UNIVERSIDADE FEDERAL DE ALAGOAS
Instituto de Computação

João Victor Ribeiro Ferro

Uma Avaliação Sistemática de Técnicas de Aprendizado de Máquina Baseadas em Ensemble para Previsão de Índices do Mercado de Ações Usando Séries Temporais Financeiras

Maceió - AL
2024

João Victor Ribeiro Ferro

Uma Avaliação Sistemática de Técnicas de Aprendizado de Máquina Baseadas em Ensemble para Previsão de Índices do Mercado de Ações Usando Séries Temporais Financeiras

Dissertação apresentada ao Programa de Pós-graduação em Informática do Instituto de Computação da Universidade Federal de Alagoas como requisito parcial para obtenção do título de Mestre em Informática.

Orientador: Prof. Dr. Evandro de Barros Costa

Maceió - AL
2024

Catalogação na fonte
Universidade Federal de Alagoas
Biblioteca Central
Divisão de Tratamento Técnico
Bibliotecária: Girlaine da Silva Santos – CRB-4 – 1127

F395u Ferro, João Victor Ribeiro.
Uma avaliação sistemática de técnicas de aprendizado de máquina baseadas em ensemble para previsão de índices do mercado de ações usando séries temporais financeiras / João Victor Ribeiro Ferro. – 2024.
150 f. : il.

Orientador: Evandro de Barros Costa.
Dissertação (Mestrado em Informática.) - Universidade Federal de Alagoas. Instituto de Informática. Programa de Pós-Graduação em Informática, Maceió, 2024.

Bibliografia: f. 132-140.
Apêndices: f. 141-150.

1. Análise de séries temporais. 2. Inteligência artificial. 3. Mercado financeiro. 4. Índices de mercados e ações. 5. Machine Learning. I. Título.

CDU: 004.8: 336.76

MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DE ALAGOAS
INSTITUTO DE COMPUTAÇÃO
Av. Lourival Melo Mota, S/N, Tabuleiro do Martins, Maceió - AL, 57.072-970
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO (PROPEP)
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

Folha de Aprovação

JOÃO VICTOR RIBEIRO FERRO

UMA AVALIAÇÃO SISTEMÁTICA DE TÉCNICAS DE APRENDIZADO DE MÁQUINA BASEADAS EM ENSEMBLE PARA PREVISÃO DE ÍNDICES DO MERCADO DE AÇÕES USANDO SÉRIES TEMPORAIS FINANCEIRAS

A SYSTEMATIC EVALUATION OF ENSEMBLE-BASED MACHINE LEARNING TECHNIQUES FOR FORECASTING STOCK MARKET INDICES USING FINANCIAL TIME SERIES

Dissertação submetida ao corpo docente do Programa de Pós-Graduação em Informática da Universidade Federal de Alagoas e aprovada em 10 de setembro de 2024.

Banca Examinadora:

Documento assinado digitalmente
gov.br EVANDRO DE BARROS COSTA
Data: 23/09/2024 16:04:45-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. EVANDRO DE BARROS COSTA
UFAL – Instituto de Computação
Orientador

Documento assinado digitalmente
gov.br ICARO BEZERRA QUEIROZ DE ARAUJO
Data: 24/09/2024 08:12:27-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. ICARO BEZERRA QUEIROZ DE ARAÚJO
UFAL – Instituto de Computação
Examinador Interno

Documento assinado digitalmente
gov.br BRUNO ALMEIDA PIMENTEL
Data: 23/09/2024 19:00:15-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. BRUNO ALMEIDA PIMENTEL
UFAL – Instituto de Computação
Examinador Interno

Documento assinado digitalmente
gov.br GEORGE DARMITON DA CUNHA CAVALCANTI
Data: 23/09/2024 18:03:37-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. GEORGE DARMITON DA CUNHA CAVALCANTI
UFPE– Universidade Federal de Pernambuco
Examinador Externo

Agradecimentos

Primeiramente, agradeço a Deus por me dar forças para continuar e superar as dificuldades.

À minha mãe, Antonia Cléria Ribeiro Ferro, ao meu pai, João Carlos Ferro da Silva, e à minha irmã, Victória Camilly Ribeiro Ferro, pelo incentivo, conselhos e apoio necessários para alcançar meus objetivos.

À UFAL e ao corpo docente do curso de Ciência da Computação por abrir oportunidades únicas para meu crescimento profissional e, em especial, ao meu orientador, Professor Evandro de Barros Costa, e aos Professores Robério José Rogério dos Santos e Roberta Vilhena Vieira Lopes, pelo tempo dedicado, por acreditarem em mim e me fazerem acreditar na minha capacidade, além do suporte dado neste trabalho.

Ao corpo técnico do Programa de Pós-Graduação em Informática (PPGI), por esclarecer dúvidas e ajudar na parte burocrática.

Aos meus companheiros, por tornar essa jornada menos difícil e mais divertida, com as brincadeiras e horas de desestresse. Em especial José Rubens da Silva Brito, Lucas A. Lisboa, Matheus Gomes de Oliveira, William Gabriel da Paz Rosendo e Arthur Sávio Bernardo de Melo.

Aos meus colegas que me ajudaram diretamente ou indiretamente neste trabalho, tornando essa jornada menos difícil e mais divertida com momentos de descontração.

Se pude enxergar mais longe, foi porque me apoiei em ombros de gigantes

(Isaac Newton)

Resumo

A previsão de índices de séries temporais nos mercados de ações tem despertado crescente interesse, fornecendo aos investidores informações valiosas sobre as tendências econômicas futuras de um país a curto, médio e longo prazo. Esse tipo de predição tem sido amplamente realizado por meio de modelos de aprendizado de máquina. No entanto, devido à volatilidade, ao ruído e à estocasticidade dos dados, os modelos individuais (*single*) frequentemente apresentam limitações em termos de precisão. Para mitigar esses problemas, surgiram as técnicas de *ensemble*, que combinam múltiplos modelos para aumentar a robustez e a acurácia das previsões. As abordagens de *ensemble* exploram a diversidade dos modelos e técnicas de *machine learning*, aproveitando as características individuais de cada um para obter resultados mais confiáveis do que aqueles gerados por modelos isolados. Dessa forma, foi realizada uma revisão sistemática da literatura voltada para a predição de índices da bolsa de valores utilizando séries temporais financeiras e abordagens de *ensemble*, com o objetivo de mapear os principais artigos, autores, tipos de técnicas utilizadas e lacunas na literatura. A revisão revela que muitas análises comparativas se limitam ao uso de métricas de desempenho tradicionais, como erro quadrático médio, raiz do erro quadrático médio, erro absoluto médio e erro percentual absoluto médio, o que pode introduzir vieses nas comparações. Além disso, essas análises frequentemente omitem testes estatísticos robustos, focam-se apenas em um tipo de mercado e adotam protocolos de comparação inadequados. Este estudo investiga como diferentes técnicas de *ensemble* podem aprimorar a previsão em séries temporais financeiras, adotando um protocolo meticuloso para eliminar vieses nos dados e padronizar comparações entre metodologias. Além das métricas tradicionais, foi introduzida uma análise de custo-benefício para uma avaliação mais abrangente das arquiteturas de *ensemble*. O teste de hipótese de Wilcoxon foi aplicado para validar as descobertas, juntamente com os testes de Friedman e Nemenyi. Os resultados destacam a importância de executar os algoritmos em diferentes ambientes, como os selecionados IBOVESPA e S&P 500, pois há divergências no desempenho das métricas de avaliação e nos testes estatísticos, o que ressalta a necessidade de utilizar a métrica de custo-benefício para verificar se há um ganho real na performance geral da arquitetura utilizada em comparação aos modelos *single*, estabelecendo assim uma estruturação para a construção e avaliação das abordagens de *ensemble*. Com estes resultados, espera-se contribuir para a construção de metodologias mais robustas e eficazes no uso de técnicas de *ensemble* para a predição de índices de ações, garantindo uma melhor compreensão das vantagens e limitações dessas técnicas em diferentes ambientes de mercado.

Palavras-chaves: Análise Comparativa, *Machine Learning*, *Ensemble*, Série Temporal Financeira, Mercado de Índices.

Abstract

The prediction of time series indices in stock markets has been garnering increasing interest, providing investors with valuable information about a country's future economic trends in the short, medium, and long term. This type of prediction has been widely performed through machine learning models. However, due to the volatility, noise, and stochasticity of the data, single models often present limitations in terms of accuracy. To mitigate these issues, ensemble techniques have emerged, combining multiple models to increase the robustness and accuracy of predictions. Ensemble approaches exploit the diversity of models and machine learning techniques, leveraging the individual characteristics of each to achieve more reliable results than those generated by isolated models. Thus, a systematic literature review was conducted focusing on stock index prediction using financial time series and ensemble approaches, with the aim of mapping the main articles, authors, types of techniques used, and gaps in the literature. The review reveals that many comparative analyses are limited to the use of traditional performance metrics, such as mean squared error, root mean squared error, mean absolute error, and mean absolute percentage error, which can introduce biases in comparisons. Furthermore, these analyses often omit robust statistical tests, focus on only one type of market, and adopt inadequate comparison protocols. This study investigates how different ensemble techniques can enhance prediction in financial time series, adopting a meticulous protocol to eliminate data biases and standardize comparisons between methodologies. In addition to traditional metrics, a cost-benefit analysis was introduced for a more comprehensive evaluation of ensemble architectures. The Wilcoxon hypothesis test was applied to validate the findings, along with the Friedman and Nemenyi tests. The results highlight the importance of running algorithms in different environments, such as the selected IBOVESPA and S&P 500, as there are divergences in the performance of evaluation metrics and statistical tests, emphasizing the need to use the cost-benefit metric to verify whether there is a real gain in the overall performance of the architecture compared to single models. This establishes a framework for constructing and evaluating ensemble approaches. With these results, the aim is to contribute to the development of more robust and effective methodologies in the use of ensemble techniques for stock index prediction, ensuring a better understanding of the advantages and limitations of these techniques in different market environments.

key-words: Comparative Analysis, Machine Learning, Ensemble, Financial Time Series, Index Market

Sumário

Lista de Figuras

Lista de Tabelas

Lista de Abreviaturas e Siglas

(AAPL)	*Apple Inc.*
(AAR)	*Annual Retorno*
(ABC)	*Artificial Bee Colony*
(ABT)	*Abbott Laboratories*
(ADF)	*Augmented Dickey Fuller*
(AE)	*Auto-encoder*
(AG)	Algoritmo Genético
(AKIMA)	*Akima Spline Interpolation Technique*
(ANN)	*Artificial Neural Network*
(AORD)	*All Ordinaries*
(AR)	*Simple Autoregressive*
(ARIMA)	*Autoregressive Integrated Moving Average*
(ARMA)	*Autoregressive Moving Average Model*
(ARMAX)	*Autoregressive-Moving Average With Exogenous Terms*
(AT)	*Análise Técnica*
(AUD)	*Australian Dollar*
(AWT)	*Adaptive Wavelet Transform*
(B3)	*Brasil, Bolsa, Balcão*
(BAC)	*Bank of America Corp*
(bagging)	*Bootstrap aggregating*
(Bi-LSTM)	*Bidirectional Long-Short Term Memory*
(BPNN)	*Back propagation neural network*
(BRL)	*Brazilian Real*
(BSd)	*B-spline Wavelet of high order d*

(BSE SENSEX) *S&P Bombay Stock Exchange Sensitive Index*

(BTC) *BITCOIN*

(BVSP) *Sao Paulo's Stock Market Index*

(CAC 40) *Cotation Assistée en Continu*

(CART) *Classification And Regression Trees*

(CD) *Critical Distance*

(CEEMD) *Complete Ensemble Empirical Mode Decomposition*

(CEEMDAN) *Complete Ensemble Empirical Mode Decomposition with Adaptive Noise*

(CHF) *Swiss Franc*

(CNN) *Convolutional Neural Network-Long Short-Term Memory*

(CPU) *Central Processing Unit*

(CS) *Crow Search*

(CSI 100) *China Securities Index 100 Index*

(CSI 1000) *China Securities Index 1000 Index*

(CSI 300) *China Securities Index 300 Index*

(CSI 500) *China Securities Index 500 Index*

(CSI 800) *China Securities Index 800 Index*

(DAE) *Deep Autoencoder*

(DAX) *Deutscher Aktienindex*

(DJIA) *Dow Jones Industrial Average*

(DT) *Decision Tree*

(DWT) *Discrete Wavelet Transform*

(EEMD) *Ensemble Empirical Mode Decomposition*

(EL) *Ensemble Learning*

(ELM) *Extreme Learning Machine*

(Elman) *Elman Neural Networks*

(EM) *MSCI Emerging Markets Index*

(EMD) *Empirical Mode Decomposition*

(EMD2FNN) *Empirical Mode Decomposition and Neural Network-based Machine Factorizatio*

(ENANFIS) *Ensemble Neuro-Fuzzy Inference System*

(EO) *Elite-based Opposition*

(ET) *Extra Trees Classifier*

(ETF) *Exchange Traded Fund*

(EU ETS) *European Union Emissions Trading System*

(EU) *MSCI European Index*

(EURONEXT 100) *EURONEXT 100 INDEX*

(EWT) *Empirical Wavelet Transform*

(FA) Função de Autocorrelação

(FAP) Função de Autocorrelação Parcial

(FFQOA) *Fast Forward Quantum Optimization Algorithm*

(FLANN) *Functional link artificial neural network*

(FMI) Função de Modo Intrínseco

(FNN) *Feedforward Neural Network*

(FQTSFM) *Fuzzy-Quantum Time Series Forecasting Model*

(FTSE) *Financial Times Stock Exchange 100 Index*

(GA) Genetic algorithm

(GARCH) *Generalized Autoregressive Conditional Heteroskedasticity*

(GB) *Gradient Boosting*

(GBP) *British Pound*

(GM) *Grey Model*

(GRNN) *General Regression Neural Network*

(GRU) *Gated Recurrent Unit*

(GSE) *Ghana Stock Exchange*

(HA) *History Attention*

(HAR) *Heterogeneous Autoregressive*

(HAR-RV) *heterogeneous autoregressive model of realized volatility*

(HCLTECH) *HCL Technologies Ltd*

(HFT) High-Frequency Trading

(HS 300) *Shanghai-Shenzhen 300 Stock Index*

(HSI) *Hang Seng Index*

(IBOVESPA) Índice da Bolsa de Valores de São Paulo

(ICEEMDAN) *Improved Complete Ensemble Empirical Mode Decompositions with Adaptive Noise*

(INR) *Indian rupee*

(IP) *US industrial production*

(ISE) *Istanbul Stock Exchange*

(ISE100_TL) *Istanbul Stock Market Index(TL)*

(ISE100_USD) *Istanbul Stock Market Index(USD)*

(IXIC) *Nasdaq Composite*

(JPY) *Japanese Yen*

(JSE) *Johannesburg Stock Exchange*

(JSX) *Jakarta Stock Exchange*

(KMX) *CarMax Inc*

(KNN) *K-Nearest Neighbors*

(KOSDAQ) *Korea Securities Dealers Automated Quotation*

(KOSPI) *Korea Composite Stock Price Index*

(KSE 100) *Pakistan Stock Exchange*

(LM) *Levenberg-Marquardt*

(LR) *Lasso Regression*

(LR) *Linear Regression*

(LRNFIS) *Locally Recurrent Neuro-Fuzzy Information System*

(LSTM) *Long Short-Term Memory*

(MAE) *Mean Absolute Error*

(MAPE) *Mean Absolute Percentage Error*

(MD) *Maximum Drawdown*

(MEMD) *Multivariate Empirical Mode Decomposition*

(ML) *Machine Learning*

(MSE) *Mean Squared Error*

(MSFT) *Microsoft Corporation*

(MXN) *Mexican Peso*

(NASDAQ) *National Association of Securities Dealers Automated Quotations*

(Nasdaq) *National Association of Securities Dealers Automated Quotations*

(NIFTY 50) *Indian stock market index*

(NIKKEI 225) *Nikkei Stock Average*

(NSE) *National Stock Exchange of India*

(NYSE) *New York Stock Exchange*

(NYSE) *New York Stock Exchange*

(ONU) Organização das Nações Unidas

(PCA) *Principal Component Analysis*

(PR) *Polynomial Regression*

(PSO) *Particle Swarm Optimization*

(RBFNN) *Radial Basis Function Neural Network*

(RF) *Random Forest*

(RidgeCV) *Ridge Regression with Built-in Cross-Validation*

(RMSE) *Root Mean Squared Error*

(RNN) *Recurrent Neural Network*

(RSL) Revisão Sistemática da Literatura

(Russell 2000) *Russell 2000 Index*

(RVFL) *Random Vector Functional Link*

(RW) *Random Walk*

(S&P 500) *Standard and Poor's 500*

(S&P/ASX 200) *Australian Australian Securities Exchange*

(SGD) *Singapore dollar*

(SME 100) *Small and Medium Board 100 Index*

(SME 300) *Small and Medium Board 300 Index*

(SME CI) *Small and Medium Board Composite Index*

(SMI) *Swiss Market Index*

(SR) *Sharpe Ratio*

(SSE 180) *Shanghai Stock Exchange 180 Index*

(SSE 380) *Shanghai Stock Exchange 380*

(SSE 50) *Shanghai Stock Exchange 50 Index*

(SSE A shares) *Shanghai A-Share Index*

(SSE B shares) *Shanghai Stock Exchange B-Share Index*

(SSE) *Shanghai Stock Exchange*

(STF) *Séries Temporais Fuzzy*

(STI) *Straits Times Index*

(STOXX 50) *EURO STOXX 50*

(SVM) *Support Vector Machines*

(SVR) *Support Vector Regression*

(SZCI) *Shenzhen Stock Exchange Component Index*

(SZSE A shares) *Shenzhen Constituent A Share Index*

(SZSE B shares) *Shenzhen Stock Exchange Constituent B-Share Index*

(SZSE) *Shenzhen Stock Exchange Index*

(TAIEX) *Taiwan Stock Exchange Capitalization Weighted Stock Index*

(TAIFEX) *Taiwan Futures Exchange*

(TATASTEEL) *Tata Steel Limited*

(TLBO) *Teaching Learning Based Optimization*

(TOPSIS) *Technique for Order of Preference by Similarity to Ideal Solution*

(TSEC) *Taiwan Stock Exchange Corporation*

(TSPI) *Time Series Prediction with Invarian*

(USD) *US Dollar*

(VMD) *Variational Mode Decomposition*

(WDBP) *Wavelet De-Noising-Based Back Propagation*

(WOA) *Whale Optimization Algorithm*

(WTI) *West Texas Intermediate*

(XG) *XGBoost Classifier*

(XOM) *Exon mobile corporation*

Capítulo 1

Introdução

Os índices de mercado de ações, que são derivados do desempenho das ações representativas listadas neles, desempenham um papel crucial ao refletir as mudanças nos preços do mercado financeiro. Frequentemente, esses índices são utilizados como indicadores-chave para antecipar o desempenho futuro da macroeconomia de um país (YAN; AASMA et al., 2020). Neste contexto, a precisão na previsão do comportamento dessas séries temporais financeiras é de suma importância, uma vez que permite aos investidores mitigar riscos na tomada de decisões e fornece informações de referência essenciais para avaliações de curto, médio e longo prazo, conferindo maior segurança ao investimento no mercado financeiro de um determinado país.

Dado este contexto, para abordar o desafio de previsão de séries temporais financeiras, visando captar as tendências do mercado, é necessário considerar características como ruído, volatilidade, estocasticidade dos dados, entre outras, as quais comprometem o desempenho dos algoritmos de *Machine Learning* – (ML) *single* (AMPOMAH; QIN; NYAME; BOTCHEY, 2021), as abordagens de *ensemble*, que fazem parte de um conjunto mais amplo de técnicas de ML, emergem com o objetivo de mitigar estes problemas (NTI; ADEKOYA; WEYORI, 2020a).

Na literatura, diversas abordagens de *ensemble*, que combinam modelos de ML e técnicas associadas, têm sido desenvolvidas com o objetivo de aprimorar as previsões, conforme apresentado no Capítulo 3. Surge, assim, a necessidade de realizar uma análise comparativa entre essas abordagens para identificar sua eficácia e limitações, conforme exemplificado nos estudos de (NTI; ADEKOYA; WEYORI, 2020a; AMPOUNTOLAS, 2023; FERROUHI; BOUABDALLAOUI, 2024; JOTHIMANI; BAŞAR, 2019). Essas análises fornecem *insights* valiosos para o avanço da pesquisa nesta área de estudo. No entanto, a maioria dos estudos apresenta vieses

nos resultados e a falta de um protocolo e testes estatísticos bem definidos, como verificado no subcapítulo 3.4.

Neste estudo, inicialmente, foi realizada uma análise sistemática e bibliométrica da literatura para identificar as principais linhas de pesquisa na predição de índices do mercado financeiro utilizando séries temporais financeiras e abordagens de *ensemble*. O objetivo foi proporcionar um entendimento geral da área e focar na análise comparativa sistemática, isto é, uma comparação padronizada dos diversos tipos de abordagens de *ensemble* por meio de um protocolo com testes estatísticos e diferentes cenários, como o mercado desenvolvido, representado pelo S&P 500, e o mercado em desenvolvimento, representado pelo IBOVESPA, para avaliar o comportamento e a confiabilidade dos algoritmos. Além disso, foi introduzida a métrica de custo-benefício para verificar se o aumento da complexidade do algoritmo é justificado pelo seu ganho em desempenho. Essa métrica consegue avaliar o desempenho do algoritmo independentemente do sistema operacional utilizado, pois é composta pelo *Instructions Retired da Central Processing Unit*, o que facilita a padronização dos resultados e sua replicação, como apresentado no subcapítulo 4.5.

Os algoritmos de *Machine Learning* selecionados para este estudo foram o CART, MLP e SVR. Na categoria de modelos estatísticos, foi escolhido o ARIMA. Para as abordagens de *ensemble*, foram selecionadas quatro áreas: *ensemble learning* composto por *Bagging* e *Stacking*, sistema híbrido residual, decomposição, completo e otimização, para serem avaliadas no estudo comparativo. No Capítulo 4 serão fornecidas as justificativas para a escolha desses algoritmos.

Os resultados desta pesquisa mostraram que tanto a abordagem de decomposição quanto o sistema híbrido residual, ambos utilizando o modelo CART como base, exibiram resultados promissores nas métricas de erro tradicionais como MSE, RMSE, MAE e MAPE. A primeira abordagem se destacou no mercado IBOVESPA e a segunda abordagem no S&P 500. No entanto, testes estatísticos como o de *Nemenyi* indicaram que esses resultados são estatisticamente similares ao desempenho do modelo *single* CART e, na métrica de custo-benefício, apresentam um baixo desempenho, o que destaca a importância de realizar um *trade-off* das métricas.

1.1 Motivação e Justificativa

Modelos de *Machine Learning single* referem-se àqueles que operam de forma isolada, ou seja, sem qualquer tipo de combinação com outros modelos ou técnicas. Esses modelos apresentam limitações em sua capacidade de adaptação a mudanças ou tendências nos dados, além de dificuldades em capturar dependências temporais e em estabelecer relações eficazes entre observações passadas e futuras. Como discutido no capítulo 1, tais limitações destacam a necessidade de desenvolver alternativas que possam mitigar esses problemas. Uma dessas alternativas é a implementação de modelos de *ensemble*, que têm como objetivo aprimorar o desempenho em face de desafios típicos dos algoritmos de ML aplicados a séries temporais financeiras. As técnicas de *ensemble* buscam combinar as forças de múltiplas abordagens, distribuindo tarefas de maneira eficaz entre diferentes modelos para aproveitar as melhores características de cada técnica empregada, resultando em previsões mais robustas e precisas (SAGI; ROKACH, 2018).

Primeiramente, observou-se, por meio das revisões sistemáticas da literatura recentes sobre *machine learning* nas séries temporais financeiras (HENRIQUE; SOBREIRO; KIMURA, 2019; KEHINDE; CHAN; CHUNG, 2023; GANDHMAL; KUMAR, K., 2019; KUMBURE et al., 2022; BUSTOS; POMARES-QUIMBAYA, 2020), que não há trabalhos voltados especificamente para as abordagens de *ensemble* na predição de séries temporais financeiras dos índices da bolsa de valores. Dessa forma, este estudo contribui para o crescente corpo de literatura no campo ao realizar uma revisão sistemática e bibliométrica. Esse mapeamento destacou os principais artigos, autores, periódicos e a classificação das abordagens conforme suas características em comum e funcionamento, salientando os principais algoritmos e técnicas utilizados nas abordagens de *ensemble*, além de identificar lacunas na literatura.

Na esfera da avaliação comparativa de técnicas de *ensemble*, como destacado em estudos anteriores na literatura científica (NTI; ADEKOYA; WEYORI, 2020a; AMPOUNTOLAS, 2023; KRAUSS; DO; HUCK, 2017; JOTHIMANI; BAŞAR, 2019), foram identificados no subcapítulo 3.4 problemas como a falta de testes estatísticos para validar os resultados obtidos nos trabalhos, o foco somente em um determinado tipo de mercado, o que pode trazer armadilhas no resultado final sobre a performance do algoritmo, e a não utilização de um protocolo coerente que tenha o objetivo de reduzir o viés dos dados, o que pode impactar negativamente os resultados obtidos e também garantir que o processo possa ser reproduzido para outros modelos e pesquisadores.

Além disso, devido aos vieses e às limitações presentes nos indicadores frequentemente observados na maioria dos estudos comparativos, a concentração unicamente na performance dos modelos, avaliada por métricas tradicionais de erro *Mean Squared Error* – (MSE), *Root Mean Squared Error* – (RMSE), *Mean Absolute Error* – (MAE) e *Mean Absolute Percentage Error* – (MAPE), negligencia uma análise crítica sobre se o aumento da complexidade do algoritmo realmente pode trazer benefícios, sendo este ponto avaliado pela métrica de custo-benefício.

Logo, este trabalho é justificado por contribuir para reduzir essas lacunas apontadas anteriormente. Ele realiza uma revisão sistemática voltada para a utilização das técnicas de *ensemble* na predição de índices da bolsa de valores, constrói uma análise sistemática comparativa que adota um protocolo para minimizar o viés nos dados, realiza testes e análises em diferentes tipos de mercados, utiliza testes estatísticos rigorosos e introduz a métrica de custo-benefício nesta área de estudo. Estas estratégias enfatizam a importância e a originalidade da contribuição desta pesquisa ao acervo de conhecimentos existente.

1.2 Objetivo Geral

O objetivo geral deste trabalho é realizar uma análise sistemática e comparativa sobre a utilização das técnicas de *Ensemble* na predição dos índices da bolsa de valores. Busca-se identificar os pontos fortes e fracos dessas abordagens, assim como o *trade-off* das métricas de erros e de custo-benefício.

1.2.1 Objetivos Específicos

Para alcançar o objetivo geral, este trabalho seguiu os seguintes objetivos específicos:

- Realizar uma revisão sistemática e bibliométrica sobre os métodos de *ensemble* aplicados na predição dos índices de mercado utilizando séries temporais financeiras.
- Mapear e categorizar os métodos identificados durante a revisão sistemática.
- Investigar as análises comparativas focadas nas abordagens de *ensemble* para predição do mercado de índices utilizando séries temporais financeiras.
- Conduzir experimentos com as técnicas levantadas na revisão sistemática da literatura para validar a metodologia proposta, utilizando métricas como MSE, RMSE, MAE, MAPE e análise de custo-benefício.

- Avaliar qual abordagem de *ensemble* apresentou o melhor desempenho.
- Examinar as limitações das técnicas de *ensemble* abordadas.

1.3 Estrutura do Trabalho

Este documento está estruturado em seis capítulos. No primeiro, foi apresentada uma introdução sobre o trabalho, contextualizando o tema da pesquisa, motivando a sua realização, justificando-a, identificando a problemática e as contribuições principais, bem como os objetivos para a sua execução. Os demais capítulos são descritos a seguir:

- **Capítulo 2**: Apresenta o conhecimento necessário para o entendimento desta dissertação.
- **Capítulo 3**: Aborda a revisão sistemática sobre as técnicas de *ensemble* utilizadas na predição de índices utilizando séries temporais financeiras e os trabalhos relacionados a esta pesquisa, destacando seus diferenciais em relação aos demais.
- **Capítulo 4**: Mostra as principais diferenças entre um mercado e outro em relação às séries temporais financeiras. Além disso, apresenta a metodologia aplicada no trabalho, incluindo o protocolo e as arquiteturas utilizadas.
- **Capítulo 5**: Apresenta os resultados deste artigo, juntamente com as análises comparativas por meio dos testes de hipóteses, além de abordar as limitações e os pontos fortes de cada modelo adotado.
- **Capítulo 6**: Contém a conclusão do trabalho, destacando os pontos importantes, as limitações desta pesquisa e os trabalhos futuros.

Capítulo 2

Fundamentação Teórica

Neste capítulo são abordados os conceitos necessários para o entendimento do problema alvo, que é a predição do índice da bolsa de valores utilizando série temporal financeira, assim como o entendimento e funcionamento dos algoritmos utilizados.

2.1 Classificação dos Países

A Organização das Nações Unidas – (ONU) passou a classificar os países de acordo com seus diferentes níveis de desenvolvimento. Segundo a ONU, apenas por conveniência, e não por juízo sobre o estágio alcançado nesse processo, as regiões podem ser denominadas de mercados desenvolvidos, em desenvolvimento e menos desenvolvido segundo (BEKAERT; HARVEY, 2003):

- **Mercado Desenvolvido**: Refere-se a um país ou região que alcançou um alto nível de desenvolvimento econômico, industrialização e avanço tecnológico. Esses mercados geralmente possuem infraestrutura bem estabelecida, estruturas regulatórias robustas e sistemas financeiros eficientes. Eles oferecem uma ampla gama de oportunidades de investimento, têm altos níveis de liquidez e são considerados relativamente estáveis e seguros para os investidores. Renda per capita alta, indústrias avançadas e um alto padrão de vida caracterizam os mercados desenvolvidos. Incluem países como os Estados Unidos, Alemanha, Japão e Reino Unido.

- **Mercado em Desenvolvimento**: Refere-se a um país ou região que está em transição de uma economia agrária de baixa renda para uma economia industrializada mais avan-

çada. Esses mercados têm menor desenvolvimento econômico, infraestrutura e avanços tecnológicos em comparação com os mercados desenvolvidos. Eles frequentemente enfrentam desafios como acesso limitado a capital, instabilidade política e sistemas financeiros subdesenvolvidos. Os mercados em desenvolvimento oferecem potencial de crescimento e oportunidades de investimento, mas também apresentam níveis mais altos de risco e volatilidade. Exemplos de mercados em desenvolvimento incluem Brasil, Índia, China e África do Sul.

- **Mercado menos desenvolvido**: Caracterizado por baixo desenvolvimento econômico, acesso limitado a capital e infraestrutura fraca. Esses mercados enfrentam desafios como instabilidade política, altos níveis de pobreza e falta de industrialização. Os mercados subdesenvolvidos têm menor renda per capita e frequentemente carecem dos recursos e instituições necessários para um crescimento econômico sustentado. Exemplos de países subdesenvolvidos incluem Somália, Sudão do Sul, Afeganistão e Mianmar.

2.2 Mercado de Capital

O Mercado de Capitais é um componente vital do sistema financeiro, em que são negociados títulos de longo, médio e curto prazo, como ações e títulos de dívida. Este mercado é fundamental para a economia de um país, pois tem o objetivo de facilitar o fluxo de recursos financeiros entre poupadores e investidores, promovendo o crescimento econômico e a criação de empregos (YAPA ABEYWARDHANA, 2017).

Os principais componentes do mercado de capitais incluem as ações, que representam a propriedade parcial de uma empresa, isto é, os investidores que compram ações tornam-se acionistas e podem ganhar dinheiro por meio de dividendos e valorização do preço das ações. Há também os títulos de dívida, que incluem debêntures, notas promissórias, entre outros, que são instrumentos de dívida emitidos por empresas ou governos, que prometem pagar uma quantia específica de juros ao longo do tempo e devolver o principal no vencimento. Estes são exemplos de alguns componentes encontrados no mercado de capitais.

Existem diferentes tipos de mercados no contexto do mercado de capitais como o mercado primário em que novos títulos são emitidos e vendidos pela primeira vez, permitindo que as empresas levantem capital diretamente dos investidores. Já no mercado secundário, os títulos

existentes são negociados entre investidores, com a bolsa de valores sendo um exemplo de mercado secundário [1].

Para o funcionamento deste sistema, várias instituições estão envolvidas, tais como as bolsas de valores, que facilitam a negociação de ações e outros títulos, como a *New York Stock Exchange* – (NYSE), *National Association of Securities Dealers Automated Quotations* – (Nasdaq) e *Brasil, Bolsa, Balcão* – (B3). Além disso, corretoras e bancos de investimento atuam como intermediários na emissão e negociação de títulos. Por fim, comissões reguladoras, como a Comissão de Valores Mobiliários no Brasil, regulam e supervisionam as atividades do mercado de capitais.

2.2.1 Índice da Bolsa de Valores

O índice da bolsa de valores é uma medida estatística que reflete a performance de um conjunto de ações representativas de uma determinada bolsa de valores. Essas ações são selecionadas de acordo com critérios predefinidos, como capitalização de mercado, liquidez, setor econômico, entre outros. O índice serve como um termômetro do mercado, indicando tendências de alta ou baixa com base no comportamento dos preços dessas ações (CAPLINGER, s.d.).

Além disso, é utilizado por investidores para tomar decisões de investimento e avaliar a saúde econômica de um setor ou país (YAN; AASMA et al., 2020). Um exemplo é o *Standard and Poor's 500* – (S&P 500), que inclui 500 das maiores empresas de capital aberto nos Estados Unidos, sendo amplamente considerado um indicador chave da economia americana. Outro exemplo é o Índice da Bolsa de Valores de São Paulo – (IBOVESPA), o principal índice da Bolsa de Valores brasileira (B3), composto pelas ações mais negociadas e representativas do mercado brasileiro.

Dessa forma, o índice pode ser utilizado como *benchmarking*, em que os investidores comparam a performance de seus portfólios com o índice. Além disso, índices servem de base para vários produtos financeiros, como *Exchange Traded Fund* – (ETF) e derivativos.

2.3 Série Temporal

Uma série temporal é uma sequência de dados ou observações coletadas ou registradas em intervalos regulares de tempo (CHATFIELD; XING, 2019). Esses intervalos de tempo podem

[1] Para mais detalhes sobre outros tipos de mercados, acessar o (TOLEDO FILHO, 2020)

ser igualmente espaçados, como medições diárias, mensais ou anuais. Em termos formais, uma série temporal pode ser definida como uma sequência ordenada de variáveis aleatórias Y_t, com $t \in \mathbb{T}$ representando os instantes de tempo:

$$Y_t = \{y_1, y_2, y_3, \dots, y_t\} \tag{2.1}$$

onde Y_t representa o valor da série temporal no tempo t e y_i é a observação no instante i. Quando os intervalos de tempo são igualmente espaçados, a série é dita de frequência regular, uma característica importante para muitas aplicações de previsão. Por outro lado, uma série temporal irregular armazena dados em uma sequência arbitrária de pontos no tempo. Séries temporais irregulares são apropriadas quando os dados chegam de forma imprevisível, como quando um aplicativo registra cada transação de ações ou quando medidores de eletricidade registram eventos aleatórios, como avisos de bateria fraca ou indicadores de baixa voltagem.

Além disso, séries temporais podem ser classificadas como:

- **Univariada**: Quando apenas uma variável é observada ao longo do tempo, formando uma única sequência temporal, como a variação diária de temperatura em uma cidade.
- **Multivariada**: Quando várias variáveis são observadas simultaneamente ao longo do tempo. Nesse caso, as variáveis podem estar correlacionadas ou influenciar-se mutuamente, como ocorre no caso do preço de ações, volume de negociações e taxa de juros sendo analisados conjuntamente.

As séries temporais possuem algumas características fundamentais que devem ser consideradas para sua análise e modelagem, conforme descrito por (VISHWAS; PATEL, 2020):

- **Tendência**: Reflete o comportamento em um determinado período de tempo. Isto é, para esse período de tempo estabelecido, a tendência indica se a série cresce, decresce ou permanece estável.
- **Sazonalidade**: Caracteriza-se por flutuações periódicas em intervalos específicos (por exemplo, semanas, meses ou anos), repetindo-se de forma previsível.
- **Estacionariedade**: Uma série é dita estacionária quando suas propriedades estatísticas, como média e variância, permanecem constantes ao longo do tempo, oscilando de maneira aleatória em torno de uma média fixa.

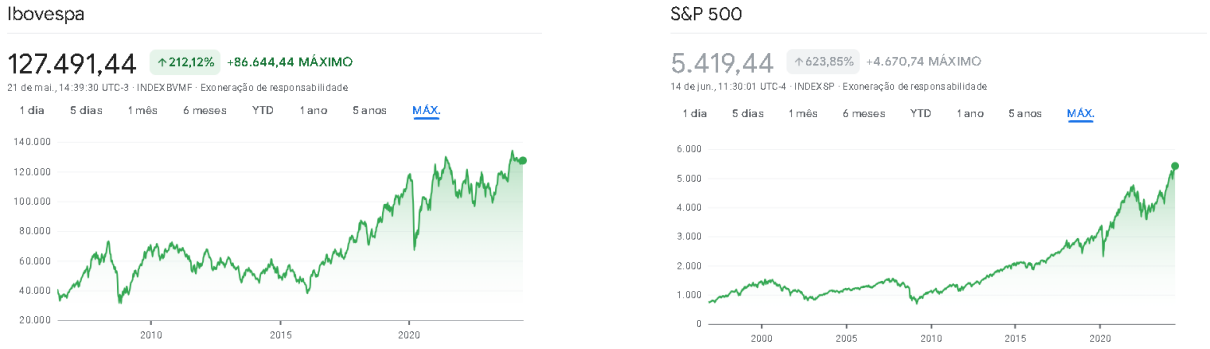
O estudo das séries temporais é fundamental, pois permite abordar uma variedade de problemas, como: investigar fatores que influenciam a série, descrever seu comportamento (identificando tendências, ciclos e sazonalidades), detectar periodicidades relevantes e prever valores futuros com base nas observações passadas.

Para problemas de previsão, especialmente em séries temporais financeiras, a análise de séries temporais se torna particularmente útil, como será discutido na seção 2.3.1 seguinte.

2.3.1 Série Temporal Financeira

Uma série temporal financeira modela os valores de um determinado indicador econômico como o índice da bolsa de valores, cotação do dolar ou o preço de uma ação ao longo do tempo. Esse tipo de série apresenta algumas características especiais, comuns também a outras séries temporais, como apresentado no trabalho de (GIACOMEL, 2016):

- **Presença de tendências:** enquanto séries temporais mais simples são estacionárias, variando em torno de um valor médio constante, séries temporais financeiras exibem tendências de alta e baixa, com durações variáveis conforme os parâmetros do mercado;
- **Sazonalidade:** os valores de uma série temporal financeira podem variar conforme a época do ano, repetindo padrões nos anos seguintes. Por exemplo, em dezembro, devido ao Natal, as ações de uma loja de presentes tendem a registrar um aumento maior do que nos outros meses do ano;
- **Pontos influentes:** são valores atípicos que se desviam do padrão da série temporal. Em séries financeiras, pontos influentes podem ser momentos de alta volatilidade no mercado, resultando em grandes altas ou quedas nos preços, que depois retornam ao patamar normal;
- **Heteroscedasticidade condicional:** a variância dos valores de entrada e saída da série temporal não é constante ao longo do tempo, o que torna o comportamento da série mais aleatório;
- **Não-linearidade:** devido à sua complexidade e comportamento estocástico, esse tipo de série temporal não pode ser modelado por uma função linear.



(a) IBOVESPA (b) S&P 500

Figura 2.1: Exemplo de séries temporais financeiras: índices IBOVESPPA e S&P 500. Fonte: Google [2024]

A Figura 2.1 apresenta as séries temporais financeiras que modelam os valores do índice IBOVESPA e do S&P 500.

2.4 Viés e Limitações dos Indicadores

Os indicadores são ferramentas essenciais para identificar e medir aspectos específicos de conceitos, fenômenos, problemas ou resultados de processos e programas. Eles desempenham duas funções principais: uma descritiva, que visa relatar informações sobre uma realidade inevitável, e uma avaliativa, que avalia a relevância ou adequação do desempenho de um programa ou projeto (FRANCESCHINI et al., 2019).

Para serem considerados adequados, os indicadores devem possuir três propriedades cruciais: validade, que é a capacidade de representar fielmente o que se pretende medir; confiabilidade, referindo-se à obtenção de fontes confiáveis e ao uso de metodologias transparentes; e mensurabilidade, relacionada à capacidade de medir de forma precisa e inequívoca, considerando a periodicidade do que está sendo medido.

No entanto, é essencial distinguir entre as limitações e os vieses associados ao uso de indicadores.

Limitações dos Indicadores:

1. **Natureza Abstrata:** Indicadores simplificam a realidade, o que pode não capturar todos os aspectos relevantes de um fenômeno.

2. **Parcimônia e Confiança:** A validação cuidadosa do processo de medição é essencial para garantir a confiabilidade e evitar complexidade desnecessária que não agregue valor.

3. **Custo de Medição:** Aumentar o número de indicadores pode elevar os custos de medição sem proporcionar benefícios proporcionais.

Vieses nos Indicadores:

1. **Viés de Seleção e Interpretação:** Indicadores estão sujeitos a vieses daqueles que os escolhem, coletam e interpretam, o que pode distorcer a representação da realidade.

2. **Distinção entre Indicador e Dimensão de Interesse:** A escolha de um indicador pode destacar certos aspectos em detrimento de outros, levando a uma interpretação enganosa do progresso ou desempenho.

3. **Interferência da Medição na Realidade:** A coleta de dados pode modificar o contexto observado, influenciando os resultados, especialmente quando conduzida sob pressões externas.

No subcapítulo 3.4.2, são observados com mais detalhes quais são os principais vieses e limitações encontrados nos trabalhos da literatura.

2.5 Abordagem *Single* e *Ensemble*

O modelo *single* em ML refere-se a um cenário em que um único modelo é construído e utilizado para realizar previsões, sem o auxílio de técnicas de *ensemble* ou de outros modelos, como apresentado nos artigos (BOX et al., 2015; BOSER; GUYON; VAPNIK, V. N., 1992; BREIMAN, 2017). Em outras palavras, o modelo *single* é um algoritmo utilizado para minerar os dados e inferir os *insights*, sendo o mais simples (MAHESH, 2020).

A definição de *ensemble* refere-se a abordagens que combinam diversas técnicas de *Machine Learning* com o objetivo de produzir uma previsão final mais robusta do que qualquer modelo *single* individual poderia oferecer. Esta abordagem baseia-se na ideia de que a combinação das forças de vários modelos pode reduzir erros e aumentar a robustez do sistema de previsão, conforme demonstrado em trabalhos como (WU, J.; ZHOU, T.; LI, T., 2020; ALHNAITY; ABBOD, 2020; ZOLFAGHARI; GHOLAMI, 2021; YAN; AASMA et al., 2020).

Os modelos *ensemble* apresentam algumas propriedades e vantagens importantes, sendo uma das principais o *trade-off* viés-variância. A ideia é que a combinação de modelos pode reduzir a variância de previsões ao agregar modelos que possuem diferentes vieses. Assim, enquanto modelos *single* podem sofrer com altas variâncias ou altos vieses, dependendo de sua configuração, o uso de *ensemble* permite balancear esses fatores, alcançando uma solução mais estável e com menor erro geral (SAGI; ROKACH, 2018).

Existem dois tipos principais de abordagens de *ensemble*:

- **Modelos em paralelo**: Neste caso, os modelos são treinados de maneira independente e, em seguida, seus resultados são combinados para produzir uma previsão final. Abordagens como *Random Forest* são exemplos típicos desse método, no qual múltiplas árvores de decisão são treinadas em subconjuntos de dados e suas previsões são combinadas, geralmente por votação ou média ponderada. Esse tipo de ensemble é eficaz na redução da variância, uma vez que agrega múltiplos modelos de alta variância para reduzir o erro final (CUTLER, A.; CUTLER, D. R.; STEVENS, 2012).

- **Modelos em sequência**: Nesse caso, os modelos são treinados sequencialmente, onde cada modelo subsequente tenta corrigir os erros cometidos pelos anteriores. Um exemplo clássico é o *Boosting*, no qual os modelos são treinados de forma iterativa, ajustando-se às falhas dos modelos anteriores. Essa abordagem é eficaz na redução de viés, ao ajustar modelos de baixo viés, mas alta variância, até que uma solução robusta seja obtida (FREUND; SCHAPIRE, 1997).

Outra vantagem dos métodos *ensemble* é sua capacidade de capturar relações complexas nos dados que modelos individuais podem não ser capazes de detectar. Isso ocorre porque diferentes modelos podem explorar diferentes aspectos dos dados, resultando em uma maior capacidade de generalização e previsões mais precisas.

2.6 Modelo Autorregressivo Integrado de Médias Móveis

O modelo Modelo Autorregressivo Integrado de Médias Móveis ou em inglês *Autoregressive Integrated Moving Average* – (ARIMA) (p, d, q), assume-se que o valor futuro de uma variável é uma função linear de várias observações passadas e erros aleatórios. Ou seja, o processo

subjacente que gera a série temporal com média μ é descrito por:

$$\phi(B)\nabla^d (y_t - \mu) = \theta(B)a_t,$$

em que y_t e a_t são o valor real e o erro aleatório no tempo t, respectivamente; $\phi(B) = 1 - \sum_{i=1}^{p} \varphi_i B^i$ e $\theta(B) = 1 - \sum_{j=1}^{q} \theta_j B^j$ são polinômios em B de grau p e q; $\phi_i (i = 1, 2, \ldots, p)$ e $\theta_j (j = 1, 2, \ldots, q)$ são parâmetros do modelo; $\nabla = (1 - B)$, em que B é o operador de defasagem; p e q são inteiros frequentemente referidos como as ordens do modelo, e d é um inteiro frequentemente referido como a ordem da diferenciação. Assume-se que os erros aleatórios, a_t, são independentemente e identicamente distribuídos com média zero e variância constante de σ^2 (KHASHEI; BIJARI, 2011).

A metodologia (BOX et al., 2015) inclui três etapas iterativas: identificação do modelo, estimativa de parâmetros e verificação diagnóstica. A ideia básica da identificação do modelo é que, se uma série temporal é gerada a partir de um processo ARIMA, ela deve exibir certas propriedades teóricas de autocorrelação. Ao combinar os padrões de autocorrelação empíricos com os teóricos, é frequentemente possível identificar um ou vários modelos potenciais para a série temporal em questão. (BOX et al., 2015) sugeriram o uso da Função de Autocorrelação – (FA) e da Função de Autocorrelação Parcial – (FAP) dos dados da amostra como ferramentas fundamentais para identificar a ordem do ARIMA.

A transformação dos dados é frequentemente necessária na fase de identificação para tornar a série temporal estacionária. A estacionaridade é uma condição necessária para a construção de um modelo ARIMA utilizado para previsão. Uma série temporal estacionária é caracterizada por características estatísticas, como média e estrutura de autocorrelação, sendo constantes ao longo do tempo. Quando a série temporal observada exibe tendência e heterocedasticidade, diferenças e transformações de potência são aplicadas aos dados para remover a tendência e estabilizar a variância antes que um modelo ARIMA possa ser ajustado. Uma vez identificado um modelo provisório, a estimativa de parâmetros é direta. Os parâmetros são estimados de modo que uma medida geral de erros seja minimizada. Isso pode ser alcançado usando um procedimento de otimização não linear. A etapa final na construção do modelo é a verificação diagnóstica da adequação do modelo. Verifica-se se as suposições do modelo sobre os erros, a_t, são satisfeitas.

Várias estatísticas diagnósticas e gráficos de resíduos podem ser usados para examinar a adequação do modelo provisoriamente obtido aos dados históricos. Se o modelo for conside-

rado inadequado, um novo modelo provisório deve ser identificado, seguido novamente pelas etapas de estimativa de parâmetros e verificação do modelo. As informações diagnósticas podem ajudar a sugerir um modelo alternativo. Este processo de construção do modelo em três etapas é tipicamente repetido várias vezes até que um modelo satisfatório seja finalmente selecionado. O modelo final selecionado pode então ser utilizado para fins de previsão.

2.7 Decomposição Empírica por Modo Completo com Ruído Adaptativo

O modelo de Decomposição Empírica por Modo Completo com Ruído Adaptativo ou em inglês *Complete Ensemble Empirical Mode Decomposition with Adaptive Noise* – (CEEMDAN) o seu principal objetivo da decomposição é simplificar a complexidade dos sinais, tornando-os mais acessíveis para avaliação e uso, esse processo é crucial para permitir que os modelos de *Machine Learning* obtenham diferentes perspectivas durante o aprendizado. Nesse contexto, surgiu o método precursor, *Empirical Mode Decomposition* – (EMD) (HUANG et al., 1998), reconhecido por sua capacidade auto-adaptativa de decompor sinais sem a necessidade de definir previamente o número ou a natureza das Função de Modo Intrínseco – (FMI), que são extraídas diretamente dos dados. No entanto, a EMD enfrenta o desafio conhecido como mistura de modos, em que uma única FMI pode misturar elementos de escalas diferentes ou semelhantes, complicando a análise (LEI; HE, Z.; ZI, 2009).

Para aprimorar a EMD, foi desenvolvida a *Ensemble Empirical Mode Decomposition* – (EEMD) (WU, Z.; HUANG, 2009). Esta técnica adaptativa é essencial para analisar sinais não estacionários, representando-os como uma soma de FMIs com parâmetros de domínio e frequência modulados. A EEMD introduz ruído branco gaussiano para mitigar o problema da mistura de modos, facilitando a identificação mais precisa das FMIs. No entanto, a EEMD ainda apresenta limitações, como a necessidade de média finita e o risco de introduzir um erro de reconstrução devido ao ruído que não pode ser eliminado. Em resposta, a CEEMDAN é definida como uma abordagem aprimorada da EEMD. No entanto, ela implica um alto custo computacional (associado à busca exaustiva) e inclui ruído residual no método EEMD. Um aumento no número de tentativas pode potencialmente aumentar o número de processos de peneiramento. A CEEMDAN foi desenvolvida para reduzir o número de tentativas enquanto mantém a capacidade de resolver o problema da mistura de modos (TORRES et al., 2011).

Em resumo, os passos da CEEMDAN são os seguintes, conforme descrito no artigo (REZAIE-BALF et al., 2019):

(1) Este método aplica-se ao cálculo da primeira função de modo da seguinte forma:

$$FMI_1(t) = \frac{1}{N}\sum_{j=1}^{N} \mathrm{FMI}_1^j(t) \tag{2.2}$$

O primeiro resíduo é dado da seguinte forma:

$$r_1(t) = x(t) - \overline{FMI_1}(t) \tag{2.3}$$

(2) Determine $\mathrm{emd}_{(t)}$ como o k-ésimo elemento FMI usando o método EMD e decomponha a sequência $+p_1 emd_1\left(n_j(t)\right)$ para alcançar o segundo componente do FMI.

$$FMI_2(t) = \frac{1}{N}\sum_{j=1}^{N} emd_1\left(r_1(t) + p_1 emd_1\left(n_j(t)\right)\right) \tag{2.4}$$

Um sinal residual é fornecido.

$$r_2(t) = r_1(t) - FMI_2(t) \tag{2.5}$$

(3) Similarmente, no passo anterior, o k-ésimo sinal residual é estimado.

$$r_k(t) = r_{k-1}(t) - FMI_k(t) \tag{2.6}$$

O componente do $k+1$-ésimo FMI é então derivado.

$$IMF_{k+1}(t) = \frac{1}{N}\sum_{j=1}^{N} emd_1\left(r_k(t) + p_k emd_k\left(n_j(t)\right)\right) \tag{2.7}$$

(4) Itere esses passos até alcançar o sinal residual. Suponha que existam L componentes. Assim, a sequência original pode ser calculada da seguinte forma:

$$y(t) = \sum_{i=1}^{L} IMF_i(t) + r(t) \tag{2.8}$$

em que $r(t)$ é o sinal residual final.

A Figura 2.2 apresenta o fluxograma dos passos do CEEMDAN:

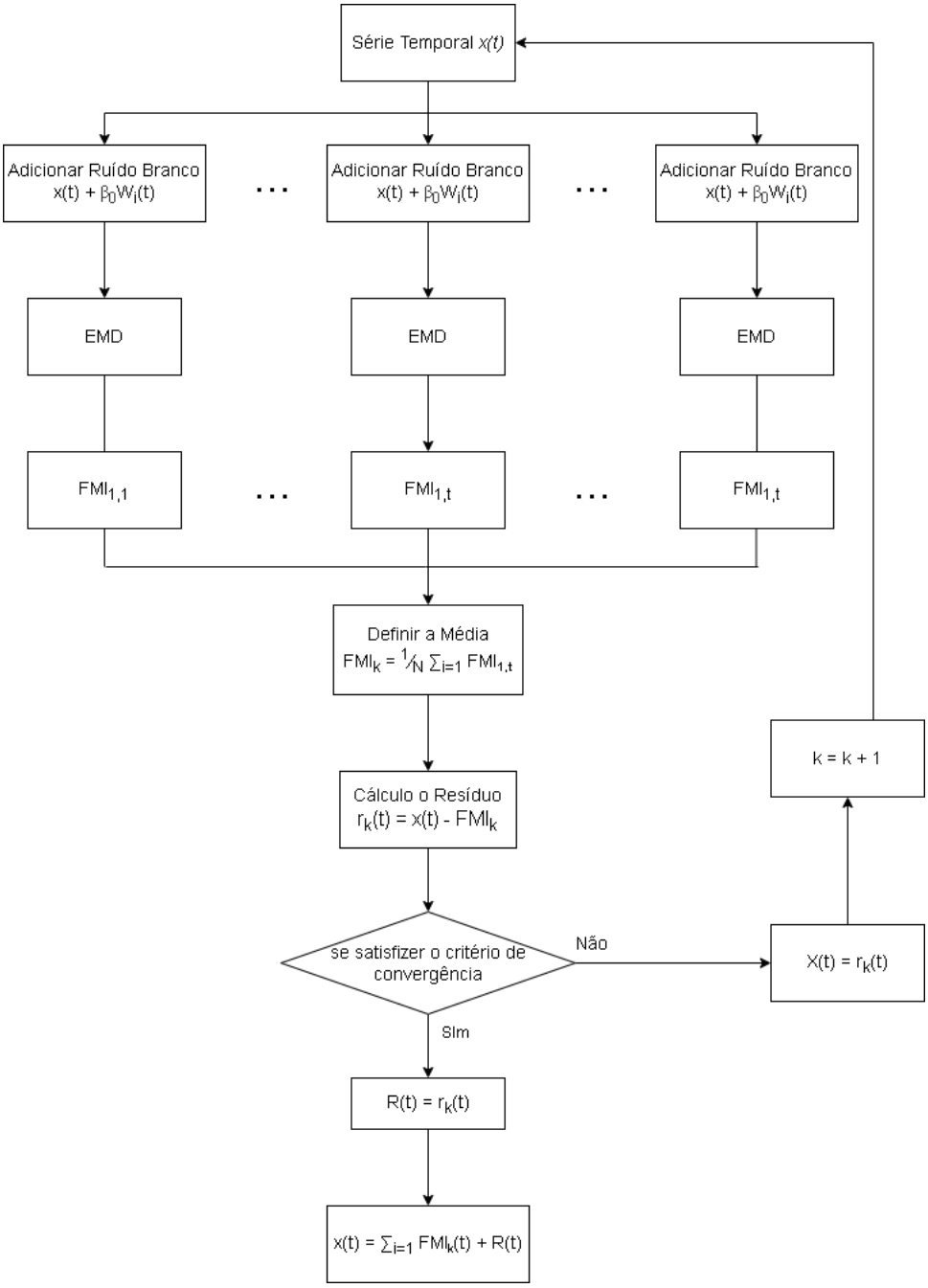


Figura 2.2: O fluxograma do algoritmo CEEMDAN, Fonte: (Autor,2024)

2.8 Regressão por Vetores de Suporte

A Máquina de Vetores de Suporte ou no inglês *Support Vector Machines* – (SVM) foi apresentada por (BOSER; GUYON; VAPNIK, V. N., 1992), essa técnica é utilizada em tarefas de classificação, reconhecimento de padrões e, para o caso de séries temporais, análise de regressão.

A forma alternativa de SVM, a Regressão por Vetores de Suporte ou em inglês *Support Vector Regression* – (SVR), apresentada inicialmente por (VAPNIK, V., 2013), é empregada no contexto de previsão de séries temporais e tem por objetivo aproximar uma função usando os dados observados que "treina" a SVM ((DRUCKER et al., 1996)). A seguir é apresentada a teoria relativa a SVR.

Dado um conjunto de dados (amostra) de treinamento $(x_1, y_1), \ldots, (x_n, y_n)\ \epsilon\ \mathbb{R}^n \times \mathbb{R}$, e que x_i é um vetor de entradas n-dimensional, y_i são as saídas e n é o número de observações no conjunto de treinamento (KANG; LI, J., 2016). A relação não linear entre entrada e saída pode ser definida como:

$$f(\mathbf{x}) = \mathbf{w}^T\mathbf{x} + b \tag{2.9}$$

onde $f(\mathbf{x})$ é o valor predito com, no máximo, um desvio ϵ de y_i (ϵ é uma constante que define a faixa de tolerância para os erros de predição na formulação do SVR); b é o viés e $\mathbf{w}$ é um vetor de pesos. Isso nos leva a um problema de otimização sujeito a restrições, em que o objetivo é encontrar um conjunto de pesos 'ótimo' (SAPANKEVYCH; SANKAR, 2009), dado por

$$\text{minimizar} \quad \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{n}\left[L\left(\xi_i\right) + L\left(\xi_i^*\right)\right], \tag{2.10}$$

$$\text{sujeito a} \quad \begin{cases} y_i - f\left(x_i\right) \leq \epsilon + \xi_i \\ f\left(x_i\right) - y_i \leq \epsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0, \end{cases} \tag{2.11}$$

na qual $C > 0$ é a constante de regularização que indica a quantidade de erros no conjunto de treinamento, $\|\cdot\|$ é a norma do vetor e $L(\cdot)$ é a função de perda ϵ-insensível,dada por 2.12:

$$L(\xi) = \begin{cases} 0 & \text{se } |\xi| \leq \epsilon \\ |\xi| - \epsilon & \text{caso contrário .} \end{cases} \tag{2.12}$$

Aplicando o método dos Multiplicadores de Lagrange ao problema de otimização anterior, temos a seguinte expansão:

$$f(\mathbf{x}) = \sum_{i=1}^{n} \left(\alpha_i - \alpha_i^*\right) \mathbf{w}^T \mathbf{x} + b \tag{2.13}$$

em que α_i e α_i^* são os multiplicadores de Lagrange.

Quando os dados não são lineares em seu espaço de entrada é necessário estender essa regressão para um caso de regressão não linear, para tal é inserida uma função Kernel, que mapeia os dados de entrada para um espaço de dimensão superior onde a regressão linear torna-se possível. Portanto, teremos:

$$f(\mathbf{x}) = \sum_{i=1}^{n} \left(\alpha_i - \alpha_i^*\right) K\left(\mathbf{x_i}, \mathbf{x_j}\right) + b. \tag{2.14}$$

Apenas para valores $|f(x_i) - y_i| \geq \epsilon$ os multiplicadores de Lagrange, α_i e α_i^*, podem ser diferentes de zero, isto é, todas as amostras que estão no ϵ-tubo possuem multiplicadores de Lagrange nulos. Os pontos com coeficientes diferentes de zero são chamados de vetores de suporte (SMOLA; SCHÖLKOPF, 2004).

2.8.1 Funções *Kernel*

Considerando $\phi : \mathbb{R}^n \to \mathbb{R}^{n+h}$, com x $\to \phi$(x), define-se o *Kernel* conforme a equação abaixo:

$$K\left(\mathbf{x_i}, \mathbf{x_j}\right) = \phi\left(\mathbf{x_i}\right)^{\mathrm{T}} \phi\left(\mathbf{x_j}\right). \tag{2.15}$$

Segundo (VAPNIK, V., 2013), uma função para ser *Kernel* precisa apresentar duas características:

1. Ser simétrica:

$$K(x, z) = (\phi(x) \cdot \phi(z)) = K(z, x)$$

2. Satisfazer as equações de Cauchy-Schwarz:

$$K(x, z)^2 = (\phi(x) \cdot \phi(z))^2 \leq \|\phi(x)\|^2 \cdot \|\phi(z)\|^2 = (\phi(x) \cdot \phi(x))(\phi(z) \cdot \phi(z)) = K(x, x) \cdot K(z, z).$$

Entretanto, essas condições não são suficientes para garantir a existência do espaço de características. Dessa forma, uma Função *Kernel* deve ainda satisfazer o Teorema de Mercer[2].

[2]Para mais destalhes ler o artigo de (MERCER, 1909)

As funções *Kernel* retornam o produto interno entre dois pontos no espaço de entrada. A expressão para a função *Kernel* pode ser escolhida de forma direta sem que a função $\Phi(\cdot)$ esteja explícita.

A quantidade de parâmetros existentes na função *Kernel* afeta muito a complexidade dos modelos (SU et al., 2014). As Funções *Kernel* mais comuns na literatura são Linear, Polinomial, Base Radial e Sigmóide.

***Kernel* Linear**

$$K\left(\mathbf{x_i}, \mathbf{x_j}\right) = \left\langle \mathbf{x_i}, \mathbf{x_j} \right\rangle . \tag{2.16}$$

Esse *Kernel* não possui nenhum parâmetro próprio para ajustar.

***Kernel* Polinomial**

$$K\left(\mathbf{x_i}, \mathbf{x_j}\right) = \left(\gamma \left\langle \mathbf{x_i}, \mathbf{x_j} \right\rangle + \text{ coef } 0\right)^{\text{degree}} , \tag{2.17}$$

o parâmetro $degree > 0$ representa o grau do polinômio e o *coef0* é o intercepto, ambos necessitam de ajuste.

***Kernel* de Base Radial**

$$K\left(\mathbf{x_i}, \mathbf{x_j}\right) = \exp\left(-\gamma \left\langle \mathbf{x_i} - \mathbf{x_j}, \mathbf{x_i} - \mathbf{x_j} \right)^2\right) \tag{2.18}$$

neste *Kernel*, $\gamma = 1/2\sigma^2$ com $\sigma > 0$ que precisa ser ajustado.

***Kernel* Sigmóide**

$$K\left(\mathbf{x_i}, \mathbf{x_j}\right) = \tanh\left(\gamma \left\langle \mathbf{x_i}, \mathbf{x_j} \right\rangle + \text{coef } 0\right), \tag{2.19}$$

O uso de Funções *Kernel* reduz a quantidade de parâmetros livres na solução de problemas que envolvem dados não lineares. Os parâmetros da SVR têm uma influência significativa nos resultados das previsões; uma escolha inadequada pode resultar em *overfitting* ou *underfitting*. Portanto, a seleção dos hiperparâmetros é uma tarefa crucial na modelagem com SVR (KANG; LI, J., 2016).

O desempenho da SVR está diretamente relacionado à combinação dos parâmetros C, ϵ e o(s) parâmetro(s) da função *Kernel*. O parâmetro C, também conhecido como penalização, é responsável pelo equilíbrio entre maximizar a margem do ϵ-tubo, aceitando valores com desvios maiores que ϵ, e minimizar o erro no conjunto de treinamento. Valores muito altos

ou muito baixos de C podem causar *overfitting* ou *underfitting* (WANG, X. et al., 2014). O parâmetro ϵ define o raio do ϵ-tubo ao redor da função de regressão, permitindo desvios da previsão em relação aos dados alvos.

2.9 Árvore de Classificação e Regressão

A Árvore de Classificação e Regressão ou em inglês *Classification And Regression Trees* – (CART) proposto por (BREIMAN, 2017), é um método estatístico não paramétrico que utiliza árvores de decisão para resolver problemas de classificação ou regressão com variáveis categóricas e contínuas.

2.9.1 Árvore de Classificação

Funciona da seguinte maneira: inicialmente, a melhor covariável será escolhida para realizar o particionamento. O ponto ótimo do particionamento será definido de modo a otimizar alguma medida de "pureza" do nó. Posteriormente, o processo será realizado recursivamente, até que ocorra uma etapa de poda. Considerando um caso de classificação (variável resposta qualitativa), utiliza-se o índice de Gini (GINI, 1921).

Índice Gini:

$$IG(S) = \sum_{i=1}^{n\ \text{Classes}} p_i \left(1 - p_i\right), \tag{2.20}$$

O S é o subconjunto no qual o índice de Gini está sendo calculado (pode ser a base de dados inteira ou apenas uma partição dela). As *nClasses* representam o número de categorias que a variável de interesse possui, enquanto o p_i é a proporção do número de elementos da categoria i da variável resposta pelo número de elementos no subconjunto S. Quanto maior for o índice de Gini, mais caótica está a partição S.

2.9.2 Árvore de Regressão

O CART consegue lidar com a variável alvo sendo quantitativa. Para entender mais sobre esse caso, é preciso definir três elementos essenciais desta árvore de decisão (BREIMAN, 2017), que são:

1. Definir um modo de particionar binariamente o nó;

2. Definir uma regra para saber quando o nó é terminal;

3. Definir um valor y para os nós filhos da árvore de regressão.

Modo de particionar binariamente o nó

Na árvore de decisão é utilizada o índice de Gini, respectivamente. Contudo, essas medidas foram idealizadas para situações com variáveis resposta qualitativas, e não quantitativas. Para contornar esse problema, utiliza-se o conceito da soma de quadrados, mas agora na árvore de decisão.

Soma de quadrados para árvore de regressão em CART

$$SQ(s,t) = SQ_{\text{Total}} - \left(SQ_{\text{Esquerda}} + SQ_{\text{Direita}}\right) \tag{2.21}$$

Em que foi definido o t é um nó da árvore de regressão, o s é um ponto de divisão do nó, então o $SQ(s,t)$ é a diferença de soma de quadrados entre o nó pai e seus nós filhos, o SQ_{Total} é a soma de quadrados calculada em todo o nó t, o SQ_{Esquerda} é a soma de quadrados calculada para os y_i "à esquerda de s", que estão em um lado da partição, $y_i < s$; estão em um outro lado da partição e o SQ_{Direita} é a soma de quadrados calculada para os y_i "à direita de s".

A soma de quadrados é a usual, ou seja,

$$SQ_{\text{grupo}} = \sum_{i=1}^{|\text{grupo}|} (y_i - \bar{y})^2, \tag{2.22}$$

em que SQ_{grupo} é a soma de quadrados calculada em "grupo", y_i é o valor do i-ésimo indivíduo na variável resposta e $\bar{y}$ é a média de todos os y_i. Dado o exposto, para decidir o melhor ponto de divisão binária do nó basta escolher o s que minimize $SQ(s,t)$.

Critério para determinar quando o nó é terminal

O processo é executado recursivamente até que a base de dados esteja completamente particionada, ou um número mínimo de observações em cada nó filho seja definido pelo usuário. Para evitar o sobreajuste, é realizada uma poda.

Determinação do valor y para os nós filhos

Em uma árvore de regressão, o nó filho contém observações com vários valores de y. Apesar desses valores serem próximos (devido ao processo de particionamento da árvore), é necessário escolher um valor como saída do nó. Para isso, utiliza-se uma estatística dos valores de

y presentes no nó filho. Esta estatística pode ser a média (usualmente empregada no CART), a moda, a mediana, etc.

2.10 Rede *Multilayer Perceptron*

Em 1958, Frank Rosenblatt introduziu o modelo *Perceptron*, a forma mais simples de rede neural artificial usada para classificação de padrões linearmente separáveis, baseada no neurônio de *McCulloch–Pitts*. A estrutura básica do *Perceptron* inclui um único neurônio com pesos sinápticos ajustáveis e um viés. Quando os padrões utilizados no treinamento do *Perceptron* são provenientes de duas classes linearmente separáveis, o algoritmo do *Perceptron* converge para um hiperplano que separa as duas classes. Este fenômeno foi posteriormente conhecido como o Teorema da Convergência do *Perceptron* (RUSSELL; NORVIG, 2022).

A Rede *Multilayer Perceptron* (MLP) é uma rede neural artificial com maior poder computacional em comparação às redes sem camadas intermediárias, pois consegue lidar com problemas não linearmente separáveis. É amplamente reconhecido na literatura que estruturas MLP com uma única camada intermediária podem aproximar qualquer função contínua, e que duas camadas intermediárias são suficientes para aproximar qualquer função matemática. No contexto de séries temporais, uma MLP com uma camada oculta geralmente é suficiente (PÁDUA BRAGA; LEON FERREIRA; LUDERMIR, 2007).

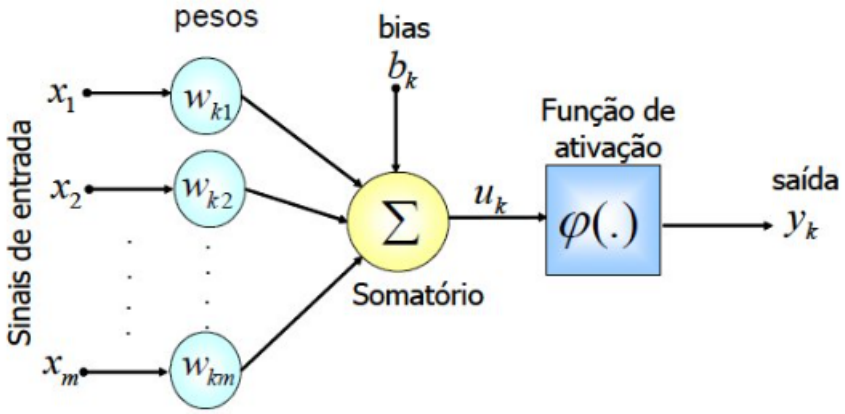


Figura 2.3: Modelagem básica do Neurônio Artificial, Fonte: (SOARES; SILVA, 2011)

De acordo com (SILVA et al., 2020), a unidade básica de processamento de informação no sistema nervoso central humano é o neurônio ($\mathcal{N}$), que é composto por três elementos: um

conjunto de dendritos ($\mathcal{I}$), um corpo celular ($\mathcal{S}$) e um axônio ($\mathcal{A}$). Essencialmente, o funcionamento de uma rede neural artificial envolve três ações realizadas pelos seus elementos básicos:

1. O conjunto de dentritos ($\mathcal{I}$) recebe os estímulos externos ao corpo celular ($\mathcal{S}$);
2. ($\mathcal{S}$) pondera esses estímulos através de operações agregativas simples;
3. O axônio ($\mathcal{A}$) processa a informação recebida de ($\mathcal{S}$), por meio de operações mais sofisticadas, gerando as respostas dos neurônios aos estímulos da camada de entrada.

Naturalmente, as saídas de um neurônio podem ser as entradas de outro, configurando assim uma rede neural. Geralmente, o neurônio ($\mathcal{N}$) é modelado por um Perceptron ($\mathcal{P}$).

A abordagem para séries temporais com redes neurais, ou seja, a previsão para o valor da série u_t, contempla a aproximação

$$\mathcal{I}_t = \left\{ u_{t-t_i} \right\}_{i \in \mathcal{I}_u}, \tag{2.23}$$

isto é, valores passados da série, e o modelo

$$\mathcal{S}\left(\mathcal{I}_t\right) = \sum \theta_i u_{t-t_i} + \theta_0, \tag{2.24}$$

na qual θ_i é o coeficiente que pondera as observações u_{t-t_i} e θ_0 é o intercepto. Dessa forma, $\mathcal{S}$ envolve uma combinação linear dos dados de entrada. Em resposta a série regressa, obtem-se a previsão

$$\hat{u}_t = \mathcal{A}\left(\mathcal{S}\left(\mathcal{I}_t\right)\right) \tag{2.25}$$

em que a aproximação $\mathcal{A}(\cdot)$ recebe o nome de Função de Ativação.

Para séries temporais têm sido comum o uso de uma rede com uma camada intermediária, uma rede MLP, levando a uma função de soma simples

$$S_h\left(I_t\right) = \sum \theta_{hi} u_{t-t_i} + \theta_{h0}, \tag{2.26}$$

com funções de ativação intermediárias dadas por

$$A_h\left(S_h\left(I_t\right)\right), \tag{2.27}$$

em que h representa o índice do neurônio.

Por fim, as funções de ativação intermediárias são operadas por uma função de ativação final que se dedica à previsão dos valores futuros da série

$$\hat{u}_t = A_0 \left(\sum_{h=1}^{H} \theta_{h0} A_h \left(S_h \left(I_t \right) \right) \right). \tag{2.28}$$

2.10.1 Funções de Ativação

As funções de ativação têm o papel crucial de informar às camadas anteriores os erros cometidos pela rede com a maior precisão possível. O cálculo para cada neurônio em uma rede MLP leva em consideração a derivada da função de ativação associada a esse neurônio. Para assegurar a existência dessa derivada, é necessário que a função de ativação seja contínua. Em essência, a diferenciabilidade é o único critério que uma função de ativação deve atender (HAYKIN, 2009). As funções de ativação mais comuns são a Linear, Sigmóide Logística e a Tangente Hiperbólica, descritas a seguir:

Função de Ativação Linear

$$y = \alpha x + \beta \tag{2.29}$$

sendo $\alpha \in \mathbb{R}$ define a saída linear para os valores de entrada x, β é o intercepto e y é a saída;

Função de Ativação Sigmóide Logística

$$y = \frac{1}{1 + \exp(-\lambda x)}, \tag{2.30}$$

em que λ determina a sensibilidade de resposta da função. Geralmente ajusta-se $\lambda = 1$. Sua maior vantagem sobre a função linear é justamente a sua não linearidade, os valores dessa função variam entre 0 e 1 .

Função de Ativação Tangente Hiperbólica

$$y = \frac{\exp(x) - \exp(-x)}{\exp(x) + \exp(-x)} \tag{2.31}$$

esta função preserva a forma sigmóide da logística, porém passa a assumir valores no intervalo aberto $(-1, 1)$.

2.11 Algoritmo Genético

O Algoritmo Genético – (AG) ou em inglês Genetic algorithm – (GA) é uma classe de algoritmos de busca estocásticos inspirados no processo natural de evolução biológica. Eles operam usando técnicas derivadas da genética, como seleção, cruzamento e mutação. Esses algoritmos são tipicamente aplicados a problemas de otimização e busca, nos quais as soluções candidatas são tratadas como indivíduos em uma população, e esses indivíduos evoluem ao longo de várias gerações para produzir melhores soluções (RUSSELL; NORVIG, 2022; HOLLAND, 1992).

Como os AG trabalham a otimização de um conjunto de soluções, então eles são mais rápidos que um algoritmo normal de busca trabalhando uma solução por vez. Segundo (LUCAS, 2002), por causa do carácter paralelo e aleatório do AG, não é garantido encontrar a melhor solução em um tempo curto, e sim soluções muito próximas a solução ótima. O cromossomo é um vetor de caracteres pertencentes a um alfabeto, usado para armazenar as características dos indivíduos da população, em que cada posição é chamada de gene, e os caracteres que podem ocupar uma determinada posição são chamados de alelos. O alfabeto usado pode ser o binário, real, inteira, letras do alfabeto, entre outras. A interpretação das características presentes em um cromossomo depende da resolução modelada para o problema proposto.

Os módulos de uma algoritmo genético padrão estão ilustrado na figura 2.4:

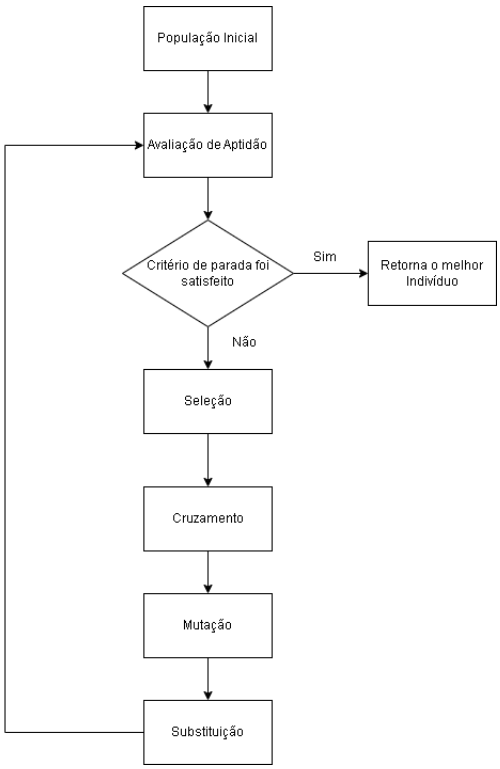


Figura 2.4: Exemplo de Estrutura Básica de um AG Simples. Fonte: (Autor, 2024)

Outro ponto importante, é a definição do tamanho da população, pois uma população muito pequena pode prejudicar o desempenho do AG, devido a baixa cobertura do espaço de busca do problema, por causa da baixa diversidade populacional. Já uma população muito grande, possui maior diversidade, mas necessita de mais tempo e recursos computacionais (KIM; HAN, 2000). Por último, a importância de gerar aleatoriamente os indivíduos da primeira população, é defendida por (OLIVEIRA ROSA; LUZ, s.d.), na qual afirma que, essa aleatoriedade garante uma exploração melhor do espaço de busca em decorrência da variedade genética, caso contrário, o algoritmo poderá gerar vários indivíduos com o material genético parecido em um curto intervalo de tempo, tendo como consequência uma convergência genética prematura em torno de uma solução ótima local. Dessa maneira, dificultando a exploração do espaço de busca.

Como mostrado, o AG é inspirado no processo evolutivo, logo, seus módulos tentam simular o comportamento dos módulos descritos pelo processo de evolução genética, que são o cruzamento, a mutação, a seleção e a substituição. A seguir, são detalhados os tipos de especificações selecionados neste trabalho para cada um deles, respectivamente.

2.11.1 Seleção

A seleção, assim como é observado na Teoria da Evolução de Charles Darwin, afirma que o indivíduo mais bem adaptado ao ambiente têm maior chance de sobreviver, e também de propagar seu material genético pelas próximas gerações através de seus descendentes (POZO et al., 2005).

Desse modo, existem vários tipos de seleção. Neste trabalho, será implementada somente a Seleção por Roleta. A escolha dessa seleção baseia-se na política de gerenciamento da diversidade da população adotada. A política é dar chances a todos os indivíduos de gerar descendentes, contribuindo para o aumento da variabilidade genética da população e, consequentemente, de sua diversidade. Ela é baseada na proporção da adaptação dos indivíduos de uma população, ou seja, a média esperada para a participação de um indivíduo nas gerações futuras é diretamente proporcional ao valor da adaptação desse indivíduos e inversamente proporcional ao valor da adaptação total da população (EIBEN; SMIT, 2011).

Segundo (GIGUERE; GOLDBERG, 1998) essa roleta pode ser implementada por meio de um vetor de tamanho m, onde cada indivíduo i de uma população p ocupa

$$\frac{\text{adaptação}(i) \times m}{\text{adaptaçãoTotal}(p)}$$

células do vetor roleta. Depois de criada a roleta, esta é girada tantas vezes, até que o número de indivíduos ancestrais seja igual ao esperado.

Na tabela 2.1, encontra-se uma população hipotética, com o ID, o seu genótipo, o score e o respectivo percentual de adaptabilidade.

Tabela 2.1: População Hipotética, Fonte (TEIXEIRA, 2005)

ID	Indivíduos	Fitness	Porcentagem do Total
1	01101	169	14,4%
2	11000	579	49,2%
3	01000	64	5,5%
4	10011	364	30,9%
Total		1170	100,0%

A figura 2.5, mostra como fica a roleta construída de acordo com o valor de score da população hipotética

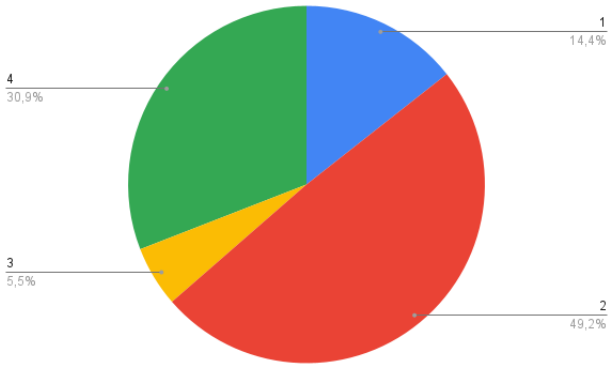


Figura 2.5: Roleta População Hipotética, Fonte: (Autor, 2024)

2.11.2 Cruzamento

O cruzamento pode ser chamado de *crossover*, ele é inspirado na reprodução sexuada, que tem o intuito de permitir a propagação de características para as gerações posteriores, que para acontecer precisa de um conjunto de progenitores cuja cardinalidade é maior ou igual a 2, para gerar um conjunto de descendentes (OLIVEIRA ROSA; LUZ, s.d.).

O funcionamento do cruzamento pode ser assim descrito, primeiro é selecionado os progenitores entre os indivíduos da população atual, para cada conjunto de progenitores uma taxa de cruzamento é gerada aleatoriamente, chamada de probabilidade de cruzamento, que irá determinar se haverá a troca ou não de características entre eles. Caso aconteça o cruzamento, pode gerar um ou mais novos indivíduos, os quais poderão ou não fazer parte da próxima população dependendo da sua adaptação e a sua validade dentro do escopo do problema, caso ele seja considerado inválido pode ser eliminado ou pode ter seus genes alterados, no intuito de promover a adaptação desses indivíduos.

No Cruzamento de Um Ponto de Corte, é gerado primeiramente um número aleatório dentro do tamanho do cromossomo um ponto de corte, ou seja, no intervalo fechado de [0, t], em seguida é copiado a primeira parte do gene do progenitor 1 da posição inicial até o ponto de corte e depois é copiado a segunda parte do progenitor 2 até o final do cromossomo, é feito então o inverso para gerar o segundo indivíduo (LUCAS, 2002), como mostrado na figura 2.6, onde a linha tracejada representa o ponto de corte escolhido aleatoriamente, nesse caso foi a posição 3.

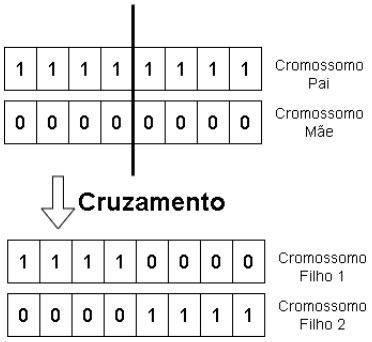


Figura 2.6: Exemplo de Cruzamento Um Ponto de Corte, Fonte: (Autor, 2024)

2.11.3 Mutação

Para produzir a variabilidade genética que ocorre no surgimento de um novo indivíduo, como no ambiente, por contato com radiação que realiza modificações no indivíduo ou por erros na

troca de informações genes durante a reprodução sexuada, o algoritmo genético faz uso da operação de mutação (NUNES, 2018).

Neste trabalho, será adotada a mutação *Flip*, que tem como ideia principal realizar a mudança de um gene aleatório de acordo com o valor específico que esse gene pode assumir. Funciona da seguinte forma: escolhe-se aleatoriamente um gene G e, em seguida, seleciona-se aleatoriamente um valor sorteado do alfabeto válido (LUCAS, 2002), conforme representado na Figura 2.7.

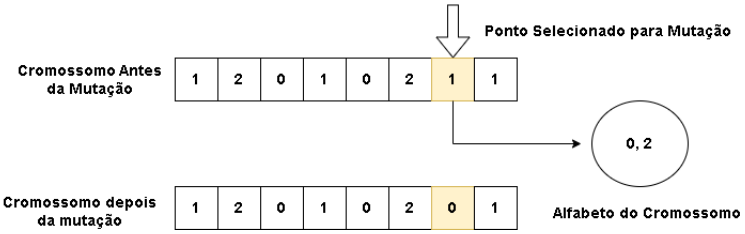


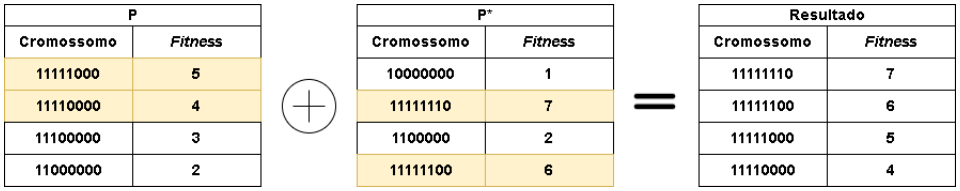
Figura 2.7: Exemplo de Mutação *Flip*, Fonte: (Autor, 2024)

2.11.4 Substituição

A substituição tem como objetivo definir quais serão os indivíduos da população que irão permanecer nela ou que serão retirados. Desse modo, é garantido uma evolução dos indivíduos dentro da população (ARABAS; MICHALEWICZ; MULAWKA, 1994).

A substituição Elitismo, segundo (EBERHART; SIMPSON; DOBBINS, 1996), afirma que o indivíduo de maior adaptação de uma geração pode não sobreviver às operações de seleção, recombinação e mutação. Sendo assim, utilizar esse tipo de substituição é uma boa forma de perpetuar um bom indivíduo nas populações futuras. Funciona da seguinte forma: ordena uma população formada pela união da população atual, população de descendentes e copia os n melhores indivíduos dessa população para fazerem parte da próxima população.

Exemplo na figura 2.8, em que ilustra uma população P de ancestrais e uma população P^* de descendentes, ambas de tamanho 4. Depois, é juntado essas populações e ordenados em ordem decrescente mantendo vivos os 4 primeiros indivíduos vivos para a próxima geração, como mostrado na população Resultante.



P	
Cromossomo	*Fitness*
11111000	5
11110000	4
11100000	3
11000000	2

P*	
Cromossomo	*Fitness*
10000000	1
11111110	7
1100000	2
11111100	6

Resultado	
Cromossomo	*Fitness*
11111110	7
11111100	6
11111000	5
11110000	4

Figura 2.8: Exemplo de Substituição Elitista, Fonte: (Autor, 2024)

Capítulo 3

Revisão da Literatura e Análise Bibliométrica

Esta Revisão Sistemática da Literatura – (RSL) tem como objetivo identificar trabalhos na literatura que utilizam abordagens de *ensemble* para a previsão de Índices do Mercado de Ações. A população de interesse neste estudo inclui Índices de Mercado de Ações, tendo processos ou modelos de *ensemble* como o contexto principal. Os trabalhos analisados abrangem o período de 2017 até o primeiro semestre de 2023, concentrando-se em artigos de periódicos e trabalhos de conferências. A escolha por artigos de periódicos e trabalhos de conferências é justificada por duas razões principais: a primeira relaciona-se ao impacto significativo que trabalhos publicados nessas categorias têm na comunidade acadêmica e na sociedade; a segunda deve-se à qualidade e ao detalhamento mais rigoroso dos artigos de periódicos, que passam por um processo de revisão editorial e avaliação mais criteriosa para publicação.

A metodologia adotada para definir o protocolo desta RSL foi estruturada em quatro etapas principais: (1) definição das questões de pesquisa; (2) estabelecimento das estratégias de busca; (3) criação do processo de seleção de artigos; e (4) extração dos dados relevantes para responder às questões de pesquisa. Nas subseções seguintes, foi detalhado cada uma dessas etapas.

3.1 Questões de Pesquisa

Para atingir o objetivo desta RSL, foi formulado a seguinte questão de pesquisa geral: Quais são as soluções associadas ao problema de prever o Índice do Mercado de Ações que empregam modelagem de *ensemble*? Para abordar essa questão de pesquisa geral, subdivide a RSL

em quatro subquestões que investigam aspectos específicos do uso de *Ensemble* no domínio financeiro. As questões de pesquisa são as seguintes:

- **RQ1:** Quais são os locais de publicação e a nacionalidade do primeiro autor?
- **RQ2:** Quais modelos de *ensemble* são implementados, quais conjuntos de dados são utilizados e quais as principais métricas de avaliação empregadas?
- **RQ3:** Qual o propósito das soluções identificadas?
- **RQ4:** Quais são as lacunas e oportunidades de pesquisa e desenvolvimento para *ensemble* no domínio financeiro?

3.2 Protocolo da Revisão Sistemática da Literatura

A revisão sistemática da literatura segue um protocolo rigoroso, conforme recomendado por várias fontes da literatura. Segundo (BRERETON et al., 2007), a revisão sistemática na área de engenharia de software precisa de adaptações para lidar com as características específicas do domínio, como a falta de padronização dos dados e a escassez de estudos empíricos em algumas subáreas. Para garantir a qualidade e relevância dos resultados, é essencial definir claramente as etapas da revisão, desde a formulação das questões de pesquisa até a seleção criteriosa dos estudos primários.

3.2.1 Termos de Pesquisa

Define o PICOC, representando *Population*, *Intervention*, *Comparison*, *Outcome*, and *Context*, com base no objetivo desta revisão. O resultado é o seguinte:

- ***Population:*** Índices do Mercado de Ações;
- ***Intervention:*** *Ensemble*;
- ***Comparison:*** Abordagens das técnicas de *Ensemble*;
- ***Outcome:*** Previsão;
- ***Context:*** Séries Temporais Financeiras.

Tabela 3.1: Palavras-chave e Sinônimos

Palavras-chave	Sinônimo	Relacionado
ensemble	*hybrid model* *hybrid system*	*Intervention*
prediction	*forecasting*	*Outcome*
stock Índice	*stock exchange*	*Population*

As palavras-chave e sinônimos, bem como, o termo relacionado aos critérios PICOC estão representados na Tabela 3.1.

A *String* de pesquisa foi gerada de acordo com os critérios PICOC: *("stock indice"OR "stock exchange") AND ("ensemble"OR "hybrid model"OR "hybrid system") AND ("prediction"OR "forecasting").*

3.2.2 Fonte de Dados

Foi Explorado três bases de dados acadêmicas para conduzir a pesquisa: Web of Science, IEEE Xplore e Scopus. Essas fontes de dados foram selecionadas com base em (HENRIQUE; SOBREIRO; KIMURA, 2019) como as principais fontes de publicações especializadas para pesquisa relacionada aos temas do estudo proposto. Publicações anteriores como as revisões apresentadas por (KUMBURE et al., 2022) e (KEHINDE; CHAN; CHUNG, 2023), também optaram por essas bases específicas como fontes de informação. A busca foi inciada por artigos em junho de 2023.

3.2.3 Processo de Seleção

Dado o propósito desta RSL, realizamos uma seleção preliminar dos artigos retornados após a execução da string de busca nas fontes de pesquisa mencionadas. Portanto, os critérios de inclusão e exclusão visam garantir que os estudos incluídos na revisão sejam relevantes para a questão de pesquisa e atendam aos padrões de qualidade desejados.

Os critérios adotados na Tabela 3.2 foram os seguintes. Primeiramente, incluímos estudos em inglês, estudos primários[1] e revisados por pares, para garantir a qualidade, credibilidade e confiabilidade dos resultados obtidos nos artigos. Nos critérios de exclusão, primeiramente foram removidos artigos duplicados, ou seja, aqueles presentes em duas ou mais plataformas

[1]Envolve a coleta de dados originais diretamente da fonte, o que significa que o pesquisador está diretamente envolvido na geração de novos dados ainda não processados (VIERA, 2023)

de indexação, o que pode ocorrer quando um periódico tem várias parcerias. Estudos não publicados, ou seja, aqueles em fase final de publicação, mas ainda não indexados nas bases de dados, e estudos anteriores a 2017 também foram excluídos. Estudos secundários[2] e terciários[3] não são o foco deste artigo, tampouco estudos de literatura cinzenta, como capítulos de livros, dissertações e teses. Por fim, foram excluídos artigos curtos, pois geralmente carecem de profundidade devido à sua brevidade, o que geralmente não permite uma exploração suficiente do tema.

Tabela 3.2: Critérios de Inclusão e Exclusão

Critérios	
Inclusão	**Exclusão**
Estudos em inglês	Estudos duplicados
Estudos primários	Estudos não publicados
	Estudos antes de 2017
	Estudos secundários e terciários
	Literatura cinzenta
	Short Papers

Para avaliar a qualidade e a relevância dos artigos selecionados, foram adotadas as seguintes perguntas:

1. Os métodos de pré-processamento de dados são claramente descritos?
2. A base de dados é acessível e o repositório do algoritmo está disponível?
3. A metodologia da abordagem é claramente definida?
4. A abordagem foi discutida no contexto dos problemas do Índice do Mercado de Ações?
5. Os métodos de divisão do conjunto de dados são descritos?
6. Funciona com conjuntos de dados de diferentes tipos de mercados?

Cada pergunta foi elaborada com o intuito de verificar aspectos específicos do estudo: desde o tratamento dos dados até a generalização da metodologia para diferentes mercados.

[2] Analisam e interpretam dados já coletados por outros. Dados secundários são derivados de estudos primários e geralmente são usados para fornecer uma perspectiva mais ampla ou reinterpretar informações existentes (VIERA, 2023).

[3] Compilam e resumem informações de fontes primárias e secundárias. Esses estudos são úteis para fornecer uma visão geral sobre um tópico, mas não apresentam novas análises ou interpretações originais (VIERA, 2023).

Um *score* foi estabelecido para cada pergunta, em que "Sim"corresponde a um ponto, "Parcial"a meio ponto e "Não"a zero. O *score* máximo possível é seis, e o mínimo é zero. Foi definido um ponto de corte de 3,5, indicando que os artigos devem atingir essa pontuação para avançar para a etapa de extração de dados.

3.2.4 Extração de Campos Relevantes

A etapa final desta Revisão Sistemática da Literatura envolve a extração de informações relevantes dos textos completos dos artigos selecionados. Nesta fase, as informações são analisadas para responder às questões de pesquisa e coletar dados demográficos dos artigos. Definimos 13 campos para a extração de dados: título, autores, ano de publicação, país do primeiro autor, nome do periódico ou conferência, palavras-chave, contexto, tipo de abordagem de *ensemble*, métricas de avaliação, experimentos, objetivos, quantidade de citações e conjuntos de dados.

3.3 Resultados e Discussões

3.3.1 Visão Geral

A Figura 3.1 ilustra o processo realizado passo a passo. A busca nas três bases de dados resultou em 203 artigos, dos quais 60 eram duplicados, restando 143 artigos submetidos à aplicação de filtros. Após essa etapa, 40 artigos foram excluídos, deixando 105 para serem avaliados com base no título e no resumo, verificando se o artigo estava dentro do escopo da pesquisa sobre o uso de abordagens de *ensemble* e índices do mercado financeiro para predição. Destes, 46 foram excluídos. Por fim, restaram 58 trabalhos para leitura detalhada e aplicação dos critérios de elegibilidade, resultando na exclusão de 5. Os 53 artigos restantes avançaram para a etapa de extração de dados.

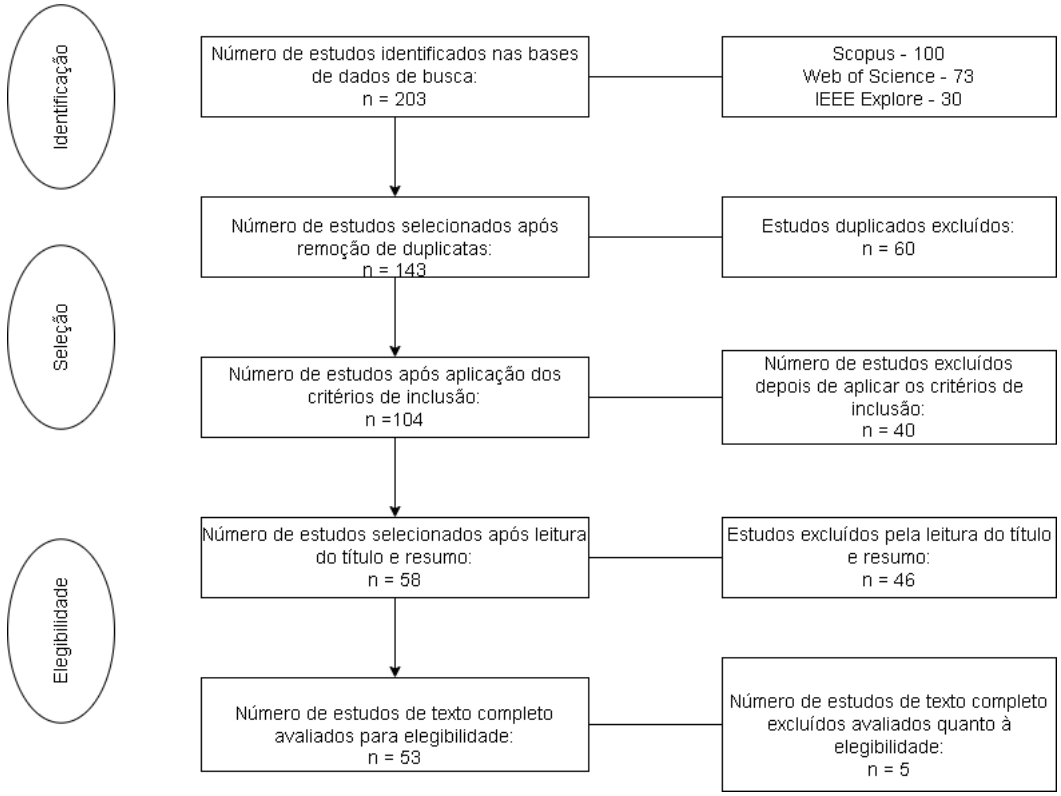


Figura 3.1: Visão geral do processo de seleção e extração de dados. Fonte: (Autor, 2024)

A Figura 3.2 apresenta uma análise de palavras-chave utilizadas pelos autores dos 53 artigos. A análise realizada com a ferramenta WordClouds indica que todos os artigos selecionados empregam abordagens de *ensemble* no contexto financeiro. Muitos artigos exploram métodos de decomposição e *Deep Learning* para analisar séries temporais financeiras, revelando quais técnicas de *ensemble*, modelos de *machine learning* e *deep lerning* são frequentemente mencionadas nas palavras-chave dos artigos.

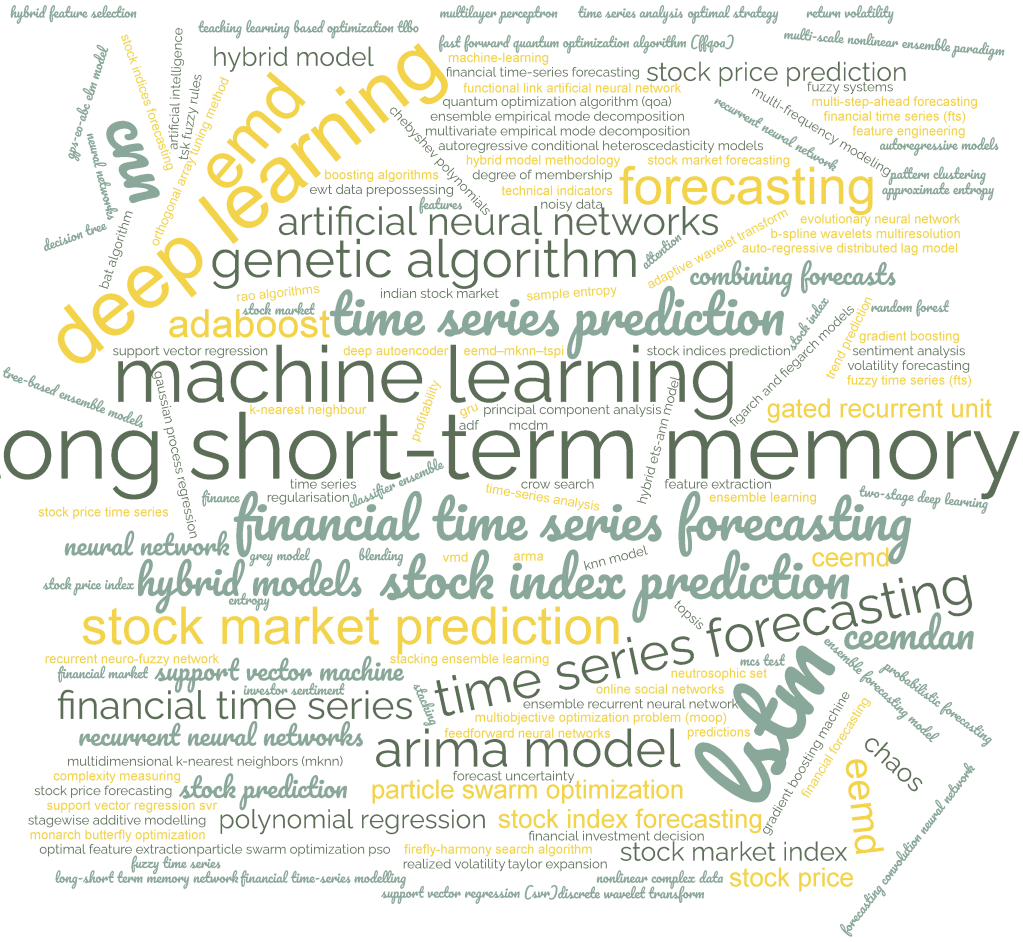


Figura 3.2: Análise das palavras-chave extraídas dos 53 artigos selecionados, Fonte: (Autor, 2024)

Os resultados da avaliação de qualidade na Figura 3.3. As pontuações de qualidade dos artigos variaram de 3,5 a 5,5 em uma escala de Likert de 6 pontos, estendendo-se de 0,0 (mínimo) a 6,00 (máximo). É observado que nenhum artigo alcançou a pontuação máxima, principalmente devido à falta de disponibilidade do repositório de algoritmos e à limitada generalização para diferentes mercados. Consequentemente, uma minoria de artigos recebeu pontuações de 3,5, enquanto a maioria apresentou pontuações de qualidade média a alta, variando entre 4,0 e 5,5 pontos, com o limite máximo sendo 6,00 pontos.

No final, 53 trabalhos foram submetidos à análise bibliométrica com o objetivo de responder às questões propostas na Subseção 3.1. Para isso, utilizou as ferramentas VOSviewer (VAN ECK; WALTMAN, 2014) e Google Sheets[4].

[4]https://www.google.com/sheets/about/

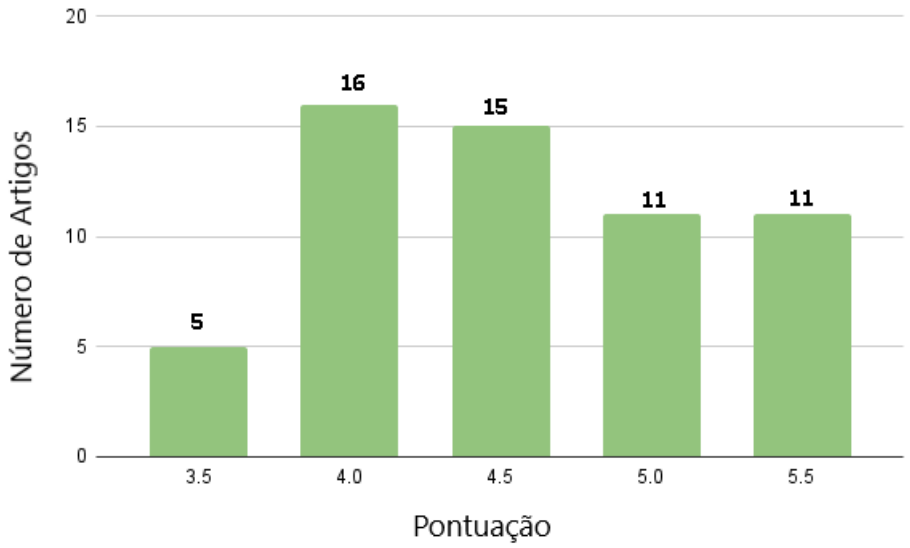


Figura 3.3: Visão geral das pontuações de avaliação de qualidade para os 58 artigos, Fonte: (Autor, 2024)

3.3.2 RQ1: Nacionalidade do primeiro autor e locais de publicação

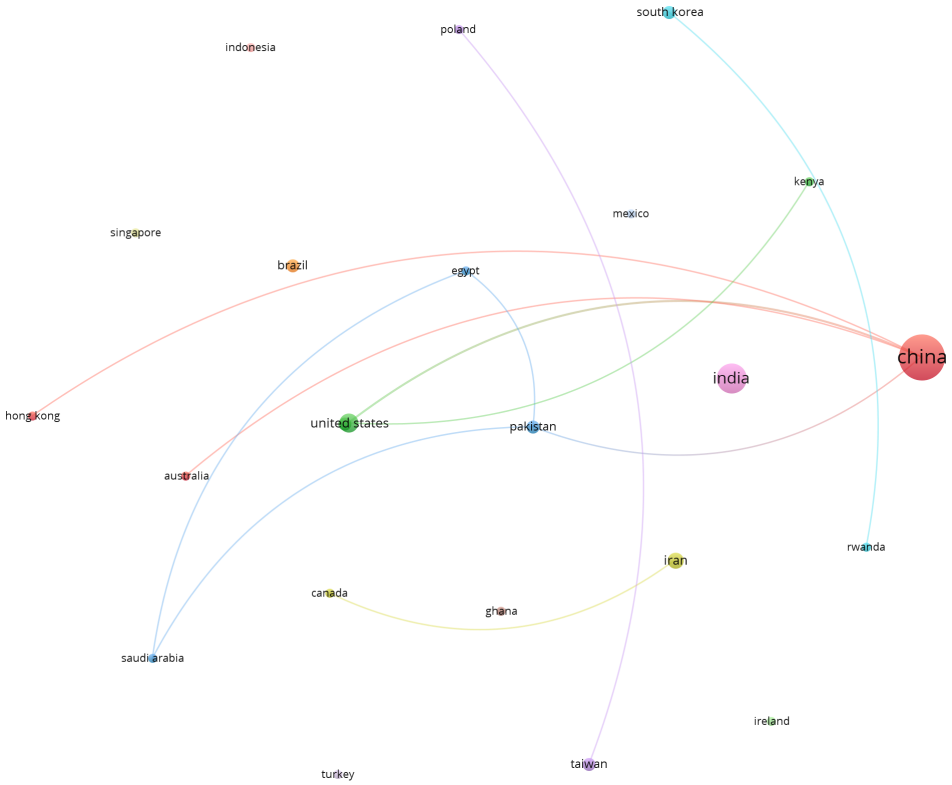


Figura 3.4: Distribuição dos estudos por país do primeiro autor. Cada balão representa o número de artigos identificados na revisão. Fonte: (Autor, 2024)

A Figura 3.4 ilustra a distribuição geográfica dos primeiros autores nos 53 estudos analisados. Cada círculo representa a quantidade de artigos identificados pela nacionalidade do pri-

meiro autor, e as conexões entre eles indicam colaborações internacionais, revelando uma forte predominância de publicações de autores da China e da Índia. Além disso, nota-se que a China é um ponto central nas colaborações, com conexões importantes com países como os Estados Unidos, Arábia Saudita, Austrália, entre outros. Esse padrão sugere que a China desempenha um papel de liderança no campo, tanto em termos de produção acadêmica quanto em parcerias internacionais. Outros países, como a Índia e os Estados Unidos, também apresentam colaborações significativas, refletindo um aumento na produção de conhecimento científico globalmente distribuído. Essa análise é crucial, pois demonstra a importância das colaborações internacionais para o avanço da pesquisa científica, com destaque para a interconexão entre países em desenvolvimento, desenvolvidos e pouco desenvolvidos.

Foi aprofundado nesta discussão a análise os conjuntos de dados utilizados para desenvolver essas soluções, visto que os dados de treinamento têm um impacto considerável na generalização dos cenários de aplicação para a previsão de séries temporais financeiras. Esta análise também proporciona *insights* sobre a nacionalidade dos pesquisadores envolvidos nesta área de estudo. Com base nessas informações, observa-se uma diversidade geográfica dos autores, contudo, as conexões entre eles não são fortemente estabelecidas, conforme evidenciado na figura.

Dos 53 artigos selecionados, foi analisado a distribuição dos estudos conforme o local de publicação, autor e número total de citações. Foi Identificado uma concentração significativa de artigos na área de inteligência artificial financeira, particularmente no periódico *Expert Systems With Applications*. A maioria dos periódicos e conferências aborda diversos cenários de aplicação, incluindo *IEEE Access* e *Complexity*, ambos com três publicações cada. Uma visão geral da distribuição dos estudos é apresentada na Tabela 3.3. Notavelmente, o trabalho de (REZAEI; FAALJOU; MANSOURFAR, 2021) destaca-se com o maior número de citações, seguido pelo artigo de (NTI; ADEKOYA; WEYORI, 2020a), que ocupa o segundo lugar e representa uma das contribuições mais relevantes no campo.

Tabela 3.3: Local de publicação e quantidade de artigos selecionados na revisão até junho de 2023

Local de Publicação	Referência	Citações	Total
Expert Systems with Applications	(REZAEI; FAALJOU; MANSOURFAR, 2021)	107	8
	(ZHOU, F. et al., 2019)	99	
	(HAJIABOTORABI et al., 2019)	51	
	(ZOLFAGHARI; GHOLAMI, 2021)	41	
	(LIN, G.; LIN, A.; CAO, 2021)	33	
	(LIN, Y. et al., 2022)	24	

Table 3.3 – continuação da página anterior

Local de Publicação	Referência	Citações	Total
	(FURLANETO et al., 2017)	17	
	(LV; SHU, Y. et al., 2022)	12	
Applied Sciences (Switzerland)	(SONG; CHOI, 2023)	1	3
	(ALI et al., 2023)	6	
	(QI; REN; SU, 2023)	0	
Complexity	(WIDIPUTRA; MAILANGKAY; GAUTAMA, 2021)	15	3
	(WU, J.; ZHOU, T.; LI, T., 2020)	12	
	(PENG et al., 2021)	4	
IEEE Access	(YU et al., 2020)	35	3
	(JI; LIEW; YANG, L., 2021)	31	
	(SHU, W.; GAO, 2020)	14	
Economic Computation and Economic Cybernetics Studies and Research	(DONGHWAN et al., 2020)	7	2
	(JUJIE; DANFENG, 2018)	8	
Entropy	(LV; WU, Q. et al., 2022)	8	2
	(LIU et al., 2022)	2	
Information Sciences	(SINGH, P., 2021)	34	2
	(DENG et al., 2022)	30	
Applied Soft Computing	(WANG, J. et al., 2021)	3	2
	(DASH et al., 2019)	35	
Algorithms	(ZHENG et al., 2021)	9	1
Applied Intelligence	(NIU; XU; WANG, W., 2020)	83	1
Computational Economics	(DAS; NAYAK; SAHOO, 2022)	5	1
Computational Management Science	(AGAPITOS; BRABAZON; O'NEILL, 2017)	8	1
Computers in Industry	(NAYAK, 2019)	6	1
Data and Knowledge Engineering	(CHEN, W. et al., 2018)	57	1
Engineering Applications of Artificial Intelligence	(ALHNAITY; ABBOD, 2020)	34	1
Evolutionary Intelligence	(PADHI; PADHY, 2021)	4	1
Expert Systems	(LUO et al., 2021)	19	1
Forecasting	(AMPOUNTOLAS, 2023)	1	1
IEEE Transactions on	(LIMA SILVA et al., 2019)	48	1
Informatica (Slovenia)	(AMPOMAH; QIN; NYAME; BOTCHEY, 2021)	9	1
Information (Switzerland)	(AMPOMAH; QIN; NYAME, 2020)	70	1
International Journal of Advanced	(SINGH, A. et al., 2020)	1	1
International Journal of Computational Intelligence Systems	(PARIDA et al., 2017)	4	1
International Journal of Intelligent Systems	(KUMAR, G.; SINGH, U. P.; JAIN, 2021)	16	1
International Journal of Intelligent Systems and Applications in Engineering	(DURAIRAJ, M.; BH, 2021)	1	1
Journal of Big Data	(NTI; ADEKOYA; WEYORI, 2020a)	103	1
Journal of Systems Science and Information	(YI SUN QINGSONG SUN, 2022)	1	1
Mathematics	(HE, K. et al., 2023)	4	1
Mobile Information Systems	(CHEN, J.; YANG, H., 2022)	1	1
Neural Computing and Applications	(DURAIRAJ, D. M.; MOHAN, 2022)	15	1
Neural Processing Letters	(CORBA; EGRIOGLU; DALAR, 2020)	13	1
Scientific Programming	(XIAO; SU et al., 2022)	7	1
Proceedings - IEEE 2nd International Conference on Artificial Intelligence and Knowledge Engineering, AIKE 2019	(PASUPULETY et al., 2019)	15	1

Table 3.3 – continuação da página anterior

Local de Publicação	Referência	Citações	Total
Open Computer Science	(NTI; ADEKOYA; WEYORI, 2020b)	49	1

3.3.3 RQ2, RQ3: Conjuntos de dados, modelos, propósitos, contexto e métricas

As Tabelas (3.5 - 3.11) fornecem informações detalhadas sobre os estudos, incluindo os conjuntos de dados utilizados, modelos empregados e métricas de avaliação aplicadas. Abreviações para algoritmos e mercados foram utilizadas para otimizar o espaço na página. Nestas tabelas, as colunas referentes a variáveis preditivas e alvo utilizam abreviações como *Análise Técnica* – (AT), em que "Preço"indica o preço de fechamento do índice, "Retorno"ao retorno do índice e "Volatilidade" à volatilidade do índice. Além disso, foram incluídas apenas informações pertinentes a produtos financeiros, ou seja, séries temporais; equações não relevantes foram omitidas durante o mapeamento dos artigos. Para auxiliar na compreensão das siglas, foi criada uma tabela relacionada aos mercados abordados A.1.

A análise destas tabelas (3.5 - 3.11) é crucial para entender como a seleção de dados para treinamento e teste pode impactar a capacidade de generalização dos modelos implementados. Pesquisas realizadas nas bases de dados *Web of Science*, *IEEE Xplore* e *Scopus* revelaram uma escassez de artigos em periódicos e conferências que focam na disponibilidade de conjuntos de dados, especialmente para mercados em diferentes estágios de desenvolvimento. Esta limitação na disponibilidade de dados mais abrangentes compromete a criação de soluções aplicáveis de forma mais ampla, particularmente ao se empregar técnicas de *ensemble*.

É essencial discutir como essas abordagens de *ensemble* são implementadas e as diversas formas de aplicação no mercado financeiro. A Figura 3.5 destaca o foco das técnicas de *Ensemble* nos 53 artigos analisados. A Tabela 3.4 apresenta uma classificação dos artigos, descrevendo os conceitos abordados e o número de trabalhos relacionados a cada categoria.

Tabela 3.4: Classificação dos artigos, conceito e número de trabalhos

Tipo de Classificação	Conceito	Número de Artigos	Referências
Análise de Sentimento	Utiliza análise de sentimento de investidores em conjunto com modelos de inteligência artificial para aumentar a robustez das previsões.	3	(CHEN, W. et al., 2018; PASUPULETY et al., 2019; YI SUN QINGSONG SUN, 2022)

Continua na próxima página

Tabela 3.4 – continuado da página anterior

Tipo de Classificação	Conceito	Número de Artigos	Referências
Completo	A abordagem proposta inclui passos como decomposição, otimização e uso de algoritmos de inteligência artificial nas séries temporais.	11	(WU, J.; ZHOU, T.; LI, T. 2020; ZOLFAGHARI; GHOLAMI, 2021; LUO et al., 2021; YAN; AASMA et al., 2020; YU et al., 2020; LIU et al., 2022; QI; REN; SU, 2023; LV; SHU, Y. et al., 2022; WANG, J. et al., 2021; LV; WU, Q. et al., 2022; ALHNAITY; ABBOD, 2020)
Decomposição	Explora abordagens de decomposição em conjunto com modelos para séries temporais.	10	(NIU; XU; WANG, W., 2020; FURLANETO et al., 2017; ZHOU, F. et al., 2019; SHU, W.; GAO, 2020; LIN, Y. et al., 2022; HAJIABOTORABI et al., 2019; LIN, G.; LIN, A.; CAO, 2021; DENG et al., 2022; ALI et al., 2023; REZAEI; FAALJOU; MANSOURFAR, 2021)
Ensemble Learning	Foca no uso de técnicas de *Ensemble Learning* nos algoritmos de inteligência artificial.	5	(NTI; ADEKOYA; WEYORI, 2020a; AMPOMAH; QIN; NYAME, 2020; PADHI; PADHY, 2021; AMPOMAH; QIN; NYAME; BOTCHEY, 2021; AGAPITOS; BRABAZON; O'NEILL, 2017)
Abordagem Fuzzy	O uso da teoria Fuzzy em conjunto com abordagens de *Machine Learning* para a predição de séries temporais financeiras.	4	(SINGH, P., 2021; ZHANG; LI, L.; CHEN, W., 2021; LIMA SILVA et al., 2019; PARIDA et al., 2017)
Otimização	Envolve técnicas de otimização aplicadas às séries temporais financeiras ou ao algoritmo para aprimorar as previsões.	10	(DURAIRAJ, D. M.; MOHAN, 2022; CHEN, J.; YANG, H., 2022; JI; LIEW; YANG, L., 2021; JUJIE; DANFENG, 2018; NTI; ADEKOYA; WEYORI, 2020b; KUMAR, G.; SINGH, U. P.; JAIN, 2021; DURAIRAJ, M.; BH, 2021; SINGH, A. et al., 2020; ZHENG et al., 2021; DAS; NAYAK; SAHOO, 2022)
Sistema Híbrido	A utilização de abordagens que combinam diretamente, como um modelo inteligente em um sistema híbrido residual, um ensemble linear de modelos, ou um modelo responsável pela seleção de características e outro para aprendizado.	10	(PENG et al., 2021; DASH et al., 2019; CORBA; EGRIOGLU; DALAR, 2020; AMPOUNTOLAS, 2023; NAYAK, 2019; HE, K. et al., 2023; DONGHWAN et al., 2020; SONG; CHOI, 2023; WIDIPUTRA; MAILANGKAY; GAUTAMA, 2021; XIAO; SU et al., 2022)

É importante destacar, conforme ilustrado na Figura 3.5, que a abordagem classificada como Completa abrange 11 artigos, representando a categoria com o maior número de estudos. Além disso, as abordagens de Decomposição, Sistema Híbrido e Decomposição ficaram empatadas com 10 artigos cada. Por outro lado, as abordagens com menor representação são o *Ensemble Learning*, Fuzzy e Análise de Sentimento.

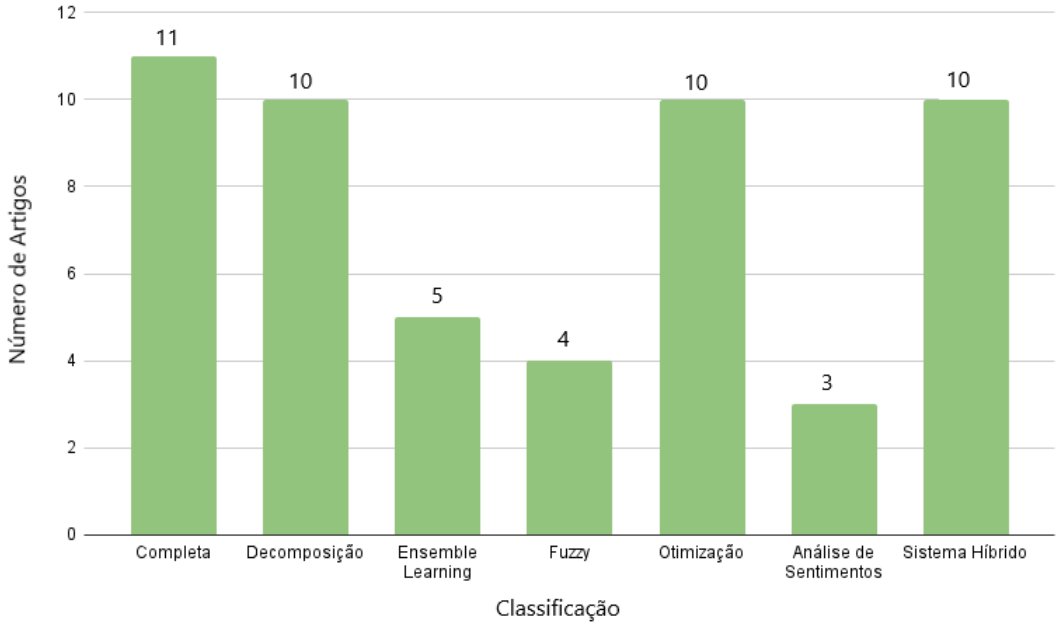


Figura 3.5: Classificação dos 53 artigos. Fonte: (Autor, 2024)

Para facilitar o entendimento das características distintas de cada uma dessas abordagens, foram estabelecidas subseções (3.4 - 3.10) dedicadas a resumir os artigos de cada macroárea. Essas subseções abordam a construção, operação e desempenho das técnicas na previsão de séries temporais financeiras.

Abordagem Completa

A Tabela 3.5 mostra que a maioria dos cenários de teste abrange mercados bem desenvolvidos como o S&P 500 e o NIKKEI 225, assim como mercados emergentes como o SSE Chinês, demonstrando a adaptabilidade dos modelos em diversos contextos econômicos. No que se refere às variável predicita, o *Preço* é predominantemente utilizada, focando principalmente na previsão de preços, embora a volatilidade e o retorno também sejam considerados. A maioria dos modelos empregados pertence à categoria de *deep learning*, com o *Long Short-Term Memory* – (LSTM) sendo o mais utilizado, seguido por *Gated Recurrent Unit* – (GRU) e *Bidirectional Long-Short Term Memory* – (Bi-LSTM). Em termos de decomposição, CEEMDAN e suas variações são as mais empregadas, com EMD e EEMD também sendo utilizadas. No aspecto de otimização, técnicas como AG e *Elite-based Opposition* – (EO) são destacadas. Modelos estatísticos como ARIMA, *Generalized Autoregressive Conditional Heteroskedasticity* – (GARCH) e suas variações, além da família *Heterogeneous Autoregressive* – (HAR) e *heterogeneous autore-*

gressive model of realized volatility – (HAR-RV), também são adotados. As principais métricas de avaliação incluem MSE, RMSE, MAE e MAPE.

Tabela 3.5: Artigos classificados como Completos

Referência	*Target*	Mercado(s)	Ativo(s)	Variável Preditiva	Prediction/s	Métodos	Métricas
(WU, J.; ZHOU, T.; LI, T. 2020)	(USD/EUR), WTI, SSE, IP	USA, China	Índice, Taxa de câmbio	Preços	Preço, Taxa de câmbio	ICEEMDAN, WOA, RVFL, RW, LSSVR, BPNN	MAPE, RMSE
(ALHNAITY; ABBOD, 2020)	FTSE, NIKKEI 225, S&P 500	USA, Japão, Inglaterra	Índice	Preço	Preço	EEMD, SA, SVR, RNN, BPNN, GA, WA, AR	MSE, RMSE, MAE, SD, R
(ZOLFAGHARI; GHOLAMI, 2021)	DJIA, IXIC	USA	Índice	AT, Dolar, Índice, WTI	Volatilidade	ANN, AWT, ARMAX, GARCH, LSTM, EGARCH, FIGARCH, HAR(3) X-RV	RMSE, MAPE, DS
(LUO et al., 2021)	S&P 500, NIKKEI 225, AORD, CSI 300	USA, China, Austrália, Japão	Índice	Preço	Preço	ARIMA, TEF, EEMD, PSR, ELM	MAE, MAPE, DS, RMSE
(YAN; AASMA et al., 2020)	SSE, S&P 500, HSI, DJIA	China, USA	Índice	Preço	Retorno	EEMD, CEEMD, PCA, LSTM, RNN	RMSE, MAE, NMSE
(LIU et al., 2022)	S&P 500, CSI 300	USA, China	Índice	Mercado de ações intradiário	Mercado de ações intradiário	FNN, LSTM, GRU, HA, CEEMD	MAE, RMSE
(QI; REN; SU 2023)	S&P 500, CSI 300	USA, China	Índice	Preço	Preço	GRU, CEEMDAN, ARIMA, Bi-LSTM, LSTM, CNN, wavelet	RMSE, MSE, MAE, R^2
(LV; SHU, Y. et al., 2022)	SSE, SZSE, HSI, Nikkei 225, DJIA, S&P 500	China, Japão, USA	Índice	Retorno	Retorno	CEEMDAN, DAE, LSTM, Bi-LSTM, EMD, PCA	RMSE, MAE, NMSE
(WANG, J. et al., 2021)	NASDAQ, DJIA, S&P 500	USA	Índice	TA	Preço	Análise de Incerteza, VMD, AE, RNN, LSTM	MAE, RMSE, MSE, MAPE, SSE
(LV; WU, Q. et al., 2022)	DAX, HSI, S&P 500, SSE	USA, Alemanha, China	Índice	AT	Preço	ADF, GRU, LSTM, EMD, ARMA, Bi-LSTM, CEEDMAN, ARIMA	MAE, RMSE, MAPE, R^2
(YU et al., 2020)	S&P 500, Letras do Tesouro, SSE	USA, China	Índice	Preço	Preço	EWT, ELM, ABC, GPS, EO, ARIMA, LSTM, ANN	RMSE, MAPE, MAE

Como representado na Figura 3.6, que visa generalizar as abordagens identificadas, variações na implementação podem ocorrer conforme o estudo específico, mas o objetivo central é facilitar a compreensão desta abrangente área. A abordagem, denominada "Completa"na Tabela 3.4, inicia com a decomposição das séries temporais, seguida por uma etapa de otimização, que pode ser aplicada diretamente na série temporal para simplificar o sinal ou para aprimorar o modelo de ML utilizado.

As séries temporais financeiras apresentam desafios significativos devido à sua natureza complexa e não estacionária, sendo influenciadas por uma variedade de fatores econômicos e financeiros (WU, J.; ZHOU, T.; LI, T., 2020). Para enfrentar esses desafios, foram introduzidos modelos híbridos inovadores. Um exemplo é o modelo MICEEMDAN-WOA-RVFL, que com-

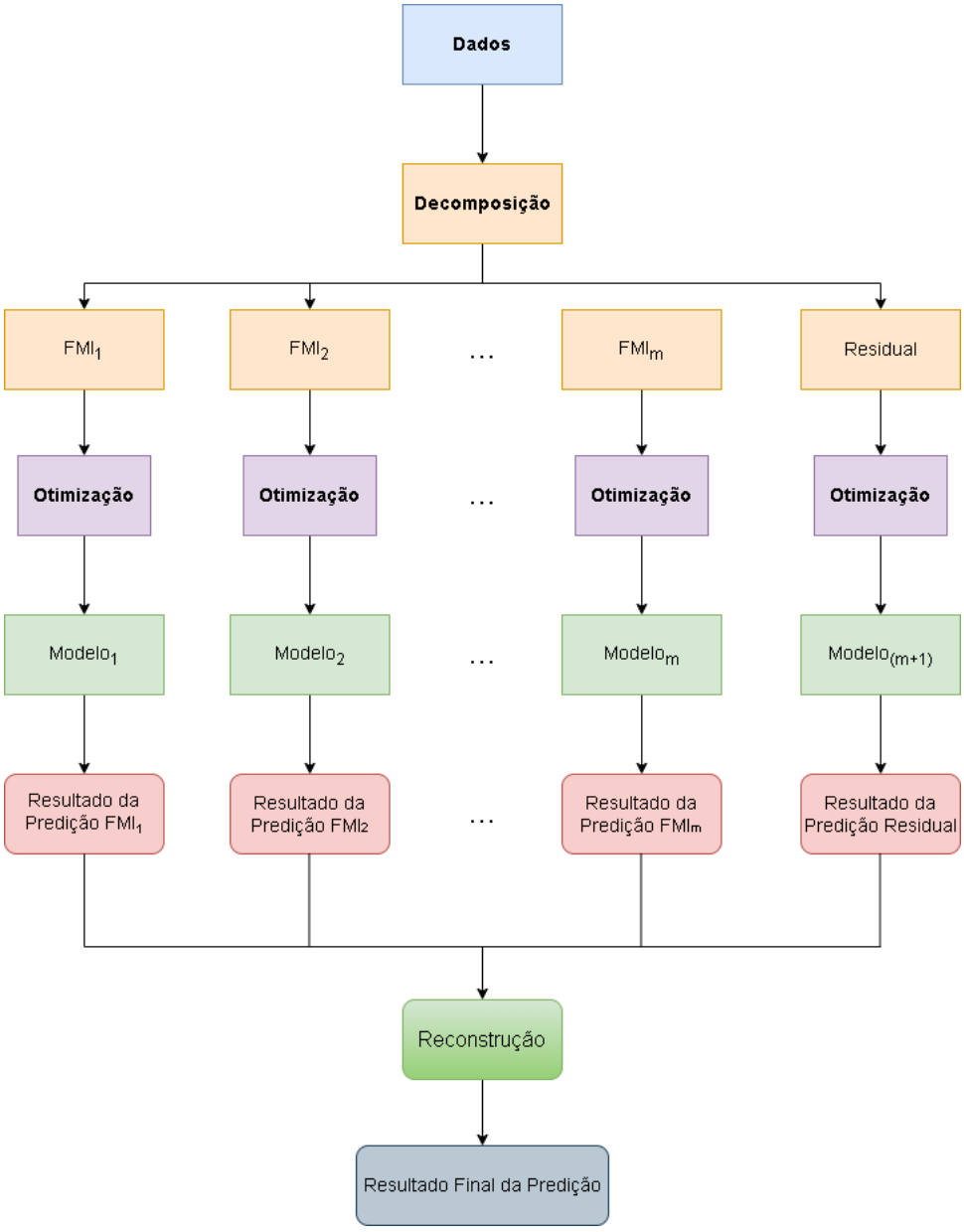


Figura 3.6: *Framework* de modelagem baseado na Abordagem Completa. Fonte: (Autor, 2024)

bina *Improved Complete Ensemble Empirical Mode Decompositions with Adaptive Noise* – (ICEEMDAN), o algoritmo *Whale Optimization Algorithm* – (WOA) e a rede *Random Vector Functional Link* – (RVFL). O ICEEMDAN decompõe a série temporal em sub-séries mais simples, as redes RVFL preveem individualmente cada sub-série, e o WOA otimiza seus parâmetros. As previsões finais são agregadas, demonstrando superioridade em precisão e estabilidade em comparação aos modelos de referência, tanto isolados quanto em conjunto, em diversas séries temporais econômicas e financeiras.

Adicionalmente, no contexto da previsão de índices de ações e volatilidade, a aplicação de modelos híbridos que incorporam variáveis financeiras tem sido explorada (ZOLFAGHARI;

GHOLAMI, 2021). O modelo AWT-LSTM-ARMAX-FIEGARCH, que combina *Adaptive Wavelet Transform* – (AWT), LSTM e modelos *Autoregressive-Moving Average With Exogenous Terms* – (ARMAX) com o GARCH, mostrou melhorias significativas na previsão de índices de ações para múltiplos horizontes temporais como 1, 10, 15, 20, 30 e 60 dias à frente. Similarmente, a combinação de AWT e LSTM tem aumentado a precisão preditiva do modelo HAR-RV para a volatilidade das ações em diversos horizontes de tempo, particularmente para DJIA e IXIC. Outro modelo de *ensemble*, que combina EEMD, ARIMA e expansão de Taylor com um diferenciador de rastreamento, foi proposto para a previsão de séries temporais financeiras (LUO et al., 2021). Esta abordagem decompõe as séries temporais financeiras em sub-séries usando EEMD, emprega modelos ARIMA para prever a parte linear de cada sub-série, e utiliza modelos de expansão de Taylor para as partes não lineares. As previsões dos modelos lineares e não lineares são então combinadas, superando os modelos de referência em previsões de séries temporais financeiras.

O trabalho de (YAN; AASMA et al., 2020) visa melhorar a precisão e a estabilidade da previsão de fenômenos financeiros. O modelo CEEMD-PCA-LSTM emprega técnicas de *Deep Learning*, *Principal Component Analysis* – (PCA) e uma versão complementar da *Complete Ensemble Empirical Mode Decomposition* – (CEEMD) para prever mercados de ações, superando consistentemente os modelos de referência em termos de precisão preditiva. Além disso, um modelo híbrido proposto por (YU et al., 2020) combina *Empirical Wavelet Transform* – (EWT), um algoritmo aprimorado por *Artificial Bee Colony* – (ABC), redes neurais de *Extreme Learning Machine* – (ELM) e o modelo ARIMA. Este modelo é projetado para eliminar o ruído e analisar as séries temporais financeiras, demonstrando um aumento no desempenho preditivo.

O estudo de (LV; SHU, Y. et al., 2022) introduz um novo modelo híbrido de *Deep Learning*, CEEMDAN-DAE-LSTM, combinando CEEMDAN com *Deep Autoencoder* – (DAE) e LSTM. A decomposição CEEMDAN organiza os dados do índice de ações em FMIs, o DAE remove dados redundantes e extrai características em níveis mais profundos, enquanto as redes LSTM preveem os retornos das ações para o próximo dia de negociação. De maneira semelhante, (QI; REN; SU, 2023) propõe uma nova estrutura que combina GRU com CEEMDAN para aumentar a precisão das previsões de índices de ações. Além disso, este modelo emprega um método de *wavelet thresholding* para remover ruídos de alta frequência em sub-sinais, mitigando assim a interferência do ruído nas previsões de dados futuros. No trabalho de (LIU et al., 2022), as séries temporais de ações intradiárias são inerentemente ruidosas e complexas, tornando-as

particularmente desafiadoras de prever. Para lidar com essa complexidade, é introduzido o modelo híbrido CEGH, composto por quatro estágios-chave. Primeiro, o CEEMD divide os dados originais do mercado de ações intradiário em diferentes FMIs. Em seguida, valores de entropia aproximada e entropia amostral são calculados para cada FMI para reduzir o ruído. As FMIs retidas são então agrupadas em quatro conjuntos, e unidades *Feedforward Neural Network* – (FNN) ou recorrentes com *History Attention* – (HA) surgindo o modelo GRU-HA, que são usadas para prever sinais abrangentes para cada grupo.

O modelo proposto por (LV; WU, Q. et al., 2022) utiliza CEEMDAN para decompor o índice de ações em FMIs com escalas características variáveis e termos de tendência. O método *Augmented Dickey Fuller* – (ADF) avalia a estabilidade de cada FMI e termo de tendência. O modelo *Autoregressive Moving Average Model* – (ARMA) é aplicado a séries temporais estacionárias, enquanto um modelo LSTM extrai características abstratas de séries temporais instáveis. Os resultados previstos de cada sequência temporal são reconstruídos para obter o valor previsto final. De maneira semelhante, a complexidade do mercado de ações como um sistema não linear levou ao desenvolvimento de um paradigma inovador de *ensemble* não linear multiescala para previsão de índices de ações e análise de incerteza. Este paradigma envolve extração de características ótimas, incluindo *Variational Mode Decomposition* – (VMD) e *Auto-encoder* – (AE), um processo de *Deep Learning* em duas etapas usando *Recurrent Neural Network* – (RNN) e LSTM, e regressão por processo gaussiano. A extração de características ótimas isola características-chave das flutuações do índice de ações e remove distúrbios. A abordagem de *Deep Learning* em duas etapas prevê sub-sinais de características individuais e os integra de forma não linear. A regressão por processo gaussiano é então usada para construir previsões intervalares para o sinal original do índice de ações e analisar incertezas do mercado (WANG, J. et al., 2021).

Em resumo, conforme demonstrado pelos estudos mencionados, a integração de várias técnicas de inteligência artificial dentro de uma única arquitetura aborda a complexidade das séries temporais financeiras, promovendo um *framework* que combina otimização, decomposição e algoritmos de ML. Embora a complexidade do modelo aumente, os resultados desses esforços indicam um ganho significativo em precisão e estabilidade das previsões.

Abordagem de Decomposição

Na Tabela 3.6, observa-se que o índice mais frequentemente utilizado é o S&P 500, seguido pelo índice DJIA, ambos representando mercados desenvolvidos. Nos mercados em desenvolvimento, o índice SSE é o mais utilizado. Em relação às variáveis preditiva, o *Preço* predomina. Quanto aos modelos de *Machine Learning*, LSTM, SVR e *Convolutional Neural Network-Long Short-Term Memory* – (CNN) são os mais utilizados. No que se refere à decomposição, o uso de EEMD e CEEMD aparece com a mesma frequência, seguidos por EMD e CEEMDAN. As principais métricas de avaliação empregadas são MSE, RMSE, MAE e MAPE.

Tabela 3.6: Artigos classificados como Decomposição

Referência	*Target*	Mercado(s)	Ativo(s)	Variável Preditiva	Prediction/s	Métodos	Métricas
(NIU; XU; WANG, W., 2020)	HSI, S&P 500, FTSE, IXIC	USA, China, Inglaterra	Índice	Preço	Preço	VMD, LSTM, ELM, CNN, BPNN, EMD	MAE, RMSE, MAPE, DS, CID
(FURLANETO et al., 2017)	ISE100_TL, ISE100_USD, S&P 500, DAX, FTSE, NIKKEI, BVSP, EU, EM	USA, Turquia, Brasil, Alemanha, Europa, Japão	Índice	Preço	Direction, Retorno	Log. Reg., EEMD, RBF, SVM, RF	Accuracy, RC
(ZHOU, F. et al., 2019)	SSE, NASDAQ, S&P 500	China, USA	Índice	Preço	Preço	ANN, FNN, EMD, WDBP	MAE, RMSE, MAPE, AAR, MD, SR, ARR/MD,
(SHU, W.; GAO, 2020)	SSE	China	Índice	Preço	Preço	LSTM, EMD, CNN, SVR	MAE, RMSE
(LIN, Y. et al., 2022)	CSI 300, S&P 500, STOXX 50	USA, China, Europa	Índice	Preço	Volatilidade	BPNN, Elman, SVR, AR, HAR, CEEMDAN	MSE, MAE, HMSE, HMAE
(HAJIABOTORABI et al., 2019)	S&P 500, NASDAQ, DJIA, NYSE	USA	Índice	Preço	Volatilidade	FNN, BSd, RNN, DWT, ARIMA, GARCH	FRMSE, FMAE, FMAPE
(LIN, G.; LIN, A.; CAO, 2021)	NASDAQ, DJIA, S&P 500, Russell 2000	USA	Índice	Preço	Preço	EEMD, KNN, TSPI	NMSE, MASE, MAPE
(DENG et al., 2022)	S&P 500, SSE, HSI	China, USA	Índice	AT	Preço	LSTM, ARIMA, BPNN, EMD, MEMD	MAPE, RMSE, DS
(ALI et al., 2023)	KSE 100	Paquistão	Índice	Preço	Preço	EMD, LSTM, EEMD, SEMD, Akima, RF, SVM, DT	RMSE, MAE, MAPE
(REZAEI; FAALJOU; MANSOURFAR, 2021)	S&P500, DJIA, DAX, Nikkei225	USA, Japão, Alemanha	Índice	Preço	Preço	CNN, LSTM, CEEMD, EMD	RMSE, MAE, MAPE

A Figura 3.7 ilustra o *framework* comumente utilizado em estudos de decomposição, em que o sinal de entrada (séries temporais financeiras) é simplificado antes de ser processado por um modelo de ML para previsão.

O estudo de (FURLANETO et al., 2017) explora o EEMD, uma abordagem de *ensemble* para o EMD, considerada uma solução potencial para melhorar a previsão de tendências. No en-

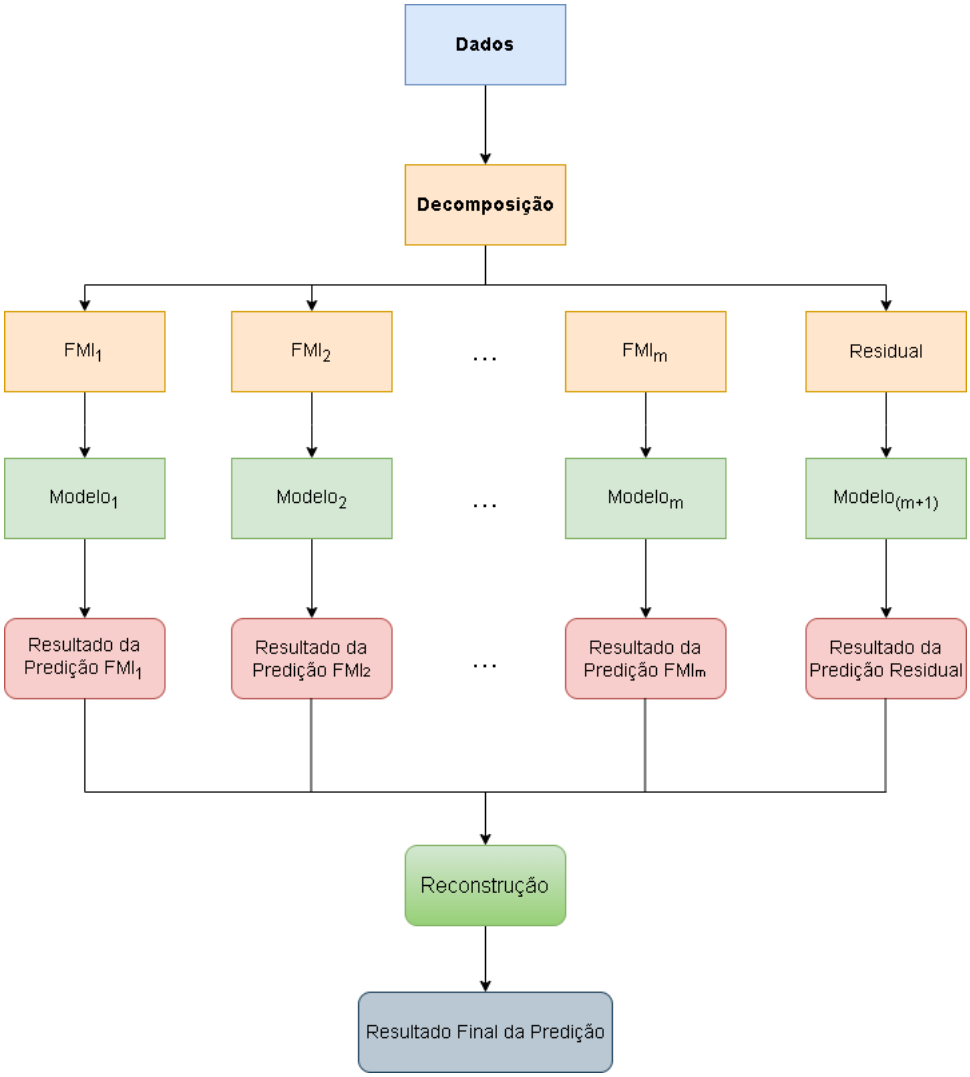


Figura 3.7: *Framework* de modelagem baseado na Abordagem de Decomposição. Fonte: (Autor, 2024)

tanto, uma avaliação crítica revela que os resultados promissores obtidos com EEMD foram influenciados por um viés de antecipação introduzido durante o pré-processamento com EMD, afetando, em última análise, os resultados das previsões. Ao contrário das conclusões anteriores, a aplicação de EMD e EEMD sozinhos não parece suficiente para melhorar a precisão ou o retorno cumulativo dos modelos na previsão de séries temporais financeiras. Por outro lado, no artigo de (ALHNAITY; ABBOD, 2020), é introduzido um novo modelo híbrido de previsão inteligente em três estágios. Este modelo combina várias técnicas de modelagem inteligente com um algoritmo de extração de características. O primeiro estágio envolve a aplicação da decomposição por modos empíricos em*ensemble* aos dados originais para facilitar o ajuste do modelo. Subsequentemente, redes neurais e SVR são usadas individualmente para modelar as características extraídas. Finalmente, uma média ponderada do *ensemble* é deter-

minada usando um algoritmo genético, contribuindo para um modelo de previsão unificado. No trabalho de (NIU; XU; WANG, W., 2020), para melhorar a precisão na previsão do índice de preço de ações, é introduzido um novo modelo híbrido, VMD-LSTM, combinando VMD e LSTM. O VMD decompõe séries complexas em sub-séries de modo de flutuação mais simples, superando problemas de mistura de modos presentes no EMD. O LSTM filtra efetivamente informações prévias cruciais, tornando-o adequado para prever séries temporais financeiras. O modelo VMD-LSTM supera significativamente modelos individuais e outros modelos híbridos baseados em EMD e VMD quando avaliados usando critérios de previsão de nível e direção, além de uma nova estatística chamada distância invariante à complexidade. Essas abordagens híbridas demonstram consistentemente superioridade na precisão da previsão em comparação com modelos únicos, com modelos baseados em VMD geralmente superando os modelos baseados em EMD.

Um modelo híbrido que combina LSTM e CEEMDAN é adotado para prever o RV do índice, e o teste de conjunto de confiança do modelo é usado. Comparativamente, *Back propagation neural network* – (BPNN), *Elman Neural Networks* – (Elman), SVR, *Simple Autoregressive* – (AR), HAR e seus híbridos com CEEMDAN também são testados. Os resultados revelam que o CEEMDAN-LSTM fornece o melhor desempenho na previsão de RV para mercados emergentes e desenvolvidos. Notavelmente, modelos híbridos superam seus equivalentes, e os resultados empíricos permanecem robustos com uma abordagem de "janela deslizante"(LIN, Y. et al., 2022). No trabalho de (SHU, W.; GAO, 2020), EMD, CNN e LSTM são combinados. O EMD decompõe a série de preços das ações originais em FMIs em diferentes frequências, e uma CNN é usada para extrair características de cada componente. Subsequentemente, uma rede LSTM modela dependências temporais entre as características, resultando em previsões. Os resultados experimentais demonstram que essa rede híbrida supera outros modelos de ponta, modelando efetivamente diferentes frequências. Por outro lado, em uma abordagem inovadora de (ZHOU, F. et al., 2019), uma nova abordagem híbrida chamada de *Empirical Mode Decomposition and Neural Network-based Machine Factorizatio* – (EMD2FNN) é introduzida. Experimentos comparativos com outros métodos, como *Artificial Neural Network* – (ANN), FNN, EMD2NN e *Wavelet De-Noising-Based Back Propagation* – (WDBP), mostram o desempenho superior do EMD2FNN. Além disso, uma análise de rentabilidade usando uma estratégia de negociação mostra que o EMD2FNN supera outros modelos em termos de e *Annual Retorno* – (AAR),*Maximum Drawdown* – (MD), *Sharpe Ratio* – (SR), and AAR/MD, tanto quanto sem

considerar os custos de transação. Esses modelos híbridos representam avanços significativos na superação dos desafios de previsão em séries temporais financeiras.

No artigo de (DENG et al., 2022), a captura das dinâmicas complexas inerentes às séries temporais do índice de preços das ações é possibilitada. A estrutura emprega uma estratégia de Múltiplas Entradas e Múltiplas Saídas, em que o *Multivariate Empirical Mode Decomposition* – (MEMD) decompõe simultaneamente as características relevantes do índice de preços das ações. Em seguida, o LSTM é usado para treinar o modelo de previsão, aproveitando os componentes extraídos pelo MEMD para fazer previsões de múltiplos passos dos preços de fechamento. Os hiperparâmetros do modelo LSTM são otimizados usando um Método de Ajuste de Array Ortogonal baseado no design experimental de Taguchi. Este modelo foi validado, demonstrando sua superioridade sobre os modelos de referência e seu aumento de precisão em previsões de múltiplos passos à frente.

Além disso, (LIN, G.; LIN, A.; CAO, 2021) propõe uma abordagem inovadora chamada EEMD–MKNN–TSPI para aprimorar a previsão de séries temporais financeiras para ações. Ao combinar o método EEMD com multidimensional *K-Nearest Neighbors* – (KNN) e *Time Series Prediction with Invarian* – (TSPI), essa abordagem melhora a funcionalidade do método KNN na previsão de séries temporais financeiras. Os resultados experimentais mostram que o EEMD–MKNN–TSPI supera tanto os modelos EEMD–MKNN quanto os MKNN–TSPI, tornando-o uma escolha promissora para prever preços de ações com benefícios potenciais para investidores e pesquisadores. Em um estudo separado, (HAJIABOTORABI et al., 2019) apresenta um modelo RNN para prever séries temporais de alta frequência, especialmente no contexto de índices de ações. O modelo combina *Discrete Wavelet Transform* – (DWT) com *B-spline Wavelet of high order d* – (BSd) para decompor séries temporais de ações intradiárias em conjuntos de dados suaves e detalhados. Esses componentes servem como entradas para o RNN. Os resultados indicam que o modelo RNN com denoising B-Spline supera outros modelos DWT-RNN comuns, e até outros modelos ANN, incluindo FNN. Isso destaca o potencial do modelo RNN com denoising B-Spline para aprimorar a capacidade preditiva dos modelos de previsão de índices de ações.

Desenvolvimentos recentes em *Deep Learning*, incluindo modelos LSTM e CNN, têm mostrado promessa em lidar com a não linearidade e alta volatilidade das séries temporais financeiras. Além disso, métodos como EMD e CEEMD, que decompõem séries temporais em diferentes espectros de frequência, oferecem ferramentas eficazes para análise de dados finan-

ceiros. Com base nesses fundamentos, novos algoritmos híbridos foram propostos, especificamente CEEMD-CNN-LSTM e EMD-CNN-LSTM. Esses algoritmos visam extrair características profundas e sequências temporais, aprimorando suas capacidades preditivas com um passo à frente. A colaboração entre esses modelos aumenta seu poder analítico, levando, em última análise, a uma maior precisão preditiva. Resultados empíricos confirmam que o uso combinado de CNN junto com LSTM e CEEMD ou EMD supera outros modelos de previsão, com o CEEMD demonstrando melhor desempenho em comparação com o EMD (REZAEI; FAALJOU; MANSOURFAR, 2021). Em uma abordagem semelhante, um novo método híbrido de *ensemble* combinando uma nova versão do EMD com LSTM foi proposto. O modelo proposto usa uma nova versão do EMD, empregando *Akima Spline Interpolation Technique* – (AKIMA), para decompor dados de ações ruidosos em vários componentes, incluindo FMIs e um único resíduo monótono. Esses subcomponentes altamente correlacionados são então usados para construir a rede LSTM. Uma avaliação abrangente do modelo híbrido em comparação com o LSTM único e outros modelos (ALI et al., 2023).

Em resumo, estudos recentes levantaram questões sobre a eficácia isolada do EMD e do EEMD em melhorar a precisão dos modelos preditivos. Por um lado, a investigação de (FURLANETO et al., 2017) sobre o EEMD sugere que seus resultados promissores foram afetados por um viés de antecipação introduzido durante o pré-processamento com EMD, lançando dúvidas sobre a eficácia isolada do EMD e do EEMD em melhorar a precisão do modelo. Por outro lado, (ALHNAITY; ABBOD, 2020) introduz um modelo híbrido inteligente de previsão em três estágios, integrando decomposição modal empírica, redes neurais, regressão por vetor de suporte e uma média ponderada do *ensemble* através de um algoritmo genético. Da mesma forma, (NIU; XU; WANG, W., 2020) propõe o modelo híbrido VMD-LSTM, abordando questões de mistura de modos no EMD e mostrando desempenho superior na previsão de preços de ações. A avaliação de (LIN, Y. et al., 2022) de modelos híbridos baseados em CEEMDAN identifica o CEEMDAN-LSTM como particularmente eficaz na previsão de volatilidade realizada em mercados emergentes e desenvolvidos. (SHU, W.; GAO, 2020) combina EMD, CNN e LSTM, destacando o desempenho superior dessa rede híbrida na modelagem de diferentes frequências dentro das séries temporais de preços de ações.

Ensemble Learning

Na Tabela 3.7, os índices estão equilibrados, com o GSE representando o mercado emergente, e S&P 500 e DJIA representando os mercados desenvolvidos. As entradas preditivas na maioria dos casos são da AT, mas a saída é focada principalmente na direção do mercado. Os principais modelos utilizados foram SVR, MLP e modelos baseados em árvores. Na parte das abordagens de *Ensemble Learning* pode ser encontrado as técnicas *Bootstrap aggregating* – (bagging), *Stacking*. As principais métricas adotadas foram Acurácia, *Recall* e Precisão.

Tabela 3.7: Artigos classificados como *Ensemble Learning*

Referência	*Target*	Mercado(s)	Ativo(s)	Variável Preditiva	Prediction/s	Métodos	Métricas
(NTI; ADEKOYA; WEYORI, 2020a)	GSE, JSE, BSE-SENSEX, NYSE	Gana, USA, África do Sul, Índia	Índice	AT	Direção, Preço	EL, DT, SVM, MLP	RMSE, MAE, R^2, RMSLE, Accuracy, Recall, Precision, AUC, MedAE, EVS, Mean, STD
(AMPOMAH; QIN; NYAME, 2020)	NYSE, NASDAQ, NSE	USA, Índia	Empresa, Índice	AT	Direção	RF, XG, EL, ET	Accuracy, precision, recall, $f1_score$, specificity
(PADHI; PADHY, 2021)	S&P 500, DJIA, HSI	USA, China	Índice	AT	Direção	SVR, RF, K-NN, LR, RidgeCv, EL	MSE, RMSE, MAE
(AMPOMAH; QIN; NYAME; BOTCHEY, 2021)	AAPL, ABT, BAC, XOM, S&P 500, MSFT, DJIA, KMX, TATASTELL, HCLTECH	USA, Índia	Índice, Empresa	AT	Direção	RF, EL, DT, ET	accuracy, f1_score, specificity, ROC curve, AUC
(AGAPITOS; BRABAZON; O'NEILL, 2017)	Gold, S&P 500, GBP/USD	USA, Inglaterra	Taxa de câmbio, Índice	Retorno	Retorno, Direção	MLP, EL, RBFNN, LSBoost, ARIMA	Sharpe ratio, retorno, variância do Retorno, matriz confusão

A Figura 3.8 exibe a construção de abordagens de aprendizado de *ensemble* juntamente com modelos de *Machine learning*. Os estudos a seguir ilustram como essas implementações ocorrem e quais mudanças podem ser feitas.

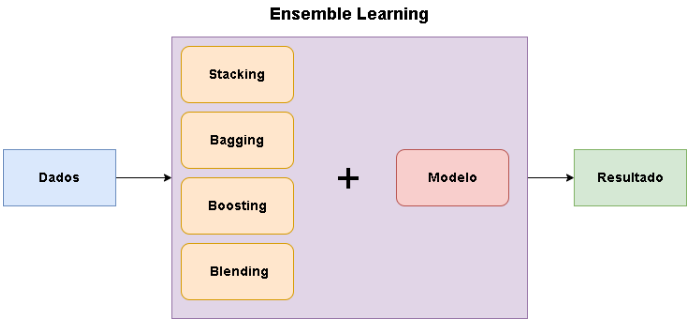


Figura 3.8: *Framework* de modelagem baseado em *Ensemble Learning*. Fonte: (Autor, 2024)

No campo da previsão do mercado de ações, os pesquisadores exploraram a utilidade dos modelos híbridos, descobrindo que a combinação de vários algoritmos de ML melhora a pre-

cisão das previsões. Um modelo de *ensemble* baseado em regressão utilizando indicadores técnicos, conhecido como *Lasso Regression* – (LR) mais o SVR e o *Ridge Regression with Built-in Cross-Validation* – (RidgeCV), demonstrou resultados impressionantes em múltiplos conjuntos de dados. Este modelo tem o potencial de fornecer *insights* valiosos para investidores e reguladores de mercado, foi trabalhado também os modelos *Random Forest* – (RF), *Linear Regression* – (LR) e técnicas de *Ensemble Learning* – (EL) (PADHI; PADHY, 2021). Outro estudo avalia modelos de ML baseados em árvores para prever movimentos de preços de ações que utilizou *XGBoost Classifier* – (XG), RF, em que o *Extra Trees Classifier* – (ET) se destacou como o melhor desempenho (AMPOMAH; QIN; NYAME, 2020). Por fim, uma análise abrangente de técnicas de *ensemble* como *boosting*, *bagging*, *blending* e *stacking*) usando modelos de *Decision Tree* – (DT), SVM e MLP, revela que *stacking* e *blending* consistentemente produzem maior precisão preditiva e valores RMSE superiores, defendendo a incorporação de métodos de ensemble na pesquisa de previsão do mercado de ações (NTI; ADEKOYA; WEYORI, 2020a). Esses estudos coletivamente destacam a importância de modelos híbridos e métodos de ensemble na melhoria da previsão do mercado de ações.

Reconhecendo a previsibilidade parcial do mercado de ações, este estudo investiga o desempenho de modelos de ML baseados em árvores em conjunto, incluindo o classificador *Bagging*, *Random Forest*, *Extra Trees*, *AdaBoost* de *Bagging*, *AdaBoost* de *Random Forest* e *AdaBoost* de *Extra Trees*. Esses modelos são avaliados usando indicadores técnicos em dados de ações de várias bolsas. O *AdaBoost* de *Bagging* emerge como o modelo de melhor desempenho no *ensemble*, aprimorando o desempenho dos modelos *ensemble* de *Bagging* (AMPOMAH; QIN; NYAME; BOTCHEY, 2021). Enquanto isso, outro estudo explora o *Gradient Boosting* – (GB) e sua suscetibilidade ao *overfitting*, mesmo com a regularização baseada em *shrinkage*, especialmente em conjuntos de dados ruidosos. A abordagem proposta introduz uma transformação baseada em função sigmoide para mitigar o impacto de *outliers* nos resíduos do GB, reduzindo assim o *overfitting* e melhorando a precisão da generalização, conforme demonstrado através de dados sintéticos ruidosos e modelagem de séries temporais financeiras do mundo real (AGAPITOS; BRABAZON; O'NEILL, 2017). Esses estudos enfatizam a importância dos modelos de *ensemble* e das técnicas inovadoras de regularização na previsão do mercado de ações, sinalizando o potencial para melhorar o desempenho e reduzir o *overfitting*.

Abordagem Fuzzy

A Tabela 3.8 apresenta como os principal índices o SSE e o S&P 500 e taxas de câmbio. Observa-se que oas variáveis preditivas predominante é o *Preço*. Entre os principais métodos utilizados, destacam-se redes neurais Fuzzy e modelos de *Machine Learning* como SVR e ANN. Em termos de métricas, as mais frequentemente selecionadas foram MSE, RMSE, MAE e MAPE.

Tabela 3.8: Artigos classificados como Fuzzy

Referência	*Target*	Mercado(s)	Ativo(s)	Variável Preditiva	Prediction/s	Métodos	Métricas
(SINGH, P., 2021)	TAIFEX, TSEC	Taiwan	Índice	Preço	Preço	GA, PSO, Fuzzy	MSE, MAPE, Theil's U statistic, Cross Entropy
(ZHANG; LI, L.; CHEN, W., 2021)	Stocks in SSE	China	Stock	AT	Preço	SVR, ANN, Fuzzy	MSE, RMSE
(LIMA SILVA et al., 2019)	S&P 500, NASDAQ, TAIEX	Taiwan, USA	Índice	Preços	Preços	Fuzzy, ARIMA, K-NN, QAR, KDE	RMSE, Winkler Score, Ranked Probability Score
(PARIDA et al., 2017)	USD, AUD, CHF, BRL, MXN, S&P 500, Nikkei 225, PJM mercado de eletricidade	USA, Japão, Brasil, Mexico, Austrália	Taxa de câmbio, Índice, Empresa	Preço, Taxa de câmbio	Preço, Taxa de câmbio	firefly-harmony search, Chebyshev polynomial functions, LRNFIS, RBFNN, Fuzzy	RMSE, MAPE, MAE, Superior Predictive Ability

A abordagem Fuzzy representada na Figura 3.9 utiliza dados de séries temporais financeiras juntamente com certos indicadores técnicos, que são processados pelo modelo Fuzzy como se fosse um especialista do mercado financeiro. Isso visa auxiliar o modelo de ML na tomada de decisões, identificando quais métricas devem ser consideradas relevantes para a previsão final.

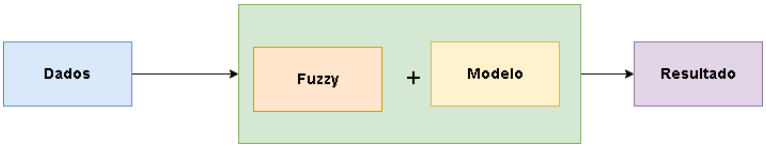


Figura 3.9: *Framework* de modelagem baseado na abordagem Fuzzy. Fonte: (Autor, 2024)

O uso de modelos de *Séries Temporais Fuzzy* – (STF) muitas vezes enfrenta desafios relacionados à seleção de um universo de discurso apropriado e à determinação do grau de relevância Fuzzy, apresentando um problema de otimização multiobjetivo. Para abordar isso, o estudo (SINGH, P., 2021) introduz um algoritmo de otimização quântica aprimorado, chamado de *Fast Forward Quantum Optimization Algorithm* – (FFQOA), para identificar de forma eficiente soluções ótimas para esses problemas. Ao integrar o FFQOA na abordagem STF, foi desenvolvido um modelo híbrido chamado de *Fuzzy-Quantum Time Series Forecasting Model* – (FQTSFM). O FQTSFM oferece uma rápida convergência em comparação com os modelos STF híbridos existentes e fornece resultados de previsão para um passo à frente.

Em outro estudo, um novo modelo de ML em dois estágios chamado SVR-ENANFIS é introduzido para previsão de preços de ações, combinando o SVR e o *Ensemble Neuro-Fuzzy Inference System* – (ENANFIS). Este modelo supera os modelos de estágio único e vários modelos de dois estágios, demonstrando sua eficácia na previsão de preços de ações (PASUPULETY et al., 2019).

O artigo de (PARIDA et al., 2017) apresenta duas abordagens inovadoras no campo da previsão de séries temporais. O modelo híbrido chamado *Locally Recurrent Neuro-Fuzzy Information System* – (LRNFIS) combina uma FNN com um sistema neural de link funcional para melhorar a previsão de séries temporais. O LRNFIS emprega funções polinomiais de Chebyshev para refinar regras Fuzzy, expandindo o espaço de entrada de forma não linear e introduzindo nós recorrentes locais para memória dinâmica e fluxo de sinal bidirecional. Ele utiliza um algoritmo de aprendizado baseado em busca harmônica Firefly aprimorada para treinamento e é validado com várias bases de dados de séries temporais do mundo real.

Além disso, o estudo de (LIMA SILVA et al., 2019) introduz uma abordagem de previsão probabilística baseada em STF que pode lidar com previsões pontuais, intervalares e de distribuição, oferecendo versatilidade, escalabilidade e baixo custo computacional. Testes empíricos com dados de séries temporais financeiras, quando comparados aos métodos STF padrão e métodos estatísticos, demonstram a eficácia dessa abordagem inovadora, tornando-a uma ferramenta valiosa para vários cenários de aplicação. Esses modelos contribuem para avanços na previsão de séries temporais e previsão probabilística, abordando as complexidades dos dados do mundo real.

Métodos de Otimização

Na Tabela 3.9, foi verificado que a maioria dos estudos utilizam os mercados americano e chinês, exemplos sendo os índices S&P 500 e SSE, focando principalmente na variável de entrada preço ou AT. Os métodos mais empregados em Otimização incluem AG, *Particle Swarm Optimization* – (PSO), Teoria do Caos. Entre os modelos adotados, destacam-se LSTM, MLP, SVR e DT. As métricas adotadas são MSE, RMSE, MAE e MAPE.

Tabela 3.9: Artigos classificados como Otimização

Referência	*Target*	Mercado(s)	Ativo(s)	Variável Preditiva	Prediction/s	Métodos	Métricas

Continua na próxima página

Tabela 3.9 – Continua na próxima página

Referência	*Target*	Mercado(s)	Ativo(s)	Variável Preditiva	Prediction/s	Métodos	Métricas
(DURAIRAJ, D. M.; MOHAN, 2022)	S&P 500, Nifty 50, SSE, Ouro, *Crude Oil*, Soja, INR/USD, JPY/USD, SGD/USD	USA, China, Índia, Japão	Taxa de câmbio, Commodities, Índice	Preço	Preço	Chaos, CNN, DT, Prophet, RF, ARIMA, PR	MSE, MAPE, Dstat, Theil's U statistic
(CHEN, J.; YANG, H., 2022)	CSI 300	China	Índice	Preço	Preço	PSO, SVR, AG, LSTM, GRNN	RMSE, MSE, MAPE, MAE
(JI; LIEW; YANG, L., 2021)	S&P/ASX 200	Austrália	Índice	Preço	Preço	PSO, LSTM, SVR	RMSE, MAE, MAPE, R^2
(JUJIE; DANFENG, 2018)	SSE, SZSE	China	Índice	AT	Preço	AG, SVR, GM, BPNN, RW	RMSE, MAE, MAPE
(NTI; ADEKOYA; WEYORI, 2020b)	GSE	Gana	Índice	AT	Preço	AG, SVM, RF, DT, MLP	Accuracy, AUC, RMSE, MAE
(KUMAR, G.; SINGH, U. P.; JAIN, 2021)	Nifty 50, SSE, S&P 500	Índia, USA,	Índice	AT	Preço	PCA, PSO, LM, FNN, DE	RMSE, MAPE, Theil's U, ARV
(DURAIRAJ, M.; BH, 2021)	INR/USD, JPY/USD, SGD/USD, S&P 500, Nifty 50, SSE, Ouro, Crude Oil, Soja	USA, China, Índia	Taxa de câmbio, Índice, Commodities	Preço	Preço	Chaos, Prophet, ARIMA, LSTM, PR	MSE, Dstat, Theil's U
(SINGH, A. et al., 2020)	BSE SENSEX	Índia	Índice	AT	Preço	PSO, SVR, TLBO	NMSE, RMSE, DS, MAE
(ZHENG et al., 2021)	SSE, SSE A shares, SSE B shares, SSE 380, SSE 180, SSE 50, CSI 300, CSI 1000, CSI 100, CSI 500, CSI 800, SZCI, SZSE B shares, SZSE, SZSE A shares, SME, SME 300, SME CI	China	Índice	AT	Preço	SVR, Bat, ANN, RF	RMSE, MAE, MAPE
(DAS; NAYAK; SAHOO, 2022)	NASDAQ, BSE, DJIA, HSI, NIKKEI 225	USA, Japão, Índia, China	Índice	Preço	Preço	FLANN, RAO, MLP, SVM, ARIMA, GA	MAPE

Como ilustrado na Figura 3.10, a literatura apresenta dois tipos de abordagens: a primeira é empregada para otimizar os hiperparâmetros do modelo de *Machine Learning* adotado, e a segunda visa processar as séries temporais financeiras para reduzir o ruído presente nelas.

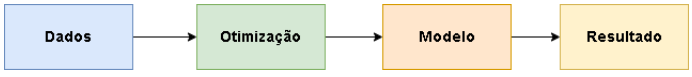


Figura 3.10: *Framework* de modelagem baseado em Otimização. Fonte: (Autor, 2024)

O estudo de (JI; LIEW; YANG, L., 2021) combina IPSO e LSTM para melhorar a previsão de preços de ações. O IPSO otimiza os hiperparâmetros do LSTM com um fator de mutação adaptativo, proporcionando resultados preditivos aprimorados que superam modelos de referência como SVR, LSTM e PSO-LSTM. O modelo introduzido por (CHEN, J.; YANG, H., 2022) apresenta uma abordagem híbrida integrando PSO, SVR e o *General Regression Neural*

Network – (GRNN). O PSO otimiza os parâmetros do SVR, e a sequência residual otimizada melhora ainda mais as previsões de séries temporais com o GRNN. Este modelo PSO-SVR-GRNN melhora significativamente a precisão da previsão em comparação com modelos individuais como PSO-SVR, GRNN, GA-SVR, LSTM, PSO-LSTM e SVR.

O trabalho de (DURAIRAJ, D. M.; MOHAN, 2022) adota um novo *framework* híbrido envolvendo Teoria do Caos, CNN e PR para prever séries temporais financeiras. A Teoria do Caos modela o caos dentro da série temporal, a CNN gera previsões iniciais, e o *Polynomial Regression* – (PR) ajusta a série de erros para gerar previsões de erro. Essas previsões de erro, combinadas com as previsões iniciais da CNN, produzem previsões finais. O híbrido Chaos CNN+PR supera outros modelos, incluindo ARIMA, Prophet, CART, RF, CNN, Chaos-CART, Chaos-RF e Chaos-CNN, em vários conjuntos de dados financeiros. Esses modelos híbridos representam abordagens inovadoras para lidar com a complexidade e volatilidade dos dados do mercado de ações, demonstrando capacidades preditivas aprimoradas em diferentes contextos.

O modelo de (KUMAR, G.; SINGH, U. P.; JAIN, 2021) combina PCA, PSO e *Levenberg-Marquardt* – (LM) para otimizar parâmetros iniciais e reduzir as dimensões das características de um FNN. Este sistema híbrido é projetado para enfrentar desafios na calibração dos parâmetros do FNN e na minimização do conjunto de características de entrada, superando modelos como PSO-FNN, FNN padrão, AG, DE e PCA combinados com um modelo autoregressivo distribuído em defasagem na previsão dos preços de fechamento do índice de ações. (NTI; ADEKOYA; WEYORI, 2020b) introduz um novo classificador *ensemble* homogêneo chamado GASVM, aprimorando o SVM com AG para seleção de características e otimização dos parâmetros do kernel do SVM. Esta abordagem híbrida aborda o *overfitting* do SVM ao lidar com conjuntos de dados de entrada ruidosos e de alta dimensionalidade, obtendo resultados notáveis na previsão de movimentos de preços de ações. GASVM supera outros algoritmos clássicos de ML, como DT, RF e ANN.

O trabalho de (JUJIE; DANFENG, 2018) utiliza AG, *Grey Model* – (GM), BPNN, *Random Walk* – (RW) e SVR para criar dois modelos híbridos, GA-SVR-GM e GA-BPNN-GM, visando melhorar as previsões de índice. O SVR ou BPNN prevê os preços das ações, e os erros são usados para construir o GM, abordando diferentes padrões que o SVR ou BPNN podem não capturar, reduzindo, em última análise, os erros. AG é utilizado para otimizar os parâmetros do modelo.

Para a previsão de STF, (DURAIRAJ, M.; BH, 2021) combina LSTM, PR e Teoria do Caos para lidar com o caos inerente em STF. A presença de caos é testada e a série temporal é modelada usando a Teoria do Caos. O modelo é então processado através do LSTM para previsões iniciais, com sequências de erros derivadas do LSTM usadas para previsões de erro. O modelo híbrido final Chaos+LSTM+PR é avaliado em três categorias: câmbio, preços de commodities e índices do mercado de ações. Ele supera modelos únicos como ARIMA, Prophet, LSTM e Chaos+LSTM, demonstrando desempenho preditivo superior em nove conjuntos de dados. (SINGH, A. et al., 2020) faz uma comparação entre dois modelos híbridos de SVR para prever preços futuros do mercado de ações indiano: um modelo de SVR otimizado por PSO e um modelo de SVR otimizado por *Teaching Learning Based Optimization* – (TLBO). O modelo PSO-SVR é baseado no comportamento de agrupamento natural, enquanto o modelo TLBO-SVR se inspira na transferência de conhecimento como o de professores e alunos.

O estudo de (DAS; NAYAK; SAHOO, 2022) explora a aplicação de algoritmos de Rao na otimização de *Functional link artificial neural network* – (FLANN), levando à criação de modelos híbridos chamados FLANNs baseados em Algoritmo Rao (RAFLANNs). Esses RAFLANNs são avaliados no contexto da previsão de cinco mercados de ações. As avaliações comparativas abrangem variações de modelos FLANN (por exemplo, descida de gradiente, otimizador de multiuniverso, otimização por borboleta e FLANNs baseados em algoritmo genético), bem como, modelos convencionais como MLP, SVM e ARIMA. Os RAFLANNs exibem desempenho superior em termos de precisão preditiva, tempo de computação e testes de significância estatística. Além disso, outro estudo (ZHENG et al., 2021) introduz o algoritmo do Bat para otimizar os parâmetros do modelo SVR, construindo um modelo híbrido conhecido como BA-SVR. Este modelo BA-SVR é aplicado para prever preços de fechamento de 18 índices de ações no mercado de ações chinês. Os resultados indicam que o modelo BA-SVR supera os modelos SVR com kernels polinomiais e sigmoid sem parâmetros iniciais otimizados, mesmo em testes de robustez usando dados de séries temporais estacionárias. A pesquisa contribui para métodos de previsão de índices de ações e destaca a aplicação do algoritmo de Bat no domínio financeiro, enfatizando o potencial para melhorar a precisão preditiva e a gestão de riscos nos mercados de ações.

Em resumo, avanços significativos foram feitos na previsão de séries temporais financeiras através da combinação inovadora de algoritmos e teorias. (JI; LIEW; YANG, L., 2021) destaca a integração de IPSO com LSTM, otimizando os hiperparâmetros do LSTM para superar modelos

de referência. (CHEN, J.; YANG, H., 2022) introduz um modelo híbrido que combina PSO, SVR e GRNN, aprimorando previsões de séries temporais com uma abordagem otimizada. Da mesma forma, (DURAIRAJ, D. M.; MOHAN, 2022) emprega um *framework* híbrido Chaos CNN+PR, aproveitando a Teoria do Caos e CNN para previsões financeiras avançadas. (KUMAR, G.; SINGH, U. P.; JAIN, 2021) propõe um modelo que combina PCA, PSO e LM para otimizar FFNN, mostrando desempenho superior na previsão de índices. (NTI; ADEKOYA; WEYORI, 2020b) apresenta o GASVM, um classificador *ensemble* que combina SVM com AG, obtendo resultados notáveis na previsão de movimentos de preços de ações. Experimentos em (JUJIE; DANFENG, 2018) revelam modelos híbridos GA-SVR-GM e GA-BPNN-GM, utilizando AG, GM, BPNN e SVR para melhorar as previsões de índices. Para a previsão de STF, (DURAIRAJ, M.; BH, 2021) combina LSTM, PR e Teoria do Caos, superando modelos únicos em várias métricas. (SINGH, A. et al., 2020) compara modelos de SVR otimizados por PSO e TLBO, focando no mercado de ações indiano. (DAS; NAYAK; SAHOO, 2022) explora algoritmos de Rao para otimizar FLANN, criando RAFLANNs que superam variações de FLANN e modelos convencionais em previsões de mercado de ações. (ZHENG et al., 2021) aplica o algoritmo do Bat para otimização de SVR, resultando em desempenho superior do modelo BA-SVR para índices de ações chineses. Esses modelos híbridos marcam progresso no enfrentamento da complexidade dos dados do mercado de ações, demonstrando capacidades preditivas aprimoradas e destacando o potencial para maior precisão e gestão de riscos nos mercados de ações.

Abordagem de Análise de Sentimento

Na análise de sentimento, pode-se observar os estudos identificados focam em mercados emergentes como os mercados chinês e indiano, usando como variável de entrada a AT como para prever a direção ou o preço. A ênfase está em modelos como LSTM, CNN, SVM e RNN. Vale ressaltar que o RMSE foi comumente utilizado como métrica.

Tabela 3.10: Artigos classificados com a utilização da Análise de Sentimento

Referência	*Target*	Mercado(s)	Ativo(s)	Variável Preditiva	Prediction/s	Métodos	Métricas
(CHEN, W. et al., 2018)	HS 300	China	Índice	AT	Preço, Direção	RNN, EL, Análise de Sentimentos	Accuracy, MAE, MAPE, RMSE
(PASUPULETY et al., 2019)	NSE	Índia	Índice	AT	Preço	SVM, ET, EL, Análise de Sentimentos	R^2, RMSE

Continua na próxima página

Tabela 3.10 – Continua na próxima página

Referência	*Target*	Mercado(s)	Ativo(s)	Variável Preditiva	Prediction/s	Métodos	Métricas
(YI SUN QINGSONG SUN, 2022)	SSE	China	Índice	AT	Preço	Análise de Sentimentos, CNN, LSTM, SVM	MSE, MAPE, R^2

Na Análise de Sentimento, como mostrado na Figura 3.11, é observado que os dados entram juntamente com a análise como uma característica adicional para auxiliar o modelo de *Machine Learning* com tendências de mercado, ou seja, para avaliar o sentimento do mercado através das mídias sociais.

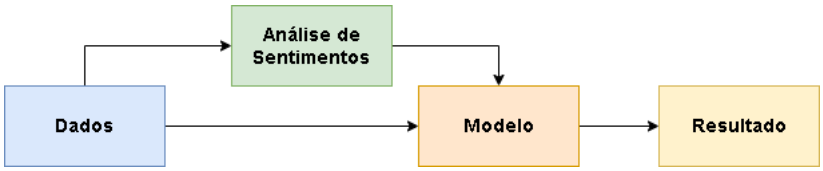


Figura 3.11: *Framework* de modelagem baseado em Análise de Sentimento, Fonte: (Autor, 2024)

No contexto da previsão do mercado de ações, três modelos inovadores são introduzidos. O primeiro modelo aproveita as redes sociais online, especificamente o Sina Weibo, para analisar o conteúdo das notícias e prever a volatilidade do mercado de ações. Este novo modelo híbrido, chamado RNN-boost, incorpora características de sentimento e *Latent Dirichlet Allocation* juntamente com indicadores técnicos, superando outros métodos e demonstrando robustas capacidades preditivas (CHEN, W. et al., 2018). O segundo modelo explora técnicas de *ensemble* usando modelos de regressão baseados em SVM e RF para prever preços de ações. Ele incorpora indicadores técnicos e introduz a análise de sentimento com base em postagens no Twitter. Os resultados mostram que o modelo de *ensemble* apresenta desempenho competitivo, com a eficácia dependendo das características dos dados de treinamento (PASUPULETY et al., 2019). O terceiro modelo, ISI-CNN-LSTM, combina LSTM e CNN para criar uma estrutura de rede de ponta a ponta para previsão de preços de ações, incorporando o sentimento dos investidores para melhorar a precisão das previsões (YI SUN QINGSONG SUN, 2022). Esses modelos, coletivamente, contribuem para a evolução do cenário de previsão do mercado de ações, demonstrando a influência das redes sociais e das técnicas avançadas de ML na melhoria das capacidades preditivas.

Abordagem de Sistema Híbrido

A Tabela 3.11 mostra uma mistura de mercados utilizados, incluindo tanto mercados emergentes quanto desenvolvidos, embora os artigos tendam a focar em apenas um tipo de mercado, seja desenvolvido ou emergente. Os índices mais comumente utilizados incluem S&P 500, DJIA, SSE e Nikkei 225. Há também variação nas variáveis de entrada, com alguns estudos utilizando preço e outros utilizando AT. Em termos de modelos, há uma preferência notável por modelos estatísticos ARIMA e GARCH, bem como, modelos de ML, incluindo LSTM, MLP, SVR e CNN. As métricas destacadas incluem RMSE, MSE, MAE e MAPE.

Tabela 3.11: Artigos classificados como Sistema Híbrido

Referência	*Target*	Mercado(s)	Ativo(s)	Variável Preditiva	Prediction/s	Métodos	Métricas
(PENG et al., 2021)	Preço das ações da India, Preços das ações do Sri Lanka, Preço das ações no Paquistão	Paquistão, Índia, Sri Lanka	Índice	Preço	Preço	ARIMA, MLP, RNN,	RMSE, MAPE, MAE
(DASH et al., 2019)	BSE SENSEX, S&P500, NIFTY 50	Índia, USA	Índice	AT	Direção	MLP, SVM, KNN, NB, DT, LR, PSO, TOPSIS, DE, RBFN, CS, WV, MV	Accuracy, Precision, Recall, f-measures
(CORBA; EGRIOGLU; DALAR, 2020)	ISE	Turquia	Índice	Preço	Preço	ES, ARIMA, MLP, AR, ARCH, GARCH, Fuzzy	RMSE, MAPE, MAE, NMAE
(AMPOUNTOLAS, 2023)	BTC-USD, DAX, FTSE, EURONEXT 100, CAC 40, SMI	França, Inglaterra, Alemanha, Suíça	Índice, Cripto	Preço	Preço	ARIMA, ETS, ANN, K-NN	MAE, RMSE, MAPE
(NAYAK, 2019)	DJIA, BSE, TAIEX, NASDAQ, FTSE, INR, JPY, SGD, AUD	USA, Índia, Taiwan, Inglaterra	Índice, Taxa de câmbio	Preço	Preço	ARIMA, RBFNN, MLP, SVM, FLANN, EL	MAPE, ARV
(HE, K. et al., 2023)	BTC, SSE, EU ETS	China	Cripto, Índice, Carbono	Preço	Preço	RW, ARMA, CNN, LSTM, MLP, EL	MAE, MAPE, RMSE, DS
(DONGHWAN et al., 2020)	KOSPI, KOSDAQ	Coreia do Sul	Índice, Empresa	AT	Preço	LSTM, RNN, BIRCH clustering	MAE, RMSE, MAPE
(SONG; CHOI, 2023)	DAX, S&P 500, DOW	USA, Alemanha	Índice	AT	Preço	RNN, CNN, LSTM, GRU, WaveNet	MSE, MAE
(WIDIPUTRA; MAILANGKAY; GAUTAMA, 2021)	HSI, Nikkei 225, STI, JSX	China, Japão, Singapura, Indonésia	Índice	Preço	Preço	CNN, LSTM	RMSE
(XIAO; SU et al., 2022)	S&P 500	USA	Índice	Empresas	Preço	ARIMA, LSTM	MSE, MAE, RMSE

A Figura 3.12 ilustra o comportamento da abordagem de sistema híbrido, que pode ser dividida em dois métodos. Um é caracterizado pelo Modelo Híbrido Residual, em que o modelo estatístico aborda a parte linear da série temporal financeira, e o modelo inteligente (representando modelos de ML) processa o residual gerado pelo modelo estatístico, obtido pela diferença

entre a previsão do modelo estatístico e o resultado real. Outro método envolve a integração com modelos inteligentes CNN usados para seleção de características para o próximo modelo inteligente, criando sinergia entre eles.

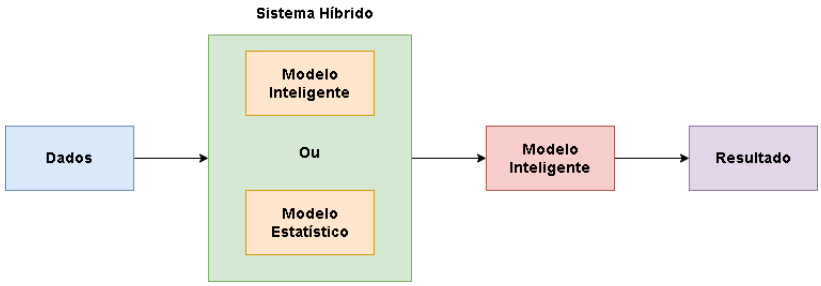


Figura 3.12: *Framework* de modelagem baseado em Sistema Híbrido, Fonte: (Autor, 2024)

A pesquisa de (AMPOUNTOLAS, 2023) foca na previsão de tendências nos mercados financeiros e na transmissão da incerteza do mercado durante vários períodos, incluindo períodos pré-pandemia, pandemia e pós-pandemia. O estudo compara diferentes modelos de previsão, abrangendo técnicas estatísticas, de ML e de *Deep Learning*. Modelos como ARIMA, o modelo híbrido ETS-ANN e kNN são avaliados usando índices diários do mercado financeiro e dados de preços. Embora prever as flutuações do mercado financeiro seja desafiador, com níveis de precisão geralmente baixos, o ARIMA e o modelo híbrido ETS-ANN apresentam melhor desempenho em períodos mais longos, com ambos os modelos demonstrando forte desempenho na maioria dos outros subperíodos. Apesar de sua precisão moderada, o modelo híbrido ETS-ANN se destaca como a abordagem mais promissora para a previsão de séries temporais financeiras, fornecendo insights valiosos para decisões de investimento.

Em outro estudo, um novo método híbrido de previsão, o modelo AR-ARCH-ANN, é proposto para lidar com séries temporais do mundo real não lineares e voláteis, especialmente aquelas com propriedades heterocedásticas. Este modelo recorrente incorpora heterocedasticidade autorregressiva condicional e utiliza a otimização por enxame de partículas para treinar redes neurais, aumentando a probabilidade de evitar armadilhas de mínimos locais (CORBA; EGRIOGLU; DALAR, 2020). Finalmente, a pesquisa de (PENG et al., 2021) visa prever o desempenho de índices do mercado de ações usando vários modelos de ML, como MLP, RNN e ARIMA. O estudo compara esses modelos com abordagens híbridas, a saber, ARIMA-MLP e ARIMA-RNN, e analisa seu desempenho com base em dados históricos da bolsa de valores. Modelos híbridos como ARIMA-MLP e ARIMA-RNN demonstram maior precisão em comparação com modelos únicos, destacando seu potencial na previsão de séries temporais financeiras. Os

resultados ressaltam a robustez e precisão desses modelos híbridos em diferentes mercados de ações, tornando-os ferramentas valiosas para extrapolar tendências do mercado financeiro.

No estudo de (SONG; CHOI, 2023), novos modelos híbridos são propostos para prever os preços de fechamento de índices proeminentes, utilizando modelos baseados em RNN, incluindo CNN-LSTM, GRU-CNN e modelos de *ensemble*. Uma característica inovadora, que envolve a média dos preços mais altos e mais baixos dos índices do mercado de ações, é introduzida. Os resultados experimentais indicam que esses modelos superam os modelos tradicionais de ML em vários casos. Em outra pesquisa (HE, K. et al., 2023), um modelo de previsão de séries temporais financeiras baseado em *ensemble* de *Deep Learning* é introduzido, combinando uma CNN, LSTM e ARMA para lidar efetivamente com características de dados espaço-temporais e autocorrelação. O modelo de *ensemble* de *Deep Learning* demonstra superior precisão preditiva e robustez em comparação com modelos de referência únicos, destacando seu potencial na previsão de séries temporais financeiras.

Além disso, (NAYAK, 2019) explora o *ensemble* de saídas de vários modelos de previsão para aumentar a precisão e mitigar o risco associado à seleção de modelos. Um combinador linear de cinco modelos de previsão, incluindo ARIMA, *Radial Basis Function Neural Network* – (RBFNN), MLP, SVM e FLANN, é usado para prever preços de fechamento em mercados de ações e câmbio. O processo de *ensemble* é suportado por quatro métodos estatísticos para determinação adequada de peso, resultando em uma precisão preditiva aprimorada e destacando a viabilidade e eficácia dessa abordagem.

O trabalho de (XIAO; SU et al., 2022) emprega modelos tradicionais, como ARIMA e LSTM, para prever preços de ações e a subcorrelação dos preços das ações. Os resultados indicam que ambos os modelos, ARIMA e LSTM, fornecem previsões precisas, com o modelo LSTM superando o ARIMA em termos de precisão preditiva. Além disso, um modelo de *ensemble* combinando ARIMA e LSTM demonstra superioridade significativa sobre outros métodos de referência, oferecendo insights valiosos e referências metodológicas para investidores no mercado de ações. Outro estudo foca no nível macroeconômico e na importância de prever índices do mercado de ações de forma confiável. A pesquisa introduz um modelo híbrido de *ensemble*, conhecido como CNN-LSTM multivariado, incorporando características dos modelos CNN e LSTM. O modelo é projetado para estimar múltiplas séries temporais financeiras em paralelo, e sua eficácia é avaliada usando índices do mercado de ações, com foco específico no período da pandemia de COVID-19. Os resultados mostram que o modelo CNN-LSTM multivariado

supera individualmente CNN e LSTM, proporcionando a mais alta precisão estatística e confiabilidade. Este achado destaca o potencial do modelo para prever vários índices do mercado de ações, tornando-o uma escolha valiosa para pesquisas de previsão de séries temporais financeiras (WIDIPUTRA; MAILANGKAY; GAUTAMA, 2021).

Por fim, o artigo de (DASH et al., 2019) apresenta um *ensemble* de classificadores de votação ponderada baseado em uma busca integrada *Crow Search* – (CS) e *Technique for Order of Preference by Similarity to Ideal Solution* – (TOPSIS) para prever movimentos de preços de índices de ações. Essa abordagem fornece uma solução eficaz para o problema de classificação binária na previsão de índices de ações. Esses modelos híbridos contribuem para previsões mais precisas e estáveis, sendo significativos para investidores e pesquisadores no campo dos mercados financeiros.

Análise Final

A maioria dos trabalhos depende de modelos de *Deep Learning* como LSTM, CNN, GRU e RNN. Em relação aos modelos convencionais de *Machine Learning*, os principais incluem MLP, Árvore de Decisão e SVM, enquanto ARIMA e GARCH representam modelos de regressão. No campo da otimização, PSO e GA são as abordagens predominantes. Finalmente, entre os modelos de decomposição, destacam-se as abordagens EMD, EEMD e CEEMDAN.

A abordagem "Completa" visa a integração de decomposição de séries temporais, otimização e modelos de ML. Reconhece-se que essa abordagem produz bons resultados, mas é notável pelo aumento do custo computacional, comprometendo sua eficiência na resolução de problemas do tipo *online*. A segunda classificação aborda "Decomposição", demonstrando a operação da decomposição de séries temporais com abordagens de ML. Isso resulta em maior precisão em comparação com um único modelo. No entanto, a sensibilidade a *outliers* e a incapacidade de capturar padrões sazonais nos dados financeiros afetam a tarefa de decomposição.

A terceira classificação é "*Ensemble Learning*", em que diferentes modelos são usados heterogeneamente. Há uma percepção diferente do algoritmo em relação ao problema, aumentando a robustez da solução e reduzindo o *overfitting*. No entanto, há desafios, como o aumento do número de modelos base, exigindo uma configuração adequada de hiperparâmetros, o que aumenta a necessidade de poder computacional e a dificuldade de interpretar os modelos. Outra abordagem envolve o uso de técnicas de *ensemble learning* em formato homogêneo, inicialmente aumentando a precisão e a robustez, semelhante ao formato heterogêneo. No entanto,

erros comuns estão presentes, pois compartilham a mesma arquitetura, resultando em uma diminuição dos benefícios de robustez para essa solução de problema.

A quarta classificação é "Fuzzy", que proporciona um tratamento adequado para o ruído nas séries temporais financeiras. Permite modelar transições suaves entre diferentes estados, útil para capturar tendências e padrões de longo prazo nos mercados financeiros. Além disso, permite acomodar a variabilidade temporal, modelando variações temporais e sazonalidade presente. No entanto, no caso de mudanças abruptas no ambiente, usar a teoria Fuzzy para capturar essas modificações se torna desafiador. Isso se deve à sua natureza intensiva em termos computacionais, decorrente da complexidade envolvida em sua modelagem para lidar com dados financeiros. A quinta classificação é "Otimização", que melhora a adaptabilidade do modelo aos dados, proporcionando maior eficiência e precisão na previsão do problema. No entanto, há um aumento risco de *overfitting*, em que *outliers* podem levar a ajustes excessivos nos parâmetros do modelo. A maioria das abordagens usa GA ou PSO. No entanto, o tempo de convergência da solução ideal para o problema pode ser dispendioso para o usuário, dificultando seu uso em um ambiente online.

Na classificação "Análise de Sentimento", há uma avaliação constante do sentimento do mercado, auxiliando na previsão para o período avaliado. Isso permite a incorporação de informações não estruturadas, como notícias e mensagens, contendo *insights* não capturados em dados estruturados. No entanto, a presença de fake news, ironia e sarcasmo, juntamente com limitações no contexto financeiro específico, torna difícil capturar adequadamente as nuances específicas desse ambiente. Em outras palavras, termos e expressões específicas do mercado podem não ser completamente compreendidos ou corretamente contextualizados. Finalmente, temos a classificação "Sistema Híbrido", que abrange vários formatos de *ensemble*, como sistemas residuais, combinações lineares de modelos ou novas propostas de *ensemble*. Isso indica que o campo está em constante evolução, pois cada abordagem traz melhorias para a previsão. No entanto, há um aumento na complexidade do sistema, tornando desafiadora a seleção de componentes adequados.

É discutido também o propósito dos estudos, os modelos implementados e o desempenho desses modelos, identificando estudos focados em cenários como prever a direção do mercado de ações, precificação, volatilidade, retornos e taxas de câmbio. A literatura apresenta várias implementações de *ensemble* de algoritmos, incluindo técnicas de *ensemble learning*, sistemas híbridos residuais, decomposição de séries temporais, análise de sentimento e otimização. Os

procedimentos de validação e as métricas de desempenho variam de acordo com o escopo do estudo. São observadas duas abordagens distintas: estudos focados em prever o movimento do mercado geralmente usam acurácia como a principal métrica, pois o problema é tratado como uma tarefa de classificação. Por outro lado, em estudos relacionados à previsão de preços de índices, métricas de regressão como RMSE, MAPE, MAE e MSE são as mais relevantes.

É notado a ausência de artigos abordando o uso de variáveis macroeconômicas como o PIB, taxa de desemprego, divida interna em conjuntos de dados, interpretabilidade do modelo, risco de mercado e análise de incertezas, bem como, o uso limitado de testes de hipótese para provar a significância do desempenho. Muitos artigos focaram apenas em variáveis técnicas ou preços em conjuntos de dados, comparando com abordagens anteriores para demonstrar melhoria significativa e direcionando países desenvolvidos como o mercado de interesse para previsão.

Alguns estudos identificados pela pesquisa apresentaram um modelo de *ensemble* implementado como parte de uma solução para auxiliar a tomada de decisão dos investidores. Outros estudos discutiram o impacto da solução de *ensemble* e a adoção dessa abordagem, conforme citado em (AGAPITOS; BRABAZON; O'NEILL, 2017; DURAIRAJ, M.; BH, 2021; NTI; ADEKOYA; WEYORI, 2020a).

3.3.4 RQ4: Lacunas e Oportunidades

Existem duas maneiras de identificar lacunas de pesquisa na literatura. A primeira é através de um pesquisador experiente na área, que lê artigos científicos e usa seu conhecimento prévio para propor novas direções na área abordada. A segunda é lendo artigos científicos e destacando pontos futuros e limitações nesses trabalhos como etapas a serem desenvolvidas. Ambos os aspectos são abordados nesta seção.

Como destacado anteriormente, a maioria dos estudos analisados durante nossa RSL obteve uma pontuação de qualidade média ou alta, variando entre 3,5 e 5,5 pontos, com um limite máximo de 6,0 pontos. Portanto, as análises são baseadas em dados confiáveis coletados desses documentos.

Considerando as perspectivas das abordagens híbridas para a previsão de séries temporais financeiras, é identificado uma oportunidade relacionada à falta de generalização entre mercados. Em outras palavras, muitos dos artigos analisados focam principalmente em mercados desenvolvidos, deixando lacunas de pesquisa em relação a mercados em desenvolvimento e

subdesenvolvidos, conforme indicado nas Tabelas (3.5 - 3.11). Vale notar que a baixa demanda nestes mercados, especialmente nos em desenvolvimento, está relacionada às oportunidades limitadas de lucro que a economia pode gerar. No entanto, nos mercados em desenvolvimento, a situação muda, pois há uma maior chance de lucro, dada a natureza transitória do mercado. Portanto, é importante testar algoritmos nesses ambientes para garantir sua adaptabilidade a várias situações transitórias, assegurando um modelo bem ajustado.

Além disso, o uso de dados macroeconômicos e indicadores técnicos juntamente com índices de mercado oferece uma oportunidade para aprofundar a compreensão da economia interna do mercado analisado (YAN; AASMA et al., 2020; LUO et al., 2021). Essa abordagem pode oferecer *insights* valiosos para a construção dos algoritmos.

Outro aspecto que merece atenção é a interpretabilidade das abordagens desenvolvidas. Frequentemente, o objetivo principal desses estudos é auxiliar os usuários do sistema na tomada de decisões financeiras. No entanto, a falta de interpretabilidade nos modelos pode gerar desconfiança entre os investidores. Portanto, é importante investigar maneiras de melhorar a interpretabilidade das abordagens propostas, tornando os resultados mais confiáveis e acessíveis aos usuários. Um exemplo que pode ser usado é a Teoria Fuzzy. No entanto, é importante notar que essa abordagem tem poucos trabalhos na literatura, como apresentado na Tabela 3.5, e uma das razões analisadas para essa falta de estudos é a necessidade de especialistas no domínio para construir o conjunto de regras Fuzzy, conforme observado em (LIMA SILVA et al., 2019). Vale ressaltar que a Teoria Fuzzy fornece regras de fácil interpretação, cruciais para melhorar a confiabilidade do modelo.

Um ponto relevante a destacar envolve a realização de um estudo comparativo empírico entre as sete classificações para identificar qual é a mais adequada para prever o índice de mercado, considerando suas características específicas. Isso porque mercados desenvolvidos e em desenvolvimento apresentam peculiaridades distintas. Tal análise fornecerá uma melhor compreensão e promoverá avanços na melhoria das classificações destacadas.

Além disso, na macroárea de análise de sentimento, foi identificado que a criação de um conjunto de dados relacionado ao mercado estudado requer bases de dados generalizadas, mas também demanda poder computacional para a execução de algoritmos inteligentes de ML, conforme mencionado em (CHEN, W. et al., 2018; PASUPULETY et al., 2019; YI SUN QINGSONG SUN, 2022). Como pesquisa futura, é relevante investigar e avaliar o impacto das políticas governamentais, eventos geopolíticos e outras notícias globais nas séries temporais financeiras.

3.4 Trabalhos Relecionados

Como apresentado no subcapítulo 3.3, foi realizada uma revisão da literatura e um levantamento bibliométrico sobre os artigos que focavam na predição do índice da bolsa de valores utilizando séries temporais financeiras. No subcapítulo 3.3.4, foi possível observar a existência de uma lacuna em relação à comparação das diferentes arquiteturas de *ensemble* para entender os pontos fortes e fracos de cada uma. Com esse objetivo, foi criada uma revisão bibliográfica focada nessa área de interesse, utilizando a mesma abordagem anteriormente realizada.

3.4.1 Protocolo da Revisão Sistemática da Literatura

Nesta subseção está o passo a passo realizado para a construção da RSL.

Termos de Pesquisa

Define o PICOC, representando *Population*, *Intervention*, *Comparison*, *Outcome*, and *Context*, com base no objetivo desta revisão. O resultado é o seguinte:

- ***Population:*** Índices do Mercado de Ações;
- ***Intervention:*** *Ensemble*;
- ***Comparison:*** Abordagens das técnicas de *Ensemble*;
- ***Outcome:*** Comparação;
- ***Context:*** Séries Temporais Financeiras.

As palavras-chave e sinônimos, bem como, o termo relacionado aos critérios PICOC estão representados na Tabela 3.12.

Tabela 3.12: Palavras-chave e Sinônimos

Palavras-chave	Sinônimo	Relacionado
ensemble	*hybrid model* *hybrid system*	*Intervention*
comparison	*comparative analysis*	*Outcome*
stock Índice	*stock exchange*	*Population*

A *String* de pesquisa foi gerada de acordo com os critérios PICOC: *("stock indice"OR "stock exchange") AND ("ensemble"OR "hybrid model"OR "hybrid system") AND ("comparison"OR "comparative analysis").*

Fonte de Dados

Foram exploradas três bases de dados acadêmicas para conduzir a pesquisa: *Web of Science*, *IEEE Xplore* e *Scopus*. Essas fontes de dados foram selecionadas com base na mesma justificativa utilizada no subcapítulo 3.2.2, sendo as principais fontes de publicações especializadas para pesquisa relacionada aos temas do estudo proposto. A busca foi iniciada por artigos em março de 2024.

Processo de Seleção

Considerando o objetivo desta RSL, foi realizada uma seleção preliminar dos artigos recuperados após a execução da *string* de busca nas bases de dados mencionadas anteriormente. Os critérios de inclusão e exclusão foram estabelecidos para assegurar que os estudos incluídos na revisão fossem pertinentes à questão de pesquisa e atendessem aos padrões de qualidade requeridos, conforme descrito no subcapítulo 3.2.3.

Extração de Campos Relevantes

A etapa final desta RSL envolve a extração de informações relevantes dos textos completos dos artigos selecionados. Nesta fase, as informações são analisadas para responder às questões de pesquisa e coletar dados demográficos dos artigos. Definimos 6 campos para a extração de dados: contexto, tipo de abordagem de *ensemble*, desempenho, experimentos, objetivos, conjuntos de dados.

3.4.2 Resultados e Discussões

A Figura 3.13 ilustra o processo passo a passo realizado. A busca nas três bases de dados resultou em artigos, dos quais 21 eram duplicatas, restando 31 artigos submetidos à aplicação de filtros. Após a aplicação dos filtros, 15 artigos foram excluídos, deixando 16 artigos para serem avaliados pelo título e resumo. Destes, 8 foram retirados. Por fim, restaram 8 trabalhos para leitura detalhada e aplicação dos critérios de elegibilidade, resultando na exclusão de 4. Os 4 artigos restantes avançaram para a etapa de extração de dados.

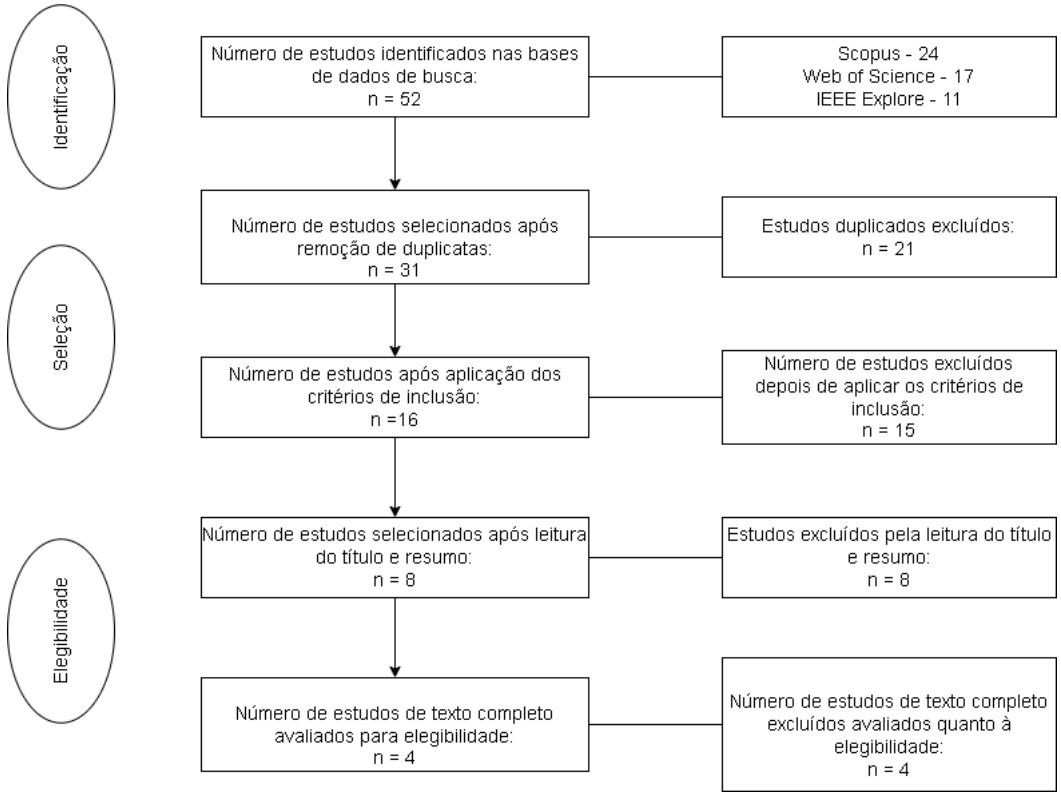


Figura 3.13: Visão geral do processo de seleção e extração de dados

Os resultados da avaliação de qualidade na Figura 3.14. As pontuações de qualidade dos artigos variaram de 3,5 a 5,5 em uma escala de Likert de 6 pontos, estendendo-se de 0,0 (mínimo) a 6,00 (máximo). É observado que nenhum artigo alcançou a pontuação máxima, principalmente devido à falta de disponibilidade do repositório de algoritmos e à limitada generalização para diferentes mercados. Logo, uma minoria de artigos recebeu pontuações de 3,5, enquanto a maioria apresentou pontuações de qualidade média a alta, variando entre 4,0 e 5,5 pontos, com o limite máximo sendo 6,00 pontos.

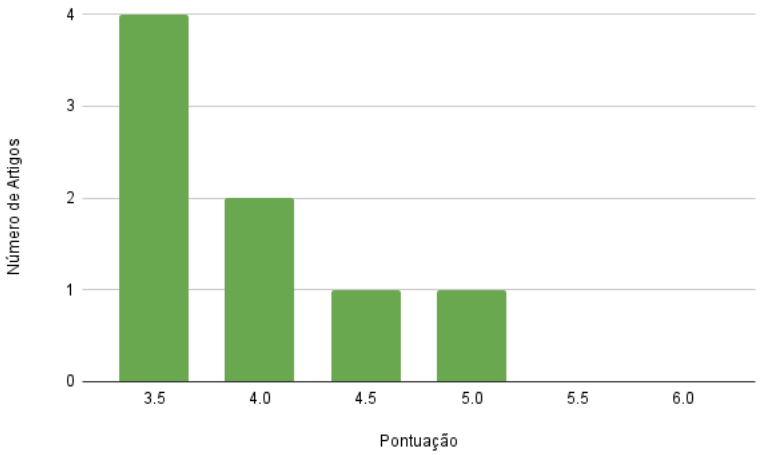


Figura 3.14: Visão geral das pontuações de avaliação de qualidade para os 6 artigos

No final, 4 trabalhos foram submetidos para análise com o objetivo de explorar as abordagens comumente adotadas e identificar lacunas. Ao analisar as metodologias e os resultados de estudos anteriores, é buscado compreender as tendências predominantes na análise comparativa de abordagens de *ensemble* em séries temporais financeiras e estabelecer uma base para o desenvolvimento de novas abordagens que possam oferecer *insights* adicionais, como mostrado nos subcapítulos abaixo:

Ensemble Learning

O estudo conduzido por (NTI; ADEKOYA; WEYORI, 2020a) investiga comparativamente quatro técnicas de *ensemble learning*: *boosting*, *bagging*, *blending* e *stacking*. Essas técnicas foram aplicadas na previsão de índices de mercado, abrangendo mercados em países desenvolvidos, em desenvolvimento e subdesenvolvidos. Os resultados revelam que a técnica de *stacking* superou as outras em termos de desempenho. No entanto, é imperativo destacar as limitações desta pesquisa, que foca exclusivamente em técnicas de *ensemble*, excluindo modelos lineares e abordagens baseadas em decomposição ou sistemas híbridos residuais. Além disso, o método adotado para comparação de desempenho, que mede o tempo de execução em segundos, pode ser suscetível a interferências de outros processos na máquina, potencialmente comprometendo a precisão dos resultados. Ademais, não foram realizados testes estatísticos apropriados para validar as comparações entre os modelos.

No estudo de (FERROUHI; BOUABDALLAOUI, 2024), o foco está na comparação das técnicas de *ensemble learning*: *Boosting*, representado pelo *Adaboost* e o XGBoost; *Bagging*, representado pelo *Random Forest* e o Bagging-LSVM; e, por último, o *Stacking*, que é a combinação

das abordagens de *ensemble learning* anteriores. Essas técnicas foram utilizadas na previsão do Índice da Bolsa de Valores de Casablanca no formato de High-Frequency Trading – (HFT) em Marrocos. A abordagem *Stacking* obteve os melhores resultados. Foi possível visualizar alguns problemas, como o foco exclusivo no mercado emergente de Marrocos, o que não oferece uma visão abrangente sobre a sua utilização nos mercados desenvolvidos. Outro ponto importante é a falta de testes estatísticos que comprovem a melhoria dos resultados, além do foco exclusivo na abordagem de *ensemble learning*, excluindo as abordagens de decomposição, otimização, completa e sistema híbrido. Além disso, o estudo não apresenta uma metodologia detalhada sobre a comparação dos modelos.

Sistema Híbrido

O estudo de (AMPOUNTOLAS, 2023) realiza uma análise comparativa entre modelos estatísticos, de Machine Learning e de Deep Learning, especificamente ARIMA, um modelo híbrido ETS-ANN, e K-NN, para avaliar mercados de índices europeus e o mercado de criptomoedas, abordando períodos antes, durante e após a pandemia (2018 - 2021). O estudo destaca a correlação da COVID-19 com o impacto no mercado financeiro, observando que o modelo híbrido demonstrou desempenho satisfatório em geral na previsão de retornos. No entanto, é essencial notar que a análise focou apenas em mercados desenvolvidos, excluindo técnicas como *ensemble learning*, decomposição, otimização ou sistemas híbridos. Além disso, as comparações de desempenho não foram suportadas por rigorosos testes estatísticos para validação, o que poderia fortalecer a robustez dos resultados.

Decomposição

O estudo conduzido por (JOTHIMANI; BAŞAR, 2019) explora uma comparação entre várias técnicas de decomposição de séries temporais financeiras, incluindo DWT, EMD e VMD, em conjunto com modelos de ML, como ANN e SVR. Este trabalho utilizou 25 índices de diversos mercados, tanto desenvolvidos quanto em desenvolvimento, além do teste estatístico de Wilcoxon para validar os resultados. No entanto, a análise deixou de abordar a relação custo-benefício dessas técnicas de decomposição, limitando uma compreensão abrangente da eficácia dessas abordagens.

Considerações Finais

Ao analisar trabalhos que comparam diferentes abordagens de *ensemble*, nota-se que o foco é predominantemente na avaliação de desempenho, quantificada através de métricas como MSE, RMSE, MAE e MAPE. Essa concentração exclusiva em métricas de desempenho introduz uma forma de viés e limitação nesse tipo de análise comparativa. A razão para isso é que essa abordagem oferece apenas uma perspectiva parcial do problema, não abrangendo uma visão holística do custo-benefício das técnicas em questão. Assim, é evidente que os trabalhos discutidos tendem a apresentar limitações classificadas nos Itens 1 e 2 e vieses enumerados nos mesmos itens apresentados na 2.4.

Este trabalho se destaca na literatura por comparar de maneira abrangente várias técnicas, incluindo *ensemble learning*, decomposição, sistemas híbridos residuais e modelos de ML individuais e lineares. Essa abordagem multifacetada visa estabelecer uma visão detalhada para determinar a aplicabilidade mais eficaz de cada técnica em diferentes tipos de mercados, considerando a volatilidade específica de cada um. Um aspecto inovador deste estudo é a avaliação de desempenho baseada nas instruções retiradas geradas pela CPU durante a execução de cada modelo, oferecendo uma métrica de desempenho unificada, pois permite uma análise da relação custo-benefício dessas técnicas para fornecer uma visão completa da viabilidade de sua aplicação. A inclusão de rigorosos testes estatísticos para validar as comparações entre modelos também reforça a robustez das conclusões.

Capítulo 4

Metodologia

Neste capítulo, são abordadas as justificativas e características das abordagens adotadas para a análise comparativa. Serão discutidas as características do ambiente de teste, descrição dos dados, a etapa de pré-processamento, a construção do protocolo utilizado, a elaboração das abordagens de *ensemble* e as métricas de avaliação empregadas.

4.1 Ambiente de Teste

O computador utilizado possui um processador Intel(R) Core(TM) i7-7700HQ CPU a 2.80 GHz, 8 GB de RAM, sistema operacional Windows 10 de 64 bits e uma placa gráfica NVIDIA GeForce GTX 1050 Ti.

A linguagem de programação *Python* 3 foi utilizada, juntamente com várias bibliotecas, tais como:

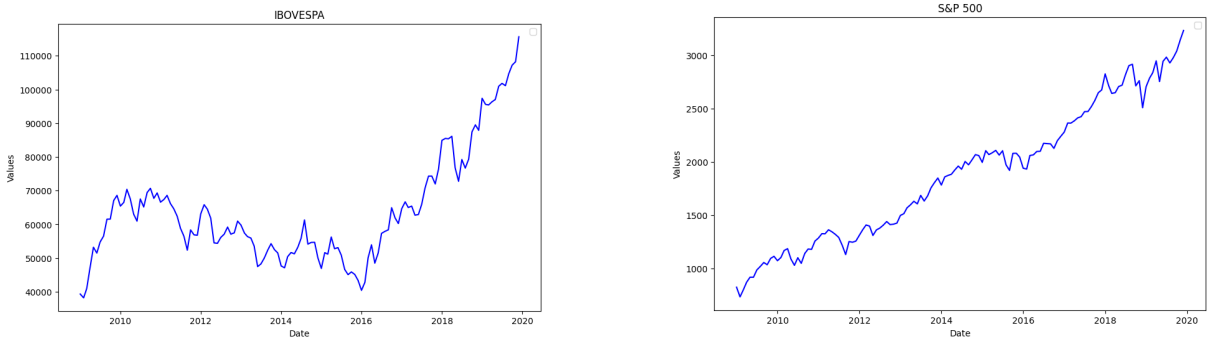
- Scikit-learn versão 1.3.1 para construção de MLP, SVR, DT, *Grid Search*, normalização de dados e abordagens de *ensemble learning*;
- Yahoo Finance versão 1.4.0 para recuperação de dados;
- Python Scipy versão 1.11.2 para obtenção de parâmetros estatísticos das séries temporais e os testes de hipótese;
- EMD-signal versão 1.5.1 para realização de decomposição de séries temporais;
- pmdarima versão 2.0.3 para utilização do modelo estatístico ARIMA;
- Intel VTune Profile versão 2024.0 para verificação das métricas de desempenho.

4.2 Descrição dos Dados

Os mercados selecionados para este estudo foram escolhidos com base em sua importância e representatividade, abrangendo tanto mercados desenvolvidos quanto mercados em desenvolvimento. O primeiro mercado considerado é o Índice Bovespa (IBOVESPA), que é o principal indicador de desempenho médio das ações negociadas na B3, representando mercados emergentes. O segundo mercado selecionado foi o dos Estados Unidos, representado pelo índice Standard & Poor's 500 (S&P 500), que compreende 500 ativos listados nas bolsas de valores NYSE e NASDAQ, sendo amplamente utilizado como um indicador dos mercados desenvolvidos. Esses índices foram escolhidos por capturarem diferentes dinâmicas econômicas globais.

Os dados históricos de ambos os índices foram coletados no período de 1º de janeiro de 2009 a 30 de dezembro de 2019. A escolha desse intervalo de tempo se justifica por capturar o período pós-crise financeira global de 2008, um marco importante para o estudo das dinâmicas dos mercados financeiros. Além disso, o período precede a pandemia de COVID-19, que impactou significativamente os mercados globais, como discutido em (FĂRCAŞ, 2023). A análise desses dez anos permite avaliar as flutuações dos mercados durante um período de relativa estabilidade econômica e crescimento gradual, sem influências externas como a crise sanitária global.

Nas Figuras 4.1a e 4.1b, são visualizadas as séries temporais financeiras dos dois índices. Nota-se que a série temporal do IBOVESPA exibe maior volatilidade, com frequentes oscilações de subida e descida, enquanto o S&P 500 demonstra uma trajetória mais estável, refletindo a maturidade do mercado norte-americano em comparação ao mercado brasileiro. A análise estatística descritiva, apresentada na Tabela 4.1, confirma essas observações.



(a) IBOVESPA (b) S&P 500

Figura 4.1: Séries Temporais Financeiras. Fonte: (Yahoo Finance, 2024)

Tabela 4.1: Estatísticas descritivas dos dados dos índices de ações

Índices de ações	**IBOVESPA**	**S&P 500**
Média	63667,984	1869,262
Desvio Padrão	15695,909	632,939
Valor Mínimo	36235	676,530
Valor Máximo	117203	3240,02
Assimetria	19,470	4,235
Jarque-Bera	640,243	153,287
p-value	0,000	0,000
Total de Observações	2716	2766

A Tabela 4.1 exibe as estatísticas descritivas dos dois índices de ações. O IBOVESPA apresenta uma média de 63.667,98 pontos, com um desvio padrão relativamente alto de 15.695,91, refletindo a alta volatilidade do mercado brasileiro. Em contraste, o S&P 500 apresenta uma média de 1.869,26 e um desvio padrão de 632,94, indicando menor variação em suas cotações. O teste de Jarque-Bera foi utilizado para avaliar a normalidade das séries temporais, com valores de 640,243 para o IBOVESPA e 153,287 para o S&P 500. Em ambos os casos, os baixos valores de p (0,000) indicam a rejeição da hipótese de normalidade, sugerindo que as distribuições dos dados não seguem uma distribuição normal.

Além disso, os coeficientes de assimetria (19,470 para o IBOVESPA e 4,235 para o S&P 500) indicam que ambas as séries são significativamente assimétricas, com o IBOVESPA apresentando uma assimetria muito maior, reforçando a ideia de que este mercado é mais instável e propenso a flutuações bruscas. Isso é particularmente relevante para a análise de séries temporais financeiras, pois a presença de assimetria e não normalidade pode impactar os modelos preditivos e as técnicas de previsão utilizadas neste estudo.

4.3 Protocolo

Este estudo segue um protocolo inspirado em (FURLANETO et al., 2017), com dois principais objetivos: minimizar a influência de vieses e garantir a padronização rigorosa dos experimen-

tos. O protocolo, ilustrado na Figura 4.2, inicia-se com a coleta de dados diários de 1º de janeiro de 2009 a 30 de dezembro de 2019, resultando em uma série temporal univariada composta por observações diárias. Esses dados passam por uma etapa de preparação, na qual o conjunto é dividido em 70% para treinamento e 30% para teste. No IBOVESPA, isso representa 1901 dias para treino e 815 dias para teste, enquanto no S&P 500 o total ficou em 1936 amostras para treino e 830 para teste.

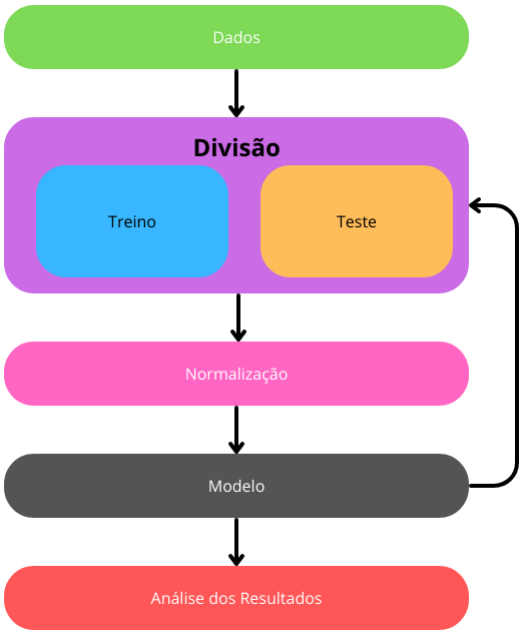


Figura 4.2: Protocolo Esquemático. Fonte: (Autor, 2024)

Para padronizar a escala das variáveis, a técnica de normalização *MinMax* foi adotada, conforme representada pela Equação 4.1. Esta técnica, que transforma os dados para um intervalo de [0,1], ajuda a evitar problemas relacionados à escala das variáveis nos algoritmos de *Machine Learning*, o que é fundamental para as abordagens de *ensemble* utilizadas no estudo, facilitando a análise comparativa com outras abordagens algorítmicas.

$$X_{scaled} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{4.1}$$

A seleção do conjunto de treinamento foi realizada utilizando *cross-validation* específica para séries temporais, conforme sugerido por (BERGMEIR; HYNDMAN; KOO, 2018). Essa técnica de validação cruzada lida com a dependência temporal dos dados ao utilizar uma janela deslizante, onde o modelo é treinado e os pontos de dados previstos são incorporados ao conjunto de treinamento subsequente. O processo é repetido para prever os pontos seguintes, conforme ilustrado na Figura 4.3. Foi utilizado $N = 30$ para o número de divisões no processo

de validação cruzada, garantindo uma avaliação robusta da generalização do modelo ao longo do tempo.

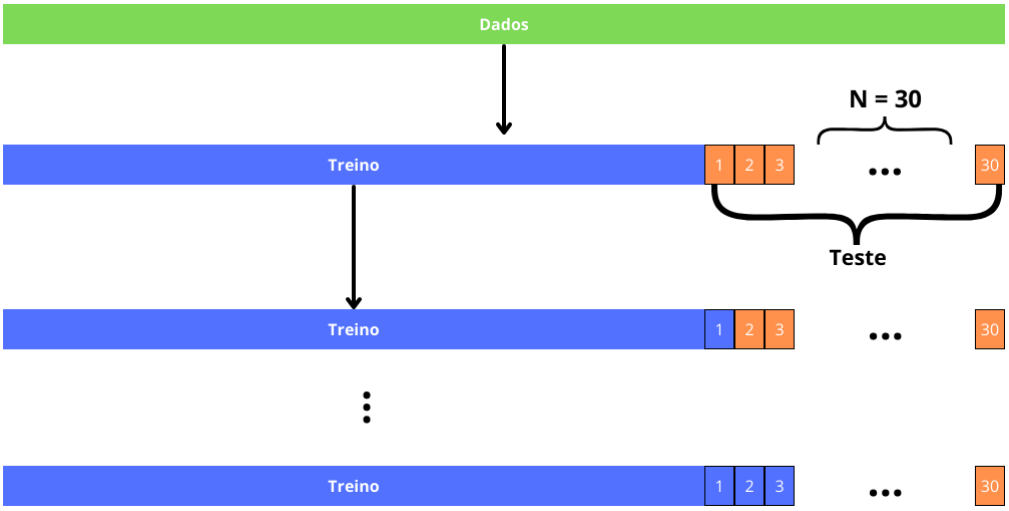


Figura 4.3: *Cross-validation* para Séries Temporais. Fonte: (Autor , 2024)

Além disso, esse protocolo tem o cuidado de evitar o problema de *look-ahead bias* (FURLANETO et al., 2017), comum em estudos de séries temporais financeiras, onde o uso inadvertido de informações futuras pode enviesar os resultados. Todos os passos de pré-processamento, incluindo normalização e divisão dos conjuntos de dados, foram realizados em conformidade com as melhores práticas para garantir a validade das previsões realizadas, como discutido em (BERGMEIR; HYNDMAN; KOO, 2018).

4.4 Construção dos Modelos de *Ensemble*

Neste estudo, a seleção dos algoritmos para avaliação foi baseada em dados coletados a partir de revisões sistemáticas realizadas no Capítulo 3, bem como, em outras revisões da literatura sobre séries temporais financeiras que empregam abordagens de *Machine Learning* (HENRIQUE; SOBREIRO; KIMURA, 2019; KEHINDE; CHAN; CHUNG, 2023; GANDHMAL; KUMAR, K., 2019; KUMBURE et al., 2022; BUSTOS; POMARES-QUIMBAYA, 2020). Por meio de uma análise meticulosa, foi possível identificar várias abordagens de *ensemble* e algoritmos específicos frequentemente aplicados, como:

- **Modelos *Single*:** Bi-LSTM, LSTM, CNN, GRU, MLP, CART, SVR, SVM.
- **Decomposição:** EMD, EEMD, CEEMD, CEEMDAN.

- **Otimização:** PSO, AG, Teoria do Caos.
- **Ensemble Learning:** Bootstrap Aggregating (Bagging), Stacking, Boosting, Blending.
- **Sistema Híbrido Residual:** ARIMA, GARCH, ARMA.

Para explorar se uma abordagem simplificada de *ensemble* poderia resultar em um aumento significativo na precisão do modelo final, foram propostas e construídas combinações específicas:

- **Modelos *single*:** CART, MLP, e SVR.
- **Modelo Linear:** ARIMA.
- **Completo:** Combinações de CEEMDAN com AG e modelos *single*.
- **Decomposição:** Uso de CEEMDAN em conjunto com modelos *single*.
- ***Ensemble Learning*:** *Bagging*, *Stacking*.
- **Sistemas Híbridos Residual:** Combinações do modelo linear com o modelo *single*.
- **Otimização:** Aplicação de GA em conjunto com modelos *single*.

Este *design* visa cobrir uma ampla gama de abordagens reconhecidas na literatura, permitindo uma análise comparativa aprofundada para determinar o impacto das estratégias de *ensemble* na melhoria da precisão dos modelos de previsão em séries temporais financeiras. Não abordamos a abordagem Fuzzy devido à necessidade de assistência especializada na área financeira para auxiliar na construção das regras. Quanto à análise de sentimento, optamos por não incluí-la devido à falta de um *dataset* adequado voltado para o mercado IBOVESPA e outro para o S&P 500. Em relação à abordagem *Blending*, decidimos não utilizá-la por ser uma derivação da abordagem *Stacking*.

4.4.1 Abordagens de *Ensemble*

A escolha dos modelos MLP, SVR e DT deve-se ao seu *status* como algoritmos clássicos na literatura sobre abordagens de *machine learning* para previsão de séries temporais financeiras. Além disso, esses algoritmos são amplamente implementados na comunidade científica e são notáveis por sua capacidade de lidar com dados não lineares. O modelo estatístico ARIMA foi

selecionado pela sua alta aparição na literatura (KUMBURE et al., 2022; BUSTOS; POMARES-QUIMBAYA, 2020).

Ensemble Learning

Técnicas de *Bagging* homogêneo e heterogêneo e *Stacking* foram selecionadas para explorar seu comportamento na previsão de séries temporais financeiras.

- ***Bagging*:** Como mostrado na figura B.1 que apresenta a construção da abordagem *bagging*, a qual envolve a utilização de vários modelos independentes por meio da amostragem com reposição do conjunto de dados original, ou seja, aplicando *Bootstrap* ao conjunto de dados. Cada modelo é treinado individualmente usando uma amostra aleatória de treinamento e pode empregar qualquer algoritmo de aprendizado, podendo consistir de modelos homogêneos ou heterogêneos. Após o treinamento, as previsões de cada modelo são combinadas por meio de uma técnica de agregação, neste caso foi adotada a média, no contexto de problemas de regressão (BROWNLEE, 2021)[1].

- ***Stacking*:** Combina as previsões de vários modelos para produzir resultados mais precisos. A ideia geral do *Stacking* é gerar um meta-modelo a partir das previsões de um conjunto de modelos, utilizando validação cruzada k-fold. Finalmente, o meta-modelo é treinado com base em um algoritmo específico. O processo geralmente consiste em dois estágios de treinamento, embora mais níveis possam ser adicionados. O objetivo do primeiro estágio é gerar os dados de treinamento para o meta-modelo, usando a validação cruzada k-fold para cada modelo base definido no primeiro passo. As previsões de cada modelo base são empilhadas para construir esse novo conjunto de dados de treinamento. No segundo estágio, o meta-modelo é treinado, como ilustrado na Figura B.2, e o detalhamento de seu aprendizado é apresentado na Figura B.3 (WOLPERT, 1992)[2].

Sistema Híbrido Residual

O Sistema Híbrido Residual é representado na Figura 4.4. A subfigura 4.4a mostra que os dados de entrada para esta fase são o respectivo conjunto de treinamento da série temporal $\left(Z_t^{tr}\right)$. A saída desta fase, por sua vez, são os modelos lineares e não lineares treinados

[1] Para mais informações sobre Bagging, consulte a documentação da biblioteca (LEARN, 2023a)
[2] Para mais informações sobre *Stacking*, consulte a documentação da biblioteca (LEARN, 2023b)

$\left(M_N^1, M_N^2, \ldots, M_N^m\right)$. Os modelos não lineares compreendem o *ensemble* usado na modelagem do residual.

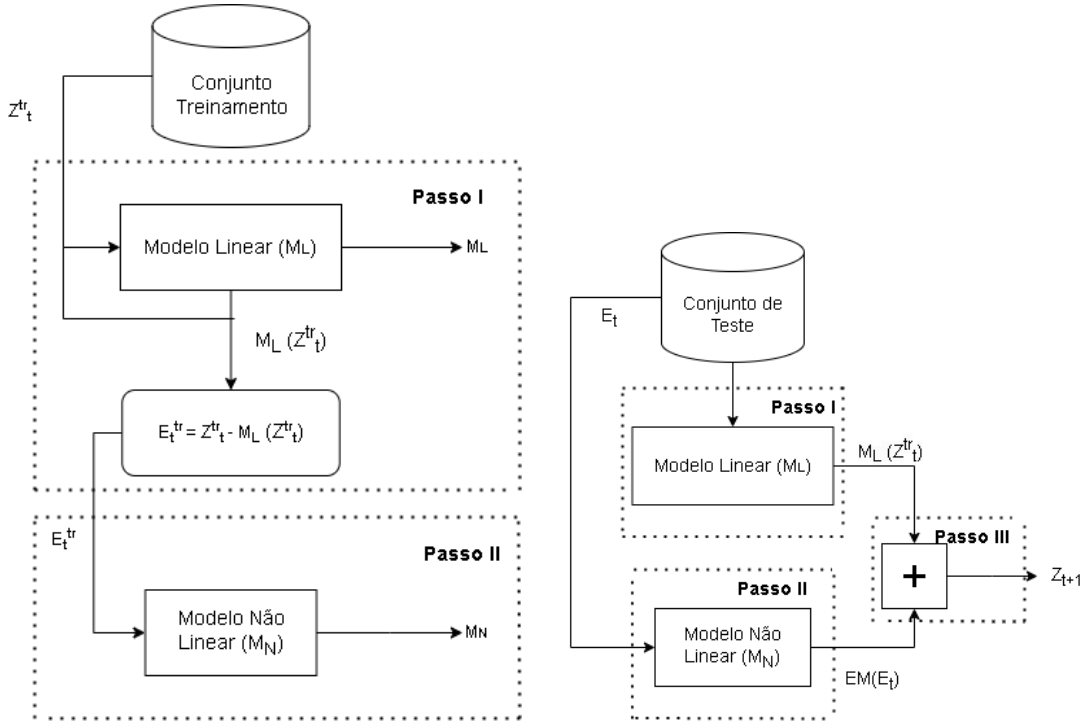
No passo (I) da fase de treinamento, o modelo M_L é criado a partir dos dados reais de treinamento da série temporal $\left(Z_t^{tr}\right)$. Nesta etapa, o treinamento de M_L visa capturar padrões lineares existentes na série temporal. Em seguida, a série residual $\left(E_t^{tr}\right)$ é gerada pela diferença entre o valor real da série temporal $\left(Z_t^{tr}\right)$ e a previsão linear $\left(M_L\left(Z_t^{tr}\right)\right)$.

No passo (II) da fase de treinamento, o modelo não linear é aplicado para capturar padrões não lineares presentes na série residual E_t^{tr}. O método proposto aplica um dos modelos base de Machine Learning. O modelo não linear é treinado e, finalmente, são realizados testes para gerar os resultados para análise.

A Figura 4.4b mostra os três passos da fase de teste do método proposto com o objetivo de gerar previsões da série temporal $\hat{Z}_{t+1}$. No passo (I), o modelo linear (M_L) realiza a previsão linear ($M_L(Z_t)$) para um novo padrão da série temporal Z_t. No passo (II), dados passados da série de resíduo (E_t) são utilizados como entrada para o modelo de *machine learning* $\left(M_N^1, M_N^2, \ldots, M_N^m\right)$. Em seguida as m previsões da série do resíduo $\left(M_N^1(E_t), M_N^2(E_t), \ldots, M_N^m(E_t)\right)$ são geradas pelo ensemble.

Como método de agregação é utilizado a média para integrar as m previsões residuais, obtendo assim a previsão do ensemble ($EM(E_t)$) para E_t. Nesse contexto, a agregação por média é obtida pelo somatório de $M_N^1(E_t), M_N^2(E_t), \ldots, M_N^m(E_t)$ dividido por m.

Finamente, no passo (III) da fase de teste, a previsão da série temporal $\left(\hat{Z}_{t+1}\right)$ é gerada por meio da soma das previsões linear ($M_L(Z_t)$) e não-linear ($EM(E_t)$).



(a) Treinamento (b) Teste

Figura 4.4: Modelo Híbrido Residual. Fonte: Adaptado de (SANTOS JÚNIOR et al., 2022)

Método de Otimização

Na otimização dos algoritmos é baseada na figura 3.10, foi utilizado a abordagem de Algoritmo Genético para a seleção de parâmetros nos modelos de ML. A escolha desta abordagem visa determinar a melhor solução global que se adapte mais efetivamente à distribuição dos dados financeiros. Em outras palavras, os cromossomos do Algoritmo Genético são configurados de acordo com os parâmetros das abordagens de ML.

Subsequentemente, a população inicial é construída e a função de aptidão é avaliada com base nas previsões sobre os dados de treinamento, utilizando validação cruzada para séries temporais financeiras com $K = 5$. Optou-se por empregar um *crossover* de ponto único com uma taxa de mutação e *crossover* de 75%, visando aumentar a diversidade da população. A seleção do *crossover* foi realizada utilizando o método da roleta e a escolha final dos indivíduos foi feita através do método elitista. Por fim, o critério de parada adotado foi se o cromossomo com o melhor desempenho sobreviver durante 5 gerações, o algoritmo para.

Abordagem de Decomposição

Como mostrado no subcapítulo 2.7 a qual mostra o funcionamento da decomposição que é ilustrado na Figura 2.2, a decomposição CEEMDAN é responsável por gerar uma quantidade M de FMIs que servirão como entrada para os M modelos determinados. Em outras palavras, cada modelo é treinado com uma FMI e, por fim, a soma das previsões de cada modelo separadamente resultará na previsão final. Dessa forma, pode avaliar o desempenho desta abordagem.

Abordagem Completo

A abordagem completa é uma extensão da abordagem de decomposição, como é mostrado no subcapítulo de Completo 3.4, em que apresenta como diferentes artigos trabalham neste campo de estudo. Primeiramente, é utilizada a abordagem de decomposição CEEMDAN 2.7, responsável por gerar uma quantidade M de FMIs, que servirão de entrada para os M modelos determinados. Juntamente com isso, a otimização dos melhores parâmetros para cada modelo usado em cada tipo de FMI ocorre por meio do AG, cuja funcionalidade é apresentada na otimização 4.4.1. O restante do algoritmo permanece inalterado: cada modelo é treinado com uma FMI e, no final, a soma das previsões de cada modelo separadamente resultará na previsão final, conforme representado na Figura 3.6.

4.4.2 Parâmetros dos Algoritmos

Foi utilizado *GridSearch* com Validação Cruzada para séries temporais com $K = 5$, usando o RMSE como métrica de pontuação para determinar os melhores parâmetros para os algoritmos MLP, SVR e CART. Para selecionar os parâmetros do ARIMA (p, d, q), foi utilizada uma abordagem automática que adotou as configurações padrão da biblioteca (PMDARIMA, 2023). A Tabela 4.2 apresenta a variação dos parâmetros que foram expostos no *GridSearch*.

Tabela 4.2: Parâmetros dos Algoritmos

Modelos	Parâmetros
CART	criterion: ['squared_error', 'friedman_mse', 'absolute_error','poisson']; splitter: ['best']; max_depth: [None, 10, 20, 30]; max_features: ['auto','sqrt', 'log2']; min_samples_split: [2, 5, 10], min_samples_leaf: [1, 2, 4]
MLP	hidden_layer_sizes: [(50, 50), (100, 100), (100, 50, 25)]; activation: ['relu', 'tanh', 'logistic']; 'solver': ['lbfgs', 'sgd','adam']; learning_rate': ['constant', 'invscaling', 'adaptive'], alpha: [0.0001, 0.001, 0.01]; batch_size: [32, 64, 128], max_iter: [200, 300, 400]

Continua na próxima página

Tabela 4.2 – continuação da página anterior

Modelos	Parâmetros
SVR	kernel: ['linear', 'rbf', 'sigmoid']; C: [0.1, 1, 10]; epsilon: [0.1, 0.2, 0.5], gamma: [0.1, 0.2, 0.5], max_iter: [1000, 10000,100000]
Bagging	Variação na quantidade de agoritmos [2, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100]
CEEMDAN	trials: [100, 200, 300]; epsilon: [0.001, 0.002, 0.005,]; noise_scalef: [0.5, 1]; noise_kind: [normal, uniform]; range_thr: 0.01; total_power_thr = 0.05; noise_seed: 42

Na otimização com o Algoritmo Genético, foram variados os seguintes parâmetros dentro de um determinado intervalo de valores, conforme apresentado na Tabela 4.3. As funções *randomInt* representam a escolha aleatória em um determinado espaço de valores inteiros, enquanto a função *randomFloat* representa a escolha aleatória em um intervalo de valores de ponto flutuante.

Tabela 4.3: Parâmetros dos Algoritmos Utilizados no Algoritmo Genético

Modelos	Parâmetros
CART	criterion: ['squared_error', 'friedman_mse', 'absolute_error','poisson']; splitter: ['best']; max_depth: [None, randomInt(10, 30)]; max_features: ['auto','sqrt', 'log2']; min_samples_split: randomInt(2, 5, 10), min_samples_leaf: randomInt(1, 2, 4)
MLP	hidden_layer_sizes: [(randomInt(1, 100), randomInt(1, 100)), (randomInt(1, 100), randomInt(1, 100), randomInt(1, 100))]; activation: ['relu', 'tanh', 'logistic']; 'solver': ['lbfgs', 'sgd','adam']; learning_rate': ['constant', 'invscaling', 'adaptive'], alpha: randomFloat(0.0001,0.01); batch_size: randomInt(32, 128), max_iter: randomInt(200, 400)
SVR	kernel: ['linear', 'rbf', 'sigmoid']; C: randomFloat(0.1, 10); epsilon: randomFloat(0.1, 0.5), gamma: randomFloat(0.1, 0.5), max_iter: randomInt(1000, 100000)

A Tabela 4.4 foi organizada com colunas representando os mercados e linhas representando os algoritmos e seus respectivos parâmetros selecionados.

Tabela 4.4: Parâmetros Selecionados nos Algoritmos

Modelos	IBOVESPA	S&P 500
ARIMA	(p = 1, d = 0, q = 1)	(p = 1, d = 0, q = 1)
CART	criterion: 'squared_error', max_depth: None, max_features: 'sqrt', min_samples_leaf: 4, min_samples_split: 2, splitter: 'best'	criterion: 'squared_error', max_depth: None, max_features: 'sqrt', min_samples_leaf: 1, min_samples_split: 10, splitter: 'best'
MLP	activation: 'tanh', alpha: 0.01, batch_size: 32, hidden_layer_sizes: (100, 100), learning_rate: 'constant', max_iter: 400, solver: 'adam'	activation: 'relu', alpha: 0.001, batch_size: 128, hidden_layer_sizes: (100, 50, 25), learning_rate: 'adaptive', max_iter: 400, solver: 'adam'
SVR	C: 0.1, epsilon: 0.5, gamma: 0.1, kernel: 'linear', max_iter: 1000	C: 0.1, epsilon: 0.1, gamma: 0.1, kernel: 'linear', max_iter: 10000
Bagging	CART: 20, MLP: 10, SVR: 10	CART: 20, MLP: 70, SVR: 30

Para o CEEMDAN, foram utilizados os seguintes parâmetros: trials = 200, epsilon = 0.005, noise_scalef = 1, noise_kind = "normal", range_thr = 0.01, total_power_thr = 0.05 e noise_seed = 42, aplicados a ambas as séries temporais. Nas abordagens *Bagging* e *Stacking*, foram adotados os parâmetros definidos por padrão da biblioteca (LEARN, 2023a) e (LEARN, 2023b), realizando modificações somente na quantidade de avaliadores que irão compor a abordagem.

Foi utilizado o método cotovelo [3] no *Bagging* homogêneo para determinar o número ótimo de avaliadores a serem incluídos na abordagem, conforme mostrado na Tabela 4.2 e na figura D.1. No *Stacking* e no *Bagging* heterogêneo, foi usado um representante de cada algoritmo.

4.5 Métricas de Avaliação

O MSE foi escolhido como métrica, pois atribui maior peso aos erros mais significativos. Isso ocorre porque, ao ser calculado, cada erro é individualmente elevado ao quadrado e, em seguida, é calculada a média desses erros quadrados. O RMSE é derivado do MSE, aplicando-se a raiz quadrada, deixando os valores na mesma escala numérica. Além disso, foi adotada a métrica do MAE, que usa o valor absoluto de cada erro para evitar subestimações, sendo menos afetada por outliers. A partir dessa métrica, deriva-se o MAPE, que representa a variação percentual do MAE no resultado.

Todas essas métricas têm múltiplas saídas, sendo adotado o parâmetro "*uniform_average*". Isso significa que os erros de todas as saídas são calculados com peso uniforme. As fórmulas para cada indicador são apresentadas nas Equações (4.2 a 4.5):

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - p_i)^2 \tag{4.2}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - p_i)^2} \tag{4.3}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|Y_i - p_i| \tag{4.4}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|Y_i - p_i|}{\max(\epsilon, |Y_i|)} \tag{4.5}$$

Em que i e n representam o índice e o número de dias de previsão, respectivamente, e Y e p denotam os valores reais e previstos, respectivamente.

Adicionalmente, a métrica de *Instructions Retired* da *Central Processing Unit* – (CPU) representa todas as instruções executadas pela CPU desde o início do algoritmo até sua conclusão, apresenta os resultados no intervalo de $[1, +\infty]$. Esta métrica é crucial, pois está relacionada ao processo de desempenho geral do algoritmo, abrangendo as fases de treinamento e teste dos

[3] Para mais informações (SCHUBERT, 2023)

algoritmos em cada execução, auxiliando na construção da métrica de custo-benefício. Essa abordagem é quantificada pela Equação 4.6, definida como $1 - Error(RMSE)$ (em que um valor mais alto indica melhor desempenho (menor erro)) sobre o $Custo$ mais e, um pequeno valor adicionado para evitar divisão por zero, garantindo que o denominador nunca seja exatamente zero. Essa métrica é útil para comparar modelos considerando o erro (RMSE) e o custo computacional (Instruções Executadas), especialmente quando há uma troca entre esses fatores.

$$Custo\_beneficio = \frac{1 - Error(RMSE)}{Custo(InstructionsRetired) + e} \tag{4.6}$$

A normalização desses valores é essencial para garantir que os custos e benefícios estejam em escalas comparáveis e evitar que um influencie desproporcionalmente o outro na formulação da métrica de eficiência. Nesse contexto, a técnica de normalização MinMax foi escolhida como método de pré-processamento para garantir que ambas as métricas contribuam igualmente para a análise. Essa abordagem redimensiona os valores para um intervalo típico [0, 1], facilitando uma comparação justa e eficaz entre custos e benefícios.

Capítulo 5

Resultados e Discussão

Este capítulo apresenta os resultados obtidos nos experimentos para cada tipo de série temporal financeira, especificamente para o IBOVESPA (5.1) e o S&P 500 (5.2). Aqui, é detalhado a Análise de Desempenho em cada cenário, ou seja, a comparação entre métodos individuais e métodos de *ensemble*, identificando as abordagens com as melhores classificações e determinando se elas são significativamente diferentes através de testes estatísticos. Finalmente, é realizada uma Análise Final (5.4), discutindo em quais cenários o uso de técnicas de *ensemble* da literatura é mais apropriado.

5.1 IBOVESPA

Os resultados mostrados nas Tabelas 5.1 e 5.2 representam a média de 30 execuções de teste usando cada um dos algoritmos para apresentar as métricas de avaliação adotadas neste estudo.

Tabela 5.1: IBOVESPA - Modelo *Single* - Resultados das métricas de avaliação. O primeiro e o segundo melhores estão destacados em negrito e sublinhado, respectivamente.

Modelos	**MSE**	**RMSE**	**MAE**	**MAPE**	***Instructions Retired***
ARIMA	<u>0,00463</u>	<u>0,06227</u>	<u>0,05383</u>	<u>6,728%</u>	<u>19.493.700.113</u>
CART	**0,00382**	**0,05640**	**0,04941**	**6,256%**	**12.320.479.943**
MLP	0,14292	0,36312	0,36082	43,438%	71.775.712.924
SVR	0,10834	0,31458	0,31199	37,291%	12.268.994.030

Como ilustrado na Tabela 5.1, o algoritmo CART demonstrou desempenho superior, registrando os menores valores para as métricas de erro: MSE, RMSE, MAE, MAPE e *Instructions Re-*

tired. Esse algoritmo foi seguido pelo modelo estatístico ARIMA. Em contraste, o modelo MLP apresentou o desempenho menos satisfatório entre os avaliados, caracterizado pela grande quantidade de instruções necessárias para sua execução e o mau desempenho nas métricas de erro.

Tabela 5.2: IBOVESPA - Abordagem *Ensemble* - Resultados de métricas de avaliação para as séries temporais usando métodos *ensemble*. O primeiro e o segundo melhores estão destacados em negrito e sublinhado, respectivamente.

Modelos	MSE	RMSE	MAE	MAPE	*Instructions Retired*
Bagging **CART**	<u>0,00395</u>	<u>0,05716</u>	<u>0,05002</u>	<u>6,264%</u>	**12.049.008.941**
Bagging **MLP**	0,14423	0,36466	0,36253	43,548%	140.231.957.752
Bagging **SVR**	0,10855	0,31483	0,31224	37,319%	<u>12.620.381.133</u>
Bagging **Heterogêneo**	0,05793	0,22851	0,22494	26,793%	109.298.840.640
Completo CART	0,01879	0,11669	0,11110	13,301%	223.640.677.943
Completo MLP	0,06096	0,22845	0,22490	27,830%	200.090.355.448
Completo SVR	0,09768	0,28284	0,27940	33,254%	192.943.863.845
Decomposição CART	**0,00376**	**0,05577**	**0,04867**	**6,164%**	123.153.141.966
Decomposição MLP	0,14718	0,36446	0,36209	43,576%	181.470.984.708
Decomposição SVR	0,12361	0,33360	0,33120	39,579%	122.375.178.270
Sistema Híbrido	0,00467	0,06223	0,05394	6,771%	19.496.429.026

Table 5.2 continuação da página anterior

Modelos	MSE	RMSE	MAE	MAPE	*Instructions Retired*
CART					
Sistema Híbrido MLP	0,00493	0,06401	0,05624	7,057%	70.037.950.118
Sistema Híbrido SVR	0,02688	0,15664	0,15210	18,483%	19.461.394.863
Otimização CART	0,01728	0,11167	0,10628	12,994%	22.194.626.270
Otimização MLP	0,10295	0,30025	0,29752	37,145%	204.404.277.564
Otimização SVR	0,12448	0,29349	0,29038	36,323%	198.348.861.617
***Stacking* CART**	0,17881	0,39750	0,39540	47,780%	129.533.082.871
***Stacking* MLP**	0,04734	0,20124	0,19695	23,297%	42.505.124.438
***Stacking* SVR**	0,10834	0,31458	0,31199	37,292%	132.099.222.878

Como ilustrado na Tabela 5.2, a abordagem de decomposição CART destacou-se ao atingir as menores métricas de erros tradicionais. Em seguida, a técnica *Bagging* CART mostrou-se a segunda mais eficiente em termos de métricas de erro tradicional, além de apresentar o melhor desempenho na métrica de *Instructions Retired*. O modelo *Bagging* SVR alcançou a segunda melhor performance em *Instructions Retired*. Uma análise detalhada dos resultados, agrupados conforme as técnicas de *ensemble* aplicadas, revelou que, dentro do grupo que utilizou *Bagging*, o modelo CART superou os demais, com o modelo Heterogêneo ficando em segundo lugar. No grupo da decomposição e completo, o modelo CART novamente se destacou, seguido pelo modelo SVR. No espectro de sistemas híbridos e técnicas de otimização, o CART demonstrou superioridade, seguido pelo MLP. Na categoria *Stacking*, o MLP alcançou a primeira posição

em desempenho, seguido pelo SVR. Esses achados indicam uma tendência de alto desempenho para configurações que incorporam o CART como modelo base, com uma exceção notável na categoria *Stacking*.

Para verificar a dispersão dos resultados, ou seja, se o modelo apresentou resultados consistentes, foi realizada uma análise utilizando a métrica MAE, que representa o erro absoluto médio, juntamente com a métrica MAPE, que permite avaliar a variação percentual do erro nos resultados. Quanto menor a porcentagem, mais consistentes são os resultados (HYNDMAN; ATHANASOPOULOS, 2018). Assim, é observável na Figura 5.1 e na Tabela 5.2 que a implementação da modelagem *Ensemble* resulta em menor dispersão dos dados, dependendo da abordagem selecionada. Um exemplo disso é a comparação entre o Modelo CART e a decomposição CART, em que se notou uma redução na variabilidade. No entanto, outras abordagens pioraram o resultado, como o *Stacking*, indicando que os resultados são menos consistentes, o que torna as previsões menos confiáveis e precisas.

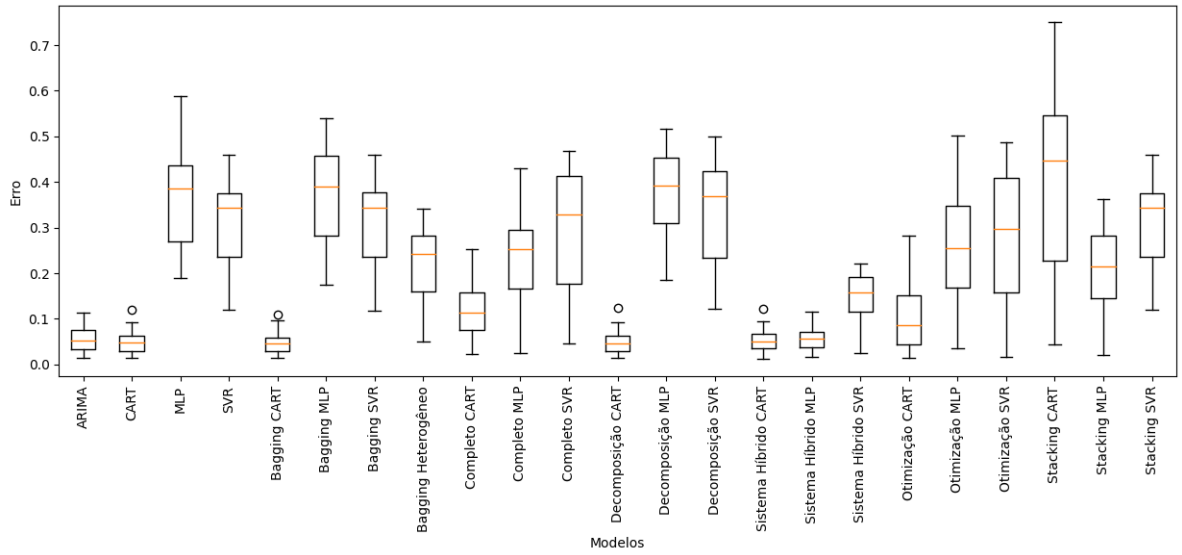


Figura 5.1: *Boxplot* dos Resultados dos Modelos *Single* e *Ensemble* da métrica MAE, Fonte: (Autor, 2024)

5.1.1 Teste de Hipótese - Métricas Tradicionais

Conforme discutido na Seção 5.1, que realiza uma comparação generalizada de cada tipo de abordagem *single* e *ensemble*, foram realizados testes de hipótese para investigar a existência de diferenças estatisticamente significativas entre as abordagens mencionadas. Para essa análise comparativa, foi adotada a métrica RMSE em conjunto com o teste de Wilcoxon, um método não-paramétrico de significância estatística, aplicado com um nível de confiança de

95%. Este teste é particularmente adequado para comparar as medianas dos modelos, aplicável mesmo quando a distribuição dos dados não segue a normalidade. As hipóteses testadas foram definidas da seguinte forma:

- H_0: Não há evidências suficientes para afirmar que existe uma diferença significativa entre as abordagens.
- H_1: Há evidências suficientes para afirmar que existe uma diferença significativa entre as abordagens.

Tabela 5.3: IBOVESPA - Modelo *Single* - Teste de hipótese de Wilcoxon com 95% de confiança. Os símbolos (+), (-) ou (=) representam se a comparação com cada abordagem pode ser melhor, pior ou equivalente, respectivamente.

Modelos	**ARIMA**	**CART**	**MLP**	**SVR**	**Total**
ARIMA	=	-	+	+	2
CART	+	=	+	+	3
MLP	-	-	=	-	0
SVR	-	-	+	=	1

A Tabela 5.3, o melhor modelo *single* foi o CART com 3 pontos positivos, o ARIMA com 2 pontos ,enquanto o MLP teve um total de 0 pontos positivos, sendo o mais mal classificado. Ao observar ambas as Tabelas 5.3 e C.1 em conjunto, o modelo individual CART alcançou 19 pontos, seguido pelo ARIMA com 17 pontos positivos.

Além disso, nas Tabelas C.1, C.2, C.3 e C.4, foram adotadas cores para representar os modelos base: o quadrado colorido em rosa ( ) representa o modelo CART, o quadrado laranja ( ) representa o modelo MLP, e o quadrado amarelo ( ) representa o modelo SVR.

Na Tabela C.1 dos Algoritmos *Ensemble*, conforme destacado na análise geral, o modelo Decomposição CART obteve o maior número de pontos positivos (+), totalizando 19. Em seguida, o *Bagging* CART alcançou um total de 18 pontos, e os modelos Sistema Híbrido CART e Sistema Híbrido MLP, 17 pontos, respectivamente. Além disso, dentro do grupo de abordagens de técnicas de *ensemble*, o Sistema Híbrido demonstrou o melhor desempenho geral, isto é, os três modelos obtiveram uma melhora equilibrada em seus pontos comparados com as outras abordagens, onde um modelo base se destacava enquanto os demais apresentavam resulta-

dos inferiores. É importante notar que os modelos CART e Decomposição CART possuem o mesmo número de pontos positivos e são estatisticamente equivalentes.

Os testes de hipóteses de *Friedman* e *Nemenyi* são utilizados para classificar todos os modelos em todas as execuções das séries temporais (HOLLANDER; WOLFE; CHICKEN, 2013). O teste de hipótese de *Friedman* compara a classificação média dos resultados dos modelos, considerando a hipótese nula de que as classificações dos modelos são equivalentes. Se a hipótese nula do teste de *Friedman* for rejeitada, conclui-se que as classificações dos modelos são significativamente diferentes. Nesse caso, é possível utilizar o teste de *Nemenyi*, que aplica uma comparação pareada dos modelos em todas as séries temporais. Esse teste considera um grupo de w modelos, c execuções nas séries temporais e o valor crítico (q_α) baseado no intervalo estatístico estudentizado (DEMŠAR, 2006). Se a diferença entre as classificações for maior que *Critical Distance* – (CD) representado na equação 5.1, os modelos são considerados diferentes.

$$CD = q_\alpha \sqrt{\frac{w(w+1)}{6c}} \tag{5.1}$$

Considerando a classificação (baseada no RMSE) dos vinte e três métodos ($w = 23$) nas 30 execuções ($c = 30$), a hipótese nula do teste de *Friedman* com 95% de confiança (p-value $= 5,48 \times e^{-87}$) pode ser rejeitada. Assim, é possível concluir que as classificações dos métodos são significativamente diferentes, permitindo prosseguir para o teste de *Nemenyi*. O teste de *Nemenyi* utiliza $q_\alpha = 3,619$ ($\alpha = 0,05$ ou 95% de confiança), resultando em $CD = 1,16$.

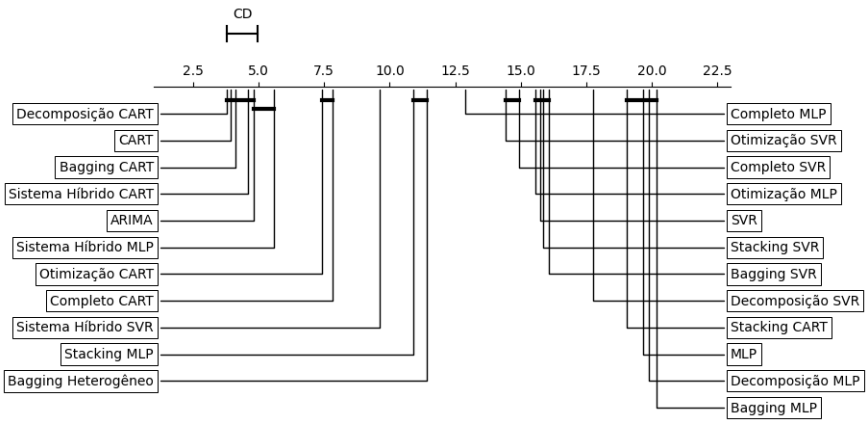


Figura 5.2: Classificação do teste de *Nemenyi* considerando o RMSE, Fonte: (Autor, 2024)

A Figura 5.2 apresenta o teste de hipótese de *Nemenyi*, que indica que a Decomposição CART, o modelo *Single* CART, *Bagging* e o Sistema Híbrido que utilizam o CART como base, juntamente com o ARIMA, não apresentam diferenças significativas. A figura também mostra

que as abordagens propostas que utilizaram o modelo CART como base atingiram classificações significativamente melhores em comparação aos modelos que utilizaram o SVR e MLP. Além disso, pode-se observar o baixo desempenho do modelo MLP, ficando estatisticamente igual às abordagens *ensemble* que o utilizaram como base, nas últimas posições. Isso demonstra a importância de utilizar um modelo base adequado, pois os resultados dependem do desempenho adquirido pelo modelo individualmente.

5.1.2 Teste de Hípotese - Custo-Benefício

A análise comparativa focada no custo-benefício destes modelos é proposta para avaliar a viabilidade da aplicação de modelos *ensemble*. Esta análise visa determinar se, além da eficácia dos algoritmos, eles são eficientemente otimizados em termos de recursos computacionais, minimizando a complexidade desnecessária em cenários que exigem soluções mais simples e rápidas. Para esse fim, é utilizada uma métrica de custo-benefício, que avalia a eficiência relativa entre diferentes alternativas, comparando seus custos e os benefícios que oferecem.

Tabela 5.4: IBOVESPA - Comparação de Custo-benefício *Single* - Teste de hipótese de Wilcoxon com 95% de confiança. Os símbolos (+), (-) ou (=) indicam se a comparação com cada abordagem pode ser melhor, pior ou equivalente, respectivamente.

Modelos	**ARIMA**	**CART**	**MLP**	**SVR**	**Total**
ARIMA	=	-	+	-	1
CART	+	=	+	+	3
MLP	-	-	=	-	0
SVR	+	-	+	=	2

Na Tabela 5.4, observa-se que o modelo CART se destaca, apresentando três aspectos positivos (+++), posicionando-se à frente dos outros modelos individuais. O modelo SVR segue, acumulando dois aspectos positivos (++). Esta análise, que abrange tanto a métrica RMSE quanto a eficiência computacional, demonstra que o modelo CART supera consistentemente os outros em termos de desempenho, marcando uma distinção significativa em comparação aos outros modelos. O SVR, embora apresente resultados inferiores em termos da métrica RMSE, é notável por sua eficiência computacional.

A análise dos dados na Tabela C.2 revela que os modelos *ensemble* do grupo *Bagging*, utilizando os algoritmos CART e SVR como bases, alcançaram o melhor desempenho em termos

de custo-benefício, totalizando 21 pontos positivos. Quando focados nas métricas de desempenho tradicional, a Decomposição CART se destacou como o melhor, embora, comparado às demais abordagens, tenha ficado em 9º lugar. Comparando os resultados entre a Tabela 5.4 e a Tabela C.2, nota-se que o modelo individual CART acumula um total de 20 pontos positivos, superando outros modelos individuais e perdendo apenas aos modelos *ensemble Bagging* CART e SVR. Essa observação destaca a eficácia dos modelos individuais quando selecionados e configurados adequadamente para o problema em questão, potencialmente superando outros algoritmos em desempenho.

Considerando a classificação (baseada no custo-benefício) dos vinte e três métodos ($w = 23$) nas 30 execuções ($c = 30$), a hipótese nula do teste de *Friedman* com 95% de confiança (p-value $= 4,08 \times e^{-117}$) pode ser rejeitada. Assim, é possível concluir que as classificações dos métodos são significativamente diferentes, permitindo prosseguir para o teste de *Nemenyi*. O teste de *Nemenyi* utiliza $q_\alpha = 3,619$ ($\alpha = 0,05$ ou 95% de confiança), resultando em $CD = 1,16$.

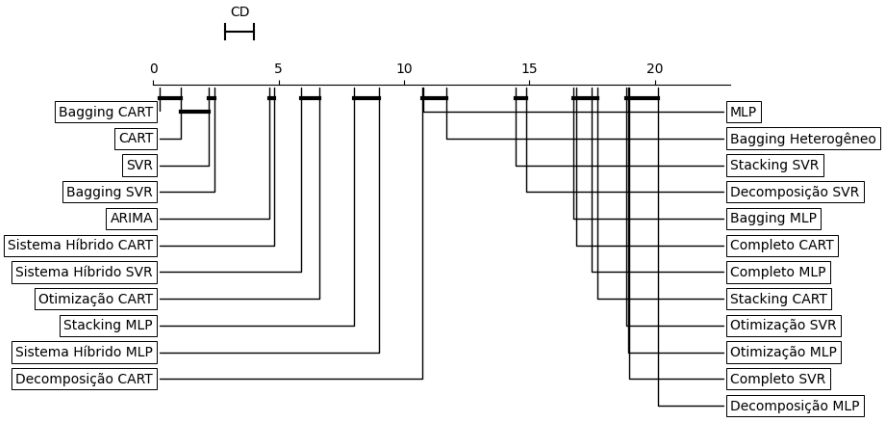


Figura 5.3: Classificação do teste de *Nemenyi* considerando o Custo-benefício, Fonte: (Autor, 2024)

Na Figura 5.3, já é possível observar uma diferença em relação à Figura 5.2, que avalia a métrica tradicional e apresenta como melhor desempenho geral a abordagem de Decomposição CART. Observando a métrica de custo-benefício, a Decomposição CART ficou em 9º lugar, o que concretiza os resultados encontrados na soma dos pontos positivos pelo teste de Wilcoxon. Dessa forma, é visto que a técnica de *ensemble Bagging* CART apresentou o melhor desempenho, seguida pelos modelos simples CART, SVR e ARIMA, como esperado dos modelos *single* por serem algoritmos base que exigem pouco custo computacional para funcionar, embora percam na entrega dos resultados. É possível analisar que a abordagem *ensemble* de Sistema Híbrido possui um equilíbrio, como apresentado nas Figuras 5.3 e 5.2. Nas últimas posições

ficaram as abordagens de Decomposição, Completo e Otimização que utilizam os modelos SVR e MLP como base.

5.2 S&P 500

Conforme realizado na Seção 5.1, a mesma análise foi conduzida em um mercado desenvolvido para observar as implicações das abordagens *single* e *ensemble*.

Tabela 5.5: S&P 500 - Modelo *Single*- Resultados das Métricas de Avaliação. O primeiro e o segundo melhores estão destacados em negrito e sublinhado, respectivamente.

Modelos	MSE	RMSE	MAE	MAPE	*Instructions Retired*
ARIMA	**0,00086**	**0,02582**	**0,02236**	**2,562%**	17.038.779.645
CART	<u>0,00091</u>	<u>0,02677</u>	<u>0,02351</u>	<u>2,701%</u>	**12.332.897.208**
MLP	0,06144	0,15448	0,15311	17,849%	15.786.669.477
SVR	0,00711	0,07545	0,07347	8,270%	<u>12.809.102.266</u>

Conforme ilustrado na Tabela 5.5, o modelo ARIMA se destacou por apresentar os menores valores nas métricas de erro MSE, RMSE, MAE e MAPE. No entanto, em relação à métrica de *Instructions Retired*, o ARIMA foi superado pelos modelos CART e SVR. O modelo CART, por sua vez, alcançou o segundo melhor desempenho nas métricas de erro mencionadas, distinguindo-se com o menor número de Instruções Executadas. Em contraste, o modelo MLP registrou o desempenho menos satisfatório entre os avaliados.

Tabela 5.6: S&P 500 - Abordagem *Ensemble*- Resultados das Métricas de Avaliação. O primeiro e o segundo melhores estão destacados em negrito e sublinhado, respectivamente.

Modelos	MSE	RMSE	MAE	MAPE	*Instructions Retired*
Bagging **CART**	0,00099	0,02764	0,02443	2,808%	36.075.288.354
Bagging **MLP**	0,00947	0,07419	0,07198	8,159%	163.516.409.251
Bagging **SVR**	0,00509	0,06477	0,06270	7,032%	**15.997.385.402**

Table 5.6 continuação da página anterior

Modelos	MSE	RMSE	MAE	MAPE	*Instructions Retired*
***Bagging* Heterogêneo**	0,18305	0,15081	0,14898	16,827%	67.509.956.442
Completo CART	0,00214	0,04021	0,03750	4,237%	224.515.380.649
Completo MLP	0,01295	0,09766	0,09615	10,907%	220.178.572.177
Completo SVR	0,03019	0,16904	0,16818	19,021%	170.628.040.505
Decomposição CART	0,00094	0,02662	0,02330	2,679%	128.690.006.620
Decomposição MLP	0,23480	0,39766	0,39714	41,909%	174.042.180.430
Decomposição SVR	0,02493	0,15333	0,15240	17,231%	128.062.464.487
Sistema Híbrido CART	**0,00082**	**0,02467**	**0,02145**	**2,450%**	17.118.060.027
Sistema Híbrido MLP	0,13101	0,15279	0,15049	16,763%	21.679.902.684
Sistema Híbrido SVR	0,03637	0,15622	0,15473	17,592%	17.738.881.416
Otimização CART	0,00230	0,04218	0,03959	4,485%	17.432.091.140
Otimização MLP	0,00732	0,06541	0,06344	7,237%	212.491.426.858
Otimização SVR	0,00824	0,07901	0,07725	8,669%	27.844.025.023

Table 5.6 continuação da página anterior

Modelos	MSE	RMSE	MAE	MAPE	*Instructions Retired*
Stacking **CART**	0,01316	0,07768	0,07434	8,342%	51.199.944.361
Stacking **MLP**	0,00194	0,03912	0,03635	4,127%	19.213.938.601
Stacking **SVR**	0,00913	0,06650	0,06434	7,363%	47.206.119.850

Conforme indicado na Tabela 5.6, a abordagem do Sistema Híbrido utilizando o modelo CART destacou-se por apresentar os melhores resultados nas métricas de erro. No entanto, ocupou a segunda posição na métrica de *Instructions Retired*. O modelo *Bagging* CART foi classificado como o segundo mais eficaz em relação às métricas de erro, enquanto o modelo *Bagging* SVR alcançou o melhor desempenho em *Instructions Retired*.

Ao agrupar os resultados por técnicas de *ensemble*, observou-se que, dentro do grupo *Bagging*, o modelo CART foi o mais eficiente, seguido pelo SVR. No grupo que adotou o método Completo, os modelos CART e MLP foram notáveis. Para os grupos focados em Decomposição e Sistemas Híbridos, o CART novamente se mostrou superior, com o SVR em segundo lugar. Nas abordagens de Otimização, o CART e o MLP apresentaram os melhores resultados. Finalmente, o MLP mostrou o desempenho mais destacado na técnica *Stacking*, seguido pelo SVR.

Esses achados sugerem que as configurações usando o modelo CART como base tiveram um desempenho notavelmente bom na maioria das variantes de ensemble analisadas, exceto na estratégia de *Stacking*.

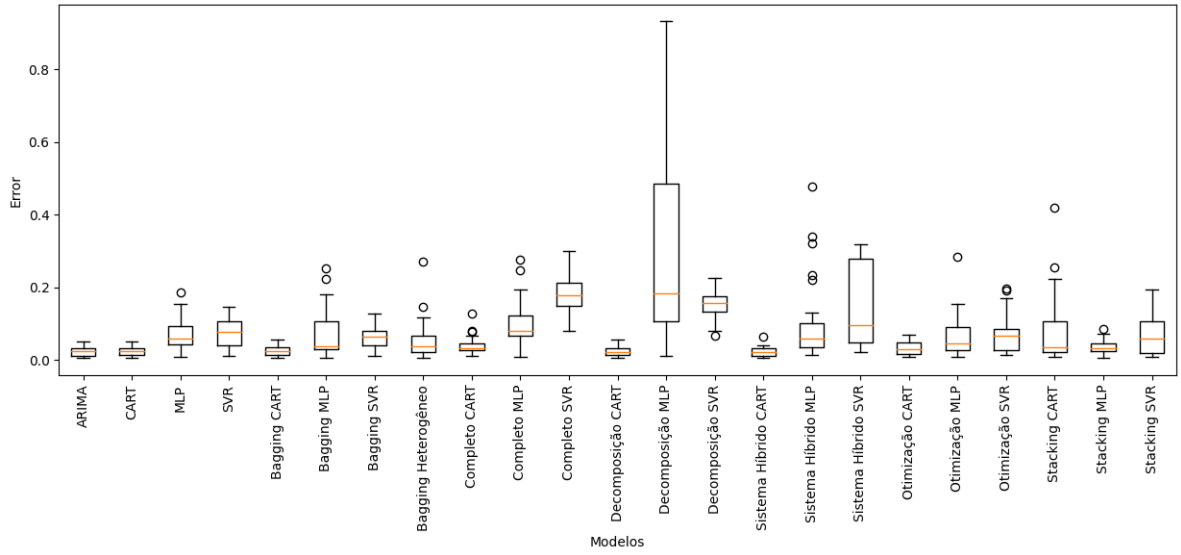


Figura 5.4: Boxplot dos Resultados dos Modelos *Single* e de *Ensemble* na métrica MAE, Fonte: (autor, 2024)

Conforme observado na Figura 5.4, e nas Tabelas 5.5 e 5.6, ao implementar a modelagem *ensemble*, os dados mostram menor dispersão, dependendo da abordagem selecionada. Um exemplo é o algoritmo MLP, cuja variação percentual do MAPE é de 17.849%, com uma redução na aplicação *Stacking* MLP, que apresenta uma variação de 4.127%. No entanto, outras abordagens pioraram os resultados, como a Decomposição, que alcançou 41.909%.

5.2.1 Teste de Hipótese - Métricas Tradicionais

Esta seção apresenta os resultados obtidos a partir do teste de hipótese de Wilcoxon, realizado com 95% de confiança. Os resultados estão mostrados na Tabela 5.7 e na Tabela C.3.

Tabela 5.7: S&P 500 - Modelo *Single* - Teste de hipótese de Wilcoxon com 95% de confiança. Os símbolos (+), (-) ou (=) indicam se a comparação com cada abordagem pode ser melhor, pior ou equivalente, respectivamente.

Modelos	**ARIMA**	**CART**	**MLP**	**SVR**	**Total**
ARIMA	=	=	+	+	2
CART	=	=	+	+	2
MLP	-	-	=	=	0
SVR	-	-	=	=	0

Na Tabela 5.7, observou-se um empate para o algoritmo com melhor desempenho, envolvendo o CART e o ARIMA, ambos com 2 pontos positivos (+). Da mesma forma, houve um empate para as piores classificações entre MLP e SVR, cada um com um total de 0 pontos positivos. Ao analisar as duas tabelas juntas (5.7 e C.3), os modelos CART e ARIMA alcançaram um total de 18 pontos positivos.

Na Tabela C.3, que apresenta os resultados do teste de Wilcoxon dos algoritmos de *ensemble*, observou-se um empate entre os modelos *Bagging*, Decomposição e Sistema Híbrido que utilizaram o modelo *single* CART como base. Cada um alcançou o maior número de pontos positivos (+), com 18 pontos cada. Por outro lado, as abordagens de Decomposição, Completa, Otimização e *Stacking*, que tiveram como base os algoritmos MLP e SVR, obtiveram os piores desempenhos.

Considerando a classificação (baseada no RMSE) dos vinte e três métodos ($w = 23$) nas 30 execuções ($c = 30$), a hipótese nula do teste de *Friedman* com 95% de confiança (p-value $= 1,28 \times e^{-50}$) pode ser rejeitada. Assim, é possível concluir que as classificações dos métodos são significativamente diferentes, permitindo prosseguir para o teste de *Nemenyi*. O teste de *Nemenyi* utiliza $q_\alpha = 3,619$ ($\alpha = 0,05$ ou 95% de confiança), resultando em $CD = 1,16$.

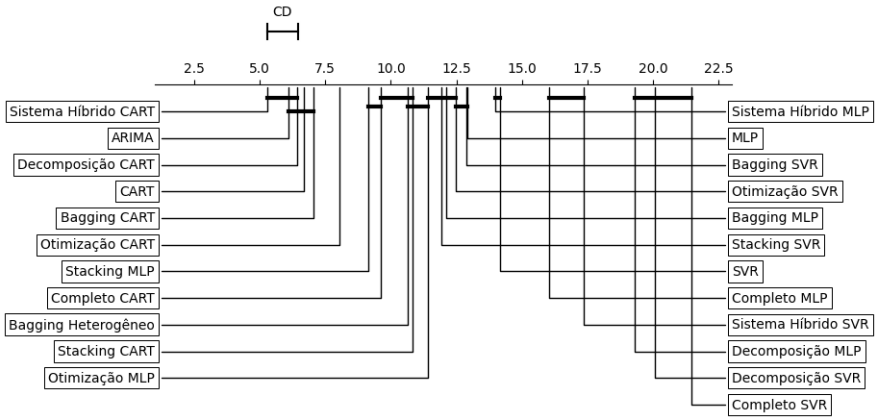


Figura 5.5: Classificação do teste de *Nemenyi* considerando o RMSE, Fonte: (Autor, 2024)

Na Figura 5.5, dentro do grupo de técnicas de *ensemble*, o modelo Sistema Híbrido demonstrou o melhor desempenho geral, devido ao bom desempenho individual dos modelos CART e ARIMA, o que reflete diretamente no resultado final. Outro aspecto importante é o baixo desempenho relacionado às abordagens que utilizam SVR e MLP como modelos base,

o que é evidenciado na figura, mostrando assim os resultados inferiores das abordagens de Decomposição e Completa.

5.2.2 Teste de Hípotese - Custo-Benefício

Tabela 5.8: S&P 500 - Comparação de Custo-benefício dos modelos *Single* - Teste de hipótese de Wilcoxon com 95% de confiança. Os símbolos (+), (-) ou (=) indicam se a comparação com cada abordagem pode ser melhor, pior ou equivalente, respectivamente.

Modelos	**ARIMA**	**CART**	**MLP**	**SVR**	**Total**
ARIMA	=	-	-	-	0
CART	+	=	+	+	3
MLP	+	-	=	-	1
SVR	+	-	+	=	2

Como mostrado na Tabela 5.8, o modelo CART se destaca ao alcançar um total de três pontos positivos (+), posicionando-se como a melhor opção custo-benefício entre os modelos *single* avaliados. Em seguida, o modelo SVR é identificado como a segunda melhor opção, acumulando dois pontos positivos (+). O pior modelo foi o ARIMA, com um total de zero pontos. Juntando as Tabelas 5.8 e C.4, observa-se que o modelo CART apresenta o melhor custo-benefício, com um total de 22 pontos positivos, seguido pelo modelo SVR, com um total de 21 pontos positivos. Em seguida, o SVR obteve 19 pontos e, por último, o ARIMA com 16 pontos.

Na Tabela C.4 que tem o foco nas abordagem de *ensemble* apresentou o *Bagging* SVR alcançou o desempenho mais notável, totalizando 19 pontos positivos. Entre as técnicas de *ensemble*, a abordagem de Sistema Híbrido provou ser a mais eficaz, oferecendo melhorias de custo-benefício para os três algoritmos selecionados.

Considerando a classificação (baseada no Custo-benefício) dos vinte e três métodos ($w = 23$) nas 30 execuções ($c = 30$), a hipótese nula do teste de *Friedman* com 95% de confiança (p-value $= 7,75 \times e^{-121}$) pode ser rejeitada. Assim, é possível concluir que as classificações dos métodos são significativamente diferentes, permitindo prosseguir para o teste de *Nemenyi*. O teste de *Nemenyi* utiliza $q_\alpha = 3,619$ ($\alpha = 0,05$ ou 95% de confiança), resultando em $CD = 1,16$.

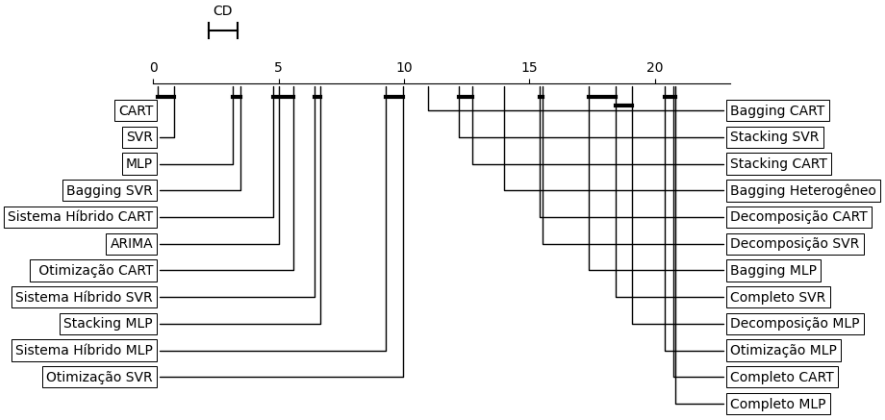


Figura 5.6: Classificação do teste de *Nemenyi* considerando o Custo-Benefício, Fonte: (Autor, 2024)

Na Figura 5.6, são destacados os modelos simples CART e SVR, que são estatisticamente iguais. Nas técnicas de *ensemble*, o *Bagging* composto pelo modelo base SVR foi o que recebeu maior destaque, seguido pelo bom desempenho equilibrado da abordagem de Sistema Híbrido nos três modelos base adotados. Já as técnicas de Decomposição e Completo apresentam um baixo desempenho, vinculado principalmente ao alto gasto de recursos computacionais para um desempenho inferior, como observado na Tabela 5.6.

5.3 Análise das Métricas de Avaliação

Neste subcapítulo, foi realizada uma visão geral das métricas de avaliação utilizadas neste estudo, abordando suas vantagens e desvantagens em diferentes contextos de mercado. As métricas consideradas incluem MSE, RMSE, MAE, MAPE e custo-benefício. As análises realizadas foram baseadas no artigo de (BOTCHKAREV, 2019).

5.3.1 MSE (*Mean Squared Error*)

O MSE é uma métrica que calcula a média dos quadrados dos erros, penalizando de forma mais significativa os grandes erros.

- **Prós:**
  - **Sensível a grandes erros:** Penaliza mais os erros grandes devido à elevação ao quadrado, o que pode ser útil quando grandes desvios são indesejáveis.

- **Contras:**
  - **Unidade ao quadrado:** pode dificultar a interpretação direta dos resultados.
  - **Penaliza fortemente *outliers*:** Como os erros são elevados ao quadrado, outliers têm um impacto desproporcionalmente grande no valor do MSE

Para mostrar essa diferença, pode-se visualizar a Tabela 5.1 e a Tabela 5.2 dos resultados dos modelos *single* e *ensemble* do IBOVESPA, que apresentam um MSE médio superior na maioria dos modelos quando comparados com as Tabelas 5.5 e 5.6 do mercado S&P 500. Isso pode ser interpretado como uma indicação de que o IBOVESPA apresenta mais *outliers* do que o S&P 500. Quando um conjunto de dados, como o IBOVESPA, contém mais *outliers*, o MSE tende a ser inflacionado, o que pode não refletir a verdadeira performance do modelo em prever a maioria dos dados, tornando-se inadequado para a comparação de modelos.

5.3.2 RMSE (*Root Mean Squared Error*)

O RMSE é a raiz quadrada do MSE, trazendo a métrica de volta para a mesma unidade dos dados.

- **Prós:**
  - **Mesma escala dos dados:** Como é a raiz quadrada do MSE, RMSE tem a mesma escala dos dados, facilitando a interpretação.
  - **Sensível a grandes erros:** Como o MSE, o RMSE também penaliza fortemente grandes desvios.
- **Contras:**
  - **Penalização de *outliers*:** Assim como o MSE, o RMSE é influenciado por *outliers*.

Isso é observado nas Tabelas 5.1, 5.2, 5.5 e 5.6. Quando comparamos os resultados MSE, observamos uma discrepância na escala dos dados. Já com as métricas RMSE, essa discrepância é mitigada, como mostrado nas Tabelas, facilitando a interpretação e a comparação dos modelos.

5.3.3 MAE (Mean Absolute Error)

O MAE calcula a média dos erros absolutos, oferecendo uma medida clara do erro médio.

- **Prós:**
  - **Fácil de interpretar:** Representa a média dos erros absolutos, tornando-a intuitiva e fácil de entender.
- **Contras:**
  - **Menos sensível a grandes erros:** Como não penaliza os *outliers*, é menos eficaz em destacar estes grandes desvios.

Não penaliza tanto os *outliers* quanto o MSE ou o RMSE, tornando-se menos informativo em certos contextos, ou seja, pode não refletir adequadamente a performance do modelo. Observando o contexto financeiro, os algoritmos precisam estar preparados para mudanças abruptas do mercado, e uma forma de verificar isso é por meio da percepção dos *outliers*. Sendo assim, não é recomendada a utilização desta métrica para comparação entre os modelos.

5.3.4 MAPE (Mean Absolute Percentage Error)

O MAPE calcula a média dos erros percentuais absolutos, proporcionando uma medida relativa do erro.

- **Prós:**
  - **Interpretação intuitiva:** Fornece erros em termos percentuais, facilitando a compreensão e comparação entre diferentes modelos.
- **Contras:**
  - **Problemas com valores próximos de zero:** MAPE pode ser indefinido ou muito grande se os valores reais forem próximos de zero.
  - **Bias para predições subestimadas:** Pode ser enviesado quando as previsões são sistematicamente menores que os valores reais.

Como o MAPE é derivado da métrica MAE, as desvantagens apresentadas no subcapítulo 5.3.3 também podem ser aplicadas ao MAPE, tornando-se não recomendada a sua utilização no contexto de comparação financeira.

5.3.5 Análise de Custo-Benefício

A análise de custo-benefício considera o impacto computacional em conjunto com o retorno das predições realizadas pelos modelos preditivos. Esta métrica permite observar se os resultados entregues pelo modelo realmente compensam o seu custo computacional, o que é essencial para avaliar a eficácia e a eficiência dos algoritmos.

- **Prós:**
  - **Relevância Computacional:** Integra diretamente os resultados do modelo com o impacto computacional, facilitando a tomada de decisões sobre a viabilidade de uso do modelo em produção.
  - **Eficiência de Recursos:** Ajuda a identificar modelos que fornecem boas predições com menor consumo de recursos computacionais, o que é crucial para aplicações em tempo real e em larga escala.
  - **Avaliação Holística:** Proporciona uma visão abrangente do desempenho do modelo, considerando tanto a precisão das predições quanto o custo associado ao treinamento e à execução.
- **Contras:**
  - **Dificuldade de Padronização:** Comparar esta métrica com outras métricas de erro pode ser desafiador, pois envolve fatores computacionais que nem sempre são diretamente comparáveis.
  - **Influência de Fatores Externos:** Pode ser influenciada por variáveis externas ao modelo, como a arquitetura do hardware, a eficiência do código e outros fatores operacionais que podem distorcer a análise.
  - **Complexidade Adicional:** Requer uma análise detalhada do custo computacional, que pode não ser trivial de medir e pode variar significativamente dependendo do ambiente de execução.

A métrica de custo-benefício é particularmente relevante no contexto financeiro, onde a eficiência computacional e a precisão das predições são cruciais. No entanto, sua aplicação deve ser cuidadosamente considerada, especialmente quando comparada com métricas de erro tradicionais como MSE, RMSE, MAE e MAPE.

Em resumo, a análise de custo-benefício pode ser uma métrica poderosa para avaliar modelos preditivos, desde que suas limitações sejam compreendidas e consideradas. Para decisões estratégicas, onde o custo computacional é uma preocupação significativa, esta métrica oferece *insights* valiosos sobre a eficiência do modelo. No entanto, para comparações puramente baseadas na precisão das predições, pode ser mais adequado usar métricas de erro tradicionais.

5.4 Considerações Finais

É crucial destacar as diferenças nos resultados observados entre os diferentes tipos de mercados. Como mostrado na comparação das métricas de avaliação entre a Tabela 5.2 e a Tabela 5.6, o mercado emergente, representado pelo IBOVESPA, apresentou erros mais significativos em comparação com o mercado desenvolvido, simbolizado pelo S&P 500. Esse contraste sublinha a relevância de considerar diferentes categorias de mercado nas análises comparativas, conforme evidenciado por estudos anteriores (HENRIQUE; SOBREIRO; KIMURA, 2019) e (KUMBURE et al., 2022). Além disso, é essencial notar que os modelos de *ensemble* que apresentaram melhor desempenho variaram significativamente entre os mercados. No IBOVESPA, o modelo Decomposição CART mostrou superioridade, enquanto no S&P 500, o Sistema Híbrido Residual Cart se destacou. Tal variação enfatiza a importância de testar um algoritmo em múltiplos ambientes para validar sua eficácia.

A aplicação de modelos como CART, SVR e MLP na previsão de séries temporais financeiras enfrenta desafios significativos devido à volatilidade e não estacionariedade inerentes desses dados. Enquanto o modelo CART é eficaz na captura de relações não lineares, ele é suscetível ao *overfitting* na presença de dados ruidosos, necessitando de estratégias rigorosas de poda e validação cruzada para mitigar este risco. Por outro lado, o SVR requer uma seleção cuidadosa das funções *kernel* e seus parâmetros associados para modelar adequadamente a dinâmica complexa dos dados financeiros, pois eles têm um forte impacto no tempo de convergência do modelo. Quanto ao MLP, este modelo pode identificar relações complexas nos dados, mas pode ter dificuldades em capturar características de dependência temporal de longo prazo, que o caso das séries temporais financeiras. Isso se deve em parte à sua estrutura de aprendizagem padrão, que carece de mecanismos específicos para reter informações temporais por longos períodos.

Além disso, é importante destacar a forte ligação entre a arquitetura de construção da abordagem *ensemble* e o modelo *single*. Conforme apresentado nas Tabelas C.2 e C.4, os resultados das métricas ficaram próximos aos das métricas de avaliação do modelo *single*.

As abordagens *bagging single* e heterogênea estão fortemente ligadas ao tipo de modelo de *machine learning* adotado. Devido a isso, a etapa de embaralhamento dos dados pode apresentar problemas, pois algoritmos como o CART podem ser beneficiados conforme sua construção, como apresentado no subcapítulo 2.9.2. Por outro lado, os algoritmos MLP e SVR não apresentam essa dinâmica, ou seja, os dados de entrada precisam ser considerados na ordem temporal para que ocorra um aprendizado coerente, como apresentado nos subcapítulos 2.10 e 2.8, respectivamente. Isso é evidenciado nos resultados apresentados na Tabelas C.1 e C.3. Por outro lado, o custo-benefício apresenta um resultado satisfatório, mas depende fortemente da quantidade de modelos adotados na construção da arquitetura.

No *stacking*, por apresentar uma arquitetura de meta-aprendizado, em que um modelo secundário aprende a partir das previsões dos modelos base. Em contextos de séries temporais, essa abordagem pode resultar na propagação e acumulação de erros dos modelos base, comprometendo a precisão das previsões finais do algoritmo de meta-aprendizagem, especialmente se os modelos base não capturarem adequadamente as dependências temporais.

A decomposição temporal de séries temporais financeiras tem como objetivo dividir os dados originais em componentes mais simples, facilitando uma representação mais precisa e útil para análises subsequentes. No entanto, o processamento do sinal em FMI apresenta um elevado custo computacional, cujo retorno está diretamente relacionado ao modelo e ao mercado adotado. Em mercados com baixa volatilidade, como o S&P 500, a construção das FMI é rápida, mas não traz um retorno significativo nos resultados finais. Já em mercados voláteis, como o IBOVESPA, a decomposição tem um impacto positivo na melhora dos resultados, conforme observado nas Tabelas C.1 e C.3. Entretanto, a integração dessa técnica com modelos de *machine learning* introduz uma complexidade considerável. Cada componente decomposto é processado e previsto individualmente, e essas previsões são então combinadas, o que pode resultar em uma acumulação de erros de previsão. Esse processo de lidar com componentes individuais e sua subsequente fusão adiciona complexidade ao modelo e pode reduzir o custo-benefício, como mostrado nas Tabelas C.2 e C.4.

Foi identificado que a abordagem de otimização, por realizar uma busca no espaço de soluções para identificar os parâmetros ideais para os modelos de *machine learning*, é essencial para

maximizar seu desempenho. No entanto, consome excessivo poder computacional e aumenta o risco de *overfitting* nos modelos, resultando em um retorno não atraente nos resultados preditos.

A junção de otimização com decomposição tem um impacto direto na abordagem de arquitetura completa, resultando em baixo desempenho nas métricas de avaliação adotadas, conforme apresentado nas Tabelas C.1, C.3, C.2 e C.4.

Os sistemas híbridos residuais são projetados para combinar um modelo linear com um modelo de *machine learning*, com o objetivo de que o último modele o resíduo não capturado pelo modelo linear, funcionando assim de forma complementar. Essa abordagem pode gerar resultados promissores, como evidenciado pelas melhorias nas métricas de erro e no custo-benefício nas Tabelas C.1, C.3, C.2 e C.4. No entanto, a integração entre os modelos deve ser cuidadosamente gerenciada, e é necessário selecionar um modelo não linear adequado para lidar com o ruído. Uma escolha inadequada do modelo de *machine learning* pode não apenas falhar em melhorar o desempenho, mas também potencialmente degradar a precisão das previsões. Como apresentado nos resultados, o IBOVESPA demonstrou um baixo desempenho do modelo linear selecionado, indicando que o comportamento da série não foi completamente capturado, gerando uma carga maior para o modelo de *machine learning*. Por outro lado, no S&P 500, o modelo linear e não linear destacaram-se, pois a série temporal apresentou menores variações ao longo do período. Consequentemente, a abordagem conseguiu obter um bom desempenho nas métricas de avaliação.

A análise comparativa indica que o modelo CART apresentou o desempenho mais destacado, tanto em termos de métricas de desempenho quanto em relação ao custo-benefício. Esse resultado é corroborado pela aplicação do teste estatístico de Wilcoxon, utilizado para comparar abordagens de *ensemble*. O teste demonstrou que, entre as metodologias consideradas, o modelo *single* CART supera os demais em custo-benefício, considerando a integração de desempenho e eficiência na utilização de recursos. Isso implica em um ponto importante: a real utilização das abordagens *ensemble*. Com o aumento da complexidade, há uma perda na interpretação dos resultados dos algoritmos e também na sua eficácia em ambientes que precisam de tomada de decisão rápida, como o HFS. Portanto, deve-se considerar a utilização das abordagens *ensemble* juntamente com a métrica de custo-benefício.

Capítulo 6

Conclusão e Trabalhos Futuros

Neste estudo, realizou-se avaliações algorítmicas em vários cenários para compreender completamente os impactos de fatores como ruído, volatilidade e as dinâmicas inerentes das séries temporais financeiras. Observou-se que a eficácia de uma abordagem específica de *ensemble* pode variar significativamente dependendo do cenário considerado. Por exemplo, no índice IBOVESPA, o método de Decomposição CART apresentou um bom desempenho em termos das métricas tradicionais, enquanto no índice S&P 500, foi o Sistema Híbrido Residual. Além disso, este trabalho ressaltou o papel crucial dos testes estatísticos na validação dos resultados. Esses testes revelaram diferenças significativas entre as abordagens; no entanto, uma análise mais detalhada indicou que, em muitos casos, os ganhos aparentes não se mantêm quando comparados a modelos *single*, como evidenciado pela equivalência dos resultados entre o Decomposição CART e o modelo CART no IBOVESPA, e entre o Sistema Híbrido Residual CART e o modelo *single* CART no S&P 500. Esses achados reforçam que o aumento na complexidade dos algoritmos nem sempre gera um ganho significativo no resultado final.

Para aprofundar a compreensão da eficácia das diferentes metodologias algorítmicas, este estudo introduziu uma análise da métrica de custo-benefício, frequentemente negligenciada em estudos comparativos. Esta métrica avaliou se os ganhos de desempenho justificam os custos de implementação de abordagens mais complexas. Os resultados indicaram que, na maioria dos casos analisados, as abordagens de *ensemble* não apresentaram vantagens significativas sobre os modelos *singles*, com o modelo CART demonstrando desempenho superior ou equivalente às metodologias de *ensemble* nos mercados considerados. Este achado sublinha a relevância desta métrica, mostrando que, em certas situações, algoritmos mais simples e compreensíveis podem ser preferíveis devido à sua eficiência e facilidade de implementa-

ção, contrariando a expectativa de que as técnicas de *ensemble*, por sua natureza, tendem a melhorar a precisão e a robustez.

Ao trabalhar com técnicas de *ensemble* em mercados financeiros, alguns passos fundamentais devem ser seguidos para garantir a eficácia das previsões. O primeiro passo envolve uma análise preliminar dos dados, identificando suas principais características, como volatilidade e sazonalidade, uma vez que essas particularidades podem influenciar diretamente a performance dos modelos. Em seguida, é essencial realizar uma divisão adequada dos dados em conjuntos de treinamento e teste, utilizando técnicas como *cross-validation* para séries temporais, conforme sugerido por (BERGMEIR; HYNDMAN; KOO, 2018), garantindo que a ordem temporal seja preservada. Posteriormente, deve-se selecionar os algoritmos base apropriados para o *ensemble*, levando em consideração as particularidades de cada mercado. Além disso, é importante otimizar os hiperparâmetros dos modelos base para maximizar seu desempenho, tomando cuidado para evitar o *overfitting*. Por fim, a avaliação dos resultados deve incluir não apenas métricas de erro tradicionais, mas também uma análise de custo-benefício, permitindo verificar se o aumento da complexidade do modelo justifica os recursos computacionais empregados. Esse processo iterativo, que integra a análise dos dados, a escolha de modelos e a avaliação detalhada, garante que a utilização de técnicas de *ensemble* traga melhorias reais e robustas nas previsões financeiras.

A revisão sistemática dos estudos produzidos de 2017 ao primeiro semestre de 2023 permitiu caracterizar o estado da arte em relação ao uso de modelos *ensemble* na previsão do Índice de Mercado de Ações. Inicialmente, os resultados da análise bibliométrica revelaram os periódicos responsáveis pela maioria das publicações, destacando-se o jornal *Expert Systems with Applications*, com 8 publicações. A análise incluiu o número de artigos publicados por cada periódico e conferência, bem como o número de citações de cada trabalho. Vale destacar que o estudo de (REZAEI; FAALJOU; MANSOURFAR, 2021) teve o maior número de citações, com 107 referências. A China é o país com o maior número de autores, contribuindo significativamente para o conhecimento e planejamento de futuras pesquisas. Em relação às abordagens de *ensemble*, a técnica mais utilizada é a denominada Completa, que combina a decomposição com a otimização dos algoritmos ou das séries temporais. As métricas de desempenho mais utilizadas são o MSE, RMSE, MAE e MAPE, e o modelo mais utilizado é o LSTM. Entre os mercados mais estudados, o S&P 500 representa o mercado desenvolvido dos Estados Unidos, enquanto o SSE representa o mercado em desenvolvimento da China.

Como limitação do estudo, notou-se que a análise está confinada a artigos encontrados em três bases de dados (*Web of Science*, *Scopus* e *IEEE Xplore*). Portanto, publicações oferecidas em outros espaços, bem como teses, dissertações e livros sobre o uso de algoritmos *ensemble* na previsão do Índice de Mercado de Ações, não foram considerados. Recomenda-se a inclusão dessas fontes para futuras revisões da literatura, permitindo a expansão do conhecimento sobre a produção científica nesta área.

Este estudo possui limitações, particularmente em relação à métrica de '*Instructions Retired*', pois não quantifica apenas as instruções processadas pela CPU; ela também pode ser influenciada por vários fatores, como latência de memória, operações de ponto flutuante e instruções não executadas devido a previsões de ramificação incorretas, introduzindo potenciais distorções nos resultados. Essas variáveis adicionam uma camada de complexidade na interpretação da eficácia computacional do programa, resultando em ruído nos dados analisados. Embora várias metodologias significativas sejam exploradas, reconhece-se que existe uma vasta gama de técnicas de *ensemble* na literatura que não foram abordadas e identificadas nesta investigação, limitando assim a generalização dos resultados para todas as configurações de *ensemble* possíveis.

Outro ponto crítico é a ameaça à validade da pesquisa, representada pela necessidade de realizar uma análise de sensibilidade dos parâmetros dos algoritmos, seja individualmente ou em conjunto. Esta análise é crucial para demonstrar como a variação dos hiperparâmetros afeta o comportamento dos modelos e, consequentemente, sua eficiência e eficácia.

Para investigações futuras, recomenda-se ampliar o escopo das comparações, incorporando metodologias como técnicas *Fuzzy*, análise de sentimento de mercado em momentos anteriores à previsão e outras abordagens de *ensemble*. Além disso, é aconselhável diversificar os algoritmos utilizados. Este estudo focou em algoritmos clássicos amplamente reconhecidos na literatura, como *CART*, *MLP* e *SVR*. Para pesquisas futuras, sugere-se incluir algoritmos de *Deep Learning*, como *LSTM*, *GRU* e *CNN*. Essa expansão permitirá uma análise mais abrangente e poderá revelar *insights* adicionais sobre a eficácia das diferentes técnicas de modelagem em séries temporais financeiras. Ademais, há um número limitado de estudos empíricos sobre a previsão do Índice de Mercado de Ações em países com mercados em desenvolvimento e subdesenvolvidos, assim como sobre seus impactos na eficácia e eficiência nesse tipo de problema. Outro ponto importante é o estudo do comportamento dos modelos de *ensemble* nos períodos anterior, durante e pós-pandemia COVID-19, para demonstrar sua capacidade de adaptação.

Por fim, recomenda-se realizar a análise de sensibilidade dos parâmetros nos modelos de *machine learning*.

Referências bibliográficas

AGAPITOS, A.; BRABAZON, A.; O'NEILL, M. Regularised gradient boosting for financial time-series modelling. **Computational Management Science**, Springer, v. 14, p. 367–391, 2017.

ALHNAITY, B.; ABBOD, M. A new hybrid financial time series prediction model. **Engineering Applications of Artificial Intelligence**, Elsevier, v. 95, p. 103873, 2020.

ALI, M.; KHAN, D. M.; ALSHANBARI, H. M.; EL-BAGOURY, A. A.-A. H. Prediction of complex stock market data using an improved hybrid emd-lstm model. **Applied Sciences**, MDPI, v. 13, n. 3, p. 1429, 2023.

AMPOMAH, E. K.; QIN, Z.; NYAME, G. Evaluation of tree-based ensemble machine learning models in predicting stock price direction of movement. **Information**, MDPI, v. 11, n. 6, p. 332, 2020.

AMPOMAH, E. K.; QIN, Z.; NYAME, G.; BOTCHEY, F. E. Stock market decision support modeling with tree-based AdaBoost ensemble machine learning models. **Informatica**, v. 44, n. 4, 2021.

AMPOUNTOLAS, A. Comparative Analysis of Machine Learning, Hybrid, and Deep Learning Forecasting Models: Evidence from European Financial Markets and Bitcoins. **Forecasting**, MDPI, v. 5, n. 2, p. 472–486, 2023.

ARABAS, J.; MICHALEWICZ, Z.; MULAWKA, J. GAVaPS-a genetic algorithm with varying population size. In: IEEE. PROCEEDINGS of the First IEEE Conference on Evolutionary Computation. IEEE World Congress on Computational Intelligence. [S.l.: s.n.], 1994. P. 73–78.

BEKAERT, G.; HARVEY, C. R. Emerging markets finance. **Journal of empirical finance**, Elsevier, v. 10, n. 1-2, p. 3–55, 2003.

BERGMEIR, C.; HYNDMAN, R. J.; KOO, B. A note on the validity of cross-validation for evaluating autoregressive time series prediction. **Computational Statistics & Data Analysis**, Elsevier, v. 120, p. 70–83, 2018.

BOSER, B. E.; GUYON, I. M.; VAPNIK, V. N. A training algorithm for optimal margin classifiers. In: PROCEEDINGS of the fifth annual workshop on Computational learning theory. [S.l.: s.n.], 1992. P. 144–152.

BOTCHKAREV, A. A new typology design of performance metrics to measure errors in machine learning regression algorithms. **Interdisciplinary Journal of Information, Knowledge, and Management**, v. 14, p. 045–076, 2019.

BOX, G. E.; JENKINS, G. M.; REINSEL, G. C.; LJUNG, G. M. **Time series analysis: forecasting and control**. [S.l.]: John Wiley & Sons, 2015.

BREIMAN, L. **Classification and regression trees**. [S.l.]: Routledge, 2017.

BRERETON, P. et al. Lessons from applying the systematic literature review process within the software engineering domain. **Journal of systems and software**, Elsevier, v. 80, n. 4, p. 571–583, 2007.

BROWNLEE, J. **Ensemble learning algorithms with Python: Make better predictions with bagging, boosting, and stacking**. [S.l.]: Machine Learning Mastery, 2021.

BUSTOS, O.; POMARES-QUIMBAYA, A. Stock market movement forecast: A systematic review. **Expert Systems with Applications**, Elsevier, v. 156, p. 113464, 2020.

CAPLINGER, D. **What is a stock market index? defined and listed**. [S.l.: s.n.]. Acessado em 20/05/2024. Disponível em: <https://www.fool.com/investing/stock-market/indexes/>.

CHATFIELD, C.; XING, H. **The analysis of time series: an introduction with R**. [S.l.]: Chapman e hall/CRC, 2019.

CHEN, J.; YANG, H. A CSI 300 Index Prediction Model Based on PSO-SVR-GRNN Hybrid Method. **Mobile Information Systems**, Hindawi, v. 2022, 2022.

CHEN, W.; YEO, C. K.; LAU, C. T.; LEE, B. S. Leveraging social media news to predict stock index movement using RNN-boost. **Data & Knowledge Engineering**, Elsevier, v. 118, p. 14–24, 2018.

CORBA, B. S.; EGRIOGLU, E.; DALAR, A. Z. AR–ARCH type artificial neural network for forecasting. **Neural Processing Letters**, Springer, v. 51, p. 819–836, 2020.

CUTLER, A.; CUTLER, D. R.; STEVENS, J. R. Random forests. **Ensemble machine learning: Methods and applications**, Springer, p. 157–175, 2012.

DAS, S.; NAYAK, S. C.; SAHOO, B. Towards crafting optimal functional link artificial neural networks with RAO algorithms for stock closing prices prediction. **Computational Economics**, Springer, v. 60, n. 1, p. 1–23, 2022.

DASH, R.; SAMAL, S.; DASH, R.; RAUTRAY, R. An integrated TOPSIS crow search based classifier ensemble: In application to stock index price movement prediction. **Applied Soft Computing**, Elsevier, v. 85, p. 105784, 2019.

DEMŠAR, J. Statistical comparisons of classifiers over multiple data sets. **The Journal of Machine learning research**, JMLR. org, v. 7, p. 1–30, 2006.

DENG, C.; HUANG, Y.; HASAN, N.; BAO, Y. Multi-step-ahead stock price index forecasting using long short-term memory model with multivariate empirical mode decomposition. **Information Sciences**, Elsevier, v. 607, p. 297–321, 2022.

DIVINA, F. et al. Stacking ensemble learning for short-term electricity consumption forecasting. **Energies**, MDPI, v. 11, n. 4, p. 949, 2018.

DONGHWAN, S.; BUSOGI, M.; CHUNG BAEK, A. M.; KIM, N. FORECASTING STOCK MARKET INDEX BASED ON PATTERNDRIVEN LONG SHORT-TERM MEMORY. **Economic Computation & Economic Cybernetics Studies & Research**, v. 54, n. 3, 2020.

DRUCKER, H. et al. Support vector regression machines. **Advances in neural information processing systems**, v. 9, 1996.

DURAIRAJ, D. M.; MOHAN, B. K. A convolutional neural network based approach to financial time series prediction. **Neural Computing and Applications**, Springer, v. 34, n. 16, p. 13319–13337, 2022.

DURAIRAJ, M.; BH, K. M. Statistical evaluation and prediction of financial time series using hybrid regression prediction models. **International Journal of Intelligent Systems and Applications in Engineering**, v. 9, n. 4, p. 245–255, 2021.

EBERHART, R.; SIMPSON, P.; DOBBINS, R. **Computational intelligence PC tools**. [S.l.]: Academic Press Professional, Inc., 1996.

EIBEN, A. E.; SMIT, S. K. Parameter tuning for configuring and analyzing evolutionary algorithms. **Swarm and Evolutionary Computation**, Elsevier, v. 1, n. 1, p. 19–31, 2011.

FĂRCAŞ, I. THE IMPACT OF THE COVID-19 PANDEMIC ON THE GOVERNMENT INTERVENTIONS ACROSS THE FINANCIAL SYSTEM. A WORLDWIDE SAMPLE. **Review of Economic and Business Studies (REBS)**, Editura Universităţii» Alexandru Ioan Cuza «din Iaşi, n. 32, p. 9–22, 2023.

FERROUHI, E. M.; BOUABDALLAOUI, I. A comparative study of Ensemble Learning Algorithms for High-Frequency Trading. **Scientific African**, Elsevier, e02161, 2024.

FRANCESCHINI, F. et al. Properties of indicators. **Designing Performance Measurement Systems: Theory and Practice of Key Performance Indicators**, Springer, p. 85–131, 2019.

FREUND, Y.; SCHAPIRE, R. E. A decision-theoretic generalization of on-line learning and an application to boosting. **Journal of computer and system sciences**, Elsevier, v. 55, n. 1, p. 119–139, 1997.

FURLANETO, D. C.; OLIVEIRA, L. S.; MENOTTI, D.; CAVALCANTI, G. D. Bias effect on predicting market trends with EMD. **Expert Systems with Applications**, Elsevier, v. 82, p. 19–26, 2017.

GANDHMAL, D. P.; KUMAR, K. Systematic analysis and review of stock market prediction techniques. **Computer Science Review**, Elsevier, v. 34, p. 100190, 2019.

GIACOMEL, F. d. S. Um método algorítmico para operações na bolsa de valores baseado em ensembles de redes neurais para modelar e prever os movimentos dos mercados de ações, 2016.

GIGUERE, P.; GOLDBERG, D. E. Population sizing for optimum sampling with genetic algorithms: A case study of the Onemax problem. **Genetic Programming**, Citeseer, v. 98, p. 496–503, 1998.

GINI, C. Measurement of inequality of incomes. **The economic journal**, Oxford University Press Oxford, UK, v. 31, n. 121, p. 124–125, 1921.

HAJIABOTORABI, Z.; KAZEMI, A.; SAMAVATI, F. F.; GHAINI, F. M. M. Improving DWT-RNN model via B-spline wavelet multiresolution to forecast a high-frequency time series. **Expert Systems with Applications**, Elsevier, v. 138, p. 112842, 2019.

HAYKIN, S. **Neural networks and learning machines, 3/E**. [S.l.]: Pearson Education India, 2009.

HE, K. et al. Financial time series forecasting with the deep learning ensemble model. **Mathematics**, MDPI, v. 11, n. 4, p. 1054, 2023.

HENRIQUE, B. M.; SOBREIRO, V. A.; KIMURA, H. Literature review: Machine learning techniques applied to financial market prediction. **Expert Systems with Applications**, Elsevier, v. 124, p. 226–251, 2019.

HOLLAND, J. H. Genetic algorithms. **Scientific american**, JSTOR, v. 267, n. 1, p. 66–73, 1992.

HOLLANDER, M.; WOLFE, D. A.; CHICKEN, E. **Nonparametric statistical methods**. [S.l.]: John Wiley & Sons, 2013.

HUANG, N. E. et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. **Proceedings of the Royal Society of London. Series A: mathematical, physical and engineering sciences**, The Royal Society, v. 454, n. 1971, p. 903–995, 1998.

HYNDMAN, R. J.; ATHANASOPOULOS, G. **Forecasting: principles and practice**. [S.l.]: OTexts, 2018.

JI, Y.; LIEW, A. W.-C.; YANG, L. A novel improved particle swarm optimization with long-short term memory hybrid model for stock indices forecast. **Ieee Access**, IEEE, v. 9, p. 23660–23671, 2021.

JOTHIMANI, D.; BAŞAR, A. Stock index forecasting using time series decomposition-based and machine learning models. In: SPRINGER. ARTIFICIAL Intelligence XXXVI: 39th SGAI International Conference on Artificial Intelligence, AI 2019, Cambridge, UK, December 17–19, 2019, Proceedings 39. [S.l.: s.n.], 2019. P. 283–292.

JUJIE, W.; DANFENG, Q. AN EXPERIMENTAL INVESTIGATION OF TWO HYBRID FRAMEWORKS FOR STOCK INDEX PREDICTION USING NEURAL NETWORK AND SUPPORT VECTOR REGRESSION. **Economic Computation & Economic Cybernetics Studies & Research**, v. 52, n. 4, 2018.

KANG, F.; LI, J. Artificial bee colony algorithm optimized support vector regression for system reliability analysis of slopes. **Journal of Computing in Civil Engineering**, American Society of Civil Engineers, v. 30, n. 3, p. 04015040, 2016.

KEHINDE, T.; CHAN, F. T.; CHUNG, S. Scientometric review and analysis of recent approaches to stock market forecasting: Two decades survey. **Expert Systems with Applications**, Elsevier, v. 213, p. 119299, 2023.

KHASHEI, M.; BIJARI, M. A novel hybridization of artificial neural networks and ARIMA models for time series forecasting. **Applied soft computing**, Elsevier, v. 11, n. 2, p. 2664–2675, 2011.

KIM, K.-j.; HAN, I. Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index. **Expert systems with Applications**, Elsevier, v. 19, n. 2, p. 125–132, 2000.

KRAUSS, C.; DO, X. A.; HUCK, N. Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500. **European Journal of Operational Research**, Elsevier, v. 259, n. 2, p. 689–702, 2017.

KUMAR, G.; SINGH, U. P.; JAIN, S. Hybrid evolutionary intelligent system and hybrid time series econometric model for stock price forecasting. **International Journal of Intelligent Systems**, Wiley Online Library, v. 36, n. 9, p. 4902–4935, 2021.

KUMBURE, M. M.; LOHRMANN, C.; LUUKKA, P.; PORRAS, J. Machine learning techniques and data for stock market forecasting: A literature review. **Expert Systems with Applications**, Elsevier, v. 197, p. 116659, 2022.

LEARN, S. **sklearn.ensemble.BaggingRegressor**. [S.l.: s.n.], 2023. Disponível em: <https://x.gd/6m8qh>. Acesso em: 30 dez. 2023.

LEARN, S. **sklearn.ensemble.StackingRegressor**. [S.l.: s.n.], 2023. Disponível em: <https://x.gd/wwNSI>. Acesso em: 30 dez. 2023.

LEI, Y.; HE, Z.; ZI, Y. Application of the EEMD method to rotor fault diagnosis of rotating machinery. **Mechanical Systems and Signal Processing**, Elsevier, v. 23, n. 4, p. 1327–1338, 2009.

LIMA SILVA, P. C. de; SADAEI, H. J.; BALLINI, R.; GUIMARÃES, F. G. Probabilistic forecasting with fuzzy time series. **IEEE Transactions on Fuzzy Systems**, IEEE, v. 28, n. 8, p. 1771–1784, 2019.

LIN, G.; LIN, A.; CAO, J. Multidimensional KNN algorithm based on EEMD and complexity measures in financial time series forecasting. **Expert Systems with Applications**, Elsevier, v. 168, p. 114443, 2021.

LIN, Y. et al. Forecasting the realized volatility of stock price index: A hybrid model integrating CEEMDAN and LSTM. **Expert Systems with Applications**, Elsevier, v. 206, p. 117736, 2022.

LIU, Y.; LIU, X.; ZHANG, Y.; LI, S. CEGH: A Hybrid Model Using CEEMD, Entropy, GRU, and History Attention for Intraday Stock Market Forecasting. **Entropy**, MDPI, v. 25, n. 1, p. 71, 2022.

LUCAS, D. C. Algoritmos genéticos: uma introdução. **Apostila referente a disciplina de Inteligência Computacional, Universidade Federal do Rio Grande do Sul, Brasil**, 2002.

LUO, Z.; GUO, W.; LIU, Q.; ZHANG, Z. A hybrid model for financial time-series forecasting based on mixed methodologies. **Expert Systems**, Wiley Online Library, v. 38, n. 2, e12633, 2021.

LV, P.; SHU, Y.; XU, J.; WU, Q. Modal decomposition-based hybrid model for stock index prediction. **Expert Systems with Applications**, Elsevier, v. 202, p. 117252, 2022.

LV, P.; WU, Q.; XU, J.; SHU, Y. Stock index prediction based on time series decomposition and hybrid model. **Entropy**, MDPI, v. 24, n. 2, p. 146, 2022.

MAHESH, B. Machine learning algorithms-a review. **International Journal of Science and Research (IJSR).[Internet]**, v. 9, n. 1, p. 381–386, 2020.

MERCER, J. Xvi. functions of positive and negative type, and their connection the theory of integral equations. **Philosophical transactions of the royal society of London. Series A, containing papers of a mathematical or physical character**, The Royal Society London, v. 209, n. 441-458, p. 415–446, 1909.

NAYAK, S. C. Escalation of forecasting accuracy through linear combiners of predictive models. **EAI Endorsed Transactions on Scalable Information Systems**, v. 6, n. 22, 2019.

NIU, H.; XU, K.; WANG, W. A hybrid stock price index forecasting model based on variational mode decomposition and LSTM network. **Applied Intelligence**, Springer, v. 50, p. 4296–4309, 2020.

NTI, I. K.; ADEKOYA, A. F.; WEYORI, B. A. A comprehensive evaluation of ensemble learning for stock-market prediction. **Journal of Big Data**, SpringerOpen, v. 7, n. 1, p. 1–40, 2020.

NTI, I. K.; ADEKOYA, A. F.; WEYORI, B. A. Efficient stock-market prediction using ensemble support vector machine. **Open Computer Science**, De Gruyter, v. 10, n. 1, p. 153–163, 2020.

NUNES, J. R. d. O. **Avaliação de taxas de cruzamento e mutação em um algoritmo genético baseado em ordem aplicado ao problema do caixeiro viajante**. 2018. B.S. thesis – Universidade Federal do Rio Grande do Norte.

OLIVEIRA ROSA, T. de; LUZ, H. S. Conceitos Básicos de Algoritmos Genéticos: Teoria e Prática.

PADHI, D. K.; PADHY, N. Prognosticate of the financial market utilizing ensemble-based conglomerate model with technical indicators. **Evolutionary Intelligence**, Springer, v. 14, n. 2, p. 1035–1051, 2021.

PÁDUA BRAGA, A. de; LEON FERREIRA, A. C. P. de; LUDERMIR, T. B. **Redes neurais artificiais: teoria e aplicações**. [S.l.]: LTC editora, 2007.

PARIDA, A.; BISOI, R.; DASH, P.; MISHRA, S. Times series forecasting using Chebyshev functions based locally recurrent neuro-fuzzy information system. **International Journal of Computational Intelligence Systems**, Atlantis Press, v. 10, n. 1, p. 375–393, 2017.

PASUPULETY, U.; ANEES, A. A.; ANMOL, S.; MOHAN, B. R. Predicting stock prices using ensemble learning and sentiment analysis. In: IEEE. 2019 IEEE second international conference on artificial intelligence and knowledge engineering (AIKE). [S.l.: s.n.], 2019. P. 215–222.

PENG, Z. et al. An application of hybrid models for weekly stock market index prediction: Empirical evidence from SAARC countries. **Complexity**, Hindawi Limited, v. 2021, p. 1–10, 2021.

PMDARIMA. **pmdarima: ARIMA estimators for Python**. [S.l.: s.n.], 2023. Disponível em: <https://alkaline-ml.com/pmdarima/index.html>. Acesso em: 30 dez. 2023.

POZO, A. et al. Computação evolutiva. **Universidade Federal do Paraná, 61p.(Grupo de Pesquisas em Computação Evolutiva, Departamento de Informática-Universidade Federal do Paraná)**, 2005.

QI, C.; REN, J.; SU, J. GRU Neural Network Based on CEEMDAN–Wavelet for Stock Price Prediction. **Applied Sciences**, MDPI, v. 13, n. 12, p. 7104, 2023.

REZAEI, H.; FAALJOU, H.; MANSOURFAR, G. Stock price prediction using deep learning and frequency decomposition. **Expert Systems with Applications**, Elsevier, v. 169, p. 114332, 2021.

REZAIE-BALF, M. et al. Forecasting daily solar radiation using CEEMDAN decomposition-based MARS model trained by crow search algorithm. **Energies**, MDPI, v. 12, n. 8, p. 1416, 2019.

RUSSELL, S.; NORVIG, P. Artificial intelligence: A modern approach, global edition. **Harlow: Pearson**, 2022.

SAGI, O.; ROKACH, L. Ensemble learning: A survey. **Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery**, Wiley Online Library, v. 8, n. 4, e1249, 2018.

SANTOS JÚNIOR, D. S. d. O. et al. Método de Ensemble para correção de modelos ARIMA: uma abordagem de sistema híbrido para previsão de séries temporais. Universidade Federal de Pernambuco, 2022.

SAPANKEVYCH, N. I.; SANKAR, R. Time series prediction using support vector machines: a survey. **IEEE computational intelligence magazine**, IEEE, v. 4, n. 2, p. 24–38, 2009.

SCHUBERT, E. Stop using the elbow criterion for k-means and how to choose the number of clusters instead. **ACM SIGKDD Explorations Newsletter**, ACM New York, NY, USA, v. 25, n. 1, p. 36–42, 2023.

SHU, W.; GAO, Q. Forecasting stock price based on frequency components by EMD and neural networks. **Ieee Access**, IEEE, v. 8, p. 206388–206395, 2020.

SILVA, T. A. d. et al. Estudo do desempenho da combinação de preditores baseados em cópulas e máquinas de vetor de suporte para séries temporais úteis ao desenvolvimento sustentável. Universidade Federal Rural de Pernambuco, 2020.

SINGH, A.; DASH, J. K.; BEHURA, B.; CHAKRAVARTY, S. Teaching Learning Based Optimized Support Vector Regression Model for Prediction of Indian Stock Market. **International Journal of Advanced Science and Technology**, v. 29, n. 5, p. 3002–3015, 2020.

SINGH, P. FQTSFM: A fuzzy-quantum time series forecasting model. **Information Sciences**, Elsevier, v. 566, p. 57–79, 2021.

SMOLA, A. J.; SCHÖLKOPF, B. A tutorial on support vector regression. **Statistics and computing**, Springer, v. 14, p. 199–222, 2004.

SOARES, P. L. B.; SILVA, J. P. da. Aplicação de redes neurais artificiais em conjunto com o método vetorial da propagação de feixes na análise de um acoplador direcional baseado em fibra ótica. **Revista Brasileira de Computação Aplicada**, v. 3, n. 2, p. 58–72, 2011.

SONG, H.; CHOI, H. Forecasting Stock Market Indices Using the Recurrent Neural Network Based Hybrid Models: CNN-LSTM, GRU-CNN, and Ensemble Models. **Applied Sciences**, MDPI, v. 13, n. 7, p. 4644, 2023.

SU, J.; WANG, X.; LIANG, Y.; CHEN, B. GA-based support vector machine model for the prediction of monthly reservoir storage. **Journal of Hydrologic Engineering**, American Society of Civil Engineers, v. 19, n. 7, p. 1430–1437, 2014.

TEIXEIRA, O. Proposta de um novo algoritmo genético baseado na teoria dos jogos. **Programa de Pós-Graduação em Engenharia Elétrica, UFPA**, 2005.

TOLEDO FILHO, J. R. de. **Mercado de capitais brasileiro: uma introdução**. [S.l.]: Cengage Learning, 2020.

TORRES, M. E.; COLOMINAS, M. A.; SCHLOTTHAUER, G.; FLANDRIN, P. A complete ensemble empirical mode decomposition with adaptive noise. In: IEEE. 2011 IEEE international conference on acoustics, speech and signal processing (ICASSP). [S.l.: s.n.], 2011. P. 4144–4147.

VAN ECK, N. J.; WALTMAN, L. Visualizing bibliometric networks. **Measuring scholarly impact: Methods and practice**, Springer, p. 285–320, 2014.

VAPNIK, V. **The nature of statistical learning theory**. [S.l.]: Springer science & business media, 2013.

VIERA, C. **Primary Sources vs. Secondary Sources vs. Tertiary Sources: What's the difference?** [S.l.: s.n.], 2023. Disponível em: <https://www.aje.com/arc/primary-secondary-tertiary-sources/>. Acesso em: 23 ago. 2024.

VISHWAS, B.; PATEL, A. **Hands-on time series analysis with Python**. [S.l.]: Springer, 2020.

WANG, J. et al. Stock index prediction and uncertainty analysis using multi-scale nonlinear ensemble paradigm of optimal feature extraction, two-stage deep learning and Gaussian process regression. **Applied Soft Computing**, Elsevier, v. 113, p. 107898, 2021.

WANG, X.; WEN, J.; ZHANG, Y.; WANG, Y. Real estate price forecasting based on SVM optimized by PSO. **Optik**, Elsevier, v. 125, n. 3, p. 1439–1443, 2014.

WIDIPUTRA, H.; MAILANGKAY, A.; GAUTAMA, E. Multivariate cnn-lstm model for multiple parallel financial time-series prediction. **Complexity**, Hindawi Limited, v. 2021, p. 1–14, 2021.

WOLPERT, D. H. Stacked generalization. **Neural networks**, Elsevier, v. 5, n. 2, p. 241–259, 1992.

WU, J.; ZHOU, T.; LI, T. A hybrid approach integrating multiple ICEEMDANs, WOA, and RVFL networks for economic and financial time series forecasting. **Complexity**, Hindawi Limited, v. 2020, p. 1–17, 2020.

WU, Z.; HUANG, N. E. Ensemble empirical mode decomposition: a noise-assisted data analysis method. **Advances in adaptive data analysis**, World Scientific, v. 1, n. 01, p. 1–41, 2009.

XIAO, D.; SU, J. et al. Research on stock price time series prediction based on deep learning and autoregressive integrated moving average. **Scientific Programming**, Hindawi, v. 2022, 2022.

YAN, B.; AASMA, M. et al. A novel deep learning framework: Prediction and analysis of financial time series using CEEMD and LSTM. **Expert systems with applications**, Elsevier, v. 159, p. 113609, 2020.

YAPA ABEYWARDHANA, D. Capital structure theory: An overview. **Accounting and finance research**, v. 6, n. 1, 2017.

YI SUN QINGSONG SUN, S. Z. Prediction of Shanghai Stock Index Based on Investor Sentiment and CNN-LSTM Model. **Journal of Systems Science and Information**, v. 10, n. 6, p. 620–632, 2022.

YU, H.; MING, L. J.; SUMEI, R.; SHUPING, Z. A hybrid model for financial time series forecasting—integration of EWT, ARIMA with the improved ABC optimized ELM. **IEEE Access**, IEEE, v. 8, p. 84501–84518, 2020.

ZHANG, J.; LI, L.; CHEN, W. Predicting stock price using two-stage machine learning techniques. **Computational Economics**, Springer, v. 57, p. 1237–1261, 2021.

ZHENG, J.; WANG, Y.; LI, S.; CHEN, H. The stock index prediction based on SVR model with bat optimization algorithm. **Algorithms**, MDPI, v. 14, n. 10, p. 299, 2021.

ZHOU, F.; ZHOU, H.-m.; YANG, Z.; YANG, L. EMD2FNN: A strategy combining empirical mode decomposition and factorization machine based neural network for stock market trend prediction. **Expert Systems with Applications**, Elsevier, v. 115, p. 136–151, 2019.

ZOLFAGHARI, M.; GHOLAMI, S. A hybrid approach of adaptive wavelet transform, long short-term memory and ARIMA-GARCH family models for the stock index prediction. **Expert Systems with Applications**, Elsevier, v. 182, p. 115149, 2021.

Apêndice A

Índeces de Bolsas Abordadas nos Artigos

A.1 Índices por nome completo, sigla e mercado

Tabela A.1: Índices por nome completo, sigla e mercado

Cripto/Índece/Empresa/Carbono/Abreviatura da taxa de câmbio	**Mercado**
Apple Inc. – (AAPL)	Estados Unidos
Abbott Laboratories – (ABT)	Estados Unidos
All Ordinaries – (AORD)	Austrália
Australian Dollar – (AUD)	Austrália
Bank of America Corp – (BAC)	Estados Unidos
BITCOIN – (BTC)	Criptomoeda
Brazilian Real – (BRL)	Brasil
Sao Paulo's Stock Market Index – (BVSP)	Brasil
S&P Bombay Stock Exchange Sensitive Index – (BSE SENSEX)	Índia
Cotation Assistée en Continu – (CAC 40)	França
Swiss Franc – (CHF)	Suiça
China Securities Index 300 Index – (CSI 300)	China
China Securities Index 1000 Index – (CSI 1000)	China
China Securities Index 100 Index – (CSI 100)	China
China Securities Index 500 Index – (CSI 500)	China
China Securities Index 800 Index – (CSI 800)	China
Deutscher Aktienindex – (DAX)	Alemanha
Dow Jones Industrial Average – (DJIA)	Estados Unidos
MSCI Emerging Markets Index – (EM)	Emergente
MSCI European Index – (EU)	Europeu
European Union Emissions Trading System – (EU ETS)	Carbono
EURONEXT 100 INDEX – (EURONEXT 100)	Europeu
Financial Times Stock Exchange 100 Index – (FTSE)	Inglaterra
British Pound – (GBP)	Reino Unido
Ghana Stock Exchange – (GSE)	Gana
HCL Technologies Ltd – (HCLTECH)	Índia
Shanghai-Shenzhen 300 Stock Index – (HS 300)	China
Hang Seng Index – (HSI)	China
Indian rupee – (INR)	Índia
US industrial production – (IP)	Estados Unidos

Table A.1 continuação da página anterior

Istanbul Stock Exchange – (ISE)	Turquia
Istanbul Stock Market Index(TL) – (ISE100_TL)	Turquia
Istanbul Stock Market Index(USD) – (ISE100_USD)	Turquia
Nasdaq Composite – (IXIC)	Estados Unidos
Japanese Yen – (JPY)	Japão
Johannesburg Stock Exchange – (JSE)	África do Sul
Jakarta Stock Exchange – (JSX)	Indonésia
CarMax Inc – (KMX)	Estados Unidos
Korea Securities Dealers Automated Quotation – (KOSDAQ)	Coreia do Sul
Korea Composite Stock Price Index – (KOSPI)	Coreia do Sul
Pakistan Stock Exchange – (KSE 100)	Paquistão
Microsoft Corporation – (MSFT)	Estados Unidos
Mexican Peso – (MXN)	México
National Association of Securities Dealers Automated Quotations – (NASDAQ)	Estados Unidos
Indian stock market index – (NIFTY 50)	Índia
Nikkei Stock Average – (NIKKEI 225)	Japão
National Stock Exchange of India – (NSE)	Índia
New York Stock Exchange – (NYSE)	Estados Unidos
Russell 2000 Index – (Russell 2000)	Estados Unidos
Singapore dollar – (SGD)	Singapura
Small and Medium Board 100 Index – (SME 100)	China
Small and Medium Board 300 Index – (SME 300)	China
Small and Medium Board Composite Index – (SME CI)	China
Swiss Market Index – (SMI)	Suiça
Shanghai Stock Exchange – (SSE)	China
Shanghai A-Share Index – (SSE A shares)	China
Shanghai Stock Exchange B-Share Index – (SSE B shares)	China
Shanghai Stock Exchange 380 – (SSE 380)	China
Shanghai Stock Exchange 180 Index – (SSE 180)	China
Shanghai Stock Exchange 50 Index – (SSE 50)	China
Shenzhen Stock Exchange Component Index – (SZCI)	China
Shenzhen Stock Exchange Index – (SZSE)	China
Shenzhen Constituent A Share Index – (SZSE A shares)	China
Shenzhen Stock Exchange Constituent B-Share Index – (SZSE B shares)	China
Straits Times Index – (STI)	Singapura
EURO STOXX 50 – (STOXX 50)	Europeu
Australian Australian Securities Exchange – (S&P/ASX 200)	Austrália
Taiwan Stock Exchange Capitalization Weighted Stock Index – (TAIEX)	Taiwan
Taiwan Futures Exchange – (TAIFEX)	Taiwan
Taiwan Stock Exchange Corporation – (TSEC)	Taiwan
Tata Steel Limited – (TATASTEEL)	Índia
US Dollar – (USD)	Estados Unidos
Exon mobile corporation – (XOM)	Estados Unidos
West Texas Intermediate – (WTI)	Estados Unidos

Apêndice B

Arquiteturas *Ensemble Learning*

B.1 Arquitetura *Bagging*

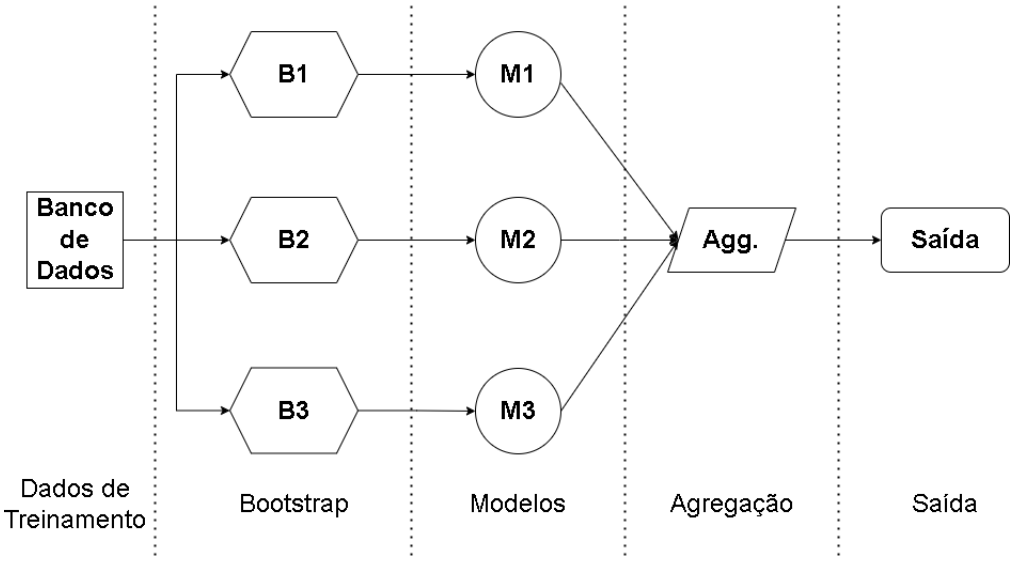


Figura B.1: *Framework* da abordagem *Bagging*. Fonte: (Autor, 2024)

B.2 Arquitetura *Stacking*

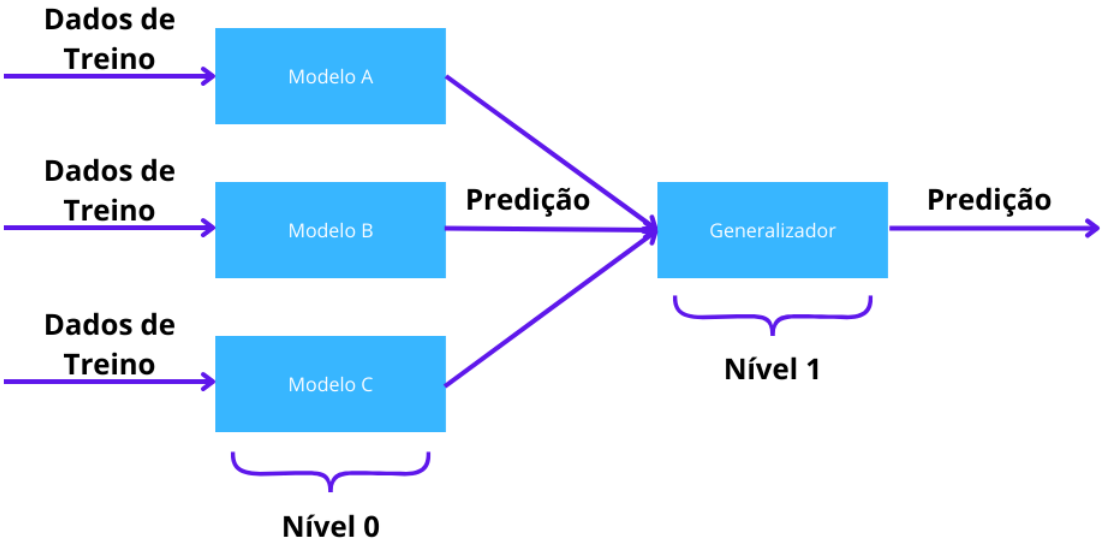


Figura B.2: *Framework* da abordagem *Stacking*. Fonte: (DIVINA et al., 2018)

B.3 Arquitetura *Stacking Detalhada*

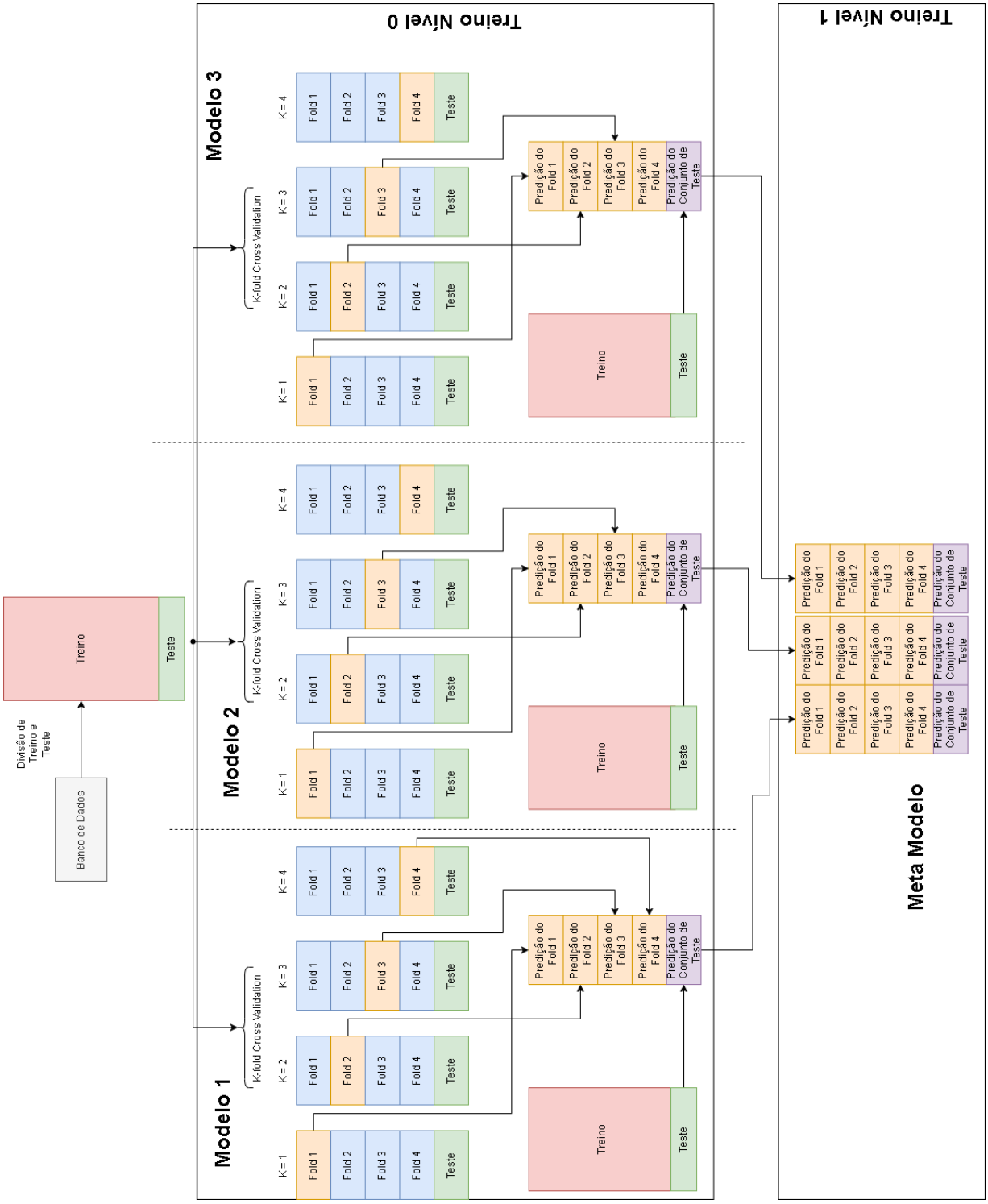


Figura B.3: *Framework* da abordagem *Stacking*. Fonte:(Autor, 2024)

Apêndice C

Tabelas do teste de Wilcoxon

C.1 IBOVESPA - Abordagem *Ensemble* (Teste de Hipótese - RMSE)

Tabela C.1: IBOVESPA - Abordagem *Ensemble* - Teste de hipótese de Wilcoxon com 95% de confiança. Os símbolos (+), (-) ou (=) indicam se a comparação com cada abordagem pode ser melhor, pior ou equivalente, respectivamente.

Modelos	***Bagging***			**Heterogêneo**	**Completo**			**Decomposição**			**Sistema Híbrido**			**Otimização**			***Stacking***		
ARIMA	=	-	-	-	-	-	-	+	-	-	=	=	-	-	-	-	-	-	-
CART	=	-	-	-	-	-	-	=	-	-	=	-	-	-	-	-	-	-	-
MLP	+	=	+	+	+	+	+	+	=	+	+	+	+	+	+	+	=	+	+
SVR	+	-	-	+	+	+	=	+	-	-	+	+	+	+	=	=	-	+	=
Bagging CART	=	-	-	-	-	-	-	=	-	-	=	-	-	-	-	-	-	-	-
Bagging MLP	+	=	+	+	+	+	+	+	=	+	+	+	+	+	+	+	=	+	+
Bagging SVR	+	-	=	+	+	+	=	+	-	-	+	+	+	+	=	=	-	+	+
Heterogêneo	+	-	-	=	+	=	-	+	-	-	+	+	+	+	=	=	-	+	-
Completo CART	+	-	-	-	=	-	-	+	-	-	+	+	-	=	-	-	-	-	-
Completo MLP	+	-	-	=	+	=	-	+	-	-	+	+	+	+	=	=	-	=	-
Completo SVR	+	-	=	+	+	+	=	+	-	-	+	+	+	+	=	=	-	+	=
Decomposição CART	=	-	-	-	-	-	-	=	-	-	=	=	-	-	-	-	-	-	-
Decomposição MLP	+	=	+	+	+	+	+	+	=	+	+	+	+	+	+	+	=	+	+
Decomposição SVR	+	-	+	+	+	+	+	+	-	=	+	+	+	+	=	=	-	+	+
Sistema Híbrido CART	=	-	-	-	-	-	-	=	-	-	=	=	-	-	-	-	-	-	-
Sistema Híbrido MLP	+	-	-	-	-	-	-	+	-	-	=	=	-	-	-	-	-	-	-
Sistema Híbrido SVR	+	-	-	-	+	-	-	+	-	-	+	+	=	+	-	-	-	-	-
Otimização CART	+	-	-	-	=	-	-	+	-	-	+	+	-	=	-	-	-	-	-
Otimização MLP	+	-	=	=	+	=	=	+	-	=	+	+	+	+	=	=	-	=	=
Otimização SVR	+	-	=	+	+	=	=	+	-	=	+	+	+	+	=	=	-	+	=
Stacking CART	+	=	+	+	+	+	+	+	=	+	+	+	+	+	+	+	=	+	+
Stacking MLP	+	-	-	-	+	=	-	+	-	-	+	+	+	+	=	-	-	=	-
Stacking SVR	+	-	-	+	+	+	=	+	-	-	+	+	+	+	=	=	-	+	=
Total	**18**	**0**	**5**	**10**	**15**	**9**	**5**	**19**	**0**	**4**	**17**	**17**	**14**	**15**	**4**	**4**	**0**	**11**	**6**

C.2 IBOVESPA - Abordagem *Ensemble* (Teste de Hipótese - Custo-benefício)

Tabela C.2: IBOVESPA - Abordagem *Ensemble* Eficiência - Teste de hipótese de Wilcoxon com 95% de confiança. Os símbolos (+), (-) ou (=) indicam se a comparação com cada abordagem pode ser melhor, pior ou equivalente, respectivamente.

Modelos	***Bagging***			**Heterogêneo**	**Completo**			**Decomposição**			**Sistema Híbrido**			**Otimização**			***Stacking***		
ARIMA	+	-	+	-	-	-	-	-	-	-	=	-	-	-	-	-	-	-	-
CART	+	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
MLP	+	-	+	-	-	-	-	=	-	-	+	+	+	+	-	-	-	+	-
SVR	+	-	=	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Bagging CART	=	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Bagging MLP	+	=	+	+	=	=	-	+	-	+	+	+	+	+	-	-	=	+	+
Bagging SVR	-	-	=	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Heterogêneo	+	-	+	=	-	-	-	+	-	-	+	+	+	+	-	-	-	+	-
Completo CART	+	=	+	+	=	=	-	+	-	+	+	+	+	+	-	-	=	+	+
Completo MLP	+	=	+	+	=	=	-	+	-	+	+	+	+	+	-	=	=	+	+
Completo SVR	+	+	+	+	+	+	=	+	-	+	+	+	+	+	=	=	=	+	+
Decomposição CART	+	-	+	-	-	-	-	=	-	-	+	+	+	+	-	-	-	+	-
Decomposição MLP	+	+	+	+	+	+	+	+	=	+	+	+	+	+	=	+	+	+	+
Decomposição SVR	+	-	+	+	-	-	-	+	-	=	+	+	+	+	-	-	-	+	=
Sistema Híbrido CART	+	-	+	-	-	-	-	-	-	-	=	-	-	-	-	-	-	-	-
Sistema Híbrido MLP	+	-	+	-	-	-	-	-	-	-	+	=	+	+	-	-	-	+	-
Sistema Híbrido SVR	+	-	+	-	-	-	-	-	-	-	+	-	=	-	-	-	-	-	-
Otimização CART	+	-	+	-	-	-	-	-	-	-	+	-	+	=	-	-	-	-	-
Otimização MLP	+	+	+	+	+	+	=	+	=	+	+	+	+	+	=	=	=	+	+
Otimização SVR	+	+	+	+	+	=	=	+	-	+	+	+	+	+	=	=	=	+	+
Stacking CART	+	=	+	+	=	=	=	+	-	+	+	+	+	+	=	=	=	+	+
Stacking MLP	+	-	+	-	-	-	-	-	-	-	+	-	+	+	-	-	-	=	-
Stacking SVR	+	-	+	+	-	-	-	+	-	=	+	+	+	+	-	-	-	+	=
Total	**21**	4	**21**	10	4	3	1	11	0	8	17	13	16	15	0	1	1	14	8

C.3 S&P 500 - Abordagem *Ensemble* (Teste de Hipótese - RMSE)

Tabela C.3: S&P 500 - Abordagem *Ensemble* - Teste de hipótese de Wilcoxon com 95% de confiança. Os símbolos (+), (-) ou (=) indicam se a comparação com cada abordagem pode ser melhor, pior ou equivalente, respectivamente.

Modelos	***Bagging***			**Heterogêneo**	**Completo**			**Decomposição**			**Sistema Híbrido**			**Otimização**			***Stacking***		
ARIMA	=	-	-	-	-	-	-	=	-	-	=	-	-	-	-	-	-	-	-
CART	=	-	-	-	-	-	-	=	-	-	=	-	-	-	-	-	-	-	-
MLP	+	=	=	=	+	=	-	+	-	-	+	=	-	+	=	=	=	+	=
SVR	+	=	+	+	+	=	-	+	-	-	+	=	-	+	=	=	=	+	=
Bagging CART	=	-	-	-	-	-	-	=	-	-	=	-	-	-	-	-	-	-	-

Table C.3 continuação da página anterior

Modelos	*Bagging*			Heterogêneo	Completo			Decomposição			Sistema Híbrido			Otimização			*Stacking*		
Bagging MLP	+	=	=	=	+	=	-	+	-	-	+	=	-	+	=	=	=	+	=
Bagging SVR	+	=	=	+	+	-	-	+	-	-	+	=	-	+	=	=	=	+	=
Heterogêneo	+	=	-	=	=	-	-	+	-	-	+	=	-	+	=	=	=	=	=
Completo CART	+	-	-	=	=	-	-	+	-	-	+	-	-	=	-	-	=	=	-
Completo MLP	+	=	+	+	+	=	-	+	-	-	+	=	-	+	+	=	=	+	=
Completo SVR	+	+	+	+	+	+	=	+	=	+	+	+	=	+	+	+	+	+	+
Decomposição CART	=	-	-	-	-	-	-	=	-	-	=	-	-	-	-	-	-	-	-
Decomposição MLP	+	+	+	+	+	+	=	+	=	+	+	+	+	+	+	+	+	+	+
Decomposição SVR	+	+	+	+	+	+	-	+	-	=	+	-	=	+	+	+	+	+	+
Sistema Híbrido CART	=	-	-	-	-	-	-	=	-	-	=	-	-	-	-	-	-	=	-
Sistema Híbrido MLP	+	=	=	=	+	=	-	+	-	+	+	=	=	+	=	=	=	+	=
Sistema Híbrido SVR	+	+	+	+	+	+	=	+	-	=	+	=	=	+	+	+	+	+	+
Otimização CART	+	-	-	-	=	-	-	+	-	-	+	-	-	=	-	-	=	=	-
Otimização MLP	+	=	=	=	+	-	-	+	-	-	+	=	-	+	=	=	=	+	=
Otimização SVR	+	=	=	=	+	=	-	+	-	-	+	=	-	+	=	=	=	+	=
Stacking CART	+	=	=	=	=	=	-	+	-	-	+	=	-	=	=	=	=	=	=
Stacking MLP	+	-	-	=	=	-	-	+	-	-	+	-	-	=	-	-	=	=	-
Stacking SVR	+	=	=	=	+	=	-	+	-	-	+	=	-	+	=	=	=	+	=
Total	**18**	4	6	7	13	4	0	**18**	0	3	**18**	2	1	14	5	4	4	13	4

C.4 S&P 500 - Abordagem *Ensemble* (Teste de Hipótese - Custo-benefício)

Tabela C.4: S&P 500 - Comparação de Custo-benefício das abordagens *Ensemble* - Teste de hipótese de Wilcoxon com 95% de confiança. Os símbolos (+), (-) ou (=) indicam se a comparação com cada abordagem pode ser melhor, pior ou equivalente, respectivamente

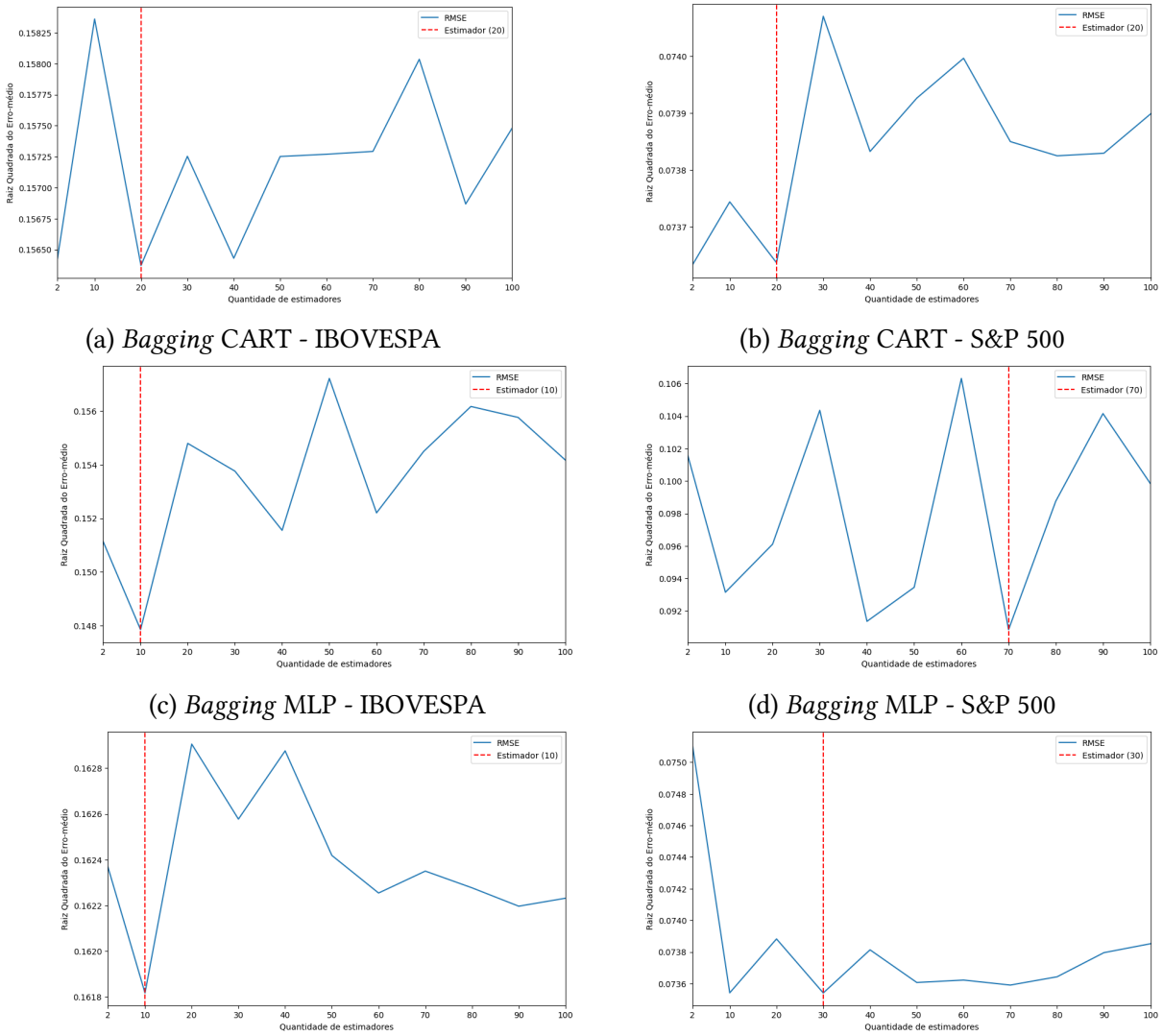
Modelos	***Bagging***			**Heterogêneo**	**Completo**			**Decomposição**			**Sistema Híbrido**			**Otimização**			***Stacking***		
ARIMA	-	-	+	-	-	-	-	-	-	-	=	-	-	=	-	-	-	-	-
CART	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
MLP	-	-	=	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SVR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Bagging CART	=	-	+	-	-	-	-	-	-	-	+	+	+	+	-	+	-	+	-
Bagging MLP	+	=	+	+	-	-	-	+	-	+	+	+	+	+	-	+	+	+	+
Bagging SVR	-	-	=	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Heterogêneo	+	-	+	=	-	-	-	-	-	-	+	+	+	+	-	+	+	+	+
Completo CART	+	+	+	+	=	-	+	+	+	+	+	+	+	+	=	+	+	+	+
Completo MLP	+	+	+	+	-	=	+	+	+	+	+	+	+	+	+	+	+	+	+
Completo SVR	+	+	+	+	-	-	=	+	=	+	+	+	+	+	-	+	+	+	+
Decomposição CART	+	-	+	+	-	-	-	=	-	-	+	+	+	+	-	+	+	+	+
Decomposição MLP	+	+	+	+	=	=	=	+	=	+	+	+	+	+	=	+	+	+	+
Decomposição SVR	+	-	+	+	-	-	-	=	-	=	+	+	+	+	-	+	+	+	+
Sistema Híbrido CART	-	-	+	-	-	-	-	-	-	-	=	-	-	=	-	-	-	-	-
Sistema Híbrido MLP	-	-	+	-	-	-	-	-	-	-	+	=	+	+	-	-	-	+	-
Sistema Híbrido SVR	-	-	+	-	-	-	-	-	-	-	+	-	=	=	-	-	-	=	-
Otimização CART	-	-	+	-	-	-	-	-	-	-	=	-	-	=	-	-	-	-	-
Otimização MLP	+	+	+	+	=	-	+	+	=	+	+	+	+	+	=	+	+	+	+

Table C.4 continued from previous page

Modelos	*Bagging*			Heterogêneo	Completo			Decomposição			Sistema Híbrido			Otimização			*Stacking*		
Otimização SVR	-	-	+	-	-	-	-	-	-	-	+	+	+	+	-	=	-	+	-
Stacking CART	+	-	+	-	-	-	-	-	-	-	+	+	+	+	-	+	=	+	+
Stacking MLP	-	-	+	-	-	-	-	-	-	-	+	-	=	+	-	-	-	=	-
Stacking SVR	+	-	+	-	-	-	-	-	-	-	+	+	+	+	-	+	-	+	=
Total	11	5	**19**	8	0	0	3	6	2	6	16	13	14	15	1	12	9	14	10

Apêndice D

Resultados do Método do Cotovelo



(a) *Bagging* CART - IBOVESPA

(b) *Bagging* CART - S&P 500

(c) *Bagging* MLP - IBOVESPA

(d) *Bagging* MLP - S&P 500

(e) *Bagging* SVR - IBOVESPA

(f) *Bagging* SVR - S&P 500

Figura D.1: Resultados do Método do Cotovelo - IBOVESPA e S&P 500. Fonte: (Autor, 2024)