi et al. \(2023\)](#) introduz uma abordagem inovadora para analisar a opinião pública do mercado de ações utilizando *web crawling* e técnicas de DL, especificamente 1DCNN e LSTM. Considerando fóruns plataformas essenciais para a comunicação entre investidores, os autores desenvolveram um *framework* que capta características dos preços das ações e também dos dados textuais. Vastas quantidades de dados textuais foram coletadas e categorizadas de acordo com o sentimento expresso. O 1DCNN foi empregado para classificação de sentimento do texto, enquanto o modelo LSTM integrou as características emocionais e técnicas visando prever as variações do mercado de ações. O modelo demonstrou uma acurácia de classificação de 74,38%, sugerindo que a integração da opinião pública da internet pode melhorar significativamente a avaliação do mercado de ações.

Em sua pesquisa, [Khalil e Pipa \(2022\)](#) exploram o potencial de usar o NLP e uma arquitetura LSTM para prever movimentos do mercado de ações. Eles introduzem um mecanismo de índice de sentimento único baseado em soma e relevância. O estudo emprega dados de preços por hora, a fim de capturar os impactos imediatos dos sentimentos nos preços das ações. Esta pesquisa se destaca por aproveitar informações textuais de 10 fontes, incluindo mídia convencional, mídia social, *blogs* e notícias internas de empresas. Os resultados revelam que as direções das ações são notavelmente influenciadas pelos sentimentos, tipicamente em uma janela de 3 a 4 horas. O estudo compara seis arquiteturas LSTM, e mostra que as arquiteturas que incorporam SA superam significativamente aquelas que não o fazem, especialmente considerando-se uma abordagem de retrocesso de 6 horas.

[Shah et al. \(2022\)](#) propõem uma metodologia de previsão do mercado de ações que combina o Modelo de Média Móvel Integrado Autorregressivo, do inglês, *Autoregressive Integrate Moving Average Model* (ARIMA), LSTM e SA para prover *insights* sobre os preços futuros das ações. O modelo híbrido de LSTM e ARIMA é integrado usando um conceito de média ponderada e o sentimento sobre as ações a partir de notícias, visando tornar as estimativas mais alinhadas com cenários do mundo real. O modelo foi testado no índice *Dow Jones Industrial Average* e os resultados mostraram que este é capaz de prever os preços com precisão notável considerando tanto o sentimento do mercado, quanto o histórico de preços. No entanto, o modelo enfrentou desafios ao prever preços de ações em dias com grandes variações devido a fatores externos. A abordagem geral oferece um método abrangente para a antecipação de preços de ações, levando em conta múltiplos aspectos determinantes, e não apenas os valores de abertura ou fechamento das ações.

A pesquisa conduzida por [Lin, Lai e Pai \(2022\)](#) destaca a influência significativa dos dados presentes nas redes sociais no comércio de ações e na confiança do mercado. Reconhecendo a importância da SA na antecipação dos mercados de ações, este estudo enfatiza a necessidade de uma abordagem multilíngue. Para isso, os pesquisadores adotaram um método que traduz textos para o inglês antes de analisá-los. Ao integrar dados não estruturados das redes sociais com dados estruturados, como dados de negociação e indicadores técnicos, o estudo visa aprimorar a capacidade de previsão do mercado de ações. Utilizando arquiteturas LSTM combinadas com o Algoritmo Genético, do inglês, *Genetic Algorithm* (GA) para seleção de parâmetros, a pesquisa foca na previsão de índices e preços de ações de empresas em regiões onde o inglês não é a língua nativa. Os resultados indicaram que ao combinar a LSTMGA com dados híbridos de SA multilíngues, a precisão da projeção superou outros modelos de ML.

Em sua contribuição para o campo, [Anese et al. \(2023\)](#) exploram a influência do sentimento das notícias financeiras públicas no índice do mercado de ações S&P 500. Utilizando técnicas avançadas de NLP, os pesquisadores empregaram métodos de classificação automática baseados em retornos de índices de ações e em dicionários. Também aplicaram redes neurais LSTM para analisar o sentimento de notícias financeiras. A pesquisa revela um impacto notável desses sentimentos no mercado dentro de uma janela de 20 minutos após o lançamento da notícia. Notavelmente, a SA baseada em dicionários produziu resultados mais significativos em comparação com aqueles baseados em retornos de índices de ações. Esta última abordagem encontrou desafios ao mapear com precisão o sentimento das notícias para retornos financeiros. O estudo destaca o potencial da SA

baseada em dicionários na previsão de movimentos de mercado e ressalta as limitações de depender exclusivamente de retornos de índices de ações.

O trabalho de [Owen e Oktariani \(2020\)](#) introduz a Rede Neural Baseada em Técnicas Combinadas para Ações, do inglês, *Stock Ensemble-based Neural Network* (SENN), uma abordagem inovadora para a previsão de tendências do mercado de ações. Esta metodologia utiliza dados históricos de ações, complementados pela análise de sentimentos extraídos de textos de *microblogs*. A SENN integra LSTM, MLP e CNN para uma extração eficaz de pontuações de sentimento. Com base nos dados da Boeing® e nas pontuações do StockTwits, a SENN demonstrou um aumento significativo na precisão preditiva, alcançando até 25% de aumento na precisão e uma redução de 48% na margem de erro em comparação com modelos tradicionais. O estudo também apresenta uma métrica de Erro Percentual Médio Absoluto Ajustado, do inglês, *Adjusted Mean Absolute Percentage Error* (AMAPE) para uma avaliação de desempenho mais aprimorada.

[Sarkar et al. \(2020\)](#) apresentam avanços no uso do modelo LSTM e Análise de Sentimentos, do inglês, *LSTM and Sentiment Analysis* (LSTMSA) para melhorar a precisão em previsões do mercado de ações. Esta abordagem combina análise técnica e de sentimentos para prever o comportamento das ações do Google® no mercado norte-americano, empregando redes neurais LSTM para analisar séries temporais e SA para interpretar dados de notícias. Essa integração resulta em previsões mais acuradas e intuitivas. O estudo demonstra a eficácia do modelo, capaz de fazer previsões generalizadas sobre tendências do mercado. A incorporação da SA de manchetes de notícias no modelo LSTM proporciona previsões mais alinhadas com os preços reais das ações, refletindo mais efetivamente as tendências de preços reais.

[Kamal et al. \(2022\)](#) exploram a combinação do Dicionário Consciente de Valência e Raciocínio de Sentimento, do inglês, *Valence Aware Dictionary and Sentiment Reasoning* (VADER) com modelos de DL para análise de dados financeiros e notícias para previsões do mercado de ações. Utiliza dados extensivos do Yahoo Finance® e da Business-Standard, empregando métodos como Lista de Palavras, do inglês, *Bag-of-Words* (BoW), Frequência do Termo-Inverso da Frequência nos Documentos, do inglês, *Term Frequency-Inverse Document Frequency* (TF-IDF), CNN e LSTM. A pesquisa demonstra que modelos como BERT e RoBERTa superam em precisão as técnicas de ML tradicionais, ressaltando a eficácia desse método em análise financeira e predição de mercado.

[Koukaras, Nousi e Tjortjis \(2022\)](#) investigam a previsão do mercado de ações utilizando SA a partir de dados obtidos em plataformas de *microblogs*, como Twitter® e StockTwits®, combinada com dados financeiros do Yahoo Finance®. Utilizando ferramentas de SA como TextBlob e VADER, juntamente com sete algoritmos de ML, incluindo Máquina de Vetores de Suporte, do inglês, *Support Vector Machine* (SVM), K-Vizinhos mais Próximos, do inglês, *K-Nearest Neighbors* (KNN), a pesquisa incorpora efetivamente o sentimento das mídias sociais na previsão do mercado de ações. Os resultados revelam que o modelo SVM, em conjunto com VADER e dados do Twitter®, supera os demais, registrando um *F-score* de 76.3% e um valor de Área Sob a Curva, do inglês, *Area Under Curve* (AUC) de 67%, indicando o potencial e a eficácia de usar esta informação na previsão do mercado de ações.

Os artigos discutidos nesta seção exploram o uso de Modelos de Linguagem de

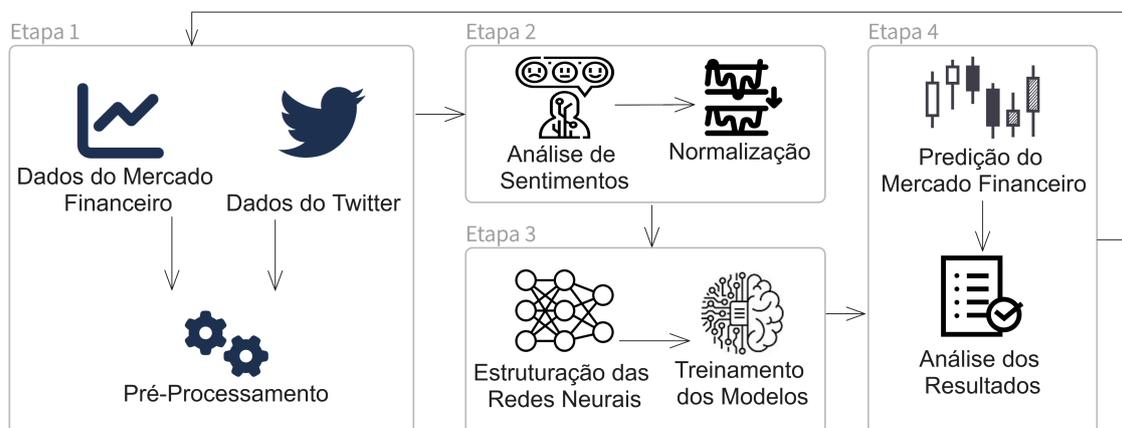
Larga Escala, do inglês, *Large Language Models* (LLMs), NLP, SA e técnicas de ML aplicados ao mercado de ações. Há um consenso geral entre os autores de que essas abordagens são promissoras para antecipar as tendências de preços das ações, embora exista a necessidade de pesquisas adicionais para aprimorar a precisão das previsões.

Nota-se que, em alguns estudos, como os de [Yi et al. \(2023\)](#), [Khalil e Pipa \(2022\)](#), [Shah et al. \(2022\)](#), [Lin, Lai e Pai \(2022\)](#), [Owen e Oktariani \(2020\)](#), [Sarkar et al. \(2020\)](#) e [Koukaras, Nousi e Tjortjis \(2022\)](#), houve a combinação de dados de sentimentos com informações referentes a preços de ações, mas essa integração, na maior parte das vezes, não é abordada de forma clara. Por outro lado, os trabalhos de [Carosia, Coelho e da Silva \(2021\)](#), [Anese et al. \(2023\)](#) e [Kamal et al. \(2022\)](#) propõem o uso exclusivo dos sentimentos na predição das tendências do mercado de ações. No trabalho atual, que será detalhado nas seções subsequentes, os sentimentos expressos em publicações em redes sociais são compilados por hora e, posteriormente, correlacionados com o preço das ações. Os pontos centrais deste estudo incluem a configuração dos parâmetros e a estruturação das ANNs, bem como a ênfase na avaliação do modelo de predição em um cenário real, proporcionando uma análise mais direcionada e específica para as tendências do mercado de ações, sendo esses, diferenciais em relação a outros trabalhos correlatos.

## 4 Método Proposto

Este estudo propõe e desenvolve um método para a interpretação dos preços das ações com os sentimentos analisados a fim de realizar a predição de tendências de preços de ações do mercado brasileiro, oferecendo suporte a investidores. O método emprega técnicas de DL voltadas para séries temporais. A [Figura 2](#) ilustra a estrutura geral do método, que é dividido em quatro etapas principais: 1) Coleta e pré-processamento dos dados; 2) Análise de sentimentos e normalização dos dados; 3) Elaboração e treinamento do modelo de predição; 4) Predição e apresentação dos resultados. Nas seguintes subseções cada uma destas etapas é detalhada, facilitando o entendimento da contribuição de cada uma para o objetivo, realizar predições no mercado de ações brasileiro.

Figura 2 – Processo geral do método proposto



Fonte: Elaborado pelo autor (2024).

## 4.1 Etapa 1: Coleta e Pré-Processamento dos Dados

A primeira etapa do método proposto consiste na obtenção e formatação dos dados para assegurar a coerência e a integridade das informações a serem analisadas. Um modelo específico em Notação de Objetos JavaScript, do inglês, *JavaScript Object Notation* (JSON) foi desenvolvido para armazenar as informações das publicações e os resultados das análises de sentimentos, previstas para a etapa subsequente. Também foi necessária uma abordagem manual para coletar, padronizar e armazenar os dados históricos de preços mais antigos, uma vez que a biblioteca utilizada para buscar esses dados de forma automática restringe a extração de dados a um período máximo de 730 dias com o detalhamento horário. Essa abordagem manual envolveu a exportação dos dados mais antigos no formato de Valores Separados por Vírgula, do inglês, *Comma Separated Values* (CSV), com a subsequente eliminação de campos excedentes, assegurando assim a padronização necessária para a análise.

## 4.2 Etapa 2: Análise de Sentimentos e Normalização dos Dados

Na segunda etapa, ocorre a aplicação de modelos de linguagem pré-treinados, que são especificamente projetados para a análise de sentimentos em textos. Esses modelos, aplicações de NLP, são utilizados para determinar as proporções de sentimentos positivo, neutro e negativo presentes em cada *tweet*. Cada proporção é determinada por um valor entre 0 e 1, de forma que a soma total para cada texto, no contexto do trabalho determinado *tweet*, seja 1. Na sequência, cada intervalo de uma hora recebe a média dos sentimentos dos textos publicados nesse período, com cada soma também resultando em 1. Paralelamente, os preços das ações foram normalizados para variarem também entre 0 e 1. Essa etapa é fundamental para a padronização e generalização dos dados de treinamento, tornando a análise temporal mais consistente.

## 4.3 Etapa 3: Elaboração e Treinamento dos Modelos de Predição

A terceira etapa envolve a preparação das redes neurais e o desenvolvimento dos modelos para a predição de valores futuros no mercado de ações brasileiro. Dois tipos de ANN foram selecionados para comparação: TCN e LSTM. A base de dados compilada anteriormente, que inclui informações financeiras e de sentimentos, é empregada no treinamento das redes. Duas arquiteturas foram propostas para cada tipo de rede: Uma incorporando sentimentos como co-variável na entrada, e outra sem sentimentos. Ambas as arquiteturas foram projetadas para receber como entradas as informações de preços de abertura, mínimo, máximo, fechamento e volume de negociações, enquanto suas saídas predizem os valores da mínima e máxima do dia seguinte.

## 4.4 Etapa 4: Predição e Apresentação dos Resultados

Na quarta etapa, os modelos desenvolvidos na etapa anterior são testados usando um conjunto de dados que não foi exposto durante o treinamento, e a precisão de cada um é avaliada por meio de métricas estatísticas apropriadas. Uma análise comparativa é então realizada para identificar qual modelo demonstra maior eficácia e qual a influência dos sentimentos nas arquiteturas que os incluem.

Por fim os modelos são aplicados em condições reais, realizando a previsão para o dia seguinte ao final de um dia útil, para possibilitar a tomada de posição no mercado antes de seu fechamento. Além disso, os resultados obtidos são submetidos a uma análise comparativa onde, para cada uma das previsões, é calculado o lucro ou prejuízo potencial resultante.

## 5 Resultados Experimentais

### 5.1 Apresentação do Cenário de Estudo

Este trabalho objetiva a análise e predição do valor das ações do Banco do Brasil<sup>®</sup>, disponíveis para negociação na B3, utilizando dados extraídos das plataformas Yahoo Finance<sup>®</sup>, Trading View<sup>®</sup> e X<sup>®</sup>. A escolha das ações do Banco do Brasil<sup>®</sup> se deve à sua relevância no cenário econômico nacional. Como empresa estatal, suas ações apresentam volatilidade ligada ao sentimento público em relação ao governo, resultando em um alto volume de negociações e dados para análise. A obtenção das publicações do X<sup>®</sup>, realizada através da biblioteca *snsrape*, abrangeu o período de 1 de janeiro de 2020 até 7 de junho de 2023, totalizando 23.240 *tweets* (como são conhecidas as publicações no X<sup>®</sup>) mencionando BBAS3, o identificador que representa a empresa alvo deste estudo. Após a extração dessas publicações, a política de aquisição gratuita de dados do X<sup>®</sup> foi alterada, proibindo tal prática.

Os dados financeiros foram coletados do Yahoo Finance<sup>®</sup>, que fornece as informações necessárias para a aplicação do método, a indicação de tempo, o volume de negociações, os preços de abertura, de fechamento, de mínimo e de máximo de cada intervalo para até 730 dias retroativos. No entanto, para abranger o período completo desde a data mais antiga cujos *tweets* foram recuperados, foi necessário integrar manualmente as informações do Trading View<sup>®</sup>, totalizando 9074 horas de informação financeira.

Além das avaliações comuns realizadas em métodos preditivos, os resultados foram testados em condições similares às de um produto comercial real. Ao final de cada dia útil, a previsão dos preços para o dia seguinte foi realizada. Para verificar essas previsões, o lucro ou prejuízo potencial foi calculado, servindo como um indicador para determinar a eficácia da proposta.

### 5.2 Materiais e Métodos

Utilizou-se a plataforma Google Colab<sup>®</sup> para a implementação e aplicação do método proposto, que oferece configurações variadas de *hardware*, incluindo os tipos CPU e T4 GPU. O tipo CPU, equipado com um processador Intel<sup>®</sup> Xeon<sup>®</sup> que opera na frequência de 2.20GHz e 12,7 GB de memória RAM, foi utilizado para desenvolvimento, testes e execução das redes. A T4 GPU, por sua vez, conta com um processador Intel<sup>®</sup> Xeon<sup>®</sup> com 2.00GHz de frequência e 12,7 GB de RAM e uma placa de vídeo NVIDIA<sup>®</sup> Tesla<sup>®</sup> T4 com 15,0 GB de GPU RAM e 2560 núcleos baseados em Arquitetura de Dispositivo de Computação Unificada, do inglês, *Compute Unified Device Architecture* (CUDA), fornecendo recursos avançados para o treinamento de redes neurais devido à compatibilidade das bibliotecas utilizadas com CUDA.

A linguagem de programação escolhida para o desenvolvimento do método proposto foi Python<sup>®</sup>, devido à sua flexibilidade e ampla adoção em projetos de ciência de dados. Para a aquisição de dados históricos de preços, utilizou-se inicialmente a biblioteca *yfinance*, que facilita a obtenção dos dados, porém os limita a 730 dias retroativos com resolução em horas. Para acessar registros mais antigos, compatíveis com as datas dos *tweets* levantados, foi necessário realizar uma coleta manual através do sistema Trading View<sup>®</sup>. Os dados coletados foram então padronizados para corresponder ao formato dos buscados do *yfinance*, garantindo a consistência do conjunto de dados.

As publicações do  $X^{\text{®}}$  foram coletadas utilizando a biblioteca *snsrape*, que, até meados de junho de 2023, permitia a extração eficiente desses dados sem a necessidade de uma Interface de Programação de Aplicações, do inglês, *Application Programming Interface* (API) oficial. A análise de sentimentos foi realizada através de uma ANN com arquitetura *transformers*, apresentada por Wolf *et al.* (2020), que permite a aplicação de modelos avançados de análise de sentimentos disponíveis no repositório de modelos da Hugging Face<sup>®1</sup>. Mais detalhes sobre o pré-processamento dos dados de sentimentos serão discutidos na [subseção 5.2.1](#).

A biblioteca Darts<sup>®</sup>, proposta por Herzen *et al.* (2022), é central para este estudo, oferecendo um conjunto robusto de funcionalidades para a modelagem e análise de séries temporais. Essencial para a integração e análise dos dados coletados, ela foi empregada na construção, treinamento e validação dos modelos preditivos que correlacionam os sentimentos extraídos das publicações com as flutuações dos preços de mercado. A capacidade da biblioteca de suportar múltiplas fontes de dados e a flexibilidade na escolha de algoritmos de DL permitiram uma análise detalhada e ajustada às necessidades específicas do trabalho.

Uma das métricas utilizadas para avaliar o desempenho dos modelos preditivos neste estudo é o Erro Médio Absoluto, do inglês, *Mean Absolute Error* (MAE), que mede a média das magnitudes dos erros entre os valores preditos e os valores observados. É uma medida de precisão útil pois indica o erro com a mesma unidade dos valores que estão sendo avaliados. A equação para calcular o MAE é:

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

onde  $y_i$  representa os valores reais e  $\hat{y}_i$ , os valores preditos pelo modelo, para  $i$  variando de 1 até  $n$ , o número total de observações.

Outra métrica utilizada foi o Erro Quadrático Médio, do inglês, *Mean Squared Error* (MSE) ([Equação 2](#)) utilizada para avaliar a qualidade dos modelos preditivos. Ele calcula a média dos quadrados das diferenças entre os valores preditos  $\hat{y}_i$  e os valores observados  $y_i$ , proporcionando uma medida de precisão que penaliza muito mais severamente diferenças com valores absolutos maiores. Baseado no MSE, a Raiz do Erro Quadrático Médio, do inglês, *Root Mean Squared Error* (RMSE) ([Equação 3](#)) fornece uma medida de erro nas mesmas unidades dos dados observados, como o MAE, porém, penalizando mais severamente erros cujos valores absolutos apresentem maior variação.

<sup>1</sup> Link para o repositório de modelos *transformers* pré-treinados filtrados para análise de sentimentos: [https://huggingface.co/models?pipeline\\_tag=text-classification&sort=trending&search=sentiment](https://huggingface.co/models?pipeline_tag=text-classification&sort=trending&search=sentiment)

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

$$\text{RMSE} = \sqrt{\text{MSE}} \quad (3)$$

O Coeficiente de Determinação ( $R^2$ ), também considerado na avaliação dos resultados neste trabalho, mede a proporção do MSE sobre a variância dos dados observados, conforme expresso na [Equação 4](#). Diferentemente das métricas anteriores, o  $R^2$  não possui unidade, o que permite avaliações independentes da escala dos dados observados sobre a eficácia do modelo. Valores de  $R^2$  mais próximos de 1 (um) indicam uma alta capacidade do modelo em prever os dados, enquanto valores mais baixos sugerem que o modelo é menos eficaz em capturar a variação dos dados. Valores de  $R^2$  inferiores a zero indicam que o modelo é menos adequado para explicar a variância dos dados do que um modelo que simplesmente retorne a média dos dados observados.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (4)$$

### 5.2.1 Pré-Processamento dos Dados de Sentimentos

Para possibilitar a análise de publicações em outras línguas por modelos pré-treinados exclusivamente em inglês, implementou-se um processo de tradução automática. Utilizou-se a biblioteca *googletrans*, na versão 4.0.0-rc1, especificando o Localizador Uniforme de Recursos, do inglês, *Uniform Resource Locator* (URL) do serviço *translate.google.com*. Este método aumenta a quantidade de modelos capazes de analisar os textos em língua portuguesa.

Os modelos *transformers* ([Figura 3](#)), utilizados no trabalho para a análise de sentimentos, possuem limitação quanto ao número máximo de *tokens* que podem analisar. Um *token* é uma unidade básica em NLP, que pode ser uma palavra, um número ou um símbolo. Devido a essa limitação, foi essencial fragmentar os textos, particionando-os na conclusão da frase mais próxima antes de exceder o limite de *tokens* de cada modelo. Após a divisão, as análises dos fragmentos foram consolidadas, calculando a média dos sentimentos de cada parte. Este método foi aplicado com o objetivo de minimizar a perda de contexto nas análises.

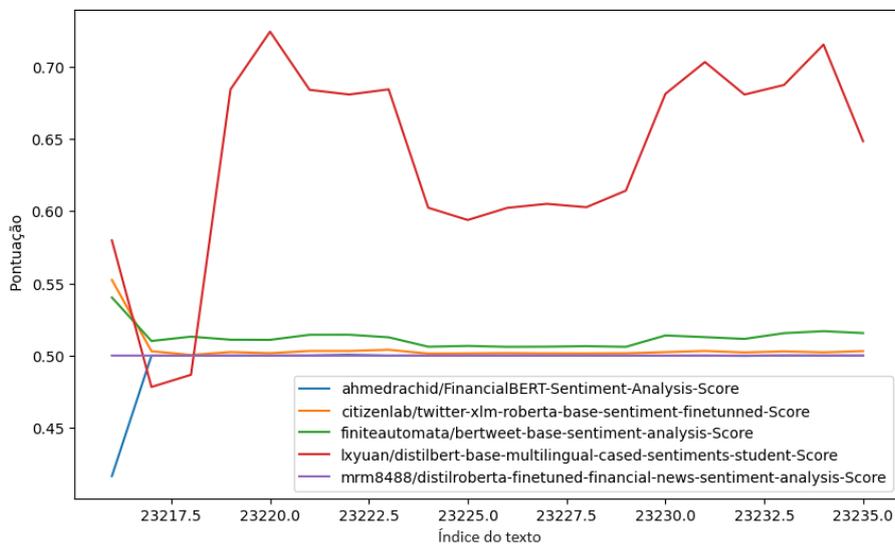
As saídas das análises de sentimentos realizadas foram simplificadas através da [Equação 5](#), onde NEG, NEU e POS representam as proporções Negativa, Neutra e Positiva de cada sentimento. Os valores possíveis para o sentimento simplificado variam de 0 (zero), representando um sentimento totalmente negativo a 1 (um), representando um sentimento totalmente positivo. Essa simplificação foi realizada para facilitar a escolha do modelo que será utilizado neste trabalho, mas os valores utilizados no modelo de predição, serão os três componentes NEG, NEU e POS.

$$\text{SentimentoSimplificado} = \frac{0.0 \times \text{NEG} + 0.5 \times \text{NEU} + 1.0 \times \text{POS}}{\text{NEG} + \text{NEU} + \text{POS}} \quad (5)$$

O modelo selecionado foi aquele que melhor enfatizou os sentimentos observados, minimizando a classificação de textos como neutros. A escolha foi baseada em uma análise gráfica, na qual se considerou a amplitude da resposta dos modelos às publicações. Modelos com maiores variações nas classificações de sentimentos foram considerados mais eficazes em destacar nuances emocionais.

A análise demonstrou uma característica distintiva do modelo *lxyuan/distilbert-base-multilingual-cased-sentiments-student*, cujas classificações dificilmente são neutras e exibe uma variação significativa nas respostas conforme o sentimento específico apresentado em cada texto. Esta propensão do modelo para detectar e enfatizar variações nos sentimentos foi crucial na decisão de adotá-lo para o desenvolvimento deste estudo. Tal escolha se baseou na superioridade em capturar a diversidade emocional expressa nas publicações, como evidenciado pelos dados comparativos apresentados na [Figura 3](#).

Figura 3 – Pontuações de sentimento por texto por modelo de análise de sentimento



Fonte: Elaborado pelo autor (2024).

O modelo adotado, *lxyuan/distilbert-base-multilingual-cased-sentiments-student*, foi projetado para operar em múltiplos idiomas e não depende da tradução prévia dos textos. Isso sugere a possibilidade de que as traduções podem ter introduzido perdas de informação. No entanto, a investigação dessa possibilidade não constituiu o foco deste estudo, e por essa razão, tal aspecto não foi explorado.

### 5.3 Análise e Discussão dos Resultados

As redes elaboradas fizeram uso dos modelos TCN, implementado na biblioteca *Darts* sob o nome *TCNModel* e LSTM que é uma das opções para RNN, identificado por *BlockRNNModel* na biblioteca. Ambos foram configurados com um *dropout* de 0,2 e programados para um treinamento de 100 épocas, processando entradas de 168 blocos horários, equivalentes a 7 dias, e produzindo saídas de 30 blocos. Além disso, no *BlockRNNModel* foi especificado o modelo LSTM com três camadas ocultas, cada uma contendo 25

neurônios. As análises realizadas a seguir detalham a utilização e o desempenho de cada rede elaborada.

Três experimentos foram realizados para avaliar o desempenho do método proposto utilizando as redes elaboradas. O primeiro utilizou dados de sentimentos até a data em que estavam disponíveis, e para datas posteriores, utilizou valores constantes. O segundo, restringiu-se ao intervalo de tempo em que os dados de sentimentos estavam disponíveis. O terceiro, avaliou os modelos em condições reais de mercado. Estes experimentos tiveram como objetivo detalhar como os modelos TCN e LSTM responderiam a diferentes condições e conjuntos de dados.

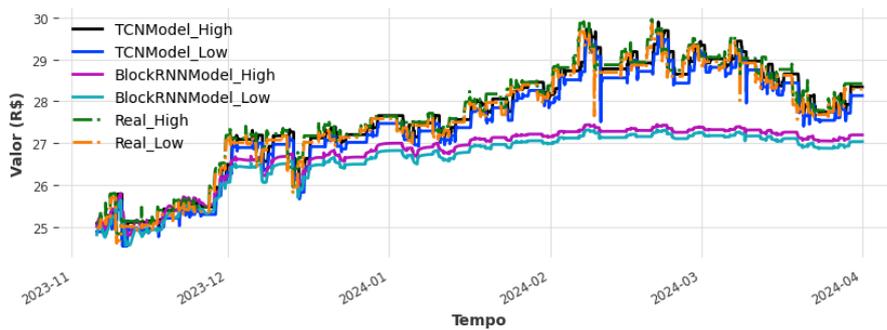
No primeiro experimento, os modelos foram submetidos a um conjunto de dados financeiros que incluía preços mínimos, máximos, de abertura e fechamento, além do volume de transações. Esse conjunto foi enriquecido com dados de sentimentos, integrados apenas na primeira parte do experimento. A particularidade desses dados de sentimentos é que eles estão disponíveis somente até a data de 07 de junho de 2023. Para os valores subsequentes, a partir dessa data, os modelos utilizaram um valor constante de sentimentos, correspondente ao último dia registrado. Este ajuste foi projetado para testar a capacidade dos modelos em adaptar suas previsões à ausência de novas informações de sentimentos, devido à restrição na obtenção das mesmas.

Os resultados revelaram diferenças significativas no desempenho entre os modelos TCN e LSTM. O modelo TCN que utilizava os dados de sentimentos como entrada demonstrou ser eficaz, com um tempo de treinamento de 1541,46 segundos e um tempo de previsão histórica de 12,83 segundos. As métricas de erro incluíram um MAE de R\$ 0,23520, indicando que, em média, as previsões do modelo estavam a aproximadamente 24 centavos do valor real, um RMSE de R\$ 0,31664, mostrando, de acordo com esta métrica, uma diferença média entre as previsões e os valores reais de aproximadamente 32 centavos, e um coeficiente  $R^2$  de 0,93678, indicando uma alta precisão nas previsões. Conforme ilustrado na [Figura 4](#), que apresenta as séries temporais de valores de negociação mínimo e máximo do ativo analisado, são duas curvas para os valores reais e mais duas para cada modelo preditivo.

Por outro lado, a LSTM exigiu um tempo de treinamento consideravelmente maior, totalizando 4814,12 segundos, e um tempo de previsão de 16,04 segundos, também ligeiramente maior do que o anterior. Este modelo apresentou um MAE de R\$ 0,97958, um RMSE de R\$ 1,15162 e um  $R^2$  de 0,16602, sugerindo uma precisão menor nas previsões em relação à TCN. Os resultados indicam uma variação considerável na eficiência e precisão dos modelos, destacando uma superioridade da TCN em termos de tempo e acurácia na previsão.

Nos resultados dos modelos que não utilizaram dados de sentimentos, as diferenças de desempenho continuaram elevadas, conforme ilustrado na [Figura 5](#). A TCN, ao ser treinada, registrou um tempo de ajuste de 1601,46 segundos e um tempo de previsão histórica de 10,16 segundos, alcançando um MAE de R\$ 0,30719, um RMSE de R\$ 0,37478, e um coeficiente  $R^2$  de 0,90731, demonstrando eficácia, porém com uma leve queda na precisão, se comparada ao cenário com sentimentos. Em contraste, a LSTM sem dados de sentimentos, teve um tempo de treinamento de 4789,90 segundos e de 17,77 segundos para previsão. As métricas para este modelo foram R\$ 0,75909 para o MAE, R\$ 0,89713

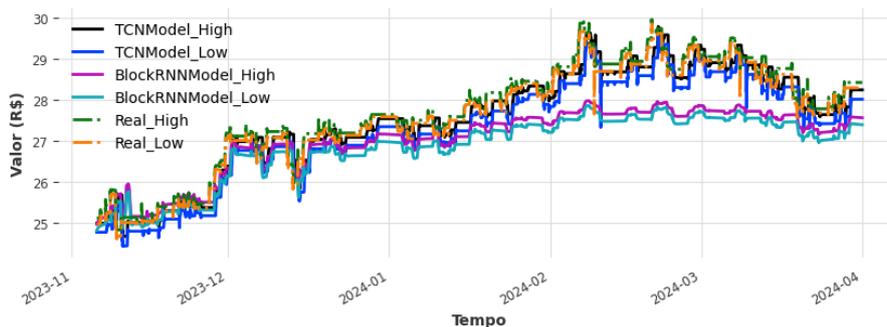
Figura 4 – Experimento 1 - Comparação entre os modelos com sentimentos



Fonte: Elaborado pelo autor (2024).

para o RMSE e 0,47290 para o  $R^2$ , indicando uma precisão significativamente menor em comparação com a TCN, porém relativamente maior que o mesmo modelo com sentimentos, o que sugere que a inclusão parcial de dados de sentimentos durante o treinamento pode ter impactado negativamente a eficácia deste modelo.

Figura 5 – Experimento 1 - Comparação entre os modelos sem sentimentos



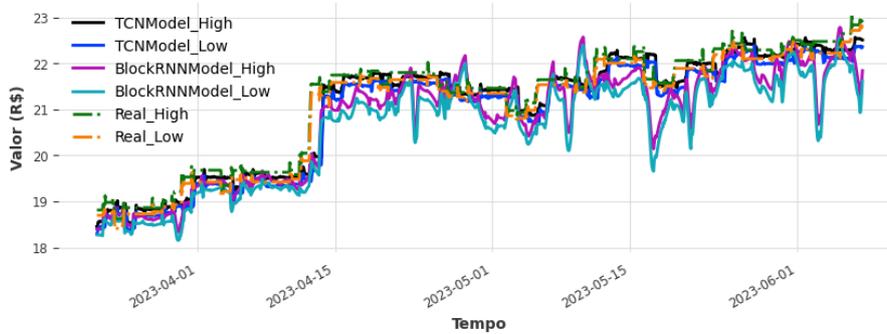
Fonte: Elaborado pelo autor (2024).

No segundo experimento, os modelos TCN e LSTM foram testados novamente com uma metodologia semelhante à do primeiro experimento, mantendo-se as mesmas arquiteturas e parâmetros das redes, porém introduzindo uma mudança significativa nos dados de entrada, que foram limitados exclusivamente ao período em que os dados de sentimentos estavam disponíveis. Esta alteração possibilitou a avaliação do impacto dos sentimentos sobre a precisão das previsões, diferentemente do primeiro experimento, que utilizou os dados de sentimentos disponíveis e os manteve constantes para os dados posteriores.

Para os dados que incluíam sentimentos contemporâneos, até 07 de junho de 2023, a TCN continuou a exibir desempenho superior, com um tempo de treinamento de 1028,16 segundos e um tempo de previsão de 3,97 segundos. As métricas de erro apontaram alta precisão deste modelo, apresentando um MAE de R\$ 0,20252, um RMSE de R\$ 0,31341 e um coeficiente  $R^2$  de 0,93352. Em contrapartida, a LSTM, totalizou 4127,63 segundos em treinamento, e 11,83 segundos para realizar as previsões. As métricas de desempenho

atingiram um MAE de R\$ 0,47466, um RMSE de R\$ 0,58349 e um coeficiente  $R^2$  de 0,76934. Embora não tenha superado a TCN, conforme ilustrado na [Figura 6](#), a LSTM teve um aumento no desempenho quando comparada com os resultados do primeiro experimento.

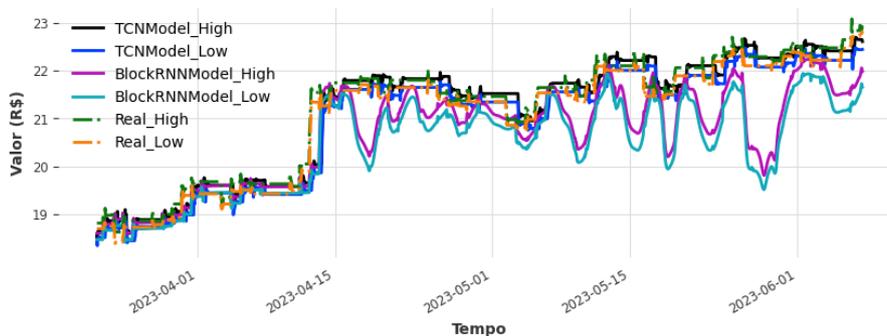
Figura 6 – Experimento 2 - Comparação entre os modelos com sentimentos



Fonte: Elaborado pelo autor (2024).

Quando avaliados sem os dados de sentimentos, a TCN registrou um tempo de treinamento de 1026,64 segundos e um tempo de previsão de 3,90 segundos, com um MAE de R\$ 0,28592, um RMSE de R\$ 0,28916 e um coeficiente  $R^2$  de 0,94341, as melhores métricas, o que sugere que este modelo não foi capaz de aprender através dos sentimentos, apresentando um melhor desempenho sem estas informações. Em oposição, a LSTM, treinada sem os dados de sentimentos, apresentou um tempo de treinamento de 4164,38 segundos e um tempo de previsão de 7,60 segundos. Este cenário resultou em um MAE de R\$ 0,58332, um RMSE de R\$ 0,80167 e um coeficiente  $R^2$  de 0,56430, mostrando uma precisão consideravelmente inferior à observada com sentimentos, apontando que este modelo foi capaz de aprender através dos sentimentos, em oposição ao modelo TCN. A [Figura 7](#) ilustra os valores preditos por cada modelo sem as informações de sentimentos até 07 de junho de 2023.

Figura 7 – Experimento 2 - Comparação entre os modelos sem sentimentos



Fonte: Elaborado pelo autor (2024).

A [Tabela 2](#) sintetiza os dados do experimento 1 e do experimento 2, revelando que o modelo TCN apresenta um desempenho consistente tanto com sentimentos, quanto sem sentimentos, indicando que não se beneficia significativamente com a inclusão dos mesmos.

Por sua vez, o modelo LSTM mostra uma capacidade de aproveitar os sentimentos, registrando seu pior desempenho quando os sentimentos são incluídos parcialmente, no primeiro experimento, e apresentando uma melhora significativa com sentimentos no segundo experimento.

Tabela 2 – Métricas de desempenho por modelo e cenário.

Experimento	Sentimentos	Modelo	MAE (R\$)	RMSE (R\$)	R2
1	Com	TCN	0.23520	0.31664	0.93678
1	Com	LSTM	0.97958	1.15162	0.16602
1	Sem	TCN	0.30719	0.37478	0.90731
1	Sem	LSTM	0.75909	0.89713	0.47290
2	Com	TCN	0.20252	0.31341	0.93352
2	Com	LSTM	0.47466	0.58349	0.76934
2	Sem	TCN	0.28592	0.28916	0.94341
2	Sem	LSTM	0.58332	0.80167	0.56430

Fonte: Elaborado pelo autor (2024).

Os resultados observados nos experimentos 1 e 2 motivaram a realização de um terceiro experimento. Este novo experimento envolve a execução dos modelos ao final de vários dias, objetivando prever os valores de mercado para o dia seguinte. Baseado nas previsões, decide-se pela aquisição de uma opção de compra (*CALL*) ou venda (*PUT*) no mercado de opções<sup>2</sup>. Este cenário, mais próximo das condições reais de operação de um sistema preditivo, proporciona uma avaliação mais robusta e prática da eficácia dos modelos estudados. A eficácia dos modelos será medida pelo lucro gerado ao final dos testes.

Para implementar este terceiro experimento, foi considerado um capital inicial de R\$1.000,00, dividido igualmente em 10 partes. Após o primeiro ciclo de 10 investimentos, o capital total foi reavaliado e redistribuído para uma nova dezena. Neste experimento, o ciclo foi realizado duas vezes, mas em análises futuras, pode ser executado de maneira contínua enquanto houver recursos financeiros. A abordagem utilizada aqui é capaz de neutralizar o risco de uma única predição malsucedida resultar em perdas substanciais ou integrais de capital, aumentando a probabilidade de demonstrar a eficácia de uma estratégia regularmente lucrativa. Cada tentativa foi constituída de um investimento previsto ao final de um dia útil, quando a B3 estava operante, totalizando 20 dias de observação, do dia 19 de abril de 2024 ao dia 21 de maio de 2024.

A [Tabela 3](#) apresenta as previsões dos modelos para o dia seguinte considerando dados com e sem sentimentos e a variação dos preços das opções. Nas células com fundo verde, a previsão indicou que o preço das ações iria subir ou se manter, sugerindo uma opção de compra (*CALL*). Por outro lado, nas células com o fundo vermelho, a previsão apontou uma queda nos preços, indicando uma opção de venda (*PUT*). Até o dia 10 de maio, as opções de compra e venda utilizadas foram *BBASE565* e *BBASQ565*, respectivamente. Após essa data, as opções utilizadas passaram a ser *BBASF560* para

<sup>2</sup> No mercado de opções, são negociados contratos que conferem o direito de comprar ou vender um determinado ativo a um preço predeterminado em uma data futura específica

compra e *BBASR560* para venda. As variações percentuais das opções são apresentadas nas duas últimas colunas da tabela.

Tabela 3 – Previsões dos modelos para o dia seguinte e variação dos preços das opções

		Previsão para o dia seguinte				Variação das opções	
		Com Sentimentos		Sem Sentimentos			
Data	Preço de Fechamento da Ação (R\$)	TCN (R\$)	LSTM (R\$)	TCN (R\$)	LSTM (R\$)	CALL (%)	PUT (%)
19/04	27,80	27,82	27,03	27,75	27,47	-29,76	35,19
22/04	27,35	27,71	26,80	27,65	27,16	15,25	-20,55
23/04	27,62	27,33	26,95	27,23	27,13	-10,29	0,00
25/04	27,31	27,58	26,92	27,47	27,00	27,08	-16,90
26/04	27,62	27,30	26,96	27,21	27,15	-13,11	-5,08
29/04	27,53	27,59	26,98	27,49	27,49	-13,21	5,36
30/04	27,45	27,56	26,96	27,45	27,09	23,91	-23,73
02/05	27,74	27,49	26,99	27,39	27,21	59,65	-48,89
03/05	28,30	27,72	27,17	27,17	27,57	2,20	-8,70
06/05	28,32	28,28	27,19	28,18	27,56	8,60	-19,05
07/05	28,44	28,37	27,22	28,26	27,61	-7,92	5,88
08/05	28,43	28,44	27,21	28,33	27,59	-81,72	211,11
09/05	27,22	28,41	26,96	28,32	27,08	47,06	-44,64
10/05	27,59	27,18	26,99	27,09	27,18	12,22	-15,56
14/05	27,96	27,85	27,11	27,74	27,38	-14,85	18,42
15/05	27,63	28,01	27,03	27,92	27,19	6,98	-17,78
16/05	27,76	27,68	27,05	27,59	27,24	-7,61	0,00
17/05	27,71	27,74	27,06	27,63	27,29	9,41	-18,92
20/05	27,87	27,72	27,09	27,63	27,39	2,15	-6,67
21/05	27,89	27,93	27,10	27,83	27,36	-30,53	53,57

Fonte: Elaborado pelo autor (2024).

Na **Tabela 4** são apresentados os retornos diários obtidos com investimentos de R\$100,00 durante o primeiro ciclo. Vale destacar que a coloração das células não reflete o resultado financeiro do investimento, mas sim a escolha pela opção de compra (*CALL*) ou de venda (*PUT*) das ações. A tabela detalha ainda os retornos totais para cada um dos modelos analisados.

Após o primeiro ciclo de investimentos, os valores iniciais para cada tentativa do segundo ciclo são recalculados somando os retornos totais ao capital inicial de cada modelo e dividindo pelo número de tentativas:

- TCN com sentimentos: Valor inicial =  $\frac{R\$1000,00 - R\$58,45}{10} = R\$94,15$ ;
- LSTM com sentimentos: Valor inicial =  $\frac{R\$1000,00 - R\$102,35}{10} = R\$89,76$ ;

Tabela 4 – Retornos do primeiro ciclo das previsões realizadas na Tabela 3

Data	Com Sentimentos		Sem Sentimentos	
	TCN (R\$)	LSTM (R\$)	TCN (R\$)	LSTM (R\$)
19/04	-29,76	35,19	35,19	35,19
22/04	15,25	-20,55	15,25	-20,55
23/04	0,00	0,00	0,00	0,00
25/04	27,08	-16,90	27,08	-16,90
26/04	-5,08	-5,08	-5,08	-5,08
29/04	-13,21	5,36	5,36	5,36
30/04	23,91	-23,73	23,91	-23,73
02/05	-48,89	-48,89	-48,89	-48,89
03/05	-8,70	-8,70	-8,70	-8,70
06/05	-19,05	-19,05	-19,05	-19,05
TOTAL	-58,45	-102,35	25,07	-102,35

Fonte: Elaborado pelo autor (2024).

- TCN sem sentimentos: Valor inicial =  $\frac{R\$1000,00 + R\$25,07}{10} = R\$102,50$ ;
- LSTM sem sentimentos: Valor inicial =  $\frac{R\$1000,00 - R\$102,35}{10} = R\$89,76$ .

A Tabela 5 revela os retornos financeiros resultantes do segundo ciclo de investimentos. Diferentemente do primeiro ciclo, no qual os retornos financeiros tinham um valor monetário igual ao percentual, devido ao investimento inicial de R\$ 100,00, neste segundo ciclo, é necessário ajustar os valores percentuais da Tabela 3 dividindo-os por 100 e multiplicando pelos valores iniciais das tentativas de cada modelo para obter os retornos reais.

Ambos os modelos LSTM, com e sem sentimentos, recomendaram opções de venda (*PUT*) para todos os dias, indicando um possível treinamento inadequado ou insuficiente, resultando em um lucro de R\$ 64,11 ou 6,51%, devido ao fato de o mercado ter demonstrado uma tendência de queda durante o período. O TCN com sentimentos variou as recomendações, porém obteve prejuízo, perdendo 10,24% (R\$ 102,45) do montante inicial, enquanto o TCN sem sentimentos, foi o modelo que concluiu o experimento apresentando o melhor retorno, um lucro de R\$ 334,51, 33,45% do investimento inicial, sugerindo que a inclusão de sentimentos somente durante uma parte do treinamento não foi benéfica.

A análise dos resultados revela que, ao longo do período do terceiro experimento, o modelo TCN apresentou um número superior de acertos em comparação ao LSTM, registrando 10 acertos com sentimentos e 13 sem sentimentos, respectivamente 50% e 65%, enquanto o LSTM obteve 40%, ou seja, 8 acertos em ambos os casos, conforme indicado nas Tabelas 4 e 5. Um acerto foi definido como um investimento que não resultou em prejuízo. No dia 26 de abril de 2024, ambas as opções resultaram em perdas, mas os modelos predisseram aquela que obteve menor impacto negativo. É importante destacar que os erros nas previsões possuem um limite de perda, enquanto os acertos não têm um teto máximo de ganhos, evidenciado pelo dia 8 de maio, onde foi possível obter um lucro

Tabela 5 – Retornos do segundo ciclo das previsões realizadas na Tabela 3

Data	Com Sentimentos		Sem Sentimentos	
	TCN (R\$)	LSTM (R\$)	TCN (R\$)	LSTM (R\$)
07/05	5,54	5,28	6,03	5,28
08/05	-76,95	189,50	216,40	189,50
09/05	44,31	-40,07	48,24	-40,07
10/05	-14,65	-13,96	-15,95	-13,96
14/05	17,34	16,54	18,88	16,54
15/05	6,57	-15,96	7,15	-15,96
16/05	0,00	0,00	0,00	0,00
17/05	8,86	-16,98	-19,39	-16,98
20/05	-6,28	-5,98	-6,83	-5,98
21/05	-28,74	48,09	54,91	48,09
<b>TOTAL</b>	<b>-44,00</b>	<b>166,46</b>	<b>309,44</b>	<b>166,46</b>

Fonte: Elaborado pelo autor (2024).

superior a 100%. Além disso, a média dos ganhos associados aos acertos superou a dos déficits causadas pelos erros, conforme observado na Tabela 3.

## 6 Considerações Finais e Trabalhos Futuros

Neste trabalho, foram explorados os modelos TCN e LSTM com o objetivo de prever os preços das ações do Banco do Brasil®, negociadas na B3 sob o código *BBAS3*, utilizando tanto dados financeiros quanto dados de sentimentos extraídos de textos publicados no X®. Para a análise das previsões, foram adotadas abordagens estatísticas e também uma abordagem prática, a fim de avaliar a eficácia do método proposto em um cenário realista.

Foram utilizados dados financeiros com resolução horária, abrangendo um período de aproximadamente 4 anos e 5 meses e dados de sentimentos de textos publicados ao longo de 3 anos e 6 meses, para treinar e avaliar os modelos preditivos. Para tal, três cenários foram elaborados, em que dois utilizaram métricas estatísticas, enquanto o terceiro avaliou o retorno financeiro dos investimentos sugeridos em uma simulação de comportamento real do mercado. Os resultados mostraram situações onde a incorporação de sentimentos aprimorou os resultados, e outras, onde a predição baseada exclusivamente nos dados financeiros galgou melhores retornos. Esses achados indicam que a eficácia da incorporação de dados de sentimentos depende da abordagem e da arquitetura de ANN utilizadas, além de sugerir que a limitação na obtenção dos dados de sentimentos prejudicou os resultados, visto que nos trabalhos correlatos, os modelos que utilizaram dados de sentimentos demonstraram desempenho superior.

Considerando a estratégia de gestão de risco adotada, que limita o prejuízo máximo causado por um único investimento malsucedido a 10% do capital e, embora atenuem os lucros, permite que eles permaneçam ilimitados. Os resultados lucrativos sugerem que

o sistema é promissor. Essa abordagem, juntamente com uma taxa de acerto de 65% alcançada pelo melhor modelo, sugere um potencial significativo de lucros elevados associado ao método proposto.

Sendo assim, tem-se a perspectiva de transformar o projeto em um produto real. Ou seja, aplicá-lo na prática repetindo o experimento 3 e realizando as operações de compra e venda manualmente de acordo com as previsões realizadas. É importante frisar que qualquer resultado positivo obtido neste trabalho não constitui garantia de lucro em uma nova sequência de testes/operações. Caso haja interesse em realizar investimentos reais baseados nos métodos propostos por este trabalho, é necessário que os riscos sejam devidamente analisados.

Apesar dos resultados promissores, este estudo apresenta algumas limitações. A coleta de dados de sentimentos foi interrompida devido às mudanças nos termos de uso do X<sup>®</sup>, o que restringiu a análise aos dados disponíveis até 7 de junho de 2023. Além disso, a inclusão de dados de sentimentos mostrou um impacto variável, destacando a necessidade de uma abordagem mais refinada para sua integração nos modelos preditivos. Ademais, a metodologia foi testada somente com as ações do Banco do Brasil<sup>®</sup>, o que limita a generalização dos resultados para outros ativos.

Os trabalhos futuros podem explorar a eficácia de outras arquiteturas de DL, como Transformadores de Fusão Temporal, do inglês, *Temporal Fusion Transformers* (TFT) ou Unidade Recorrente com Portas, do inglês, *Gated Recurrent Unit* (GRU), visando aprimorar a precisão das previsões. Além disso, testar diferentes tipos de ativos pode validar a generalização e adaptabilidade dos modelos. Incorporar técnicas de análise de sentimentos baseadas em contexto, que consigam captar nuances detalhadas dos sentimentos específicos acerca da ação cujo preço está sendo estimado, pode melhorar o impacto e a relevância desses dados. Implementar métodos para acessar em tempo real as publicações ou outras notícias relevantes pode permitir a utilização dessas informações na previsão diária, possibilitando, além de melhores previsões, uma análise mais precisa da influência dos sentimentos nos modelos preditivos. Finalmente, implementar e testar estratégias automatizadas de *trading* baseadas nos modelos preditivos desenvolvidos e avaliar o desempenho em condições reais do mercado contribuirá para a robustez e aplicabilidade prática do método proposto.

## Referências

ANESE, G. *et al.* Impact of public news sentiment on stock market index return and volatility. *Computational Management Science*, v. 20, n. 1, p. 20, Apr 2023. ISSN 1619-6988. Disponível em: <<https://doi.org/10.1007/s10287-023-00454-2>>. Citado 2 vezes nas páginas 18 e 20.

ANGAITKAR, P.; JANGHEL, R. R.; SAHU, T. P. DI-tcnn: Deep learning-based temporal convolutional neural network for prediction of conformational b-cell epitopes. *3 Biotech*, v. 13, n. 9, p. 297, Aug 2023. ISSN 2190-5738. Disponível em: <<https://doi.org/10.1007/s13205-023-03716-7>>. Citado na página 14.

ARAÚJO, R. C. d. C.; MACHADO, M. A. V. Book-to-market ratio, return on equity and brazilian stock returns. *RAUSP Management Journal*, Emerald Publishing Limited, v. 53, n. 3, p. 324–344, 2018. ISSN 2531-0488. Disponível em: <<https://www.emerald.com/insight/content/doi/10.1108/RAUSP-04-2018-001/full/html>>. Citado na página 9.

ARAYA, I. A.; VALLE, C.; ALLENDE, H. A multi-scale model based on the long short-term memory for day ahead hourly wind speed forecasting. *Pattern Recognition Letters*, v. 136, p. 333–340, 2020. ISSN 0167-8655. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0167865519302922>>. Citado na página 15.

BATOOL, I.; KHAN, T. A. Software fault prediction using data mining, machine learning and deep learning techniques: A systematic literature review. *Computers and Electrical Engineering*, v. 100, p. 107886, 2022. ISSN 0045-7906. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0045790622001744>>. Citado na página 13.

BENGIO, Y. *Learning Deep Architectures for AI*. [s.n.], 2009. Disponível em: <<https://doi.org/10.1561/22000000006>>. Citado na página 12.

BOLLEN, J.; MAO, H.; ZENG, X. Twitter mood predicts the stock market. *Journal of Computational Science*, v. 2, n. 1, p. 1–8, 2011. ISSN 1877-7503. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S187775031100007X>>. Citado na página 7.

BUSTOS, O.; POMARES-QUIMBAYA, A. Stock market movement forecast: A systematic review. *Expert Systems with Applications*, v. 156, p. 113464, 2020. ISSN 0957-4174. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0957417420302888>>. Citado na página 7.

CAROSIA, A. E. de O.; COELHO, G. P.; da Silva, A. E. A. Investment strategies applied to the brazilian stock market: A methodology based on sentiment analysis with deep learning. *Expert Systems with Applications*, v. 184, p. 115470, 2021. ISSN 0957-4174. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0957417421008812>>. Citado 3 vezes nas páginas 10, 17 e 20.

CASTRO, F. H. *et al.* Fifty-year history of the ibovespa. *Revista Brasileira de Finanças*, v. 17, p. 47–65, 2019. ISSN 1679-0731. Disponível em: <<https://periodicos.fgv.br/rbfin/article/view/80028>>. Citado na página 9.

CASTRO, R. D. d.; MARQUES, V. A. Relevância da informação contábil para o mercado de capitais: evidências no mercado brasileiro. *Enfoque: Reflexão Contábil*, v. 32, n. 1, p. 109–124, maio 2013. Disponível em: <<https://periodicos.uem.br/ojs/index.php/Enfoque/article/view/15362>>. Citado na página 10.

CHEN, S.; ZHOU, C. Stock prediction based on genetic algorithm feature selection and long short-term memory neural network. *IEEE Access*, v. 9, p. 9066–9072, 2021. Citado na página 8.

CHONG, E.; HAN, C.; PARK, F. C. Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies. *Expert Systems with Applications*, v. 83, p. 187–205, 2017. ISSN 0957-4174. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0957417417302750>>. Citado na página 7.

DEAN, J. A Golden Decade of Deep Learning: Computing Systems & Applications. *Daedalus*, v. 151, n. 2, p. 58–74, 05 2022. ISSN 0011-5266. Disponível em: <[https://doi.org/10.1162/daed\\_a\\_01900](https://doi.org/10.1162/daed_a_01900)>. Citado na página 13.

DOGGART, P. *et al.* Identifying noisy ecg signals in large datasets using a temporal convolutional neural network trained to estimate pseudo-snr. In: *2023 Computing in Cardiology (CinC)*. [s.n.], 2023. v. 50, p. 1–4. Disponível em: <<https://ieeexplore-ieee-org.ez46.periodicos.capes.gov.br/document/10363865>>. Citado na página 14.

GANDHI, A. *et al.* Multimodal sentiment analysis: A systematic review of history, datasets, multimodal fusion methods, applications, challenges and future directions. *Information Fusion*, v. 91, p. 424–444, 2023. ISSN 1566-2535. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1566253522001634>>. Citado na página 11.

GOMES, F. R. A bolsa de valores brasileira como fonte de informações financeiras. *Perspectivas em Ciência da Informação*, v. 2, n. 2, nov. 2007. Disponível em: <<https://periodicos.ufmg.br/index.php/pci/article/view/23238>>. Citado na página 8.

GROSSMAN, S. J.; STIGLITZ, J. E. On the impossibility of informationally efficient markets. *The American Economic Review*, American Economic Association, v. 70, n. 3, p. 393–408, 1980. ISSN 00028282. Disponível em: <<http://www.jstor.org/stable/1805228>>. Citado na página 10.

HARROU, F. *et al.* Chapter 7 - unsupervised recurrent deep learning scheme for process monitoring. In: HARROU, F. *et al.* (Ed.). *Statistical Process Monitoring Using Advanced Data-Driven and Deep Learning Approaches*. Elsevier, 2021. p. 225–253. ISBN 978-0-12-819365-5. Disponível em: <<https://www.sciencedirect.com/science/article/pii/B9780128193655000139>>. Citado na página 16.

HASSAIRI, S.; EJBALI, R.; ZAIED, M. Supervised image classification using deep convolutional wavelets network. In: *2015 IEEE 27th International Conference on Tools with Artificial Intelligence (ICTAI)*. [s.n.], 2015. p. 265–271. Disponível em: <<https://doi.org/10.1109/ICTAI.2015.49>>. Citado na página 12.

HERZEN, J. *et al.* Darts: User-friendly modern machine learning for time series. *Journal of Machine Learning Research*, v. 23, n. 124, p. 1–6, 2022. Disponível em: <<http://jmlr.org/papers/v23/21-1177.html>>. Citado na página 23.

HOCHREITER, S.; SCHMIDHUBER, J. Long Short-Term Memory. *Neural Computation*, v. 9, n. 8, p. 1735–1780, 11 1997. ISSN 0899-7667. Disponível em: <<https://doi.org/10.1162/neco.1997.9.8.1735>>. Citado na página 15.

HUANG, A. C.; MENG, S. H.; HUANG, T. J. A survey on machine and deep learning in semiconductor industry: methods, opportunities, and challenges. *Cluster Computing*, v. 26, n. 6, p. 3437–3472, Dec 2023. ISSN 1573-7543. Disponível em: <<https://doi.org/10.1007/s10586-023-04115-6>>. Citado na página 16.

JAHIN, M. A. *et al.* Big data—supply chain management framework for forecasting: Data preprocessing and machine learning techniques. *Archives of Computational Methods in Engineering*, Mar 2024. ISSN 1886-1784. Disponível em: <<https://doi.org/10.1007/s11831-024-10092-9>>. Citado na página 14.

JIM, J. R. *et al.* Recent advancements and challenges of nlp-based sentiment analysis: A state-of-the-art review. *Natural Language Processing Journal*, v. 6, p. 100059, 2024. ISSN 2949-7191. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2949719124000074>>. Citado na página 11.

JOSHI, M.; GOEL, G. Stock market prediction approach: An analysis. In: *2023 International Conference on Artificial Intelligence and Smart Communication (AISC)*. [s.n.], 2023. p. 945–948. Disponível em: <<https://ieeexplore-ieee-org.ez46.periodicos.capes.gov.br/document/10085262>>. Citado na página 7.

KALOEV, M.; KRASSTEV, G. Comparative analysis of activation functions used in the hidden layers of deep neural networks. In: *2021 3rd International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*. [s.n.], 2021. p. 1–5. Disponível em: <<https://ieeexplore.ieee.org/document/9461312>>. Citado na página 13.

KAMAL, S. *et al.* Trading stocks based on financial news using attention mechanism. *Mathematics*, v. 10, n. 12, 2022. Cited by: 7; All Open Access, Gold Open Access. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85132404425&doi=10.3390%2fmath10122001&partnerID=40&md5=ddd55d674df1993de6f815c1eb2ef284>>. Citado 2 vezes nas páginas 19 e 20.

KARIRI, E. *et al.* Exploring the advancements and future research directions of artificial neural networks: A text mining approach. *Applied Sciences*, v. 13, n. 5, 2023. ISSN 2076-3417. Disponível em: <<https://www.mdpi.com/2076-3417/13/5/3186>>. Citado na página 12.

KHALIL, F.; PIPA, G. Is deep-learning and natural language processing transcending the financial forecasting? investigation through lens of news analytic process. *Computational Economics*, v. 60, n. 1, p. 147–171, Jun 2022. ISSN 1572-9974. Disponível em: <<https://doi.org/10.1007/s10614-021-10145-2>>. Citado 2 vezes nas páginas 18 e 20.

KHATTAK, B. H. A. *et al.* A systematic survey of ai models in financial market forecasting for profitability analysis. *IEEE Access*, v. 11, p. 125359–125380, 2023. Disponível em: <<https://ieeexplore.ieee.org/document/10309124>>. Citado na página 8.

KIM, H. E. *et al.* Transfer learning for medical image classification: a literature review. *BMC Medical Imaging*, v. 22, n. 1, p. 69, Apr 2022. Disponível em: <<https://link.springer.com/article/10.1186/s12880-022-00793-7>>. Citado na página 14.

KIM, T. S.; REITER, A. Interpretable 3d human action analysis with temporal convolutional networks. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. [s.n.], 2017. p. 1623–1631. Disponível em: <<https://ieeexplore-ieee-org.ez46.periodicos.capes.gov.br/document/8014941>>. Citado na página 14.

KISHTAINY, N.; KINDERSLEY, D. *The Economics Book*. [S.l.]: DK, 2013. ISBN 9781740339612. Citado na página 9.

KOUKARAS, P.; NOUSI, C.; TJORTJIS, C. Stock market prediction using microblogging sentiment analysis and machine learning. *Telecom*, v. 3, n. 2, p. 358 – 378, 2022. Cited by: 16; All Open Access, Gold Open Access. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85134408271&doi=10.3390%2ftelecom3020019&partnerID=40&md5=7c8309fd2f6173230e9529e21fa23986>>. Citado 2 vezes nas páginas 19 e 20.

KUMAR, S. K. *et al.* Stock price prediction using optimal network based twitter sentiment analysis. *Intelligent Automation & Soft Computing*, v. 33, n. 2, p. 1217–1227, 2022. ISSN 2326-005X. Disponível em: <<http://www.techscience.com/iasc/v33n2/46775>>. Citado na página 10.

LAU, M. M.; LIM, K. H. Review of adaptive activation function in deep neural network. In: *2018 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES)*. [s.n.], 2018. p. 686–690. Disponível em: <<https://ieeexplore.ieee.org/document/8626714>>. Citado na página 13.

LEA, C. *et al.* Temporal convolutional networks for action segmentation and detection. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. [s.n.], 2017. p. 1003–1012. Disponível em: <<https://ieeexplore-ieee-org.ez46.periodicos.capes.gov.br/document/8099596>>. Citado na página 14.

LIN, Y.-L.; LAI, C.-J.; PAI, P.-F. Using deep learning techniques in forecasting stock markets by hybrid data with multilingual sentiment analysis. *ELECTRONICS*, MDPI, v. 11, n. 21, NOV 2022. Disponível em: <<https://doi.org/10.3390/electronics11213513>>. Citado 2 vezes nas páginas 18 e 20.

MAHMUD, M. S. *et al.* A systematic literature review on deep learning applications for precision cattle farming. *Computers and Electronics in Agriculture*, v. 187, p. 106313, 2021. ISSN 0168-1699. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0168169921003306>>. Citado na página 13.

MALKIEL, B. G.; FAMA, E. F. Efficient capital markets: A review of theory and empirical work\*. *The Journal of Finance*, v. 25, n. 2, p. 383–417, 1970. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-6261.1970.tb00518.x>>. Citado na página 10.

MAQSOOD, H. *et al.* A local and global event sentiment based efficient stock exchange forecasting using deep learning. *International Journal of Information Management*, v. 50, p. 432–451, 2020. ISSN 0268-4012. Disponível em:

<<https://www.sciencedirect.com/science/article/pii/S0268401218313008>>. Citado na página 9.

MINTARYA, L. N. *et al.* Machine learning approaches in stock market prediction: A systematic literature review. *Procedia Computer Science*, v. 216, p. 96–102, 2023. ISSN 1877-0509. 7th International Conference on Computer Science and Computational Intelligence 2022. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1877050922021937>>. Citado na página 10.

MUB3, M. da Bolsa do B. B3, n.d. Acessado em: 30-01-2024. Disponível em: <<https://mub3.org.br/acervo/historia-da-bolsa>>. Citado na página 9.

MUELLER, J. P.; MASSARON, L. *Deep learning for dummies*. [S.l.]: John Wiley & Sons, 2019. ISBN 978-1-119-54304-6. Citado na página 15.

NOLASCO, D.; OLIVEIRA, J. Subevents detection through topic modeling in social media posts. *Future generation computer systems*, Elsevier B.V, v. 93, p. 290–303, 2019. ISSN 0167-739X. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0167739X18307611>>. Citado na página 11.

NTI, I. K.; ADEKOYA, A. F.; WEYORI, B. A. A systematic review of fundamental and technical analysis of stock market predictions. *Artificial Intelligence Review*, v. 53, n. 4, p. 3007 – 3057, 2020. Cited by: 173. Disponível em: <<https://link.springer.com/article/10.1007/s10462-019-09754-z>>. Citado na página 7.

ORTIZ, D. P. Economic policy statements, social media, and stock market uncertainty: An analysis of donald trump’s tweets. *Journal of economics and finance*, Springer US, New York, v. 47, n. 2, p. 333–367, 2023. ISSN 1055-0925. Disponível em: <<https://doi.org/10.1007/s12197-022-09608-5>>. Citado na página 11.

OWEN, L.; OKTARIANI, F. Senn: Stock ensemble-based neural network for stock market prediction using historical stock data and sentiment analysis. In: *2020 International Conference on Data Science and Its Applications (ICoDSA)*. [s.n.], 2020. p. 1–7. Disponível em: <<https://doi.org/10.1109/ICoDSA50139.2020.9212982>>. Citado 2 vezes nas páginas 19 e 20.

PAGOLU, V. S. *et al.* Sentiment analysis of twitter data for predicting stock market movements. In: *2016 International Conference on Signal Processing, Communication, Power and Embedded System (SCOPE5)*. [s.n.], 2016. p. 1345–1350. Disponível em: <<https://ieeexplore-ieee-org.ez46.periodicos.capes.gov.br/document/7955659>>. Citado na página 7.

PERROT, N. M. *et al.* Chapter 10 - machine learning for agri-food processes: learning from data, human knowledge, and interactions. In: SIROHI, R. *et al.* (Ed.). *Current Developments in Biotechnology and Bioengineering*. Elsevier, 2022. p. 261–286. ISBN 978-0-323-91167-2. Disponível em: <<https://www.sciencedirect.com/science/article/pii/B978032391167200006X>>. Citado na página 16.

REN, R.; WU, D. D.; LIU, T. Forecasting stock market movement direction using sentiment analysis and support vector machine. *IEEE Systems Journal*, v. 13, n. 1, p.

760–770, 2019. Disponível em: <<https://ieeexplore-ieee-org.ez46.periodicos.capes.gov.br/document/8326522>>. Citado na página 7.

SAMAL, K. K. R.; BABU, K. S.; DAS, S. K. Multi-directional temporal convolutional artificial neural network for pm2.5 forecasting with missing values: A deep learning approach. *Urban Climate*, v. 36, p. 100800, 2021. ISSN 2212-0955. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2212095521000304>>. Citado na página 14.

SAMANT, R. M. *et al.* Framework for deep learning-based language models using multi-task learning in natural language understanding: A systematic literature review and future directions. *IEEE Access*, v. 10, p. 17078–17097, 2022. Citado na página 13.

SARKAR, A. *et al.* Lstmsa: A novel approach for stock market prediction using lstm and sentiment analysis. In: *2020 International Conference on Computer Science, Engineering and Applications (ICCSEA)*. [s.n.], 2020. p. 1–6. Disponível em: <<https://doi.org/10.1109/ICCSEA49143.2020.9132928>>. Citado 2 vezes nas páginas 19 e 20.

SEZER, O. B.; GUDELEK, M. U.; OZBAYOGLU, A. M. Financial time series forecasting with deep learning : A systematic literature review: 2005–2019. *Applied Soft Computing*, v. 90, p. 106181, 2020. ISSN 1568-4946. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1568494620301216>>. Citado na página 12.

SHAH, D. V. *et al.* Stock price prediction using lstm-arima hybrid neural network model with sentiment analysis of news headlines. In: *2022 International Conference on Futuristic Technologies (INCOFT)*. [s.n.], 2022. p. 1–4. Disponível em: <<https://doi.org/10.1109/INCOFT55651.2022.10094422>>. Citado 2 vezes nas páginas 18 e 20.

SIREGAR, S.; WANTO, A. Analysis of artificial neural network accuracy using backpropagation algorithm in predicting process (forecasting). *IJISTECH (International Journal of Information System and Technology)*, v. 1, n. 1, p. 34–42, 2017. ISSN 2580-7250. Disponível em: <<https://ijistech.org/ijistech/index.php/ijistech/article/view/4>>. Citado na página 13.

VANAMBATHINA, S. D. *et al.* Real time speech enhancement using densely connected neural networks and squeezed temporal convolutional modules. *Multimedia Tools and Applications*, Nov 2023. ISSN 1573-7721. Disponível em: <<https://doi.org/10.1007/s11042-023-17492-2>>. Citado na página 15.

WANG, Z. *et al.* Fine-grained sentiment analysis of social media with emotion sensing. In: *2016 Future Technologies Conference (FTC)*. [s.n.], 2016. p. 1361–1364. Disponível em: <<https://doi.org/10.1109/FTC.2016.7821783>>. Citado na página 11.

WOLF, T. *et al.* Transformers: State-of-the-art natural language processing. In: LIU, Q.; SCHLANGEN, D. (Ed.). *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*.

Online: Association for Computational Linguistics, 2020. p. 38–45. Disponível em: <<https://aclanthology.org/2020.emnlp-demos.6>>. Citado na página 23.

YANG, F. *et al.* A systematic literature review of deep learning approaches for sketch-based image retrieval: Datasets, metrics, and future directions. *IEEE Access*, v. 12, p. 14847–14869, 2024. Disponível em: <<https://ieeexplore.ieee.org/document/10413476>>. Citado na página 12.

YI, J. *et al.* Analysis of stock market public opinion based on web crawler and deep learning technologies including 1dcnn and lstm. *Arabian Journal for Science and Engineering*, v. 48, n. 8, p. 9941–9962, Aug 2023. ISSN 2191-4281. Disponível em: <<https://doi.org/10.1007/s13369-022-07444-7>>. Citado 2 vezes nas páginas 17 e 20.

YI, J. *et al.* Sentiment analyzer: extracting sentiments about a given topic using natural language processing techniques. In: *Third IEEE International Conference on Data Mining*. [s.n.], 2003. p. 427–434. Disponível em: <<https://doi.org/10.1109/ICDM.2003.1250949>>. Citado na página 10.

ZAINI, N. *et al.* A systematic literature review of deep learning neural network for time series air quality forecasting. *Environmental Science and Pollution Research*, v. 29, n. 4, p. 4958–4990, 1 2022. ISSN 1614-7499. Disponível em: <<https://doi.org/10.1007/s11356-021-17442-1>>. Citado na página 16.

ZAYTAR, M. A.; AMRANI, C. E. Sequence to sequence weather forecasting with long short-term memory recurrent neural networks. *International Journal of Computer Applications*, Foundation of Computer Science, v. 143, n. 11, p. 7–11, 2016. Disponível em: <[https://scholar.google.com/scholar\\_lookup?title=Sequence%20to%20sequence%20weather%20forecasting%20with%20long%20short-term%20memory%20recurrent%20neural%20networks](https://scholar.google.com/scholar_lookup?title=Sequence%20to%20sequence%20weather%20forecasting%20with%20long%20short-term%20memory%20recurrent%20neural%20networks)>. Citado na página 16.

ZUO, K. Integrated forecasting models based on lstm and tcn for short-term electricity load forecasting. In: *2023 9th International Conference on Electrical Engineering, Control and Robotics (EECR)*. [s.n.], 2023. p. 207–211. Disponível em: <<https://ieeexplore-ieee-org.ez46.periodicos.capes.gov.br/document/10149951>>. Citado na página 8.