

UNIVERSIDADE DE SÃO PAULO

Instituto de Ciências Matemáticas e de Computação

Análise de oportunidades no mercado de ações brasileiro usando aprendizado de máquina

Leandro Sampaio Peçanha

Monografia - MBA em Inteligência Artificial e Big Data

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: _____

Leandro Sampaio Peçanha

Análise de oportunidades no mercado de ações brasileiro usando aprendizado de máquina

Monografia apresentada ao Departamento de Ciências de Computação do Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo - ICMC/USP, como parte dos requisitos para obtenção do título de Especialista em Inteligência Artificial e Big Data.

Área de concentração: Inteligência Artificial

Orientador: Prof. Ivan José dos Reis Filho

Versão original

São Carlos

2024

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTA TRABALHO,
POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO PARA FINS DE ESTUDO E
PESQUISA, DESDE QUE CITADA A FONTE.

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi, ICMC/USP, com os dados
fornecidos pelo(a) autor(a)

S856m	<p>Peçanha, Leandro Sampaio Análise de oportunidades no mercado de ações brasileiro usando aprendizado de máquina / Leandro Sampaio Peçanha ; orientador Ivan José dos Reis Filho. – São Carlos, 2024. 55 p. : il. (algumas color.) ; 30 cm.</p> <p>Monografia (MBA em Inteligência Artificial e Big Data) – Instituto de Ciências Matemáticas e de Computação, Universi- dade de São Paulo, 2024.</p> <p>1. LaTeX. 2. abnTeX. 3. Classe USPSC. 4. Editoração de texto. 5. Normalização da documentação. 6. Tese. 7. Disserta- ção. 8. Documentos (elaboração). 9. Documentos eletrônicos. I. Reis Filho, Ivan José dos, orient. II. Título.</p>
-------	--

Leandro Sampaio Peçanha

**Análise de oportunidades no mercado de ações brasileiro
usando aprendizado de máquina**

Monograph presented to the Departamento de Ciências de Computação do Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo - ICMC/USP, as part of the requirements for obtaining the title of Specialist in Artificial Intelligence and Big Data.

Concentration area: Artificial Intelligence

Advisor: Prof. Ivan José dos Reis Filho

Original version

São Carlos

2024

AGRADECIMENTOS

Agradeço à minha mãe e ao meu pai, exemplos de dedicação à família e ao trabalho, às minhas irmãs e ao meu irmão pelo sempre presente amparo e fonte de inspiração.

Agradeço à minha esposa, companheira de vida e de *home office*, com quem convivo harmoniosamente vinte e quatro horas por dia, sete dias por semana. E ao meu filho, exemplo de criança sociável, curiosa e inteligente.

Agradeço também ao corpo docente do ICMC pela organização do curso e ao professor Ivan Filho pela orientação neste trabalho.

*“O destino guia o homem que assim deseja,
e arrasta os que não desejam”
Cleanthes de Assos*

RESUMO

Peçanha, L. S. **Análise de oportunidades no mercado de ações brasileiro usando aprendizado de máquina**. 2024. 55p. Monografia (MBA em Inteligência Artificial e Big Data) - Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2024.

Há diversas abordagens possíveis para o investidor do mercado de ações que busca alcançar um desempenho superior ao mercado de ações com uma devida exposição mais controlada ao risco. Esse trabalho trata de avaliar uma abordagem quantitativa específica, que é busca de oportunidades de dias com alta probabilidade de serem positivos (valor de fechamento superior ao valor da abertura). A vantagem estaria em não precisar operar todos os dias, e nem permanecer comprado por muito tempo ao papel, mas sim comprar e vender em dias determinados pelo modelo. Neste trabalho, selecionou-se dados históricos de cinco ações relevantes que fazem parte da composição do índice Ibovespa, que são a Vale, a Petrobrás, o Itaú, a Wege e a Eletrobrás, pertencentes a diferentes segmentos, como mineração, petróleo, banco, indústria e energia. Junto a informações macroeconômicas como cotação do dólar e índice Ibovespa, foram submetidos a treinamento de aprendizado de máquina como Gated Recurrent Units (GRU), Convolutional Neural Network (CNN), Recurrent Neural Network (RNN) e Long Short-Term Memory (LSTM) para tentar extrair esse padrão dos dias que serão positivos, e alcançar uma alta precisão. Os resultados se focaram principalmente em avaliar nível de precisão e construir uma matriz de resultados para demonstrar também a quantidade de dias em que o modelo indicou a compra. Com base no resultado, cada investidor pode traçar sua devida estratégia para tentar extrair vantagem dos resultados.

Palavras-chave: oportunidades; ações; bovespa.

ABSTRACT

Peçanha, L. S. **Oportunidades**. 2024. 55p. Monograph (MBA in Artificial Intelligence and Big Data) - Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2024.

There are various possible approaches for stock market investors seeking to achieve superior performance with a more controlled exposure to risk. This paper evaluates a specific quantitative approach that aims to identify days with a high probability of being positive (i.e., closing values higher than opening values). The advantage of this method lies in not needing to trade every day or hold positions for extended periods but rather buying and selling on specific days identified by the model. For this analysis, historical data from five significant stocks that are part of the Ibovespa index were selected: Vale, Petrobrás, Itaú, Weg, and Eletrobrás, which belong to different sectors such as mining, oil, banking, industry, and energy. Along with macroeconomic information such as the exchange rate of the dollar and the Ibovespa index, these data were subjected to machine learning training using Gated Recurrent Units (GRU), Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and Long Short-Term Memory (LSTM) to identify patterns of positive days and achieve high accuracy. The results focused primarily on assessing accuracy levels and constructing a result matrix to also demonstrate the number of days the model indicated a buy. Based on the results, each investor can devise their strategy to leverage the findings.

Keywords: opportunities; stocks; bovespa.

LISTA DE FIGURAS

Figura 1 – Dados de preço e volume do índice Ibovespa ao longo de alguns meses do ano de 2023 e 2024.	28
Figura 2 – Combinações de tendência, ciclo e ruído (Fonte: Elaborado pelo autor através de biblioteca <i>matplotlib.pyplot</i>).	29
Figura 3 – Marcadores de mínima, máxima, abertura e fechamento representando em candle stick	40

LISTA DE TABELAS

Tabela 1 – Ações escolhidas para esse trabalho. Participação referente à carteira do ibovespa em 01/04/2024	27
Tabela 2 – Dados extraídos da plataforma TradingView para insumo desse trabalho	39
Tabela 3 – Dados das ações após incorporações	41
Tabela 4 – Dados das ações após normalizações	41
Tabela 5 – Contraste percentual de dias positivos e negativos para cada ação, divididos em treinamento e teste.	43
Tabela 6 – Comparação de Desempenho de Modelos de Aprendizado de Máquina - VALE3	44
Tabela 7 – Comparação de Desempenho de Modelos de Aprendizado de Máquina - PETR4	44
Tabela 8 – Comparação de Desempenho de Modelos de Aprendizado de Máquina - ITUB4	45
Tabela 9 – Comparação de Desempenho de Modelos de Aprendizado de Máquina - WEG3	45
Tabela 10 – Comparação de Desempenho de Modelos de Aprendizado de Máquina - ELET3	45
Tabela 11 – Quantidade de vezes em que o modelo teve desempenho superior aos demais	45
Tabela 12 – Desempenho dos Modelos para diferentes cortes de Sigmoide - quantidade e precisão - VALE3	46
Tabela 13 – Desempenho dos Modelos para diferentes cortes de Sigmoide - quantidade e precisão - PETR4	46
Tabela 14 – Desempenho dos Modelos para diferentes cortes de Sigmoide - quantidade e precisão - ITUB4	47
Tabela 15 – Desempenho dos Modelos para diferentes cortes de Sigmoide - quantidade e precisão - WEG3	47
Tabela 16 – Desempenho dos Modelos para diferentes cortes de Sigmoide - quantidade e precisão - ELET3	48
Tabela 17 – Corte sigmoide que alcançou o melhor resultado em termos de precisão	48

LISTA DE ABREVIATURAS E SIGLAS

IA	Inteligência Artificial
RNN	Recurrent Neural Network
LSTM	Long Short-Term Memory
GRU	Gated Recurrent Units
CNN	Convolutional Neural Network
HME	Hipótese dos Mercados Eficientes
B3	Brasil, Bolsa, Balcão

SUMÁRIO

1	INTRODUÇÃO	23
1.1	Motivações e lacunas	24
1.2	Objetivo e Questões de Pesquisa	24
1.3	Possibilidades e Hipóteses em Caso de Sucesso	25
1.4	Organização do Trabalho	25
2	FUNDAMENTOS E REFERENCIAL TEÓRICO	27
2.1	Mercado Financeiro	27
2.2	Séries Temporais	28
2.2.1	Classificação	28
2.3	Modelos de Classificação	30
2.3.1	Redes Neurais Recorrentes	30
2.3.2	Convolutional Neural Network	31
2.3.3	Long-Short Term Memory	32
2.3.4	Gated Recurrent Unit	32
2.4	Métricas de avaliação	33
2.4.1	Acurácia	33
2.4.2	Precisão	34
2.4.3	Recall	35
2.4.4	F1-Score	36
2.4.5	Definição de oportunidade para esse trabalho	36
2.5	Trabalhos relacionados	37
3	PROPOSTA E DESENVOLVIMENTO	39
3.1	Obtenção dos dados	39
3.2	Pré processamento dos dados	40
3.3	Configuração experimental	42
3.4	Resultados	44
3.4.1	Avaliação geral do modelos	44
4	CONCLUSÃO	51
	Referências	53

1 INTRODUÇÃO

O mercado financeiro é um importante elemento das economias capitalistas modernas, onde cotas patrimoniais de empresas são negociadas entre investidores através de pregões eletrônicos, intermediada por bolsas de valores. A função principal deste mercado é a de possibilitar às empresas a captação de recursos entre investidores para o financiamento das suas operações e das suas atividades de expansão. Segundo Pagano (1993), a criação de um mercado de ações aumenta a produtividade do investimento e a taxa de crescimento.

Pela perspectiva dos investidores, proporciona a possibilidade da participação nos lucros e no crescimento das empresas, seja pela valorização de suas ações ou pela distribuição de dividendos. O mercado de ações é vital para a dinâmica econômica, uma vez que fomenta o investimento direto nas empresas, o crescimento econômico e a distribuição de riqueza, o que contribui para a estabilidade financeira e para o desenvolvimento do mercado de capitais. De acordo com Bond, Edmans and Goldstein (2012), muitos especuladores com diferentes perspectivas e informações se encontram nesse mercado para negociar, tentando lucrar com suas informações. Os preços agregam essas diversas perspectivas de informação e, em última instância, refletem uma avaliação precisa do valor da empresa.

A teoria clássica de Fama (1970), conhecida por Hipótese dos Mercados Eficientes (HME), afirma que os preços dos ativos já refletem todas as informações conhecidas pelos investidores, o que tornaria frustrante qualquer tentativa de buscar uma predição a fim de obter alguma vantagem e retorno acima da média do mercado. Porém, empiricamente, sempre foi possível observar investidores que fizeram boas escolhas e colheram bons frutos; alguns, certamente, foram “iludidos pelo acaso”, fenômeno descrito por Taleb (2004) no qual alguns arquétipos de investidores obtiveram sucesso creditado mais por um conjunto de acasos e uma imprudência excessiva na exposição ao risco, que por uma robusta e verdadeira estratégia vencedora em si.

Portanto é necessário um cuidado extra e um ceticismo adicional para avaliar a existência real de estratégias capazes de superar consistentemente a média do mercado, sendo este um dos maiores desafios dos investidores e analistas. Alguns investidores prosperaram por longos períodos e passaram pelo "teste do tempo", como Benjamin Graham (defensor do famoso Value Investing como estratégia de investimento e mentor de Warren Buffet), Peter Lynch que durante 13 anos (1977-1990) bateu todos os índices e fundos de investimento, Philip Arthur Fisher considerado o pai do investimento em crescimento, entre outros (Rankia Portugal, 2021).

Essa divergência entre a HME e a realidade empírica, impeliu a comunidade a estudar contrapontos relevantes à teoria, sendo uma delas a Hipótese dos Mercados Adap-

tativos (HMA), Lo (2004) argumenta que os mercados financeiros não são completamente eficientes, pois evoluem de acordo com o comportamento dos investidores e as mudanças externas, sugerindo que as estratégias adaptativas baseadas em modelos preditivos podem explorar ineficiências.

1.1 Motivações e lacunas

À medida que as tecnologias avançam, é de se esperar que sejam utilizadas novas abordagens para problemas complexos contemporâneos, como a busca por métodos de análise e previsão mais eficientes, robustos e adaptáveis, em especial num mercado complexo e volátil como o financeiro, que apresenta um desafio significativo para a previsão de tendências utilizando métodos convencionais. Como avaliado por Araujo and Gaglianone (2023), construir previsões precisas geralmente não é algo trivial, pois requer uma abordagem complexa o suficiente para incorporar variáveis relevantes, mas também focada em excluir dados irrelevantes. Nesse sentido, os métodos de aprendizado de máquina, em geral, são capazes de identificar padrões e relacionamentos dentro dos dados que podem ser ignorados por métodos tradicionais.

Dentro do escopo da Inteligência Artificial (IA), a *Recurrent Neural Network* (RNN), *Gated Recurrent Units* (GRU), *Convolutional Neural Network* (CNN) e *Long Short-Term Memory* (LSTM) emergem como alternativas avançadas capazes de superar as limitações das técnicas tradicionais de análise financeira. Segundo Fischer and Krauss (2018), as redes LSTM são especificamente projetadas para aprender dependências de longo prazo e são capazes de superar os problemas anteriormente inerentes às RNNs, ou seja, gradientes desaparecendo e explodindo, e que é possível demonstrar com sucesso que uma rede LSTM é capaz de extrair informações significativas de dados de séries temporais financeiras ruidosas. Estudos como Fischer and Krauss (2018), obtiveram um certo sucesso ao utilizar LSTM para prever movimentos direcionais do S&P 500, superando os métodos de classificação sem memória.

Este trabalho entende como lacuna e oportunidade de estudo o foco na utilização de aprendizado de máquina para extração de dias com alta probabilidade de serem positivos, uma vez que essa busca pode conferir uma vantagem competitiva relevante ao investidor. Neste contexto, a IA se apresenta como uma ferramenta poderosa para analisar sequências temporais de dados financeiros, potencializando essa identificação de padrões.

1.2 Objetivo e Questões de Pesquisa

A questão central deste trabalho está limitada em entender e mensurar o potencial e as limitações do uso de redes neurais recorrentes, especialmente as LSTM, para serem capazes de apontar com eficiência dias com alta probabilidade de retorno financeiro no mercado de ações brasileiro.

O crescente interesse em modelos de aprendizado profundo para análise financeira é evidenciado por estudos como o de Yoshinaga and Castro (2023), onde destaca que "a indústria financeira vem passando por uma transformação profunda com o surgimento das tecnologias de IA. A IA revolucionou diversos aspectos das finanças, incluindo estratégias de investimento, avaliação de riscos, tomada de decisões financeiras e finanças corporativas", também é destacado que "algoritmos de machine learning (aprendizado de máquina) podem processar dados históricos de mercado e identificar padrões que analistas humanos não conseguem captar. Esses algoritmos permitem desenvolver modelos para prever movimentos de mercado e auxiliar a otimização da alocação de ativos"(YOSHINAGA; CASTRO, 2023). Estas perspectivas ressaltam a relevância do uso de tecnologias avançadas de IA para superar os desafios impostos pela previsão de movimentos de mercado.

Este trabalho pesquisa não tem por objetivo obter uma alta taxa de acurácia no entendimento dos movimentos de preço, focará apenas na exploração de dias com sinais de alta probabilidade de subida do preço. Há estudos como Nelson (2017) que tem o foco na predição do período seguinte, mas não entram no mérito de filtrar apenas os melhores dias para o devido *trading system*

Será abordado a capacidade das RNRs, com maior ênfase nas LSTMs, para prever tendências curtas de mercado no contexto do ações brasileiras selecionadas. A seleção de sinais que apresentam maior confiabilidade é relevante para um sistema de trading eficaz. A implementação de filtros baseados em critérios rigorosos de confiança pode significativamente melhorar a performance de modelos preditivos, minimizando riscos e maximizando retornos potenciais. Essa seleção utilizará níveis elevados de confiança através de um corte mais rigoroso numa função de ativação.

1.3 Possibilidades e Hipóteses em Caso de Sucesso

A hipótese por trás deste estudo é demonstrar, em especial, que além da potencial captura de padrões complexos em dados de séries temporais do mercado, as redes neurais são capazes também de quantificar estatisticamente de alguma forma esse nível de certeza, possibilitando assim a extração de um modelo específico contendo apenas dias selecionados com sinais de melhor qualidade.

Em caso de sucesso, os resultados deste estudo poderão contribuir para a área de análise financeira, oferecendo uma nova abordagem para investidores e analistas no mercado de ações brasileiro.

1.4 Organização do Trabalho

O seguinte trabalho apresentará no capítulo 1 um referencial teórico que foi utilizado como embasamento. Em seguida, no capítulo 2, será descrito as diferentes fases do

experimento: desde a extração dos dados, seu tratamento, suas incorporações, e sua utilização como insumo para o treinamentos.

Após a fase de experimentos, os resultados serão organizados e analisados individualmente para cada ação no capítulo 3, a fim de extrair conclusões gerais e conclusões específicas com foco nos objetivos desse trabalho.

Por fim, no capítulo 4 haverá a conclusão onde os resultados serão analisados em seu todo, em termos de resultados alcançados e potencialidades dos modelo.

2 FUNDAMENTOS E REFERENCIAL TEÓRICO

Este capítulo se dedica a oferecer um panorama dos princípios fundamentais desse trabalho, explorando tanto os conceitos elementares do mercado de ações, das séries temporais, dos modelos de previsão, das métricas de avaliação e trabalhos relacionados, estabelecendo assim a base teórica para o modelo que será proposto.

2.1 Mercado Financeiro

O mercado financeiro é um setor onde se pode comprar e vender bens como títulos, moedas, ações, derivativos, mercadorias, commodities entre outros bens e ativos com algum valor financeiro, segundo Levine and Zervos (1996) o desenvolvimento do mercado de ações está positivamente associado ao crescimento econômico. Normalmente cada país possui o seu próprio ambiente financeiro, embora a negociação não seja restrita apenas de valores oriundos do seu mercado interno. O mercado de ações no Brasil, operado pela B3 (Brasil, Bolsa, Balcão), abriga uma variedade de empresas que vão desde multinacionais até médias e pequenas empresas com alto potencial de crescimento. Essa diversificação oferece um ambiente aos investidores com acesso a diferentes perfis de empresas, cada uma apresentando um conjunto específico de riscos e oportunidades de investimento.

Para o propósito deste trabalho, selecionou-se um conjunto de cinco ações por sua relevância e diversidade de segmento no mercado brasileiro, denominadas como “blue chips”, um termo que se refere a ativos de empresas com presença a longo prazo no mercado, sólido histórico de desempenho financeiro, estabilidade e pagamento regular de dividendos.

A tabela 1 apresenta essas ações, o percentual de participação no índice Ibovespa e seu segmento:

Tabela 1 – Ações escolhidas para esse trabalho. Participação referente à carteira do Ibovespa em 01/04/2024

Ação	Participação Ibovespa (%)	Segmento
VALE3	11,81	Mineração
PETR4	7,89	Petróleo
ITUB4	7,69	Bancário
WEGE3	2,62	Industrial
ELET3	3,82	Energia

O índice Ibovespa é composto por uma carteira teórica de ativos, selecionados com base em sua liquidez e representatividade. A composição deste índice é revisada

trimestralmente, para que permaneça atualizada e reflita as variações do mercado. As empresas incluídas no índice representam, em conjunto, a maior parte do volume negociado e do valor de mercado da bolsa.

2.2 Séries Temporais

Séries temporais são uma sequência de dados coletados em intervalos sequenciais específicos, normalmente uniformes, usados frequentemente com o objetivo de monitorar, modelar ou prever fenômenos ao longo do tempo. Tem por característica uma dependência entre os valores conexos, e no contexto financeiro, esses dados costumam incluir histórico de preços e volumes de negociação, acrescido de indicadores técnicos, que costumam ser funções matemáticas derivadas dos dados sequenciais de preço e volume. A importância das séries temporais na análise de mercados financeiros está na sua capacidade potencial de capturar a dinâmica temporal das variáveis, permitindo a modelagem e previsão de tendências de mercado, algo relevante para identificar potenciais oportunidades de investimento. (BOX *et al.*, 2016).

A Figura 1 ilustra um exemplo do histórico de preço e volume do índice Ibovespa.

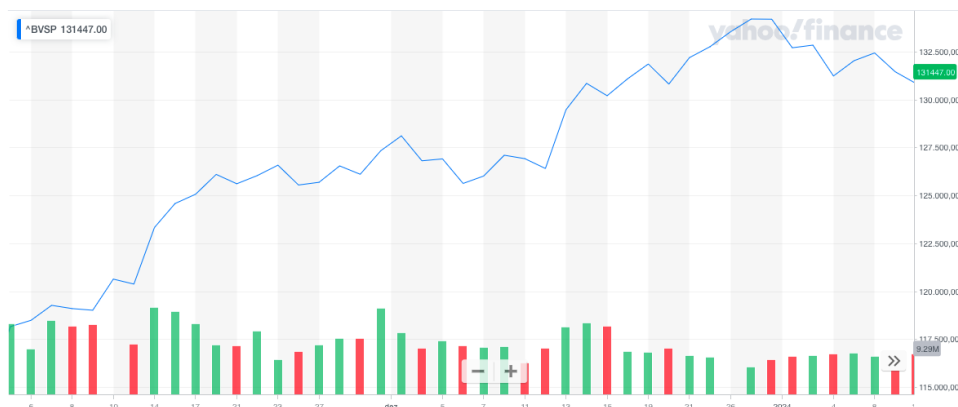


Figura 1 – Dados de preço e volume do índice Ibovespa ao longo de alguns meses do ano de 2023 e 2024.

2.2.1 Classificação

Quando se trata de análise de séries temporais, particularmente no contexto do mercado de ações, essas séries são frequentemente classificadas de acordo por três componentes: tendência, sazonalidade (ciclo) e a variação aleatória denominada de ruído (LATORRE; CARDOSO, 2001). A tendência é definida como a direção geral de longo prazo dos dados, ou seja, um movimento sustentado em direção ascendente ou descendente. A sazonalidade refere-se a padrões que se repetem em um período específico, como mensal e anual, e é comum a séries e dados econômicos e financeiros. O ruído é uma componente

estocástica de todas as séries temporais e é definido como erro, ou seja, as variações aleatórias que não podem ser explicadas pelos economistas.

A Figura 2 ilustra a evolução no preço ao longo do tempo com foco na tendência, nos ciclos, no ruído e suas combinações.

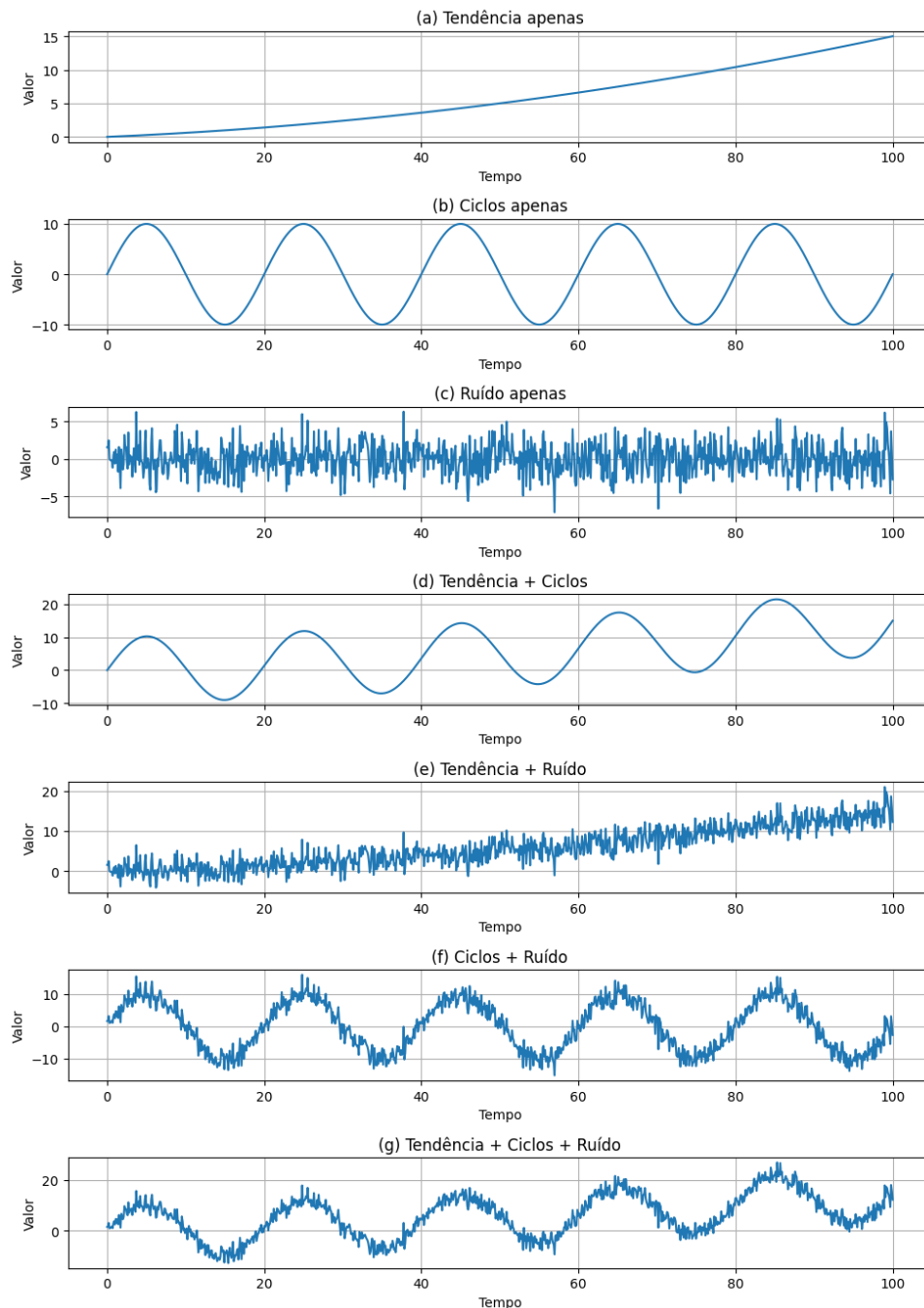


Figura 2 – Combinações de tendência, ciclo e ruído (Fonte: Elaborado pelo autor através de biblioteca *matplotlib.pyplot*).

A Figura 2(a) ilustra uma tendência isolada, ilustrando uma série temporal com uma clara progressão unidirecional ao longo do tempo, sem apresentar flutuações cíclicas ou ruído aleatório. Já a Figura 2(b) destaca ciclos isolados, exibindo uma série que

manifesta padrões cíclicos ou sazonais que se repetem em intervalos regulares, porém sem uma tendência de longo prazo ou interferências aleatórias. A Figura 2(c) representa exclusivamente ruído, caracterizando uma série totalmente aleatória, sem tendências ou ciclos, conhecida como ruído branco, onde cada ponto é independente dos demais.

A Figura 2(d) combina tendência e ciclos, apresentando uma série que não apenas segue uma direção ao longo do tempo, mas também exibe flutuações regulares. A Figura 2(e) une tendência e ruído, mostrando uma série com uma tendência evidente que é afetada por variações aleatórias, ilustrando como o ruído pode obscurecer uma tendência subjacente.

Por sua vez, a Figura 2(f) mescla ciclos e ruído, demonstrando ciclos regulares que são interrompidos por variações aleatórias, destacando como o ruído pode complicar a identificação de padrões cíclicos definidos. Finalmente, a Figura 2(g) exibe a combinação mais complexa: tendência, ciclos e ruído. Este gráfico ilustra a complexidade encontrada em dados reais, frequentemente observada em aplicações econômicas ou financeiras, onde todos esses elementos coexistem.

2.3 Modelos de Classificação

Na análise de séries temporais financeiras, especialmente na predição de movimentos de mercado no curto prazo, os modelos de classificação baseados em IA têm demonstrado eficácia significativa. Entre as arquiteturas de redes neurais, as RNN, CNN, LSTM, e GRU são frequentemente aplicadas devido à sua capacidade de processar sequências de dados temporais.

2.3.1 Redes Neurais Recorrentes

As Redes Neurais Recorrentes (RNRs) são um tipo de arquitetura de rede neural projetada para reconhecer padrões em sequências de dados, como séries temporais (SCHMIDT, 2019). Sua capacidade de reter informações de estados anteriores as torna ideais para prever eventos futuros em mercados voláteis, como o mercado de ações. As RNNs possuem loops em suas conexões, permitindo que informações de estados anteriores sejam passadas adiante, o que as torna ideais para tarefas que envolvem dependências temporais ou sequenciais.

A estrutura básica de uma RNN pode ser descrita por suas equações de estado e saída. Considerando x_t como o input no tempo t , h_t o estado oculto no mesmo instante, e y_t a saída, as equações fundamentais de uma RNN simples são:

$$h_t = \sigma(W_{hh}h_{t-1} + W_{xh}x_t + b_h) \quad (2.1)$$

$$y_t = W_{hy}h_t + b_y \quad (2.2)$$

Em que W_{hh} , W_{xh} , e W_{hy} representam, respectivamente, os pesos das conexões do estado oculto anterior para o atual, do input para o estado oculto, e do estado oculto para a saída; b_h e b_y são os termos de bias para o estado oculto e a saída, respectivamente; e σ denota a função de ativação, comumente uma sigmóide ou tanh.

As RNNs são capazes de capturar dependências temporais de diferentes escalas, tornando-as adequadas para uma ampla gama de aplicações, incluindo processamento de linguagem natural, previsão de séries temporais, e análise de sequência de genomas.

2.3.2 Convolutional Neural Network

A Rede Neural Convolutiva (CNN) é uma das redes mais significativas no campo da aprendizagem profunda. Como a CNN obteve conquistas impressionantes em muitas áreas, incluindo, entre outras, visão computacional e processamento de linguagem natural, atraiu muita atenção tanto da indústria quanto da academia nos últimos anos (LI *et al.*, 2021). A CNN é composta por várias camadas que processam a entrada de forma sequencial. Inicialmente, as camadas convolucionais utilizam filtros para extrair características locais relevantes da entrada. Esses filtros, ou conjuntos de pesos W , são aplicados através de operações de convolução que produzem mapas de características. A convolução para uma entrada unidimensional $x[t]$ com um filtro W de tamanho k é definida como:

$$C[t] = (W * x)[t] = \sum_{i=0}^{k-1} W[i] \cdot x[t - i]$$

Após a convolução, geralmente se utiliza uma função de ativação como a ReLU, $f(x) = \max(0, x)$, para introduzir não-linearidades ao modelo.

Segue-se com camadas de agrupamento, como a de Max Pooling, que reduz a dimensionalidade dos mapas de características ao manter apenas os valores máximos dentro de uma janela definida p . A operação de max pooling é dada por:

$$P[t] = \max(C[t \cdot p : (t + 1) \cdot p - 1])$$

Esta etapa ajuda a tornar a representação mais compacta e invariante a pequenos deslocamentos e distorções nos dados.

As camadas densas vêm após as camadas de convolução e de agrupamento. Os mapas de características são achatados e alimentados em camadas densas que combinam as características aprendidas para formar previsões. Os pesos nessas camadas são ajustados durante o treinamento para maximizar a precisão das previsões finais.

Consideremos a aplicação de uma CNN para analisar os preços de fechamento diários de ações. A entrada para a CNN seria uma série temporal dos preços, e a rede

aprenderia a identificar padrões que precedem aumentos ou quedas significativas nos preços. A partir dos dados de entrada, a rede pode prever, por exemplo, se o preço aumentará ou diminuirá no dia seguinte.

2.3.3 Long-Short Term Memory

Redes Long-Short Term Memory (LSTM) são um tipo avançado de redes neurais recorrentes, introduzidas por Hochreiter e Schmidhuber em 1997. Estas redes foram projetadas para superar o problema do desaparecimento do gradiente que ocorre com sequências longas em redes neurais recorrentes convencionais (HOUDT; MOSQUERA; NÁPOLES, 2020). A estrutura das LSTMs inclui células de memória equipadas com três portas: de entrada, de esquecimento e de saída. Essas portas regulam o fluxo de informações, permitindo que a rede retenha ou descarte dados de forma eficaz.

O comportamento interno de uma célula LSTM é determinado por um conjunto de equações que descrevem como as informações são processadas e transmitidas ao longo do tempo. As equações principais são:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (\text{porta de esquecimento})$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (\text{porta de entrada})$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (\text{candidato ao estado da célula})$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (\text{atualização do estado da célula})$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (\text{porta de saída})$$

$$h_t = o_t * \tanh(C_t) \quad (\text{estado oculto atual})$$

A porta de esquecimento decide quais informações anteriores serão descartadas. A porta de entrada e o candidato ao estado da célula determinam quais novos dados serão adicionados ao estado da célula. O estado da célula atual reflete a combinação de novas entradas com os dados retidos. A porta de saída e o estado oculto atual determinam a saída da célula LSTM no momento t . Estes mecanismos permitem à célula decidir de forma inteligente como integrar, armazenar e descartar informações ao longo do tempo, tornando as LSTMs ideais para tarefas que requerem memória de longo prazo e discernimento contextual.

2.3.4 Gated Recurrent Unit

A Gated Recurrent Unit (GRU) é uma abordagem que simplifica a estrutura de uma LSTM ao eliminar a porta de saída e combinar a porta de entrada com a porta de esquecimento numa única porta de atualização. Esta simplificação reduz a complexidade do modelo sem comprometer significativamente o desempenho (NOSOUHIAN; NOSOUHIAN;

KHOSHOUEI, 2021). Em uma GRU, existem dois componentes principais que controlam o fluxo de informações: a porta de atualização e a porta de reinicialização. A porta de atualização z_t é calculada a cada momento t pela função sigmóide, conforme expresso na fórmula:

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t])$$

Simultaneamente, a porta de reinicialização r_t determina quanto do passado será esquecido, usando uma fórmula similar.

O estado da memória h_t é atualizado por uma combinação ponderada entre o estado anterior e o estado candidato, que é modulado pela porta de reinicialização e transformado através de uma função tangente hiperbólica. Este mecanismo é expresso pela seguinte fórmula:

$$h_t = z_t \circ h_{t-1} + (1 - z_t) \circ \tanh(W_h \cdot [r_t \circ h_{t-1}, x_t])$$

Os pesos W_h são ajustados para transformar o estado candidato de acordo com a entrada e o estado modificado pela porta de reinicialização. As GRUs são particularmente eficazes para modelar séries temporais financeiras, como preços de ativos, onde é relevante capturar padrões complexos e dependências de longo prazo. Sua capacidade de manter informações relevantes ao longo do tempo e resistir ao problema do desvanecimento do gradiente as torna valiosas para análises preditivas em finanças. Utilizando séries temporais de preços de ativos $P[t]$, onde t varia de 1 até T e T representa o total de períodos observados, as GRUs podem efetivamente prever direções de preços, volatilidades e outras métricas financeiras importantes, adaptando-se dinamicamente às mudanças do mercado.

2.4 Métricas de avaliação

Avaliar a eficácia de modelos de classificação é importante para entender sua capacidade de generalização e precisão em tarefas de previsão. As métricas mais comuns incluem Acurácia, Precisão, Recall (Sensibilidade) e o F1-Score. Cada uma dessas métricas fornece perspectivas diferentes sobre o desempenho do modelo.

2.4.1 Acurácia

A acurácia é uma métrica fundamental para avaliar o desempenho de modelos de classificação, representando a proporção de previsões corretas (verdadeiros positivos e verdadeiros negativos) em relação ao total de previsões realizadas (HANDELMAN *et al.*, 2019). A fórmula da acurácia é dada por:

$$\text{Acurácia} = (TP + TN) / (TP + TN + FP + FN)$$

Em que o TP (Verdadeiros Positivos) e TN (Verdadeiros Negativos) são as previsões corretamente identificadas como positivas e negativas, respectivamente, enquanto FP

(Falsos Positivos) e FN (Falsos Negativos) representam as previsões incorretamente identificadas. Esta métrica é altamente valorizada por ser intuitiva e fácil de comunicar, especialmente útil em cenários onde as classes são equilibrada.

Apesar de suas vantagens, a acurácia possui limitações, particularmente em situações de desbalanceamento de classe. Por exemplo, em um conjunto de dados onde 95% das instâncias pertencem à classe A e apenas 5% à classe B, um modelo que prediz sempre a classe A atingiria uma acurácia de 95%, mesmo sem ter aprendido a classificar a classe B corretamente. Além disso, a acurácia não distingue entre tipos de erros, o que pode ser problemático em contextos onde os custos associados a falsos positivos e falsos negativos são diferentes.

Para ilustrar, considere um modelo de classificação que prediz se o preço de ações vai subir ou descer no próximo dia. Se testado com 100 previsões, onde 50 são corretas de subida (TP), 30 são corretas de descida (TN), 10 são subidas incorretas (FP) e 10 são descidas incorretas (FN), a acurácia seria:

$$Acurácia = (50 + 30)/(50 + 30 + 10 + 10) = 0.80$$

Este valor indica que o modelo foi capaz de prever corretamente o movimento do preço das ações 80% das vezes, refletindo um bom desempenho dependendo do contexto específico do problema e da natureza das classes envolvidas.

2.4.2 Precisão

A precisão, também conhecida como Valor Preditivo Positivo, é uma métrica utilizada em classificação que indica a proporção de identificações positivas que foram corretamente classificadas. Em outras palavras, é a probabilidade de que uma previsão positiva seja realmente positiva (HANDELMAN *et al.*, 2019). É calculada pela seguinte fórmula:

$$Precisão = TP/(TP + FP)$$

TP representa os Verdadeiros Positivos, ou instâncias corretamente identificadas como positivas, e FP representa os Falsos Positivos, ou instâncias erroneamente identificadas como positivas. A precisão é especialmente importante em contextos onde os Falsos Positivos têm um alto custo associado, como por exemplo, em testes de doenças raras onde um resultado falso positivo pode levar a tratamentos desnecessários e dispendiosos.

A importância da precisão se destaca em situações onde é preferível evitar um Falso Positivo, mesmo que isso possa resultar em um aumento no número de Falsos Negativos. Ela é particularmente eficaz em datasets onde uma das classes é muito menos comum do que a outra e ajuda a minimizar os custos associados a erros de classificação. No entanto, a precisão pode ser enganosa se o número de Falsos Positivos é baixo, e não considera os

Falsos Negativos, o que pode ser problemático em situações onde estes últimos têm um alto custo associado.

Considere, por exemplo, um modelo que prediz se transações financeiras são fraudulentas ou não. Suponha que, em um teste com 1000 transações, 50 foram corretamente identificadas como fraudulentas, 950 foram corretamente identificadas como não fraudulentas, 5 foram incorretamente identificadas como fraudulentas, e 20 como não fraudulentas. A precisão deste modelo seria de aproximadamente 90.9%, calculada pela seguinte fórmula:

$$\text{Precisão} = \frac{50}{50 + 5} = 50/55 \approx 0.909$$

indicando um alto grau de acerto nas identificações positivas.

2.4.3 Recall

Recall, também conhecido como Sensibilidade ou Taxa de Verdadeiros Positivos, é uma métrica essencial em modelos de classificação que indica a proporção de casos positivos reais corretamente identificados pelo modelo. É calculado pela seguinte fórmula:

$$\text{Recall} = TP / (TP + FN)$$

onde TP representa os Verdadeiros Positivos, ou casos corretamente identificados como positivos, e FN representa os Falsos Negativos, ou casos positivos que o modelo erroneamente identificou como negativos. O Recall é relevante em situações onde os Falsos Negativos carregam um alto custo, como na detecção de doenças graves, onde um diagnóstico perdido pode ser fatal.

Nessas condições, a maximização do Recall é prioritária para garantir que todos os casos positivos sejam detectados. Embora seja importante para segurança em áreas como medicina e segurança pública, onde um alto Recall significa que poucos casos positivos são perdidos, esta métrica pode sacrificar a Precisão ao aumentar o número de Falsos Positivos. Em contextos como o diagnóstico de doenças raras, considerando um sistema testado em 1000 pacientes, se 100 realmente têm a doença e 90 são corretamente diagnosticados, com apenas 10 não detectados pelo sistema, o Recall seria de 90%.

Essa alta taxa de Recall é benéfica, mas é importante também considerar outras métricas como Precisão e F1-Score para manter um equilíbrio adequado e evitar decisões baseadas em informações potencialmente enganosas. Em conclusão, Recall é uma métrica indispensável para avaliar a eficácia de modelos de classificação, especialmente em campos onde é vital capturar todos os casos positivos, apesar de que sua maximização pode reduzir a Precisão, é essencial para garantir que nenhum caso positivo seja omitido em situações críticas.

2.4.4 F1-Score

O F1-Score, que é a média harmônica da Precisão e do Recall, oferece um balanço eficaz entre essas duas métricas críticas. É uma função da taxa de verdadeiros positivos e do valor preditivo positivo para fornecer uma indicação geral do desempenho do classificador (HANDELMAN *et al.*, 2019).

Ele é particularmente útil em situações onde as classes são desbalanceadas ou quando os custos associados aos Falsos Positivos e Falsos Negativos são comparáveis. O F1-Score é calculado pela seguinte fórmula:

$$F1 - Score = 2 \times (Precisão \times Recall) / (Precisão + Recall)$$

que pode ser reescrita como:

$$F1 - Score = 2TP / (2TP + FP + FN)$$

onde TP representa os Verdadeiros Positivos, FP os Falsos Positivos, e FN os Falsos Negativos.

A importância do F1-Score reside na sua capacidade de avaliar a qualidade de modelos de classificação, fornecendo um equilíbrio entre Precisão e Recall. Isso é relevante em aplicações como modelos de classificação de doenças, onde minimizar tanto diagnósticos falsos quanto omissões é vital, e em sistemas de detecção de fraudes, onde é essencial equilibrar a captura de todas as transações fraudulentas com a minimização de alarmes falsos. O F1-Score é vantajoso por fornecer uma métrica balanceada que ajuda a reduzir o viés em favor de uma classe ou outra e por combinar dois aspectos importantes da performance do modelo, tornando-se útil para comparar modelos e para ajustes finos.

No entanto, o F1-Score tem limitações. Ele presume igual importância para Precisão e Recall, o que pode não ser adequado para todos os contextos de aplicação, e pode ser menos intuitivo de entender e comunicar para não especialistas, comparado a métricas como a Acurácia. Apesar desses desafios, o F1-Score é uma ferramenta valiosa para garantir que um modelo de classificação não seja tendencioso em direção a uma classe específica, especialmente em situações críticas onde os tipos de erros têm implicações significativas.

2.4.5 Definição de oportunidade para esse trabalho

Embora a função de ativação sigmoide seja tradicionalmente utilizada em problemas de classificação binária devido à sua capacidade de mapear saídas variadas para um intervalo estrito entre 0 e 1, representando a probabilidade de pertencimento a uma das duas classes, em nosso estudo, entretanto, buscamos uma abordagem diferenciada, focando não na classificação binária tradicional, mas na identificação de sinais que estejam mais próximos de 1, indicando uma alta probabilidade de ganho para o período seguinte.

O foco que prioriza valores próximos de 1 significa que será priorizado a precisão em detrimento da quantidade de sinais de negociação gerados. Essa estratégia visa maximizar o retorno sobre os sinais identificados, mesmo que isso resulte em um número muito menor de oportunidades de negociação. A premissa considerada é que ao filtrar apenas aqueles sinais com maior confiança, o modelo possa oferecer previsões de melhor qualidade, embora em menor quantidade. O desempenho geral do modelo será avaliado com base em métricas tradicionais de acurácia, precisão, recall, F1-Score. O desempenho específico do modelo será avaliado principalmente em razão de sua precisão.

2.5 Trabalhos relacionados

Na literatura possui uma variedade de estudos sobre rede neurais e aprendizado de máquina aplicado ao mercado financeiro, aplicando metodologias diferentes e focando em domínios diversos, ao longo do tempo. A seguir, é dada uma breve descrição de alguns dos estudos relevantes, todos contribuindo para o corpo de conhecimento disponível.

Kamijo and Tanigawa (1990), e, em seguida, Wang and Leu (1996) são os primeiros a elaborar redes neurais recorrentes para dados de preços e indicadores técnicos, respectivamente, e estes estudos descrevem a evolução dos atributos de modelagem e um sistema recorrente para respectivos dados do mercado de ações. Allen and Karjalainen (1999) em seu trabalho sobre algoritmos genéticos recorre a um domínio muito mais amplo, permitindo fazer a primeira avaliação das perspectivas de mecanização nas estratégias de negociação. Este trabalho, portanto, define um novo caminho para melhor entender a aplicação de habilidades computacionais no sistema financeiro.

Kim (2003) considera a utilização de máquinas de vetor de suporte para análise de dados financeiros e aborda o uso de métodos semi-automatizados para explorar grandes volumes de informações e tomar decisões informadas. Pommerenzenbaum (2014) usa um modelo preditivo, que se concentra no preço diário do Ibovespa na bolsa de valores brasileira, com uma análise do comportamento histórico de várias bolsas de valores, servindo de dados de treinamento.

Bebarta, Biswal and Dash (2015) utilizam a *Polynomial Based Functional Link Artificial Recurrent Neural Network (PFLARNN)* para prever preços de ações, mostrando que o modelo otimizado com *evolução diferencial (DE)* supera métodos tradicionais como *ARMA* e outras redes neurais, como *MLP* e *RBFNN*. Yoshihara *et al.* (2014) exploram o uso de RNN para modelar os efeitos temporais de eventos passados e prever tendências de mercado, demonstrando a validade da abordagem em dados reais de empresas da *Nikkei*.

Jaybhay, Argiddi and Apte (2012) apresentam modelos que combinam *mineração de dados* de preços e textos de notícias para prever o preço de fechamento do mercado de ações. Eles utilizam uma *rede neural feed-forward* treinada com o *algoritmo de retropropagação*,

integrando cotações de ações e frases-chave extraídas de notícias financeiras, melhorando a precisão das previsões.

Persio, Honchar *et al.* (2016) comparam diferentes arquiteturas de *redes neurais* para prever movimentos de preços de mercado. Eles investigam *perceptrons multicamadas (MLP)*, CNN e LSTM, concluindo que a combinação de *wavelets* com CNN supera outras abordagens, destacando a importância da seleção e pré-processamento corretos das características de entrada.

Nos últimos anos, o desenvolvimento de técnicas mais avançadas tem ampliado as fronteiras do conhecimento em análise financeira. Bichara (2019) explorou o uso de redes neurais profundas para prever movimentos do mercado de ações no Brasil, destacando a superioridade dessas técnicas em relação aos modelos tradicionais. Liu *et al.* (2018) investigaram a aplicação de aprendizado profundo em conjunto com análise de sentimentos extraídos de notícias financeiras, demonstrando que tais modelos podem significativamente aumentar a precisão das previsões de mercado.

Esses estudos fornecem um panorama da evolução do uso de técnicas de aprendizado de máquina e redes neurais no mercado financeiro. Desde os primeiros esforços com redes neurais recorrentes para analisar preços de ações e indicadores técnicos, até a utilização de notícias e sentimentos, a literatura mostra um avanço contínuo, e forneceram uma base teórica para embasar esse trabalho.

3 PROPOSTA E DESENVOLVIMENTO

Este capítulo se dedica a apresentar todos os processos da proposta do experimento em todas as suas fases, desde a obtenção dos dados, o tratamento, a incorporação, a análise de características, configuração experimental e resultados.

3.1 Obtenção dos dados

As séries temporais foram obtidas por meio da plataforma *TradingView*, que é uma plataforma de análise e gráficos financeiros amplamente utilizada por *traders* e investidores para acompanhar os mercados financeiros em tempo real. Fundada em 2011, a plataforma se destaca por oferecer uma variedade de ferramentas de desenho técnico, indicadores de mercado e opções de personalização que auxiliam na análise técnica e na tomada de decisões de investimento. Além disso permite a exportação dos dados em formato *Comma-Separated Value* (CSV), que trata-se de valores separados por vírgula, que podem ser usadas como entrada em programas python para ser devidamente tratada e processada.

Os dados foram selecionados no tempo gráfico diário, a partir de janeiro de 2017 até fevereiro de 2024, o que permite uma análise detalhada das tendências do mercado ao longo de um período significativo, mas evita também que o período seja excessivamente longo, pois poderia ocasionar perda de qualidade em virtude da ação inflacionária. A escolha do intervalo diário deve-se à necessidade de equilibrar detalhe e abrangência temporal, permitindo identificar padrões de médio e longo prazo sem perder a sensibilidade a eventos específicos de mercado. A Tabela 2 apresenta a granularidade e o período dos dados definidos para esse trabalho.

Tabela 2 – Dados extraídos da plataforma TradingView para insumo desse trabalho

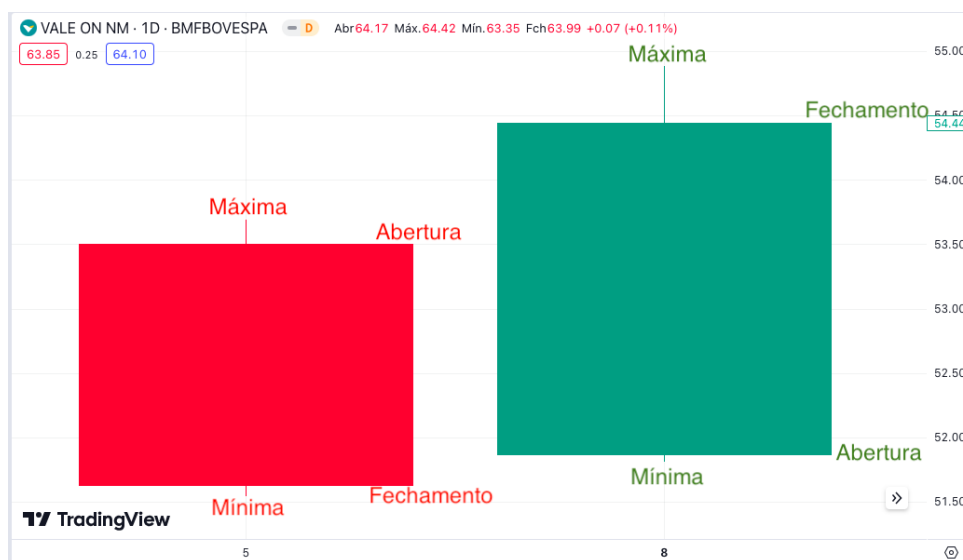
Ticker	Granularidade	Data Início	Data Fim
VALE3	Diário	02/02/2017	08/02/2024
PETR4	Diário	02/02/2017	08/02/2024
ITUB4	Diário	02/02/2017	08/02/2024
WEGE3	Diário	02/02/2017	08/02/2024
ELET3	Diário	02/02/2017	08/02/2024
IBOV	Diário	02/02/2017	08/02/2024
USDBRL	Diário	02/02/2017	08/02/2024

Fonte: TradingView (2024)

Para cada registro na granularidade diária, além da devida data há mais quatro informações: valor de abertura, valor mínimo, valor máximo e valor de fechamento. A Figura 3 ilustra um exemplo real de cotação diária da VALE3, sendo que em um dia o

valor de fechamento foi menor que o valor de abertura (candle vermelho), e no outro dia o valor de fechamento foi superior ao valor de abertura (candle verde).

Figura 3 – Marcadores de mínima, máxima, abertura e fechamento representando em candle stick



Fonte: TradingView (2024)

No exemplo ilustrado pela Figura 3, no primeiro candle que representa o dia negativo, a abertura ocorreu a R\$53,50, o fechamento a R\$51,59, a mínima foi R\$51,54 e máxima R\$53,69. No segundo candle, que representa o dia positivo, a abertura ocorreu a R\$51,89, o fechamento a R\$54,44, a mínima foi R\$51,84 e máxima R\$54,81.

3.2 Pré processamento dos dados

Dados de séries temporais das empresas citadas na Tabela 2 são o foco deste estudo. Duas variáveis econômicas foram incorporadas duas variáveis econômicas que influenciam diretamente o mercado de ações. Especificamente, a variação percentual diária do dólar (em relação ao real) e a variação do índice Bovespa (Ibovespa). A inclusão desses indicadores é relevante, pois a variação cambial pode afetar diretamente o valor das ações das empresas com operações significativas no exterior, enquanto o Ibovespa oferece uma visão abrangente do desempenho do mercado acionário brasileiro. A utilização desses dados adicionais visa enriquecer as séries temporais com dados mais amplos da economia, permitindo uma compreensão mais completa das forças que moldam os mercados financeiros. A Tabela 3 apresenta amostras com dados da VALE3 após incorporação da variação percentual diária do dólar (em relação ao real) e do índice Ibovespa.

Tabela 3 – Dados das ações após incorporações

Data	Abertura	Fechamento	Máxima	Mínima	Dólar	Ibovespa
23/03/2023	R\$ 76,78	R\$ 74,95	R\$ 77,40	R\$ 74,23	-0,09%	-2,29%
24/03/2023	R\$ 75,06	R\$ 74,10	R\$ 75,22	R\$ 73,55	1,16%	0,92%
27/03/2023	R\$ 74,89	R\$ 73,97	R\$ 75,42	R\$ 73,49	-0,97%	0,85%
28/03/2023	R\$ 74,14	R\$ 74,75	R\$ 75,25	R\$ 74,02	-0,92%	1,52%
29/03/2023	R\$ 75,34	R\$ 75,83	R\$ 76,14	R\$ 75,20	-0,63%	0,60%

Para realizar normalização dos dados, em análises financeiras, especialmente aquelas que envolvem aprendizado de máquina, o pré-processamento dos dados é importante para o desempenho dos modelos. Para dados de abertura, mínima, máxima e fechamento de ações, a transformação logarítmica é frequentemente aplicada para estabilizar a variância e normalizar a distribuição dos retornos financeiros. Conforme descrito por Tsay (2010), a transformação logarítmica é útil para modelar séries temporais financeiras, pois ela ajuda a reduzir as variações inconsistentes nos dados, o que é comum em séries de preços financeiros.

Para dados agregados como variação do dólar e variação do Ibovespa, a normalização usando a média e o desvio padrão é geralmente preferida por uma melhor uniformidade dos dados. Esta técnica ajusta os dados para ter uma média zero e desvio padrão de um, o que é particularmente útil para dados que são combinados de várias fontes e possuem diferentes escalas. Segundo Pedregosa *et al.* (2011), essa forma de normalização ajuda na convergência de algoritmos que são sensíveis à escala dos dados, como é típico em muitos algoritmos de aprendizado de máquina. A Tabela 4 apresenta o mesmo conjunto de amostras da séries temporais da Tabela 3, porém com os dados devidamente normalizados.

Tabela 4 – Dados das ações após normalizações

Data	Abertura	Fechamento	Máxima	Mínima	Dólar	Ibovespa
23/03/2023	3.5775	3.0210	3.5775	3.0906	-2.2892	-0.0916
24/03/2023	3.1999	3.2794	3.0806	3.2098	0.9220	1.1573
27/03/2023	3.2695	3.3788	3.2396	3.3191	0.8512	-0.9742
28/03/2023	3.3092	3.3986	3.2198	3.3589	1.5196	-0.9151
29/03/2023	3.3986	3.4384	3.1900	3.3291	0.6003	-0.6292

De acordo com a Tabela 4, foi exposto um conjunto de exemplos para demonstrar os efeitos da normalização logarítmica para os dados de abertura, fechamento, máxima e mínima, e normalização usando média zero e desvio padrão um para os dados agregados de dólar e Ibovespa.

Para este trabalho é proposto a criação de uma variável binária, que servirá como o alvo (*target*) para o nosso modelo preditivo. O propósito desta variável é prever a direção

do mercado no dia subsequente, baseando-se na comparação entre os preços de fechamento e de abertura das ações.

Considere uma série temporal $S = \{s_1, s_2, \dots, s_n\}$, em que s_i representa a i ésima amostra da série. A série temporal pode ser considerada multivariada, isto é, cada amostra s_i^j pode representar um conjunto de atributos j . Neste trabalho, os atributos j representam os valores de abertura (a), fechamento (f), entre outros apresentados na Tabela 4. A amostra s_i^t , em que t representa uma variável definido como 0 ou 1, representando um dia posterior negativo ou positivo, respectivamente. Um valor 0 indica que o preço de fechamento do dia seguinte é menor do que o preço de abertura do dia seguinte, sugerindo uma performance negativa do mercado nesse dia. Inversamente, um valor de 1 indica que o preço de fechamento do dia seguinte é maior do que o preço de abertura do dia seguinte, refletindo um dia positivo no mercado de ações. A variável pode ser formalmente definida como:

$$d_t = \begin{cases} 1 & \text{se } s_{i+1}^f > s_{i+1}^a \\ 0 & \text{se } s_{i+1}^f \leq s_{i+1}^a \end{cases} \quad (3.1)$$

Onde s_{i+1}^f representa o preço de fechamento do dia seguinte e s_{i+1}^a representa o preço de abertura (a) do dia seguinte ($i+1$). Dessa forma, a variável d_t nos permite identificar se o mercado teve um desempenho positivo ou negativo no dia subsequente.

3.3 Configuração experimental

Nesta seção, é apresentado a configuração experimental dos modelos de aprendizado de máquina utilizados no estudo, a estrutura, os parâmetros de compilação, a preparação dos dados e os detalhes do treinamento. A seguir, será descrito as configurações dos modelos Long Short-Term Memory (LSTM), RNN, Convolutional Neural Network (CNN) e Gated Recurrent Unit (GRU).

Segue detalhes dos parâmetros definidos:

- **Configuração Comum:** Utilizamos o otimizador *adam*, a função de perda *binary_crossentropy*, 200 épocas de treinamento e um tamanho de lote de 256. Os dados de entrada (X) excluíram a última coluna (*Candle Status*) e o valor alvo (y) correspondia ao status do candle no último dia da janela. A janela de entrada (*window_size*) foi definida com tamanho 3. A camada densa final foi configurada com uma unidade de saída e função de ativação *sigmoid*.
- **Long Short-Term Memory (LSTM):** O modelo LSTM foi configurado com uma camada LSTM contendo 80 unidades e função de ativação *tanh*.

- RNN: O modelo RNN foi estruturado com uma camada *SimpleRNN* de 80 unidades e função de ativação *tanh*, seguido por uma camada densa com uma unidade de saída e função de ativação *sigmoid*.
- **Convolutional Neural Network (CNN)**: O modelo CNN foi implementado com uma camada *Conv1D* contendo 64 filtros, função de ativação *relu* e tamanho de kernel de 3, seguida por uma camada *Flatten* para achatar a entrada, e uma camada densa com 50 unidades e função de ativação *relu*.
- **Gated Recurrent Unit (GRU)**: O modelo GRU incluiu uma camada *GRU* com 50 unidades, configurada para retornar sequências (*return_sequences=True*) e *input_shape* de (*window_size, X_train.shape[2]*). Esta camada foi seguida por uma camada *Flatten* para achatar a entrada, e uma camada densa com 50 unidades e função de ativação *relu*.

Após definir as configurações específicas de cada modelo, foi realizado o treinamento utilizando os dados históricos de preços de ações, no tempo gráfico diário. A combinação da função de ativação *tanh* nas camadas internas com a função *sigmoid* na camada de saída permite que os modelos capturem as dependências temporais e forneçam previsões probabilísticas adequadas para classificação binária.

Em projetos de aprendizado de máquina, a divisão dos dados em conjuntos de treino e teste é um passo para garantir a validade e a generalização do modelo. A divisão adequada dos dados permite avaliar a capacidade do modelo de generalizar para novos dados, evitando *overfitting* e *underfitting*.

Para esse trabalho optou-se pela proporção 80% treino e 20% teste, pois entende-se oferecer um equilíbrio eficiente entre a quantidade de dados utilizados para treinar o modelo e a necessidade de avaliar sua capacidade de generalização.

Tabela 5 – Contraste percentual de dias positivos e negativos para cada ação, divididos em treinamento e teste.

Ação	Treinamento (%)		Teste (%)	
	Dias Negativos	Dias Positivos	Dias Negativos	Dias Positivos
VALE3	52.20%	47.80%	56.09%	43.91%
PETR4	51.00%	49.00%	47.86%	52.14%
ITUB4	52.34%	47.66%	50.42%	49.58%
ELET3	55.61%	44.39%	52.12%	47.88%
WEGE3	53.05%	46.95%	51.27%	48.73%

Como observado Tabela 5, a distribuição de dias negativos e positivos é bastante similar entre os conjuntos de treinamento e teste. Essa similaridade garante que o modelo

não será influenciado por um desequilíbrio significativo nas classes, permitindo uma avaliação mais justa e precisa de seu desempenho.

3.4 Resultados

A seguir são demonstrado avaliações gerais dos modelos em suas capacidades preditivas, e avaliações específicas com foco na busca por maior precisão.

3.4.1 Avaliação geral do modelos

Nesta seção, será apresentado uma avaliação geral dos modelos em termos gerais de suas capacidades preditivas. A função sigmoide é utilizada com um corte de 0.5, que é o padrão em classificações binárias, para determinar as predições.

Como pode ser observado na Tabela 6 (VALE3), a precisão e a acurácia mais alta foi alcançada com LSTM, com 0,54% e 0,57%, respectivamente, já os melhores indicadores de Recall e F1-Score ficaram com CNN.

Tabela 6 – Comparação de Desempenho de Modelos de Aprendizado de Máquina - VALE3

Modelo	Precisão	Acurácia	Recall	F1-Score
LSTM	0,54	0,57	0,25	0,34
RNN	0,50	0,56	0,27	0,35
CNN	0,50	0,56	0,45	0,47
GRU	0,47	0,54	0,37	0,41

Segundo a Tabela 7 (PETR4), a precisão e a acurácia mais alta também foram alcançadas com LSTM, com 0,57% e 0,57%, respectivamente, já os melhores indicadores de Recall e F1-Score ficaram com RNN.

Tabela 7 – Comparação de Desempenho de Modelos de Aprendizado de Máquina - PETR4

Modelo	Precisão	Acurácia	Recall	F1-Score
LSTM	0,57	0,53	0,36	0,44
RNN	0,56	0,53	0,40	0,47
CNN	0,55	0,52	0,35	0,42
GRU	0,55	0,51	0,26	0,36

Como indicado na Tabela 8 (ITUB4), a precisão mais alta foi alcançada com LSTM 0,49%, porém a melhor acurácia ficou com RNN (0,5%). O melhor recall e F1-Score ficaram com RNN.

De acordo com Tabela 9 (WEG3), o melhor desempenho para precisão e acurácia ficaram com CNN (0,57% e 0,53%, respectivamente) e o melhor desempenho para recall e F1-Score ficaram com LSTM.

Tabela 8 – Comparação de Desempenho de Modelos de Aprendizado de Máquina - ITUB4

Modelo	Precisão	Acurácia	Recall	F1-Score
LSTM	0,49	0,49	0,28	0,35
RNN	0,49	0,50	0,28	0,36
CNN	0,45	0,47	0,25	0,32
GRU	0,44	0,48	0,17	0,24

Tabela 9 – Comparação de Desempenho de Modelos de Aprendizado de Máquina - WEG3

Modelo	Precisão	Acurácia	Recall	F1-Score
LSTM	0,50	0,51	0,34	0,41
RNN	0,52	0,52	0,31	0,39
CNN	0,57	0,53	0,17	0,26
GRU	0,48	0,50	0,23	0,32

De acordo com a Tabela 10 (ELET3), o melhor desempenho para precisão e acurácia ficaram com RNN (0,57% e 0,53%, respectivamente) e o melhor desempenho para recall e F1-Score ficaram com CNN.

Tabela 10 – Comparação de Desempenho de Modelos de Aprendizado de Máquina - ELET3

Modelo	Precisão	Acurácia	Recall	F1-Score
LSTM	0,48	0,51	0,19	0,27
RNN	0,55	0,54	0,25	0,34
CNN	0,52	0,53	0,46	0,48
GRU	0,52	0,53	0,28	0,36

Segue a Tabela 11 com uma compilação da quantidade de vezes que cada modelo foi superior aos demais, conforme pode ser observado o modelo que melhor performou na precisão foi o LSTM, e na acurácia LSTM e RNN empataram, no quesito recall o LSTM foi o que melhor performou e no quesito F1-Score foi o RNN.

Tabela 11 – Quantidade de vezes em que o modelo teve desempenho superior aos demais

Modelo	Precisão	Acurácia	Recall	F1-Score
LSTM	3	2	2	1
RNN	1	2	1	2
CNN	1	1	1	1
GRU	0	0	1	1

Observamos que o modelo LSTM frequentemente obteve a maior precisão e acurácia, enquanto os melhores indicadores de recall e F1-Score variaram entre os modelos, dependendo do conjunto de dados específico. A tabela de compilação final mostra que o LSTM foi o modelo mais frequentemente superior em precisão, empatou com o RNN

em acurácia, e teve um bom desempenho em recall. O RNN, por sua vez, destacou-se no F1-Score. Isso indica que, embora o LSTM seja geralmente confiável para precisão e acurácia, o RNN pode ser mais vantajoso em termos de equilíbrio entre precisão e recall.

O principal propósito desse trabalho é encontrar dias com boas oportunidades de negociação, e o principal indicador para tal é a precisão do modelo. Na tentativa de melhorar a precisão, optou-se por aumentar o corte sigmoide para avaliar as mudanças percentuais de precisão, a cada sigmoide mais alta é de esperar que a quantidade de sinais diminua, criando uma relação inversa (quanto maior o corte da sigmoide, menor a quantidade de sinais).

Pode ser observado na Tabela 12 uma tendência geral de melhora da precisão à medida que o corte sigmoide aumenta para os modelos, à exceção foi para o modelo GRU que apresentou precisão superior a 0.70 do que 0.75 ou 0.80, mostrando que apesar de apresentarem uma tendência geral, pode haver exceções.

Tabela 12 – Desempenho dos Modelos para diferentes cortes de Sigmoide - quantidade e precisão - VALE3

Sigmoide	LSTM		RNN		CNN		GRU	
	Qtd	Precisão	Qtd	Precisão	Qtd	Precisão	Qtd	Precisão
0,50	72	54%	83	50%	140	50%	121	47%
0,55	33	60%	41	48%	85	51%	75	45%
0,60	13	76%	15	60%	41	63%	46	56%
0,65	6	83%	6	60%	26	65%	21	66%
0,70	1	100%	2	100%	16	75%	12	83%
0,75	0	-	0	-	9	77%	9	77%
0,80	0	-	0	-	5	80%	4	75%

Na Tabela 13 (PETR) pode ser observado a mesma tendência geral para LSTM e RNN, para CNN e GRU o sistema apresentou melhora até certo ponto.

Tabela 13 – Desempenho dos Modelos para diferentes cortes de Sigmoide - quantidade e precisão - PETR4

Sigmoide	LSTM		RNN		CNN		GRU	
	Qtd	Precisão	Qtd	Precisão	Qtd	Precisão	Qtd	Precisão
0,50	114	57%	129	56%	114	55%	86	55%
0,55	75	62%	78	60%	90	55%	70	55%
0,60	46	65%	42	57%	64	57%	48	56%
0,65	24	70%	12	57%	39	56%	30	58%
0,70	8	62%	6	66%	21	52%	17	58%
0,75	3	66%	0	-	0	0%	8	50%
0,80	1	100%	0	-	0	0%	1	0%

Foi observado na Tabela 14 (ITUB), a tendência de melhora se manteve também para LSTM e RNN; nesse caso GRU também obteve uma precisão maior, CNN melhorou até certo ponto.

Tabela 14 – Desempenho dos Modelos para diferentes cortes de Sigmoides - quantidade e precisão - ITUB4

Sigmoides	LSTM		RNN		CNN		GRU	
	Qtd	Precisão	Qtd	Precisão	Qtd	Precisão	Qtd	Precisão
0,50	100	49%	101	49%	99	45%	70	44%
0,55	63	46%	55	51%	62	45%	55	47%
0,60	29	51%	27	56%	41	46%	39	44%
0,65	14	50%	9	56%	25	36%	25	48%
0,70	5	60%	3	66%	13	30%	15	46%
0,75	1	100%	0	-	6	0%	8	50%
0,80	0	0%	0	0%	0	0%	2	50%

Na Tabela 15 (WEG3), além de não haver uma melhora, houve uma queda da precisão, contrariando os resultados anteriores.

Tabela 15 – Desempenho dos Modelos para diferentes cortes de Sigmoides - quantidade e precisão - WEG3

Sigmoides	LSTM		RNN		CNN		GRU	
	Qtd	Precisão	Qtd	Precisão	Qtd	Precisão	Qtd	Precisão
0,50	118	50%	105	52%	52	57%	84	48%
0,55	79	45%	61	49%	24	54%	60	50%
0,60	70	45%	26	50%	11	54%	40	45%
0,65	32	46%	7	50%	5	40%	19	36%
0,70	16	50%	2	0%	3	33%	8	38%
0,75	0	-	2	-	2	0%	5	40%
0,80	0	-	2	-	1	0%	3	0%

Como indicado na Tabela 16 (ELET3), pode ser observado a mesma tendência da WEG3 de não haver melhorias significativas da precisão.

Tabela 16 – Desempenho dos Modelos para diferentes cortes de Sigmoides - quantidade e precisão - ELET3

Sigmoides	LSTM		RNN		CNN		GRU	
	Qtd	Precisão	Qtd	Precisão	Qtd	Precisão	Qtd	Precisão
0,50	68	48%	77	55%	150	52%	92	52%
0,55	33	54%	24	50%	83	49%	61	54%
0,60	18	50%	10	50%	42	47%	38	55%
0,65	10	33%	0	50%	25	36%	28	50%
0,70	4	0%	0	0%	14	42%	17	47%
0,75	4	-	1	-	8	50%	11	54%

Conforme observado nas tabelas, foi avaliada a precisão dos modelos utilizando diferentes cortes sigmoides. O propósito foi identificar a configuração ideal que maximiza a precisão dos modelos na previsão de oportunidades de negociação. Observou-se uma tendência geral de melhora na precisão para a maioria dos modelos com o aumento do corte sigmoide, embora algumas exceções tenham sido notadas.

Os resultados apresentados demonstram que, ao aumentar o corte sigmoide, geralmente há uma tendência de melhora na precisão dos modelos. Esta tendência foi observada consistentemente para os modelos LSTM e RNN nos conjuntos de dados VALE3, PETR4 e ITUB4, para os modelos CNN e GRU observa-se também uma tendência de melhora, mas ocorreu até certo ponto.

Nos conjuntos de dados WEG3 e ELET3, a precisão dos modelos não apresentou melhoras significativas ou até diminuiu, contrariando a tendência geral observada nos outros conjuntos de dados. Estes resultados sugerem que, enquanto o aumento do corte sigmoide pode ser benéfico para a precisão em alguns contextos, pode não ser eficiente para outros conjuntos de dados.

A Tabela 17 demonstra qual foi o corte sigmoide que trouxe a melhor precisão para o modelo e cada ação:

Tabela 17 – Corte sigmoide que alcançou o melhor resultado em termos de precisão

Ação	LSTM	RNN	CNN	GRU
VALE3	0.7	0.7	0.8	0.7
PETR4	0.8	0.7	0.6	0.7
ITUB4	0.75	0.7	0.6	0.8
WEG3	0.5	0.5	0.5	0.55
ELET3	0.55	0.5	0.5	0.6

É importante destacar que encontrar o melhor corte sigmoide é um desafio, pois à medida que aumentamos o corte, a quantidade de sinais disponíveis diminui, o que diminui o aproveitamento prático do modelo. Embora os modelos apresentados sejam úteis

para identificar oportunidades em alguns conjuntos de dados, suas limitações devem ser consideradas ao aplicá-los em diferentes contextos.

4 CONCLUSÃO

Este trabalho foi realizado no contexto de extrair algum conhecimento estatístico a partir da análise de séries temporais financeiras, com foco específico em cinco ações negociadas na Bovespa e utilizando-se de aprendizado de máquina. Qualquer investidor que pretenda realizar a compra de ações por um curto período, pode extrair nos resultados apresentados algum conhecimento que dê suporte às suas estratégias.

O objetivo do trabalho foi encontrar algum padrão e estratégia capaz de apontar antecipadamente algumas oportunidades no mercado financeiro, que no contexto desse trabalho, se definem como a capacidade de apontar de antemão alguns dias com alta probabilidade de serem dias positivos (preço de fechamento maior do que o preço de abertura). Essa estratégia, por sua natureza, priorizou qualidade a quantidade, focando a busca pelo maior nível de certeza possível.

O experimento realizado incluiu uma fase de treinamento dos dados para modelos de aprendizado de máquina tentarem extrair padrões de como seria o dia seguinte positivo, e uma segunda fase de usar uma estratégia para extrair do modelo somente os dias com maior grau de certeza.

Foram utilizados dados históricos como preços de abertura, fechamento, máximo e mínimo, também foram incluídos variáveis macroeconômicas para os dados de treinamento, como a variação do índice Ibovespa e valor do dólar e utilizado modelos como LSTM, RNN, CNN e GRU. Para refinar ao máximo possível essa busca por dias positivos, foi proposto uma alteração na função de ativação, ao contrário do tradicional ponto de corte padrão, foi proposta e realizado o teste com pontos de corte acima desse padrão, para se focar nos sinais em que havia um nível maior de confiança do modelo.

Como ponto positivo destaca-se que foi possível demonstrar, em parte das amostras, uma capacidade do modelo aumentar consideravelmente o nível de precisão. Como ponto negativo, pode-se destacar que em parte das amostras não houve nenhuma melhora significativa, e mesmos nas amostras em que foi possível demonstrar esse aumento da precisão, destaca-se que o número de sinais diminuíram bastante.

Como o foco maior desse trabalho esteve em alcançar o máximo de precisão possível, não foi realizado um sistema de *trading system* para avaliar o desempenho do modelo, e também não está totalmente claro o motivo do sistema ter sido bastante eficiente para algumas ações, porém modestos em outras, algumas hipóteses podem ser levantadas, como volume de negociação, segmento ou outras variáveis desconsideradas.

REFERÊNCIAS

- ALLEN, F.; KARJALAINEN, R. Using genetic algorithms to find technical trading rules. **Journal of financial Economics**, Elsevier, v. 51, n. 2, p. 245–271, 1999.
- ARAÚJO, G. S.; GAGLIANONE, W. P. Machine learning methods for inflation forecasting in brazil: New contenders versus classical models. **Latin American Journal of Central Banking**, Elsevier, v. 4, n. 2, p. 100087, 2023.
- BEBARTA, D.; BISWAL, B.; DASH, P. Polynomial based functional link artificial recurrent neural network adaptive system for predicting indian stocks. **International Journal of Computational Intelligence Systems**, v. 8, 11 2015.
- BICHARA, G. L. G. Redes neurais profundas para auxílio à tomada de decisão no mercado de ações. Universidade Federal do Rio de Janeiro, 2019.
- BOND, P.; EDMANS, A.; GOLDSTEIN, I. The real effects of financial markets. **Annu. Rev. Financ. Econ.**, Annual Reviews, v. 4, n. 1, p. 339–360, 2012.
- BOX, G. E. *et al.* **Time Series Analysis: Forecasting and Control**. 5th. ed. Hoboken, NJ: Wiley, 2016. cm p. ISBN 978-1-118-67502-1.
- FAMA, E. F. Efficient capital markets: A review of theory and empirical work. **The Journal of Finance**, Wiley, v. 25, n. 2, p. 383–417, 1970.
- FISCHER, T.; KRAUSS, C. Deep learning with long short-term memory networks for financial market predictions. **European journal of operational research**, Elsevier, v. 270, n. 2, p. 654–669, 2018.
- HANDELMAN, G. S. *et al.* Peering into the black box of artificial intelligence: evaluation metrics of machine learning methods. **American Journal of Roentgenology**, Am Roentgen Ray Soc, v. 212, n. 1, p. 38–43, 2019.
- HOUDT, G. V.; MOSQUERA, C.; NÁPOLES, G. A review on the long short-term memory model. **Artificial Intelligence Review**, Springer, v. 53, n. 8, p. 5929–5955, 2020.
- JAYBHAY, K. M.; ARGIDDI, R. V.; APTE, S. Stock market prediction model by combining numeric and news textual mining. **International Journal of Computer Applications**, Foundation of Computer Science, v. 57, n. 19, 2012.
- KAMIJO, K.-i.; TANIGAWA, T. Stock price pattern recognition-a recurrent neural network approach. *In*: IEEE. **1990 IJCNN international joint conference on neural networks**. [S.l.: s.n.], 1990. p. 215–221.
- KIM, K. jae. Financial time series forecasting using support vector machines. **Neurocomputing**, v. 55, n. 1–2, p. 307–319, 2003. ISSN 0925-2312.
- LATORRE, M. d. R. D. d. O.; CARDOSO, M. R. A. Análise de séries temporais em epidemiologia: uma introdução sobre os aspectos metodológicos. **Revista brasileira de epidemiologia**, SciELO Brasil, v. 4, p. 145–152, 2001.

LEVINE, R.; ZERVOS, S. Stock market development and long-run growth. **The world bank economic review**, Oxford University Press, v. 10, n. 2, p. 323–339, 1996.

LI, Z. *et al.* A survey of convolutional neural networks: analysis, applications, and prospects. **IEEE transactions on neural networks and learning systems**, IEEE, v. 33, n. 12, p. 6999–7019, 2021.

LIU, Y. *et al.* Stock price movement prediction from financial news with deep learning and knowledge graph embedding. *In*: SPRINGER. **Knowledge Management and Acquisition for Intelligent Systems: 15th Pacific Rim Knowledge Acquisition Workshop, PKAW 2018, Nanjing, China, August 28-29, 2018, Proceedings 15**. [S.l.: s.n.], 2018. p. 102–113.

LO, A. The adaptive markets hypothesis: Market efficiency from an evolutionary perspective. **Journal of Portfolio Management**, **Forthcoming**, v. 30, p. 15–129, 2004.

NELSON, D. M. Q. Uso de redes neurais recorrentes para previsão de séries temporais financeiras. Universidade Federal de Minas Gerais, 2017.

NOSOUHIAN, S.; NOSOUHIAN, F.; KHOSHOUEI, A. K. A review of recurrent neural network architecture for sequence learning: Comparison between lstm and gru. Preprints, 2021.

PAGANO, M. Financial markets and growth: An overview. **European economic review**, North-Holland, v. 37, n. 2-3, p. 613–622, 1993.

PEDREGOSA, F. *et al.* Scikit-learn: Machine learning in python. **Journal of Machine Learning Research**, v. 12, p. 2825–2830, 2011.

PERSIO, L. D.; HONCHAR, O. *et al.* Artificial neural networks approach to the forecast of stock market price movements. **International Journal of Economics and Management Systems**, v. 1, n. Anno 2016, p. 158–162, 2016.

POMMERANZENBAUM, I. **Redes neurais artificiais na predição das principais séries do índice Ibovespa e suas aplicações em sistemas automatizados de negociação**. 114 p. 2014. Tese (Doutorado) — Dissertação (Mestrado)-COPPE, Programa de Engenharia Elétrica, Universidade . . . , 2014.

Rankia Portugal. **Os maiores investidores do mundo**. 2021. <<https://www.rankia.pt/fundos-de-investimento/os-maiores-investidores-do-mundo/>>. Uma visão geral sobre os investidores mais bem-sucedidos de todos os tempos, incluindo Benjamin Graham, Peter Lynch, e Philip Arthur Fisher. Este artigo destaca suas estratégias de investimento e sucessos no mercado.

SCHMIDT, R. M. Recurrent neural networks (rnns): A gentle introduction and overview. **arXiv preprint arXiv:1912.05911**, 2019.

TALEB, N. N. **Iludidos pelo acaso: A influência oculta da sorte nos mercados e na vida**. 1. ed. Rio de Janeiro: Best Seller, 2004.

TSAY, R. S. **Analysis of Financial Time Series**. [S.l.: s.n.]: John Wiley & Sons, 2010.

WANG, J.-H.; LEU, J.-Y. Stock market trend prediction using arima-based neural networks. *In: IEEE. Proceedings of International Conference on Neural Networks (ICNN'96)*. [S.l.: s.n.], 1996. v. 4, p. 2160–2165.

YOSHIHARA, A. *et al.* Predicting stock market trends by recurrent deep neural networks. *In: PHAM, D.-N.; PARK, S.-B. (ed.). PRICAI 2014: Trends in Artificial Intelligence*. [S.l.: s.n.]: Springer, Cham, 2014. (Lecture Notes in Computer Science, v. 8862), p. 759–769.

YOSHINAGA, C. E.; CASTRO, F. H. Inteligência artificial: a vanguarda das finanças. **GV-EXECUTIVO**, v. 22, n. 3, 2023.